

# Integrating and Managing HP ProLiant Servers in the Enterprise

CSG18406SG0405





## Integrating and Managing HP ProLiant Servers in the Enterprise

CSG18406SG0405

HP Training

# Student guide

© Copyright 2004 Hewlett-Packard Development Company, L.P.

The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

This is an HP copyrighted work that may not be reproduced without the written permission of HP. You may not use these materials to deliver training to any person outside of your organization without the written permission of HP.

Printed in USA

**Integrating and Managing HP ProLiant Servers in the Enterprise**

Student Guide

May 2004

HP Restricted

## Overview

Introduction .....	1
Course objectives.....	1
Course module overview .....	2
Scenarios .....	2
Learning checks .....	2
Prerequisites .....	3
Prerequisite certifications .....	3
HP Certified Professional program .....	4
Classroom facilities and guidelines .....	5

## Module 1 — Heterogeneous server deployment and integration

Objectives .....	1
Introduction .....	2
Planning the enterprise environment .....	3
Service level agreements.....	4
Planning for availability.....	5
Planning storage capacity.....	7
Choosing the storage configuration .....	7
Data migration .....	7
Planning for performance.....	8
Planning an operating system migration.....	10
Solution Sizers .....	11
HP Services .....	11
Choosing the appropriate ProLiant servers .....	12
ProLiant server family.....	12
Deploying ProLiant servers.....	14
ProLiant server deployment tools .....	15
SmartStart .....	16
SmartStart Scripting Toolkit.....	17
Rapid Deployment Pack .....	18

Deploying multiple servers.....	23
RDP installation methods.....	23
Simple installation .....	24
Custom installation .....	25
Enterprise installation .....	27
Predeployment configuration.....	29
Configuring PXE to process new computers automatically .....	29
Synchronizing the console name with the operating system name .....	30
Changing the primary lookup key to the serial number .....	30
Configuring ProLiant Support Packs.....	30
Creating PXE boot images and diskettes.....	30
Configuring ProLiant BL server blade enclosures .....	31
Remotely installing Deployment Agent for Windows .....	31
Configuring the database .....	32
File sharing in a multiple operating system environment .....	33
Downloading Samba .....	33
Installing Samba using the RPM.....	33
Using Samba .....	34
Deploying a web server.....	35
HP recommended configurations .....	36
Operating system integration in heterogeneous environments.....	38
Integrating Active Directory support.....	40
Planning and resources.....	41
Naming conventions .....	42
Sample format.....	42
Establishing an IP addressing scheme .....	43
DNS considerations.....	45
DNS forwarding.....	46
Tools .....	47
Common problems.....	47
Infrastructure subsystems .....	48
Windows 2000 DNS configuration.....	51
Subdomains and DNS.....	54
DHCP address pools.....	56
Final network topology.....	57
Summary.....	58
Learning check .....	59

## Module 2 — Enterprise management

Objectives .....	1
Introduction .....	2
Managing the enterprise .....	3
Systems Insight Manager architecture .....	5
Web browser interface layer .....	6
CMS layer .....	6
Agent layer .....	7
Systems Insight Manager usage scenarios .....	8
Branch office .....	8
Data center .....	9
Enterprise .....	10
Installing Systems Insight Manager .....	11
CMS requirements for Windows .....	11
CMS requirements for Linux and HP-UX .....	12
Systems Insight Manager database options .....	13
PostgreSQL 7.2 .....	13
MSDE 2000 .....	13
Console browser requirements .....	14
Systems Insight Manager installation tasks .....	15
WMI mapper .....	15
Viewing the Systems Insight Manager Home page .....	16
Customizing the Home page .....	17
Configuring Systems Insight Manager .....	19
Creating and assigning user rights .....	19
Setting up the next discovery .....	20
Browsing discovered devices .....	21
Identity tab .....	21
Links tab .....	22
Events tab .....	22
Setting up and customizing status polling .....	23
Receiving notification of a problem .....	23
Managing numerous devices .....	24
Handling numerous events .....	25

Monitoring managed systems .....	26
System Type Manager .....	26
Rules list .....	27
SNMP categories .....	27
DMI categories (Windows only) .....	28
Managing hosts files .....	29
MIB importing and compiling .....	30
MIB importer CLI .....	30
MIB compiler .....	31
MIB control layer .....	31
MIB configuration .....	32
Tasks .....	33
Task rules .....	34
Predefined tasks .....	35
Task Wizard .....	37
Scheduling a task .....	38
Events .....	40
Displaying events .....	40
Managing events .....	43
SNMP Trap Settings .....	44
Tools .....	45
Application plug-ins .....	46
Creating reports in Systems Insight Manager .....	47
Managing reports .....	48
Creating reports .....	49
Copying reports .....	49
Editing reports .....	49
Deleting reports .....	49
Viewing reports .....	50
Export to file in CSV format .....	50
Viewing SQL queries used to generate a report .....	51
Developing a system software maintenance strategy .....	52
Version Control Repository Manager .....	52
Version Control Repository .....	53
Maintaining a repository .....	54
Discovering software status .....	55
Advanced Search .....	55
Contacting the VCA .....	56
Creating a task to update a managed system .....	58
Running the Update Software and Firmware task .....	59
Updating software and firmware status .....	59
Managing the Systems Insight Manager database .....	60
Database configuration .....	60
CLI configuration .....	60

Using Lights-Out devices to manage remote servers .....	61
Integrating management processors .....	61
Authentication .....	62
Directory Service objects .....	63
Upgrading management processor firmware .....	64
Administering management processors .....	67
Management snap-in installer .....	68
Directory Service administration .....	69
Virtual power .....	70
Virtual Media .....	70
Virtual Floppy .....	70
Remote console .....	71
iLO Remote Console (dual cursor) .....	71
Program Remote Console Hot Keys .....	71
Terminal Services .....	72
Adding new servers to the managed enterprise .....	73
Summary .....	74
Learning check .....	75

## Module 3 — Security

Objectives .....	1
Introduction .....	2
Securing enterprise servers .....	3
Physical .....	4
Network .....	5
Host .....	6
Applications .....	7
Websites .....	7
Email .....	8
Security vulnerabilities .....	9
Common attacks .....	10
Weak passwords .....	10
Application code .....	10
Denial of service .....	11
Data interception .....	12
Open ports .....	12
Information security policies .....	13
Hierarchy of information security policies .....	14
Application and network security policies .....	15
Account policies .....	15
Security patch management .....	16
Auditing .....	17

Configuring security with Systems Insight Manager .....	18
Network security .....	19
Password security .....	20
Program access .....	20
Authentication .....	21
Authentication between the CMS and the managed device .....	21
Authentication between the CMS and applications .....	24
Authentication between the CMS and the database .....	24
Authentication between the CMS and the browser .....	24
Storing credentials .....	25
Installation considerations .....	25
Vulnerability assessment and patching .....	26
Infrastructure .....	27
PVA Interface .....	28
Antivirus strategy .....	29
Summary .....	31
Learning check .....	32

## Module 4 — Performance management

Objectives .....	1
Introduction .....	2
Using the HP troubleshooting methodology .....	3
Step 1 — Collecting data .....	4
Selecting monitoring tools .....	4
Performance Management Pack .....	5
Features .....	6
Competitive tool comparison .....	8
Architecture .....	9
Licensing and license management .....	13
Licensing monitored servers .....	15
Administration .....	16
Cost of ownership and ROI calculator .....	17
Insight Manager 7 console integration .....	18
Proactive performance monitoring .....	19
Logging and logged data manipulation .....	20
PMP usage scenarios .....	23
Data collection, processing, and presentation .....	25
HP OpenView performance tools .....	26
Operating system tools .....	27
Windows monitoring tools .....	27
Red Hat Linux monitoring tools .....	29
NetWare monitoring tools .....	31
Application monitoring tools .....	35
HP integration with third-party tools .....	36
Creating a performance baseline .....	38
Creating a performance baseline using PMP .....	39
Creating a performance baseline using other tools .....	39

Step 2 — Evaluating the data .....	40
Using PMP to detect and analyze performance bottlenecks .....	41
Updating HP Management Agents .....	41
Updating PMP .....	42
Configuring monitoring behavior .....	42
Monitoring overhead .....	43
Performing static analysis .....	44
Detecting and analyzing performance bottlenecks .....	44
Logging and manipulating logged data .....	45
Comparing current performance with baseline performance .....	46
Step 3 — Developing an action plan .....	47
Scaling up performance .....	47
Processors .....	48
Memory .....	49
Disk subsystems .....	51
Host buses .....	55
Network .....	56
Tuning the operating system and applications .....	70
Scaling out performance .....	73
Distributed applications .....	73
Clustering .....	74
Optimizing remote management performance .....	76
Systems Insight Manager client configuration .....	76
Systems Insight Manager database .....	76
Lights-Out graphical remote console .....	77
Remote console hot keys .....	78
Lights-Out device group administration .....	78
Step 4 — Executing the action plan .....	80
Reducing service time .....	80
Reducing usage .....	80
Step 5 — Determining the effectiveness of your actions .....	81
Step 6 — Implementing preventive measures .....	82
Keeping system software current .....	82
Maintaining the performance baseline .....	82
Configuring performance-based alerting .....	83
Using PMP .....	83
Using other tools .....	83
Implementing fault tolerance and failure recovery .....	83
Consolidating workloads .....	84
Models for workload consolidation .....	84
Resource Partitioning Manager .....	85
Summary .....	89
Learning check .....	90

## Module 5 — High availability and clusters

Objectives .....	1
Introduction .....	2
ENSA .....	3
ENSAextended architecture .....	5
ENSAextended capabilities .....	6
DtS technology .....	7
Storage area networks .....	8
SAN benefits .....	9
SAN components .....	10
Server layer .....	11
Fabric layer .....	16
Storage layer .....	19
HP SAN Solution Service offerings .....	22
HP position in the SAN marketplace .....	23
Market opportunities .....	24
HP target markets .....	25
Planning and designing a SAN solution .....	27
Configuring a SAN .....	28
Documenting the SAN .....	28
Designing for availability .....	29
Designing SAN management .....	29
Configuration utilities .....	30
Validating the design .....	31
Using HP sizing and planning tools .....	31
What is a cluster? .....	32
Cluster-aware applications .....	34
Cluster models .....	34
Advantages of clustering .....	35
High availability of resources .....	35
Scalability for growth .....	36
Centralized administration .....	36
Load balancing .....	36
Determining the need for clusters .....	37
Decision points .....	37
HP cluster solutions .....	38
ProLiant cluster family .....	39
Windows clustering technologies .....	40
Network load balancing .....	41
Microsoft Cluster Service .....	43
Choosing network load balancing or MSCS .....	45
Serviceguard for Linux .....	46
LifeKeeper .....	47
ProLiant clusters for NetWare .....	48

Deploying and managing clusters .....	49
Cluster monitoring.....	50
Cluster Monitor .....	52
Configuring Cluster Monitor .....	53
Cluster resources.....	54
Smart Array Multipath software .....	55
Microsoft feature support .....	55
Linux feature support.....	55
Multiple operating system support .....	55
Recovery Server Option.....	56
Cluster technologies .....	58
Distributed Lock Manager .....	58
Non-Uniform Memory Access .....	58
NUMA-CC.....	59
Virtual Interface Architecture.....	59
Remote Direct Memory Access.....	60
HSx technology .....	61
Troubleshooting SAN and cluster problems .....	62
Identifying points of failure .....	63
Troubleshooting shared storage .....	64
HP troubleshooting utilities .....	66
Summary.....	70
Learning check .....	71

## **Module 6 — Business continuity planning and disaster recovery**

Objectives .....	1
Introduction .....	2
Business continuity.....	3
Continuity planning.....	4
Beginning the continuity planning process.....	5
Documentation and standards.....	9
Writing the continuity plan.....	11
Exercising the plan .....	11
HP Services.....	12
HP solutions for fault tolerance.....	13
Thermal management.....	13
Determining the HVAC requirements .....	14
Determining the placement of HVAC units .....	15
Airflow distribution .....	16
Configurations for high-density datacenters.....	18
Power protection .....	20
Extended runtime modules .....	20
Enhanced battery management.....	21
Load sharing .....	22
Unity power rating.....	23

Rack stability.....	24
Leveling jacks.....	24
Ballast kits .....	25
Fixed stabilizer.....	26
Additional guidelines.....	26
HP solutions for disaster recovery — Enterprise Backup Solutions.....	27
Recognizing a suitable environment for EBS .....	27
Performing a needs analysis .....	28
Determining the backup strategy .....	35
Refining your backup and restore process.....	37
Selecting backup devices .....	38
Designing an Enterprise Backup Solution .....	42
EBS drivers.....	44
Supported backup software.....	45
Maximum supported configuration .....	45
OpenView Storage Data Protector .....	47
Key features.....	47
Storage Data Protector topology.....	48
Other HP backup tools .....	50
Business Copy .....	50
StorageWorks Virtual Replicator .....	50
Microsoft Exchange backup/restore solution .....	51
HP management support.....	51
Replication .....	52
NAS Data Copy .....	52
Continuous Access EVA and XP .....	53
StorageWorks Data Replication Manager .....	53
Comparing DRM and Continuous Access.....	54
CASA.....	54
Summary .....	55
Learning check .....	56

## **Learning check answers**

## **Appendix — Documenting the EBS solution**

## **Glossary**

## Introduction

The Integrating and managing HP ProLiant servers in the enterprise course provides in-depth training on key systems technologies, HP technology differentiators in HP X86-based servers, integration of different operating systems, and systems management in an enterprise environment. Discussion topics include HP server architecture, industry technologies, and high-availability technologies. Hands-on activities include system configuration using advanced utilities and HP Systems Insight Manager, performance management, and deploying two-node clusters.

## Course objectives

After completing this course, you should be able to:

- Plan an enterprise environment
- Deploy multiple servers using HP ProLiant Essentials Rapid Deployment Pack (RDP)
- Integrate a multiple operating system environment
- Use HP Systems Insight Manager to manage network systems
- Use HP best practices to secure servers in an enterprise environment
- Develop an information security policy that outlines requirements, roles, and responsibilities to determine access to network devices and applications
- Detect and analyze performance bottlenecks
- Plan and design an HP storage area network (SAN)
- Identify HP cluster solutions
- Identify HP solutions for fault tolerance and disaster recovery

## Course module overview

This course comprises the following modules:

- Module 1 — Heterogeneous server deployment and integration
- Module 2 — Enterprise management
- Module 3 — Security
- Module 4 — Performance management
- Module 5 — High availability and clusters
- Module 6 — Business continuity planning and disaster recovery

In addition, there is a glossary of terms and acronyms in the appendix in this Student Guide. Although acronyms are spelled out throughout the Student Guide the first time they are mentioned in each module, the glossary provides a convenient reference.

### Scenarios

Each module begins with an introduction that presents a hypothetical scenario. These sample scenarios are designed to help you apply HP technology to a production environment. They are not designed to limit you to specific hardware, software, or any other supported option.

In the introduction to each module, you will read of the growth of RC Engineering, a fictional company that has grown from the small and medium business level in the HP Accredited Integration Specialist (AIS) course to enterprise level in this course. You will follow Bob, the CEO of RC Engineering, as his business requirements expand and help him and his staff make smart decisions regarding his IT infrastructure.

### Learning checks

Each module ends with questions that review the material covered in the module. These questions, together with the Exam Preparation Guide that is published with this course, are designed to prepare you for the certification exam. Answers to the questions are provided in the back of this Student Guide.

## Prerequisites

This course is an Accredited Systems Engineer (ASE) level course. It is intended to teach you to design, support, and integrate platform, operating system, storage, network, and option components and solutions at the enterprise level.

This course is designed for students who have the required certifications or equivalent knowledge and experience. It is directed toward both a sales and a non-sales audience.

The Student Guide and Lab Guide designed for this course, combined with other information you receive from HP, will help you prepare for the ASE certification exam.

## Prerequisite certifications

HP recommends and the instructor assumes that before taking this course, you will have successfully completed *Implementing HP ProLiant servers*, the HP AIS course. In addition, you must have attained at least one of the following operating system vendor administrator certifications:

- Microsoft Certified Systems Engineer (MCSE) certification
- Red Hat Certified Engineer (RHCE) certification
- SAIR Linux Certified Engineer (LCE) certification
- Certified Novell Engineer (CNE) certification



### Important

This course builds on knowledge gained in the prerequisite certifications and training. If you do not meet the prerequisites, this course can be extremely difficult or impossible to complete. The course is written, and will be taught, as though you have met the prerequisites.

---

## HP Certified Professional Program

The HP Certified Professional Program is a world-class certification program recognized as the benchmark in training excellence. This program provides a globally consistent framework that significantly enhances our overall capability to sell and support HP technologies, products, and solutions. It achieves this by aligning our technical capabilities—from Pre-Sales architecture and solution design to warranty and after-sales support—to our marketing, sales, and services strategies.

HP certification is beneficial because:

- Certification provides access to higher-level technical support.
- Certification demonstrates your knowledge and ability to support ProLiant technology.
- Certification makes you part of a global sales and technical community with access to benefits such as HP tools and utilities, priority hotline access, training and event opportunities, and much more.
- The dominance of HP in the server market compels resellers to be knowledgeable about HP products to garner support revenue.

Customer-employed Accredited Systems Engineers (ASEs) are not only highly involved in implementation and support for the hardware, they are also often involved in setting the strategic IT directions for their company.

The program makes a significant contribution in delivering the benefits customers typically look for in an IT solution:

- Customer satisfaction
- Lowest total cost of ownership
- Managed risk
- Certified competence when delivering projects
- Confidence of a successful project outcome

## Classroom facilities

The instructor will give you detailed information concerning:

- Location of restrooms
- Class hours
  - Class start time
  - Scheduled breaks
  - Class stop time

## Classroom guidelines

Use the following guidelines:

- Do not interfere with other students' learning.
  - Be on time for class.
  - Turn all mobile phones and pagers to *off* or silent setting.
  - Be professional in your speech and actions.
  - Do not change or modify lab equipment, passwords, or software configurations unless told to do so by the instructor.
- Do not smoke in the classroom.



### **Important**

You may be removed from the classroom and not allowed to return if you fail to follow the classroom guidelines.

---



---

# Heterogeneous server deployment and integration

## Module 1

### Objectives

After completing this module, you should be able to:

- Plan the IT infrastructure for an enterprise environment
- Choose the HP ProLiant server appropriate for a given enterprise environment
- Deploy multiple ProLiant servers using the HP ProLiant Essentials Rapid Deployment Pack (RDP)
- Install Samba for file sharing in a multiple operating system environment
- Integrate an operating system in a heterogeneous environment
- Integrate Active Directory support

## Introduction

RC Engineering started out several years ago as a small company with only ten employees and no network infrastructure. In the past several years, the company has been awarded multiple government contracts and has more than doubled in size.

When they acquired an engineering company in Silicon Valley a few years ago, RC Engineering centralized the management of its Microsoft Windows, Novell NetWare, and Linux servers, and implemented remote management using HP Remote Insight Lights-Out Edition (RILOE) II technology. Now RC Engineering is expanding into the e-commerce arena and Bob, the CEO of the company, wants to be sure he has the infrastructure he needs.

The Houston headquarters presently employs 500 people, including a small IT staff. Yesterday Bob invited you to meet with him and Jackie, his new network administrator. Bob needs your advice for determining exactly which HP solutions will meet his business requirements. Bob understands that proper planning will minimize downstream problems, ensure long-term reliability, and maximize performance. Now that RC Engineering has grown to the enterprise level, you will also need to educate Bob and Jackie about HP best practices for an enterprise environment.

One of Bob's main concerns is that there is not enough floor space at the company's headquarters to house an e-commerce system. Bob would like you to tell him how he can fit all the equipment he needs into the space he has. You can take this opportunity to familiarize Bob with the HP family of ProLiant servers, including the BL server blades, and how to deploy them using the HP ProLiant Essentials RDP.

In addition, RC Engineering is rolling out a document management system in both locations, with servers at each location to support the new system. As a part of the document management system, the Linux server must be able to share information with the other servers. Jackie, the IT administrator, wants to implement file sharing using Samba.

Bob, Jackie, and the IT staff have scheduled a meeting to plan the rollout of the document management system. On the agenda is a conversation about the deployment tools they should use to facilitate the multiple server rollout. Your job is to prepare them to plan the rollout intelligently.

## Planning the enterprise environment

In IT terms, an enterprise is any large business organization that uses computer systems as an essential part of its business process. In practice, the term *enterprise* refers to a large corporation using shared files across a private and public network.

Making the transition from a small or medium business to an enterprise is often a process of real-time evolution, with little time for planning. Proper planning, however, can provide a key to success by reducing design problems, speeding time to production, and laying the groundwork for long-term reliability and high performance. Enterprise planning, where effective performance management begins, combines people, processes, and technology in preparation for an infrastructure that will perform optimally over time.

The first step in planning the enterprise environment is to perform a needs analysis to determine the customer's current environment and business requirements, availability of resources, and future plans. The next step is to conduct a site survey to assess the customer's facility and its suitability for the proposed infrastructure. Based on this information, you can begin to make recommendations for new equipment and configurations and for modifications to existing systems.

As part of your plan to install a new system or change an existing system, you should:

- Learn as much as possible about the organization and the physical arrangement, features, capabilities, and configurations of the appropriate servers, storage systems, and optional components.
- Decide what equipment and supporting software must be purchased.
- Determine which components require changes to standard procedures for online tasks.
- Check whether a set of completed installation forms describing the current system exists.

- Plan for the environmental requirements of the system. This involves:
  - Selecting the rooms that will contain system equipment
  - Arranging for the installation of a raised floor, if desired
  - Installing any necessary air conditioning and other environmental controls
- Plan for the power requirements of the system equipment. Make arrangements for power installation early in the planning process because it can take a long time to get new electrical circuits installed.
- Arrange for telecommunication lines. As with power installation, you should make arrangements to install telecommunications lines early in the planning process because it can take a long time.
- Prepare the delivery route from the building entrance to the installation site.

## Service level agreements

Enterprise environments are designed to meet certain service level agreements (SLAs) between the IT department and the infrastructure users. SLAs specify in measurable terms what the service consists of and how it is measured, justified, and maintained. SLAs typically prescribe the:

- Percentage of time the service is available
- Number of concurrent users
- Performance baseline for the service
- Maintenance schedule and notification of reduced SLA
- Help desk response time for problem resolution

The challenge when planning an enterprise environment is to define the architecture and determine its maximum levels of availability. An SLA must be designed with the infrastructure in mind.

## Planning for availability

HP believes that the measurement of availability should be from the user's perspective. Simply recording that a certain hardware or software component is operating is not enough; you must also consider the user's ability to access the service, the quality of the service provided, and the acceptability of the response time to the user.

Although major changes—such as installing a new operating system—obviously affect availability, the effect of other types of changes might be less apparent. For example, changing the characteristics of a communications line could cause response time to become unacceptable to a user who is trying to use that line to access a file on a remote system.

By taking the time to anticipate and plan for changes, you can avoid taking your system down for unnecessary planned outages. You can take action now to prevent planned outages in the future in several areas.

- **System performance and growth** — Evaluating system performance and growth involves tracking and anticipating growth and then establishing plans to accommodate that growth.

Task	Definition	How it helps
Application sizing	Using models to determine how well new applications will handle their intended workloads	Helps you plan for growth in system workloads caused by new applications
Capacity planning	Forecasting future capacity needs based on performance trends and the growth in users, applications, and the company's business	Helps you plan for growth in system workloads based on business growth
Performance analysis and tuning	Measuring system performance and acting on the results of the measurements	Improves system performance and availability
Usage accounting	Tracking the use of system resources for accounting purposes	Helps you plan for system growth if user activity is known in sufficient detail

- **Computer room resources** — Some changes require more power and air conditioning. You can avoid unnecessary downtime by ensuring that you have enough physical space, power capacity, and cooling capacity for additional equipment.

- **Future changes** — Most changes can be performed while a system is operational (online). Some changes that require the system to be offline can be performed online if you have configured your system to allow room for growth.
- **Formal change-control processes to manage change** — Change control is the process for proposing, planning, implementing, and testing change and is a key requirement for minimizing the duration of planned outages. Change control ensures the successful migration of a system or application from one stable configuration to another by:
  - Ensuring that the scope and ramifications of the change are fully understood
  - Providing a recovery plan
  - Ensuring that problems and errors are anticipated and reacted to appropriately
  - Maintaining the security of your system and applications

## Planning storage capacity

Capacity planning helps you calculate how much hardware is necessary to handle the demand on the system. To plan for storage capacity, you must consider an interconnected set of processes, components, and management applications.

It can also help you identify the components that are causing performance degradation. By adding hardware or by reconfiguring processes and systems, you can resolve these weak links.

## Choosing the storage configuration

One of the first steps in planning the storage system is to choose what type of storage configuration is dictated by the data. HP provides an overall architecture and a growing portfolio of hardware, software, and services to meet the capacity requirements of any enterprise environment.

Storage systems can be configured as follows:

- **Direct attached storage (DAS)** — Ideal for smaller deployments, such as small data centers and remote office locations, these affordable solutions can scale to a networked storage environment.
- **Network attached storage (NAS)** — HP StorageWorks NAS solutions use HP ProLiant servers running Microsoft Windows Storage Server 2003 to deliver file and print capabilities, enhanced network manageability, and accelerated data availability.
- **Storage area network (SAN)** — In a SAN, multiple compute servers and backup servers can access a common storage pool. Servers can be added and removed from a SAN while data remains in the SAN. Multiple servers can access the same storage for more consistent and rapid processing. The storage itself can be easily increased, changed, or reassigned.

## Data migration

The DAS to SAN (DtS) architecture is an exclusive HP feature that provides a quick and easy way to migrate stored data protected by Smart Array controllers to an HP StorageWorks Modular Smart Array (MSA) 1000 storage system.

By removing the drives from the older systems and inserting them into the MSA1000, existing data, RAID sets, and configuration information remain intact, allowing data migration to be completed in minutes, not hours.

---

**INTERNET**

For more information on DtS migration, refer to the *DtS Data Migration to the MSA1000* white paper, which is available for Microsoft Windows, Novell NetWare, and Linux environments. You can download the appropriate white paper from:  
<http://h18006.www1.hp.com/storage/arraywhitepapers.html>

---

## Planning for performance

Planning for performance involves building a server configuration that best meets the performance requirements of the computing environment. It includes these steps:

1. Identify the application environment and its performance requirements. For example, for a database server, determine whether the server will be used primarily for online transaction processing or decision support.
2. Determine the application resource utilization characteristics. For example, the new database server processor long-term utilization should be more than 85% but should not exceed 95%.
3. Match the application resource utilization with the type and number of hardware resources needed to match the resource utilization requirements. For example, ensure that the online transaction processing database server has a sufficient number of processors to meet given performance and utilization requirements.
4. Allow for performance fluctuations and future growth in terms of capacity, performance, and upgradeability. For example, if the previous steps yield a five-processor configuration at 90% utilization, plan for a six-processor server or a server capable of supporting more than five processors.
5. Match the required hardware resources with a server configuration that best meets the requirements.

Often, a part of the performance planning process involves deciding whether the future expansion will occur inside the server (scaling up) or outside the server (scaling out). The HP virtualization strategy enables dynamic allocation of computing resources, such as servers, storage, network, or applications, based on demand and computing needs. Deciding early on your approach will save you time and money in the long run.

As a result of expansion, mergers, and acquisitions, companies can experience rapid growth that negatively impacts system performance. Neglecting performance considerations until after the system begins to show signs of poor performance often only aggravate the problem.

When planning for performance, pose what-if scenarios to compare the cost, performance, and risks involved with implementing one option or another. Ask question such as:

- What is the maximum workload increase that the current configuration can handle?
- Would a new computer purchase improve performance? How many new servers should be purchased? Which models?
- How much impact would an application software upgrade have?
- Would more software licenses for existing applications be beneficial?
- How would overall performance change if more computers were added at the same time that the network was upgraded?
- How would present system performance be affected by unplanned downtime?

## Planning an operating system migration

Changing the operating system platform involves planning for performance, capacity, and data and application availability, because when you migrate to another operating system, often you deploy a different server. As a result, you must plan in advance for all three major planning areas.

In addition, planning a migration or new implementation of operating systems and platforms requires proactive change management plans, timely communication, and preparation for contingencies. These change plans must be accurate, disciplined, and tuned to the specific service levels required by your business.

The first step in planning a platform change is to create a profile that identifies the current operating system and platform environment, as well as the change objectives. Document the plan and recommend needed change management processes that include (but are not limited to) hardware, applications, and data.

To plan an operating system migration, follow these steps:

1. Understand your current operating system and platform environment, identify your objectives and business needs, and create a customer profile.
2. Match your requirements to a specific migration service and verify that your system environment meets the specific service prerequisites.
3. Develop and document the migration target and provide a viable migration checklist.
4. Identify the change management process that includes but is not limited to identifying needed changes, the people affected, required training, as well as hardware, data, and application concerns. This process also includes defining the migration team, risks to be resolved, and contingency plans.
5. Deliver the migration plan and conduct a review with your IT management staff.
6. Present a summary report on the migration plan, as well as an executive summary of the engagement, and recommendations for additional HP products and services that can benefit the migration effort.

## Solution Sizers

Solution Sizers are automated tools for sizing new IT environments and planning for future growth. These tools cover a range of applications, from simple to complex, from an easy to use Windows-based sizing tool to an enterprise-level capacity planner. The sizing information and algorithms have been developed using testing and performance data on a wide range of HP servers running solutions from HP partners such as Citrix, Lotus, Microsoft, and Oracle.

Solution Sizers take into consideration resource availability, recovery, capacity, and performance when determining the optimal solution for a given application environment. They take the guesswork out of planning and determining the best configuration that meets the customer requirements. Whenever possible, you should use an appropriate solution sizer before the actual deployment to:

- Determine what is the best server and its configuration
- Validate a given configuration against a known set of requirements

---

**INTERNET** For a list of HP ProLiant sizing tools that help match computing resources with business requirements, visit: <http://activeanswers.compaq.com/>

---

## HP Services

HP Services can help determine the enterprise architecture that is right for a given environment. HP Services provides a range of services, including operating system and platform migration, consulting services, and customized, turnkey rack solutions with preloaded software.

---

**INTERNET** For more information, visit: <http://www.hp.com/go/4service>

---

The HP Global Method for IT Strategy and Architecture helps create an overarching roadmap that facilitates investment priorities, technology choices, and the evolution of existing enterprise applications and infrastructure. HP Services consultants guide key solution stakeholders through the architecture process, looking at the chosen solution from a business, functional, technical, and implementation view.

---

**INTERNET** HP partners with Cognos, Inc. to offer the Cognos Enterprise Planning Series. This application provides a solution for planning, budgeting, forecasting, consolidating, and financial reporting on HP ProLiant servers on a Microsoft Windows platform. For more information, visit:  
[http://h21007.www2.hp.com/dspp/mop/mop\\_partner\\_product\\_detail\\_IDX/1,1331,1162,00.html](http://h21007.www2.hp.com/dspp/mop/mop_partner_product_detail_IDX/1,1331,1162,00.html)

---

## Choosing the appropriate ProLiant servers

Visit the HP website for information about ProLiant servers. For enterprise-class ProLiant servers, click the *Large Enterprise Business* link; then click *Servers*. Click one of the ProLiant server lines under *HP ProLiant and TC series servers* to view information on:

- ProLiant ML and DL servers
- ProLiant server blades
- ProLiant and tc entry-level servers
- Accessories and options

For each server line, you can find specific product information including QuickSpecs, benchmarks, and software and driver support. You can also click links for related information such as technical support, management, and solutions.

### ProLiant server family



ProLiant BL40p servers

The ProLiant family of ML, DL, and BL servers is ideal for enterprise environments of all sizes.

ProLiant ML servers feature tower and rack deployment options. The ProLiant ML servers are optimized for:

- Maximum internal storage and I/O flexibility
- Remote and branch offices and data centers

ProLiant DL servers feature a design that combines high server density and high levels of expansion. ProLiant DL servers are:

- Density-optimized for flexibility and manageability
- Ideal for multiserver deployments

ProLiant BL servers provide the ideal solution for an enterprise requiring maximum server density and minimal power consumption. HP developed the ProLiant BL line of modular server blades to address the need for:

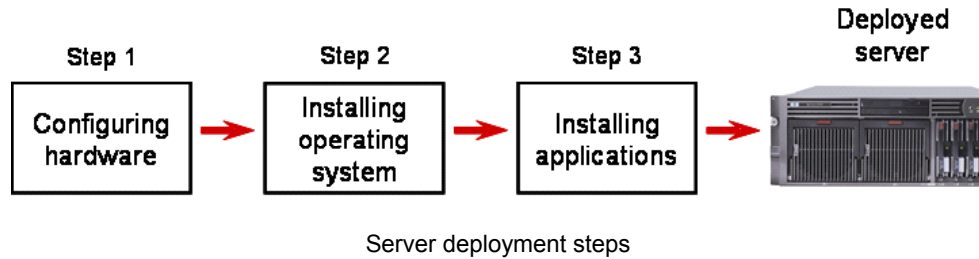
- Increased density
- Rapid deployment and server software provisioning
- Remote manageability
- Centralized storage
- Effective price:performance
- Industry-standard compatibility

Server blades enable horizontally scaled server virtualization environments, where large server farms can be deployed and managed to support a dynamic computing infrastructure such as:

- Distributed application servers
- Clustered databases
- Load balancers
- Multiple firewalls
- Caching and web servers

Server blades deliver sustained performance in scalable environments, enabling customers to use resources more efficiently to accommodate scale-out growth.

## Deploying ProLiant servers



The deployment of ProLiant servers includes three basic steps:

1. **Configuring hardware** — Before you can install an operating system or application software on the server, you must configure the hardware. RDP provides the tools and scripts to automate server hardware configuration using the SmartStart Scripting Toolkit.
2. **Installing the operating system** — After the hardware is configured, you can install the operating system on the server using either of the following methods:
  - **Scripted installation** — Required for first-time installation
  - **Imaging** — Used to replicate the configuration of one server onto other servers
3. **Installing applications** — After the operating system installation is complete, you can install additional applications on the server using either of the following methods:
  - Scripted installation
  - RapidInstall Packages (RIPs)

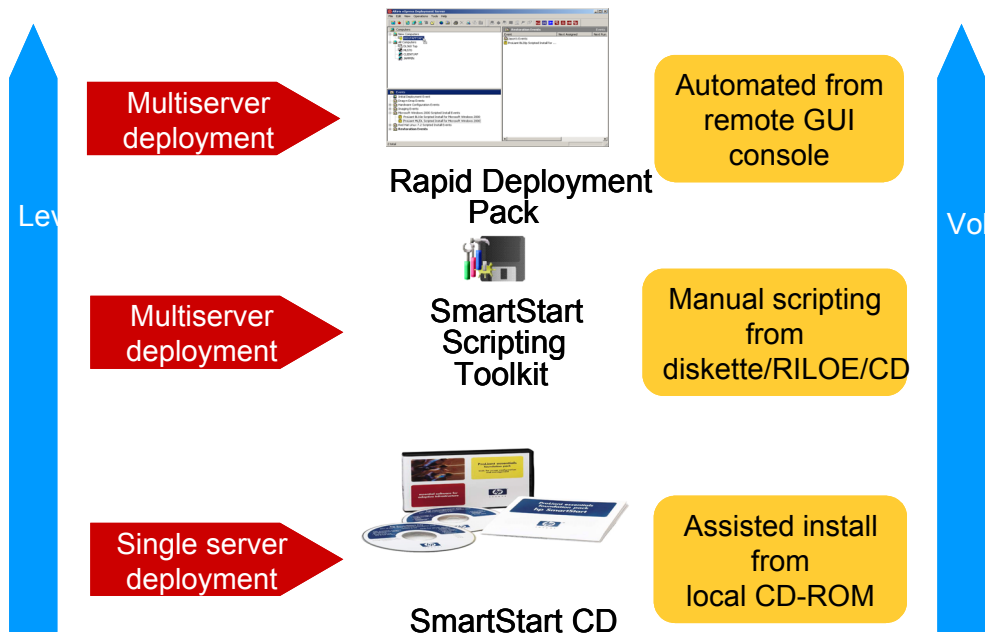
---

**Note**

A deployment job can contain both scripting and imaging tasks.

---

## ProLiant server deployment tools



HP provides several deployment solutions for ProLiant servers, including:

- SmartStart
- SmartStart Scripting Toolkit
- RDP

Each of these solutions is powerful, easy to use, and designed to meet specific customer requirements.

SmartStart provides reliable and consistent server configurations with functionality for integrating operating system installations to achieve optimum reliability and performance. SmartStart is the best solution for simplified single-server deployment. For deployments requiring remote or network-type installations, use RDP or the SmartStart Scripting Toolkit, which is a component of RDP.

## SmartStart

The SmartStart process is easy to use and is ideal for deployments of a small number of servers. SmartStart requires user attention throughout the process and can be performed on only one server at a time.

The SmartStart CD:

- Is delivered in the HP ProLiant Essentials Pack
- Is an ideal solution for single-server deployment
- Supports all HP ProLiant ML and DL servers
- Is helpful for both novice and advanced users

SmartStart 5.5 and earlier versions include assisted, manual, and replicated installation paths. SmartStart 6.00 and later offers only the assisted installation path. The assisted installation prepares the server hard drives by:

- Erasing the drives
- Creating a boot partition
- Preparing for the file system
- Installing server support software, including the HP ProLiant Support Pack (PSP)

For manual installations, use vendor-supplied operating system media and the HP ROM-based utilities:

- ROM-Based Setup Utility (RBSU) to configure the hardware
- Option ROM Configuration for Arrays (ORCA) to set up the hard drives and array controller

Server systems that support RBSU and ORCA feature maintenance utilities and automatic configuration operations that enable you to install the operating system from the operating system CD and then manually install server support software from the SmartStart CD.

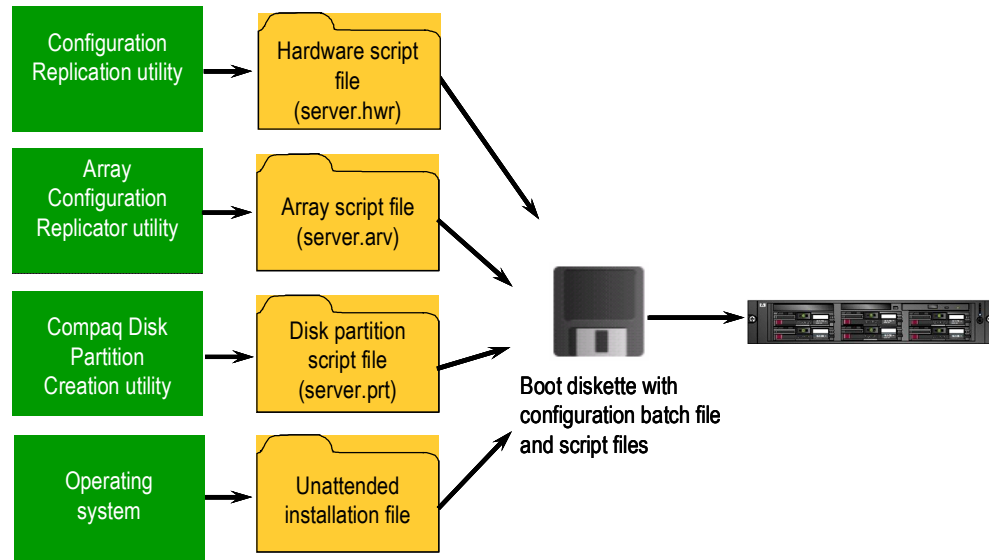
---

### INTERNET

For a complete list of ProLiant servers and the versions of SmartStart they support, visit: [http://h20000.www2.hp.com/bc/docs/support/UCR/SupportManual/TPM\\_na\\_rev1\\_us/TPM\\_na\\_rev1\\_us.pdf](http://h20000.www2.hp.com/bc/docs/support/UCR/SupportManual/TPM_na_rev1_us/TPM_na_rev1_us.pdf)

---

## SmartStart Scripting Toolkit



The SmartStart Scripting Toolkit uses a combination of DOS utilities and batch files to automate operating system installation. It is an excellent solution for large server deployments. However, it requires a high degree of user knowledge to configure. Additionally, maintaining installation batch files for multiple server types requires detailed scripting knowledge.

The SmartStart Scripting Toolkit:

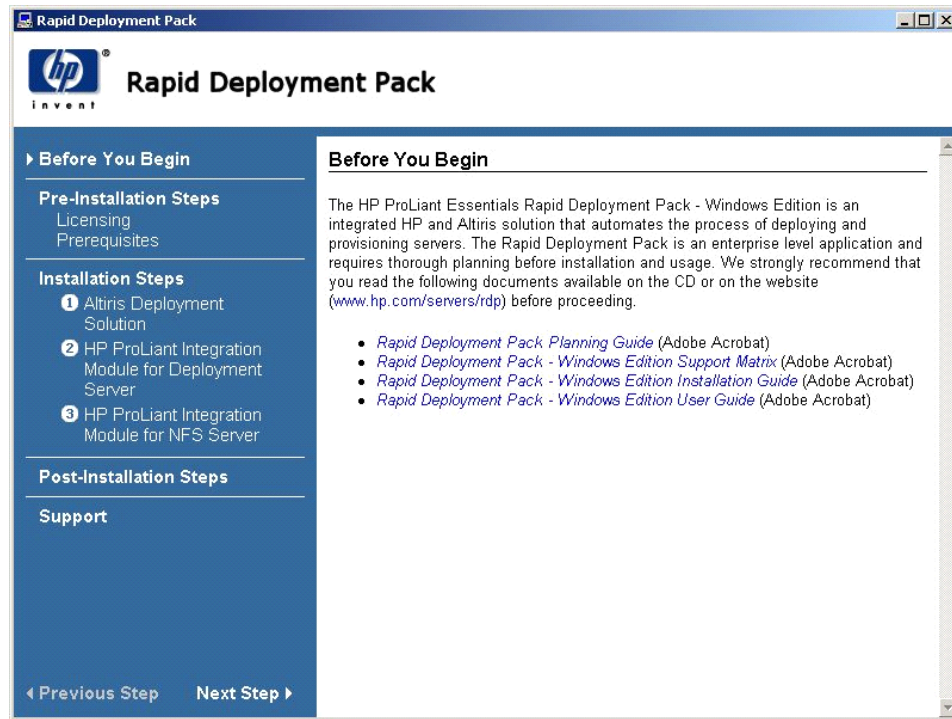
- Is delivered only through web download
- Provides advanced user setup and novice user deployment
- Supports all HP ProLiant DL and ML servers and selected legacy server models
- Performs an unattended install
- Can perform remote installations with the Remote Insight Lights-Out Edition (RILOE) Virtual Floppy (VFLOP)

### INTERNET

For newer servers, use the SmartStart Scripting Toolkit 2.2; for older systems, use 1.7. Refer to these websites for the respective versions of the SmartStart Scripting Toolkit:

- SmartStart Scripting Toolkit 2.2  
<http://h18004.www1.hp.com/products/servers/management/toolkit/index.html>
- SmartStart Scripting Toolkit 1.7  
<http://h18004.www1.hp.com/manage/toolkit.html>

## Rapid Deployment Pack



RDP autorun home page

RDP is an integrated HP and Altiris solution that automates the process of deploying and provisioning server software in Windows and Linux environments. It enables unattended, headless installations for intermediate user setup and novice deployment. It consists of two main components:

- The **HP ProLiant Integration Module** enables you to deploy ProLiant servers easily, without spending weeks developing your own infrastructure. It contains:
  - Sample configuration events
  - Batch files
  - Support software
- The **Altiris Deployment Server** enables you to store complete on-site server images. If a machine fails, you can reimage and restart it in minutes using the Deployment Server image. Deployment Server encapsulates all the steps for configuring and deploying ProLiant servers in easy-to-use configuration jobs that you can drag and drop to deploy one or many servers. The Deployment Server software also provides remote console capabilities for imaging and scripting. It integrates with RILOE and ProLiant BL servers.

RDP enables unattended, headless installations for intermediate user setup and novice deployment. It supports ProLiant DL, ML, and BL servers and is delivered in HP ProLiant BL kits. RDP is available for both Windows and Linux environments.

## Benefits of RDP

With RDP you no longer have to individually set up and configure every server that you deploy. Although using SmartStart is easy and effective for single server deployments, RDP enables large-scale deployments without the need for extensive manual scripting. The SmartStart Scripting Toolkit and Deployment Solution provide the tools and methods to enable efficient, large-scale deployments with minimal development.

The Deployment Server console, combined with the ProLiant Integration Module, gives even a novice administrator extensive, flexible control to manage changes. Servers can be deployed and upgrades such as applications or service packs can be pushed across the network from a single console. Deployments and changes are made quickly and easily, lowering the total cost of ownership.

RDP enables you to:

- **Deploy service packs** — Using RDP conditional events enables you to deploy executables and run scripts easily to patch thousands of servers in minutes.
- **Enforce software baselines with RDP** — RDP helps achieve software baseline enforcement, keeping unwanted software off critical servers and ensuring that all software current.
- **Deploy applications with RDP** — You can use RDP to deploy various applications. Options include:
  - Installing SQL Server, virus scanners, and so on
  - Installing to many servers simultaneously
  - Deploying application farms quickly

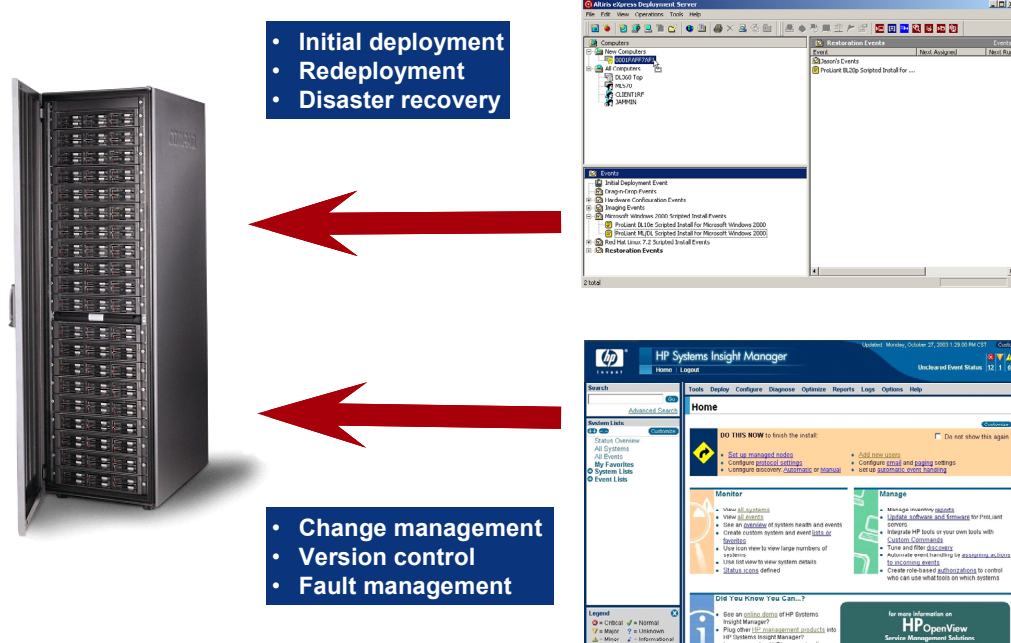


### Important

To optimize and scale RDP components, HP recommends that you limit the number of open consoles to avoid an adverse affect on management performance working with large numbers of clients. You should also use one Deployment Server per site or network segment to optimize bandwidth, Preboot eXecution Environment (PXE) connections, and other limiting factors.

---

## RDP and Systems Insight Manager



When implemented together, RDP and HP Systems Insight Manager provide you with a complete deployment and change management solution.

Systems Insight Manager extends Insight Manager 7 features with capabilities for managing ProLiant servers and server blades, tracking inventory and assets, and controlling device discovery through discovery filters. Systems Insight Manager simplifies server blade management by providing graphical representations of ProLiant BL e-Class server blades and their location within server blade enclosures and racks.

Systems Insight Manager is fully integrated with HP Lights-Out technology to provide in-depth fault, configuration, and performance monitoring from a single management console.

**Important**

Configuration changes and software updates made through Systems Insight Manager are **not** recorded in the history within the Deployment Server application and therefore cannot be re-created through rip-and-replace. For changes to be recorded, you must make them through the Deployment Server.

---

**INTERNET**

---

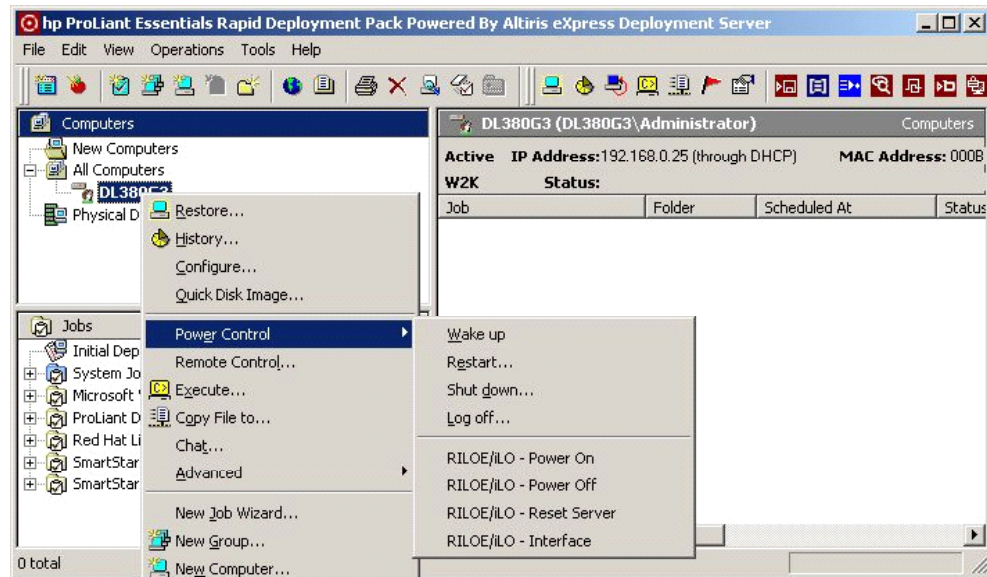
For additional information on Systems Insight Manager, visit:

**<http://h18013.www1.hp.com/products/servers/management/hpsim/index.html>**

---

RDP and Systems Insight Manager can share a database service (Microsoft Data Engine [MSDE], Microsoft SQL Server 7.0, or Microsoft SQL Server 2000); however, they must use the most current supported database version for both. Each service uses unique tables, so no conflict results. However, in a large network with many managed servers, it is best to install the services on separate servers to avoid placing additional network traffic and processor usage on a single server.

## Integration with Lights-Out management



When HP Integrated Lights-Out (iLO) or RILOE is installed in the server, you can manage remote servers and perform remote console operations regardless of the state of the operating system or hardware on the unit that you are managing.

The Deployment Server software uses the power management feature of the RILOE board to access the target computer to:

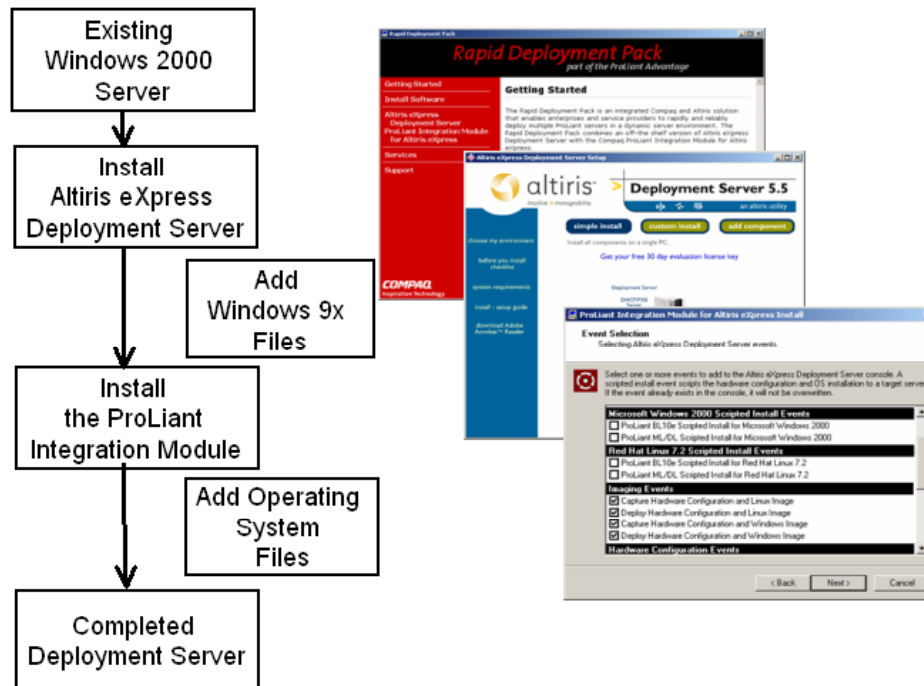
- Power on the server
- Power off the server
- Reset the server
- Access the RILOE interface

Each time a computer connects to Deployment Server, the software polls the computer to see if a RILOE board is installed. If it is, Deployment Server gathers information about the board, including the Domain Name Service (DNS) name, IP address, and first user name. To maintain security, Deployment Server requires the user to enter the correct password for that user name.

## Deploying multiple servers

Server deployment can be complicated and time-consuming, especially if you need to deploy hundreds of servers quickly and reliably. RDP facilitates the installation, configuration, and deployment of high-volumes of servers through a graphical user interface (GUI)-based console using either scripting or imaging technology. RDP reduces server configuration time, making it possible to scale server deployments to high volumes quickly.

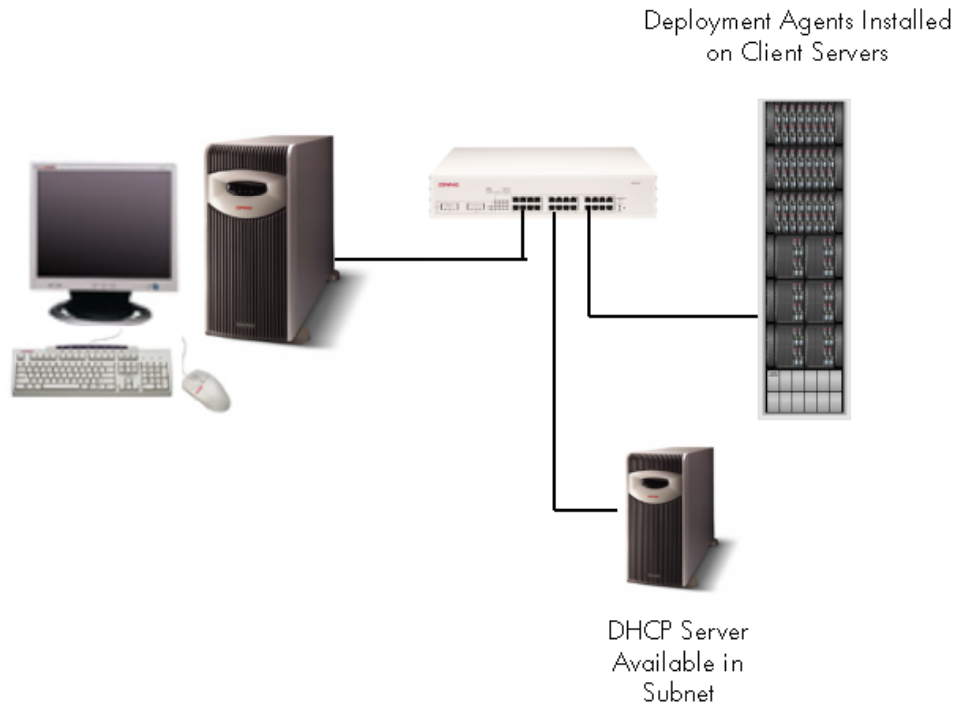
### RDP installation methods



The three main configuration and installation methods for RDP are:

- Simple
- Custom
- Enterprise or distributed

## Simple installation



HP recommends the simple installation method for a first-time installation on a simple infrastructure involving a group of servers on a single subnet. This configuration method places the following components on a single deployment server:

- Deployment Server software
- Deployment Server console
- Deployment Server database
- File server
- PXE Server
- SQL Server
- Windows services

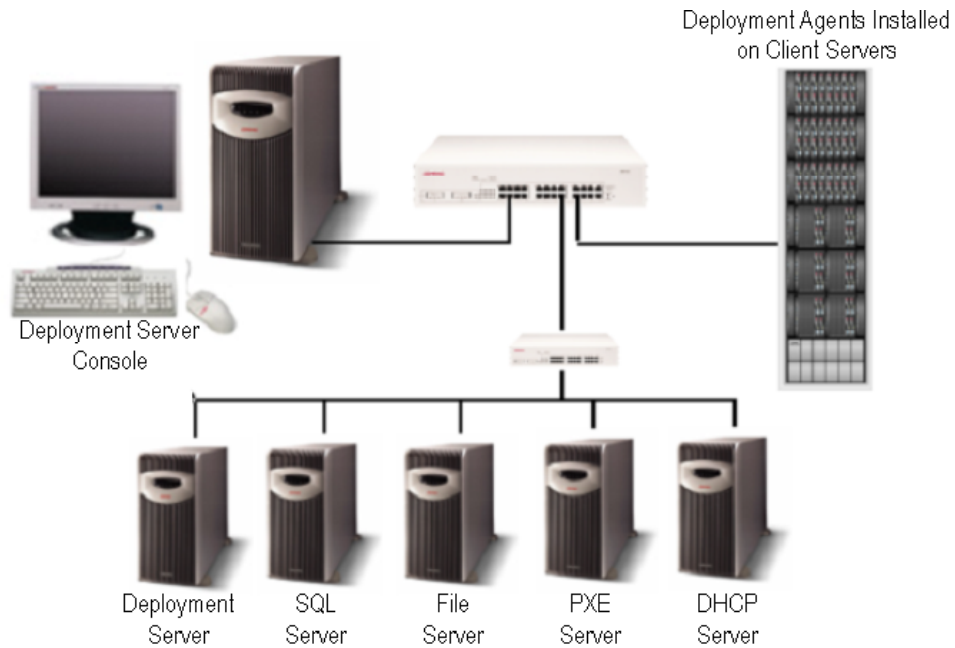
---

### **Note**

The Deployment Server web console installs automatically during a simple install (and during a silent install) if the Microsoft Internet Information Server (IIS) services are running on the selected computer.

---

## Custom installation

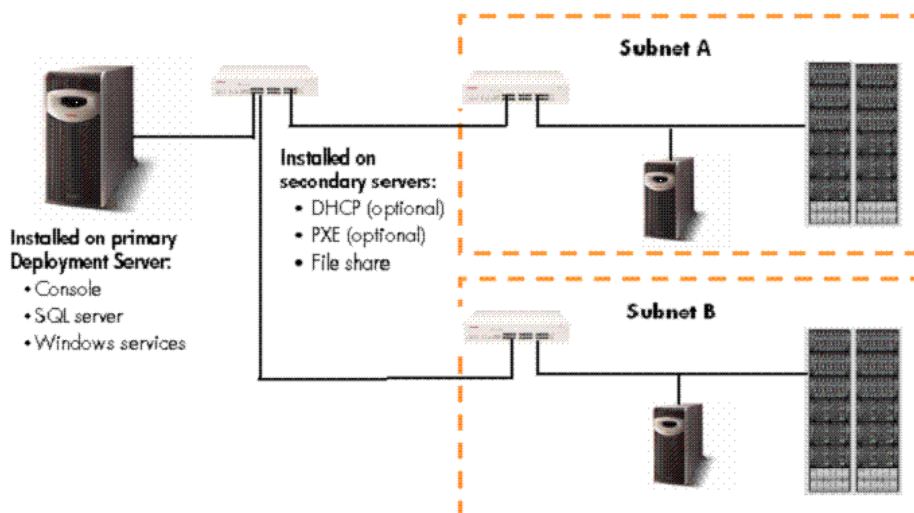


The custom installation method allows you to choose the server name and path for each component of the Deployment Solution suite. This method is more complex than a simple installation, but is required if you want to modify your configuration. The custom install method is recommended when a company has multiple subnets but would prefer to have a single deployment server managing the deployment activities.

To use this method, follow these steps:

1. Allocate a server to be the deployment server.
2. Install RDP on the deployment server.
3. Using the custom install method of the Deployment Server installation wizard, install the PXE Server component on the Dynamic Host Configuration Protocol (DHCP) server in each subnet.
4. Install the remaining components on the deployment server:
  - Deployment Server software
  - Deployment Server console
  - Deployment Server database
  - File server

## Adding additional components



When performing a custom installation, you can install an additional PXE Server, Deployment Server console, or Client Access Point (file server) after your initial installation is complete. You can install the additional components using the Add Component feature of the installation wizard.

### INTERNET

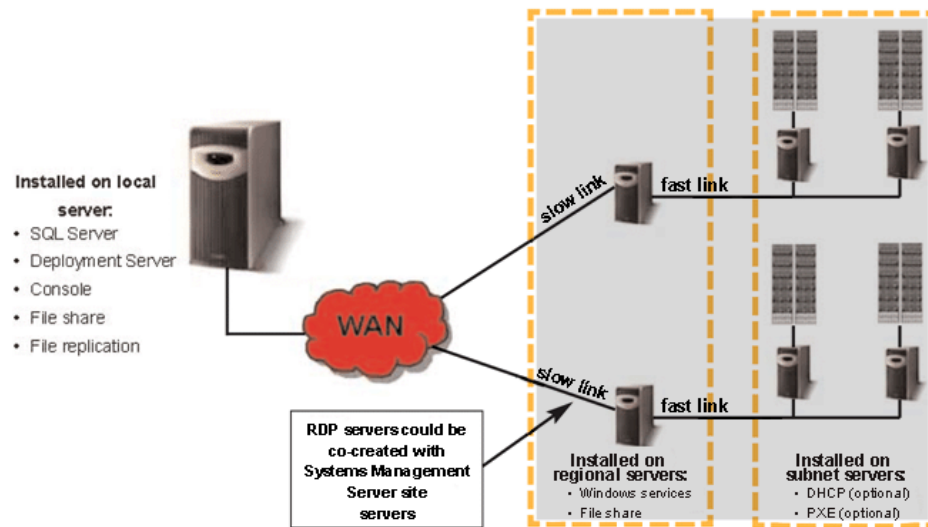
For more information about performing a custom installation, see the *Deployment Server User Guide* at: <http://h18013.www1.hp.com/products/servers/management/rdp/documentation.html>

## Comparing a simple and custom installation

Use the following table as a guide for selecting an installation method.

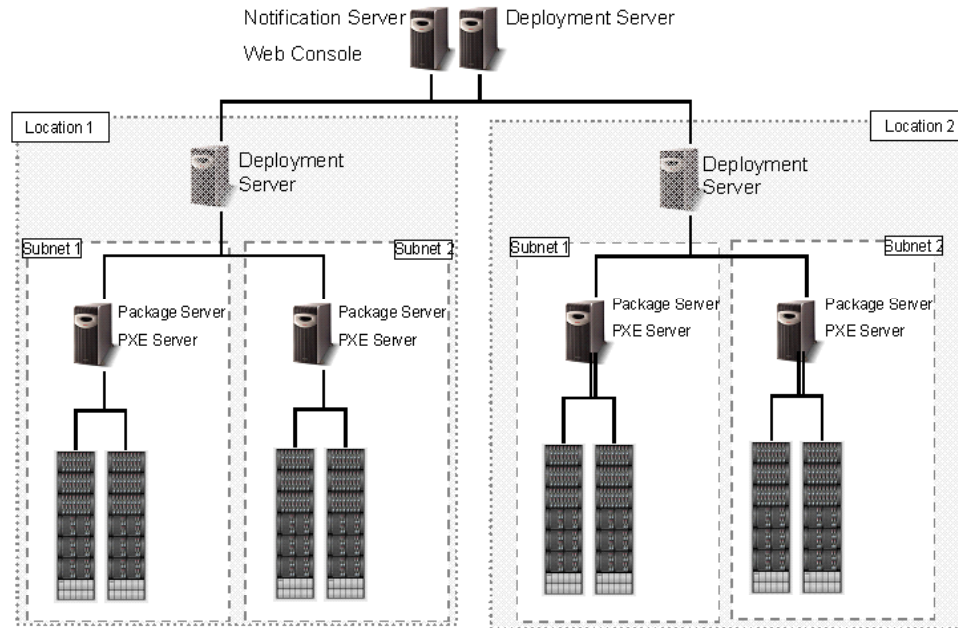
Use the simple install method when:	Use the custom install method when:
You do not have SQL Server and want the Microsoft SQL Server 2000 Desktop Engine (MSDE 2000) installed on the deployment server	You already have SQL Server or MSDE 2000 installed on the intended deployment server or you want to use an existing SQL Server somewhere else in the network
Your intended deployment server has one NIC or you want to use the first NIC as the deployment NIC	Your intended deployment server has multiple NICs, and you want to specify which NIC the deployment server will use (other than the first NIC)
You want PXE Server installed on the deployment server	You do not want PXE Server or you want PXE Server installed on another computer
You want the Client Access Point (file server) on the deployment server	You want to install the Client Access Point (file server) on another computer

## Enterprise installation



An enterprise or distributed installation typically would be used by a geographically distributed company that needs to control server deployment from a central location.

## Distributed deployment architecture



Adopting an enterprise or distributed deployment solution requires the following equipment at each remote site:

- The console at a central location
- The installation of a PXE Server on each subnet
- One or more
  - Deployment servers
  - SQL databases
  - File servers

## Predeployment configuration

Before using RDP, you must make the following configuration modifications to ensure that your deployment server functions optimally:

- Configure PXE to process new computers automatically.
- Synchronize the console name with the operating system name.
- Change the primary lookup key to the serial number.
- Configure PSPs.
- Create PXE images, boot diskettes, and multiple NIC boot images.
- Configure ProLiant BL enclosures (if applicable).
- Remotely install the Deployment Agent for Windows (formerly *AClient*).

### Configuring PXE to process new computers automatically

By default, when a new server (one not listed in the Deployment Server database) boots PXE, the PXE Server gives it the PXE menu and then waits for you to select the Initial Deployment boot option. This user interaction prevents destructive events from running on the new server because the deployment server does not recognize it. This safety feature is intended for desktop computers that might be discovered by an existing deployment server for the first time when an initial deployment job is set up to deploy a server using imaging.

However, in some cases (especially with server blades), this user interaction is unwanted because it requires the user to be at the new server to press a key. Therefore, sometimes it is advantageous to configure PXE to automatically process new computers. To configure the PXE Server to choose the Initial Deployment menu item automatically and continue without user interaction:

1. Click *Start* → *Programs* → *Altiris* → *PXE Services* → *PXE Configuration Utility*.
2. Click *Altiris Bootworks (Initial Deployment)*.
3. Click *Edit* to display the Menu Item Properties window.
4. Select *Execute Immediately*.
5. Click *OK* to close both windows.



---

#### Caution

Do **not** reorder the boot menu on the Menu Items list on the Boot Configuration tab. Deployment Server selects the top menu item as its default action when there is no task for a computer to perform. Moving another selection, such as Initial Deploy, to the top of the list causes the server to never boot locally and to cycle into an endless loop of reboots.

---

## **Synchronizing the console name with the operating system name**

Deployment Server allows you to specify an alias, which is a name that displays differently in the Deployment Server console from the one actually used by the operating system or NetBIOS. However, you can change a setting to ensure that the Deployment Server console always reflects the name used by the operating system.

## **Changing the primary lookup key to the serial number**

Deployment Server uses the primary lookup key to determine if a server is already in the database. By default, Deployment Server uses the media access control (MAC) address of the NIC as its primary means of identifying computers. If you change the NIC in a computer, Deployment Server treats it as a new computer.

HP recommends setting the primary lookup key as the serial number because doing so:

- Enables servers to be imported by their serial number, instead of by difficult to determine keys such as the MAC address
- Avoids duplicate database entries that might occur with servers with two or more NICs

## **Configuring ProLiant Support Packs**

HP Management and Foundation Agents, which are downloadable as PSPs, collect information to measure key parameters of the server. The web-based management portion of the HP ProLiant Foundation Agents requires that you configure a password in the Smart Component before installation. This password is also used by several other components in the PSP. Without the password, the web-based management portion of the Foundation Agents will install but will not function correctly and will not be accessible.

## **Creating PXE boot images and diskettes**

After you install the PXE Server software, you can create new images specific to your environment for both new and managed machines. The default images will work in most environments; HP recommends that you use the default image unless you have a reason to change it.

Two basic images are required for use with PXE Server—one for the initial deployment job on a new computer and another for managed computers already in the database.

## Configuring ProLiant BL server blade enclosures

The Physical Devices View in the Deployment Server console displays the rack name and enclosure name for each ProLiant BL server blade. The default name for the server rack is *UnnamedRack* and the default name for the BL e-Class server enclosure is the MAC address of the NIC associated with Integrated Administrator.

After the ProLiant BL server blades are powered up for the first time and the rack and enclosure names are recorded in the Deployment Server console, the server blades must be restarted to change the rack and enclosures name. Therefore, HP recommends that you set the rack and enclosure names before the first server blade in an enclosure connects to Integrated Administrator.

## Remotely installing Deployment Agent for Windows

The Deployment Agent for Windows is a small utility that runs as a service on a computer, enabling it to communicate with the deployment server and process commands sent to it by the server.

The Deployment Agent for Windows:

- Captures the hardware configuration and disk image from a reference server and deploys them to target servers
- Uses automatic post-image configuration to customize the computer name, IP address, domain, and so on

By installing the deployment agent, you can redeploy or manage existing computers in your infrastructure, as well as perform pre- and post-imaging configuration and application installation.

If you do not install the deployment agent as part of the deployment process, you lose these capabilities. Scripted installation events provided by RDP install the deployment agent by default on deployed servers.

---

### Note

HP recommends that you leave the agent on the system after initial installation.

---

Deployment Server enables you to install the deployment agent remotely on Windows systems provided you have the appropriate administrative permissions on those systems. You also can perform this installation in batch mode to install the agent to multiple systems simultaneously. This means that you do not have to visit each system individually to install the agent manually from a diskette.

## Configuring the database

The Deployment Server database is the heart of the Deployment Solution system. The database must be either:

- **Microsoft SQL Server 7 or later** — If the computer on which you are installing the Deployment Server database has an existing SQL Server database, the Deployment Server database is added to SQL Server database. Use SQL Server 2000 for installations with more than 2,500 clients per server.
- **MSDE 2000 running on a computer with Windows 98 or later** — If you do not have access to SQL Server, you can use the free version of MSDE 2000 that is included in the Deployment Solution on installations with 2,500 or fewer clients per server.

A deployment server can communicate with only one Deployment Server database, so you can install only one database per Deployment Solution. A deployment server can communicate with a database server in another system, but the deployment server must have a separate installation of the Deployment Server database. However, you can install multiple Deployment Solutions to manage different sets of computers. Multiple Deployment Solutions cannot manage the same set of computers at the same time.

The Deployment Server database contains information about the managed computers, including:

- **Hardware information** — RAM, asset tag, and serial numbers
- **General information** — Computer names and MAC addresses
- **Configuration information** — TCP/IP addresses, Microsoft networking information, and user information
- **Application information** — Information about the applications installed such as the name of the application, publisher, and product ID
- **Device information** — Windows devices installed, such as network adapter, keyboard, and monitors
- **Service information** — Windows services installed
- **Location information** — Contact name, phone number, email address, department, mail stop, and site

The database also contains jobs and other information used to manage computers.

## File sharing in a multiple operating system environment

Servers running versions of the Linux and Windows operating systems must be able to share files across the network with minimal compatibility issues. This interoperability must allow seamless data exchange in a heterogeneous environment.

Samba, an open source implementation of the Common Internet File system protocol, allows connectivity between Linux and Windows computers by using the Windows native file sharing protocol Server Message Block (SMB). Samba enables Linux servers to function as file and print servers for PCs running Windows.

SMB is a simple client-server protocol that is available for any Windows platform. Any Windows operating system later than Windows 3.11 ships with SMB already installed. SMB works over Token Ring, Ethernet, and even serial hardware using the TCP/IP, NetBEUI, or IPX/SPX protocols. SMB uses NetBIOS to function properly over TCP/IP and NetBEUI.

Because Samba uses SMB, it displays on the network as another network node. Windows is not aware that the machine running Samba is also running another operating system.

Because many companies running Windows 2000 and Windows Server 2003 are migrating to a Linux platform, it is helpful know how to integrate the two servers into a heterogeneous network environment.

### Downloading Samba

After determining that Samba will meet your resource sharing needs, download the proper file. You can download a tar.gz file or the latest Red Hat Package Manager (RPM) file.

---

<b>INTERNET</b>	Download the latest version of Samba from: <a href="http://www.samba.org">http://www.samba.org</a>
-----------------	--

---

### Installing Samba using the RPM

A typical Samba RPM file will resemble the following:

```
samba-2.2.5-1.i386.rpm
```

To install this RPM, execute the following command line as root:

```
rpm -ivh samba-2.2.5-1.i386.rpm
```

This command installs Samba with most functionality. If the RPM does not install all required features, installing Samba from source might be required.

## Using Samba

The operation of Samba is straightforward. The client machine sends a request to the server, which provides the requested resource to the client if permissions allow.

The following table lists the Samba commands and their descriptions.

Command	Description
smbmount	<p>Enables you to mount a Windows share from a computer running Linux. The command syntax is: <code>smbmount //remotecomputer/sharename /directory</code> where <code>remotecomputer</code> is the NetBIOS name or IP address of the remote computer, <code>sharename</code> represents the share to mount, and <code>/directory</code> represents the directory in which to mount the share.</p> <p>For example, to mount a share called <code>C</code> on the Windows computer named Jake, default into the directory <code>/home/jake/remote</code>, and execute the following command: <code>smbmount //default/c /home/jake/remote</code>.</p> <p>If user level security is enabled, <code>smbmount</code> will ask for the username and password to access the share.</p>
smbd	Provides access to the Samba computer from a Windows machine. You can invoke the command <code>smbd</code> with the following arguments: <code>start</code> (starts <code>smbd</code> ), <code>stop</code> (stops <code>smbd</code> ), and <code>restart</code> (stops then starts <code>smbd</code> ).
nmbd	Provides shares to the Windows computer. You can invoke the command <code>nmbd</code> with the same arguments as <code>smbd</code> .
smbclient	Provides access to shared resources on a Windows computer. The <code>smbclient</code> command is similar to an FTP and is used to connect to Windows shares. Its usage is similar to <code>smbmount</code> . To gain access to a Windows computer from the Samba machine, execute: <code>smbclient //remotecomputer/sharename</code> , where <code>remotecomputer</code> is the IP address and <code>sharename</code> is the name of the share to access.
testparm	Tests the validity of the <code>smb.conf</code> file. It looks for the <code>smb.conf</code> file in either <code>/usr/local/samba/lib</code> or <code>/etc</code> .
testprns	Tests the validity of a shared printer configuration file on the Samba machine. Execute <code>testprns</code> to see if a printer is configured correctly for sharing.
smbstatus	Gives an overview of the status of a running Samba server.
smbtar	Backs up a Windows share. Execute <code>smbtar//remotecomputer/sharename</code> where <code>//remotecomputer</code> is the name or IP address of the remote computer and <code>sharename</code> is the name of the share to back up.
nmblookup	Matches the NetBIOS name of a computer to its IP address. Execute <code>nmblookup hostname</code> , where <code>hostname</code> is the name of the computer to be found. For example, a command of <code>nmblookup Test2</code> should return <code>192.168.2.3 Test2</code> .

## Deploying a web server

Although some sites might require a single application web server, other sites might require a multitiered, fault-tolerant solution. Before deciding which configuration provides the best solution in a given environment, ask the following questions:

- How busy will the site be? (concurrent transaction or hits per second)
- How large will the site content be? (memory and hard drive requirements)
- Will the site be serving heavy static content? Dynamic content? Mixed content?
- Can the site be shut down completely for any reason? (availability of the site)
- Should the individual web server have fully redundant components? (power supplies, hard drives, NICs, fans)
- Should multiple, load-balanced, single or dual-processor web servers be used or should a single web server with four or eight processors be used?

Selecting the configuration that best fits a specific site will be easier after you understand the answers to these questions.

---

**INTERNET**

For additional information, refer to the HP ActiveAnswers website for a performance characterization of Microsoft IIS in Windows Server 2003 on ProLiant servers.

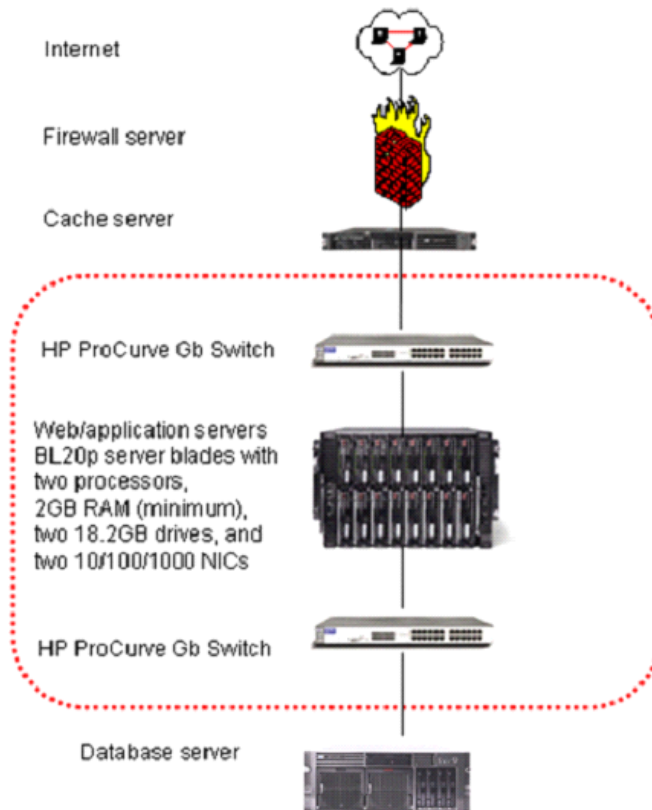
---

The performance of a web server is measured by the number of requests it can handle per second and the total data rate the server can provide, or its *throughput*. This performance is dependent on a number of highly variable conditions that include:

- Popularity of the website
- The number of static pages (.html, .gif, .jpg) compared to dynamically generated pages
- Complexity of the server-side scripting code
- Size and design of the database tables
- Search engine usage
- Usage of encryption technology such as Secure Sockets Layer (SSL)

Performance guidelines are based on the documentation provided with the open source software and on testing performed by HP.

## HP recommended configurations



HP recommended web configuration based on ProLiant BL20p servers

HP recommends the following configurations for web servers:

- **Firewall servers** — Firewall security should include intrusion detection, virus protection, and if required, controlled access to the site using Virtual Private Networks (VPNs). ProLiant DL360 servers and BL server blades are well suited for this layer. HP recommends that a firewall contain two processors (fastest available speed), 1 to 2GB RAM, and sufficient disk space.

Always determine security policies and site requirements before selecting a firewall product.

- **Caching servers** — A caching server caches information for faster retrieval. By handling a request for a specific file or information at the highest level, a caching server improves performance.

The ProLiant DL380 server is an excellent choice for this layer because it supports dual processing, 12GB RAM, and up to six disks (876GB) for caching site information. Caching servers use both disk and memory for caching content, so carefully plan memory caching and disk definitions within the application.

- **Web and application servers** — The servers at this layer provide web services and perform the primary work of the applications for the site. The applications and site content are commonly replicated on each server in this layer. The content and applications on every website are unique. Therefore, the amount and type of content that can be placed on both the web/application resource layer and data resource layer can vary.
- **Database servers** — The data resource layer servers contain the data for the applications installed on the web and application resources layer. Highly available, centralized file systems can store critical data on clusters, SANs, or Fibre Channel storage solutions. The data can consist of database information, email folders, streaming video, or e-commerce information, among many other options.

Several ProLiant products and configurations can be used in the database resource layer. Considerations include desired availability, data size, performance, and so forth.

### Content types

The recommended configurations are categorized into three content types:

- **Static** — Performance is maximized with static content because the web page requires no generation within the server and little processor time before it is sent as a response to a request.
- **Dynamic** — Dynamic content requires heavy processor utilization, so faster processors may be required to produce optimum results.
- **Mixed** — A typical website has a mix of static and dynamic content.

---

**INTERNET**

For more information, review the *Technical Overview of Internet Information Services (IIS) 6.0*, which is located at:  
**[http://www.microsoft.com/  
windows.netserver/techinfo/overview/iis.mspx](http://www.microsoft.com/windows.netserver/techinfo/overview/iis.mspx)**

---

## Operating system integration in heterogeneous environments

For best results when integrating multiple operating systems in a heterogeneous environment, apply HP adaptive infrastructure design principles to the IT infrastructure with an incremental, balanced, and focused approach. The relevant HP adaptive design principles are:

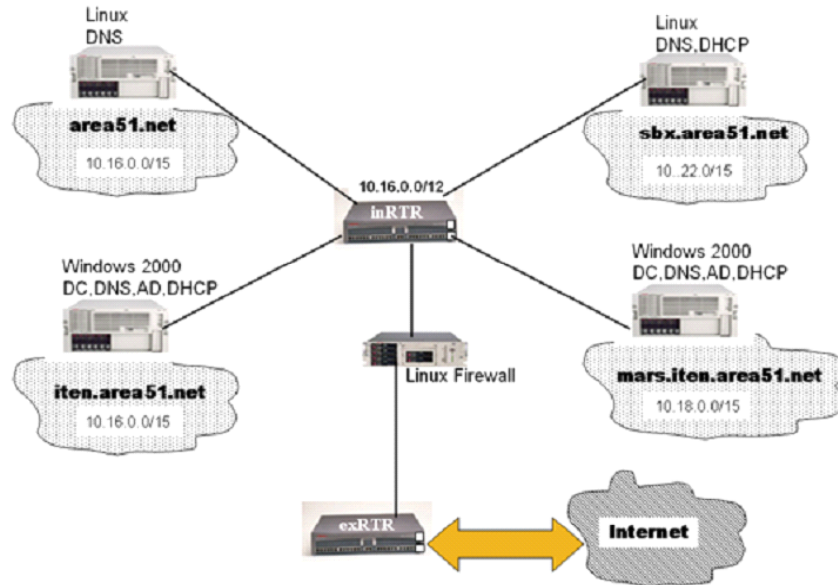
- **Simplification** — Reduces the complexity and risk that can increase costs and reduce the return on investment. Simple applications and systems are easier to adopt, use, connect, manage, and modify. Simplification can be achieved in a number of ways, including:
  - Consolidation — Streamlines and updates the infrastructure. Fewer elements lead to greater simplicity, which in turn yields greater speed and ease when making changes.
  - Outsourcing — Simplifies the IT environment the business is responsible for managing.
  - Application integration — Simplifies connections between applications for improved interoperability.
- **Standardization** — Helps expedite change and reduces its cost and risk. When changes are standardized, they can be applied across different processes, procedures, technologies, or applications. Standardization can be achieved by:
  - Using industry-standard interfaces, platforms, and software development techniques.
  - Establishing common processes and policies.
  - Defining common requirements for manageability, security, version control, configuration management, capacity and performance management. This approach enables outsourcing of the lower layers of the solution stack (hardware, software, and middleware).

- **Modularity** — Provides the ability to change one aspect of a system without affecting other components. A modular approach to resources increases flexibility. Resources such as storage and computing power can be assigned to applications or business processes as needed.

The HP adaptive network architecture uses modularity to improve the responsiveness of the network to business needs. You can group systems in predefined network compartments based on common access requirements; you can construct compartments to connect or disconnect in near real-time; and you can modify any compartment without changing the others.

- **Integration** — Facilitates through a uniform system of relationships that is easy to understand, manage, and modify. In application development, enterprise application integration simplifies connectivity by using pervasive and widely adopted standards.

## Integrating Active Directory support



Many companies are either in the planning or implementation stages of some type of Linux migration, moving from a uniform platform to a heterogeneous platform consisting of Windows 2000, UNIX, Linux, Solaris or a combination of these operating systems.

Such a migration becomes complex when trying to merge the Linux Berkeley Internet Name Domain (BIND) 9.x for DNS with Windows Active Directory. The BIND DNS server provides standard APIs to map domain names to IP addresses; integrating BIND with Active Directory reduces administrative overhead and increases security.

HP has developed a seven-step process for simplifying the Linux BIND DNS 9.x/ Windows 2000 Active Directory integration:

1. Planning and resources
2. Naming conventions
3. Establishing an IP addressing scheme
4. DNS considerations
5. Infrastructure subsystems
6. Linux DNS configuration
7. Windows 2000 DNS configuration

## Planning and resources

When planning a Linux/Windows 2000 integration, consider the technologies needed, available resources, common problems, and network topology.

To integrate Linux BIND 9.x, include the following supporting technologies, which are used to seamlessly integrate BIND 9.x with Active Directory:

- TCP/IP
- Routing and subnetting
- Linux DNS
- Windows 2000 DNS
- Windows 2000 DHCP
- Linux DHCP

---

**Note**

Only use Linux DHCP on a single network interface host.

---

- DNS forwarding
- Lightweight Directory Access Protocol (LDAP) (implemented for single logon for Windows and Linux users)

Refer to the following resources for more information on Linux BIND and Windows Active Directory integration.

- URLs
  - <http://itccweb.ccm.cpqcorp.net/TechDocs>
  - <http://itccweb.ccm.cpqcorp.net/LinuxFirewall.html>
- HP Education and Training classes
  - Designing a Microsoft Windows Server 2003 Active Directory and network infrastructure (MOC 2282)
  - Planning, implementing, and maintaining a Microsoft Windows Server 2003 Active Directory infrastructure (MOC 2279)

## Naming conventions

It is important to have a consistent naming convention for all the devices on your network. Devices include servers, routers, web appliances, switches, hubs, and so on. The naming conventions:

1. Consistently identify all devices within the heterogeneous infrastructure
2. Can quickly provide the server location and operating system or function type to the system administrator
3. Help identify invaders to the heterogeneous network
4. Facilitate the deployment of Systems Insight Manager

### Sample format

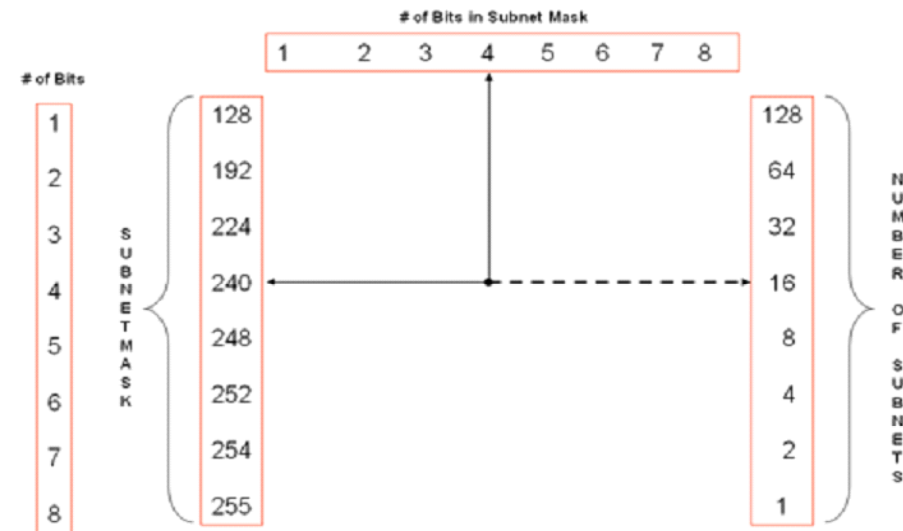
All device names are case-sensitive; for example, HOST1 is different from host1. The following sample shows naming conventions for devices and operating systems.

#### Example

AAA-B-XXXX

- AAA — Location, site, or function prefix:
  - inf — Infrastructure servers and devices
  - sbx — Sandbox servers and devices
  - test — Temporary devices
  - spc — Special servers, reserved category
  - cip — Cluster IP address
  - prd — Production systems and devices
- B — Operating system type:
  - l — Linux
  - u — Tru64 Unix, Solaris, HP-UX, AIX
  - w — Windows 2000, Windows NT 4.0
  - v — Novell
  - o — Other
- XXXX — Name of device defined by user; no more than 12 characters

## Establishing an IP addressing scheme



Subnet calculator

The next step in the integration of BIND 9.x on Linux with Windows 2000 and Active Directory is to establish an IP addressing scheme.

- Given IP: 10.16.0.0/12
- Option to use one large address space or subdivide it

For manageability and isolation, the administrator must subdivide the 10.16.0.0/12 address space. Subdividing also minimizes the problem areas to the subnet segment only.

### Example

10.16.0.1 through 10.16.255.254 (255.240.0.0) is the useable address space. To divide this address space into eight subnets, you must borrow three bits from the left, resulting in a netmask that is now 255.240.0.7 (15 bits). This will result in each new subnet segment being 2\*\*1 bits wide:

0 = + 0.0.0.1 to + 0.0.255.254

1 = + 0.1.0.1 to + 0.1.255.254

The previous subnet had the following parameters:

- Range: 10.16.0.1 to 10.31.255.254
- Netmask: 255.240.0.0
- Broadcast: 10.31.255.255

The new subnets have the following parameters:

- Subnet A
  - ♦ Range: 10.16.0.1 to 10.17.255.254
  - ♦ Broadcast: 10.17.255.255
- Subnet B
  - ♦ Range: 10.18.0.1 to 10.19.255.254
  - ♦ Broadcast: 10.19.255.255
- Subnet C
  - ♦ Range: 10.20.0.1 to 10.21.255.254
  - ♦ Broadcast: 10.21.255.255
- Subnet D
  - ♦ Range: 10.30.0.1 to 10.31.255.254
  - ♦ Broadcast: 10.31.255.255

## DNS considerations

To integrate BIND 9.x with Active Directory, you must address the following DNS considerations:

- Routing
- Domain name
- Forwarding
- Tools

Routing is important in a multidomain heterogeneous environment. Without routing, there is no DNS. The routes must work independently before implementation of a DNS Active Directory infrastructure. This requires:

- Routes between subnets
- Atomic routers

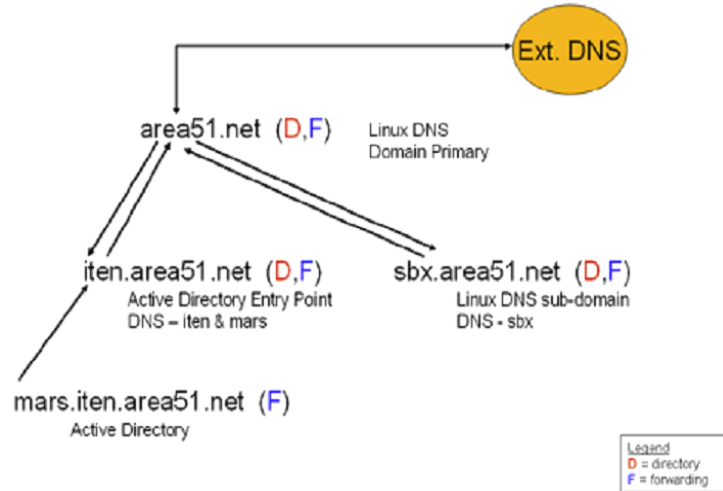
---

**Note**

Atomic routers are defined as routers that know all the routes on the network.

---

## DNS forwarding



In the preceding figure, the mars subdomain is hosted on the iten Active Directory server. In this example, the forwarders also cache the DNS entries. To enhance network performance, disable recursion at lower levels. This minimizes network traffic.

On the LAN connection between mars.iten and iten, forward lookup from mars.iten to iten is enabled. However, recursion is disabled between these two connections, which places recursion responsibilities on the edge-of-the-network DNS servers, such as Ext. DNS.

## Tools

Some of the tools needed to successfully integrate BIND 9.x with Active Directory are:

- `dig` — Linux utility to query DNS server; it provides more information than `nslookup`
- `traceroute` — Displays route information as a packet proceeds to its destination

---

**Note**

When probing external sources using `ping`, the utility does not work all the time. Some external DNS servers drop or reject Internet Control Message Protocol (ICMP) packets associated with `ping`.

---

- `nslookup` — Traditional DNS server query utility
- `bindconf` — Linux DNS configuration utility to modify `named.conf`

## Common problems

Usually the problems associated with a BIND 9.x/Active Directory implementation are caused by one of the following:

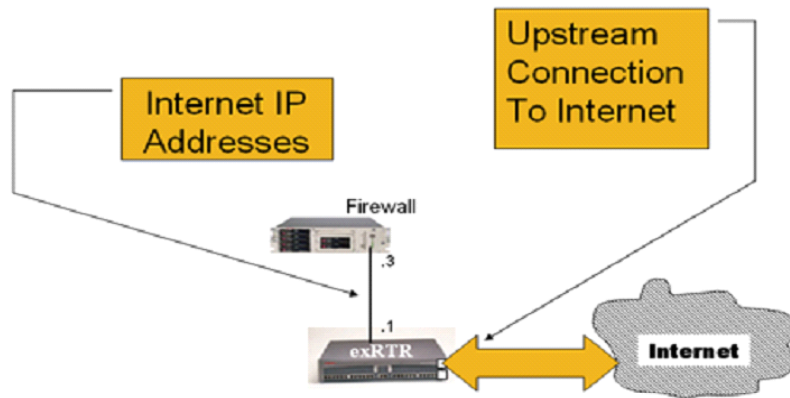
- No plan
- Routes
- DNS forwarding
- Names
- No reverse IP lookup
- No DNS “stub” for subdomains
- Vi or PICO Editor

## Infrastructure subsystems

To implement a successful BIND 9.x with Active Directory integration, the following hardware is needed for infrastructure subsystems:

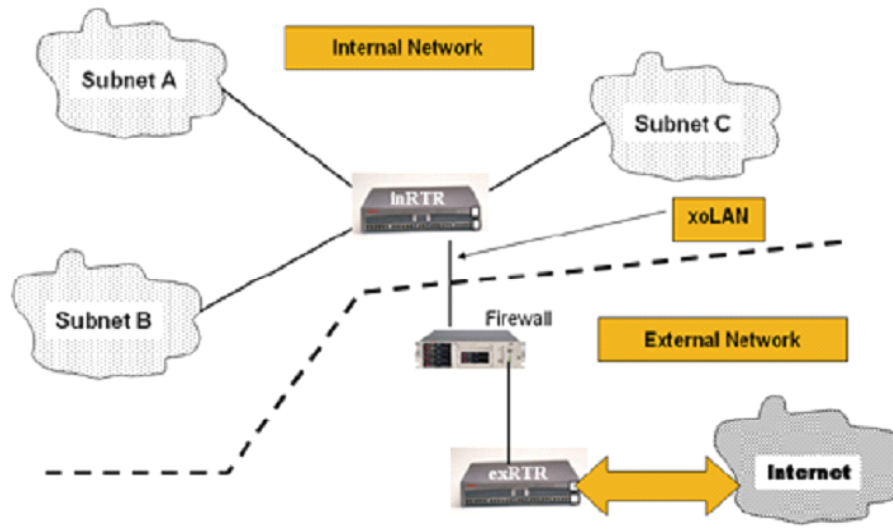
- External router
- Gateway and firewall
- Internal router (atomic)
- Primary domain server (Linux BIND 9.x)
- Active Directory entry point server (Windows 2000)
- Active Directory subdomain server

### External router



The placement of infrastructure subsystems is critical to the success of an implementation. The sample project shown in the preceding graphic depicts an actual implementation running in an HP lab. In this graphic, the external router is needed if the internal domain accesses the Internet. The external router points to the external DNS server. All traffic to and from the Internet is handled by the router.

## Internal router



The preceding figure shows the model for an internal router. With an internal router, you must have a 10/100 switch with Gigabit uplinks capable of creating VLANs. HP also suggests that you develop a good static routing table.

## Linux DNS server

The following requirements are essential for the Linux DNS server:

- The BIND server must be running BIND 9.x or later.
- All domain planning must be complete.
- The Linux server must include the following files and directories:
  - /etc/named.conf
    - ◆ “named” daemon startup file (main DNS configuration file)
  - /var/named (directory)
  - /var/named/YourZone (directory)
  - Domain zone files
    - ◆ named.ca
    - ◆ named.local
    - ◆ zone.rev
    - ◆ domainFile (must be created)

---

### Note

The preceding files are the minimum files necessary to establish a working DNS server.

---

HP recommends creating your own directory. It is easier to back up and you can create multiple distinct directory trees.

## Linux DHCP

The DHCP server is not installed by default. If you have installed all the RPM packages in a custom installation or if you have chosen `dhcpd` as one of the packages, then you do not have to install the DHCP server. Otherwise, you must install the DHCP server.

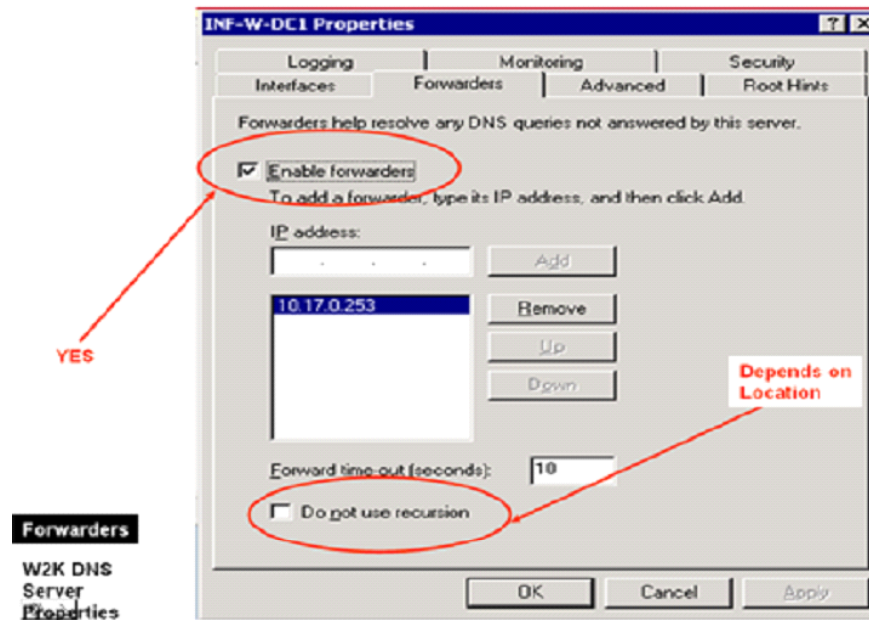
To install the DHCP server, install the `dhcpd-*.i386.rpm` package. After you have installed the DHCP server, perform the following steps:

1. At a shell prompt command line, start the DHCP service by entering:  

```
service dhcpd start
```
2. Create or modify the `/etc/dhcpd.conf` file.
3. Restart the service by entering:  

```
/etc/init.d/dhcpd restart
```

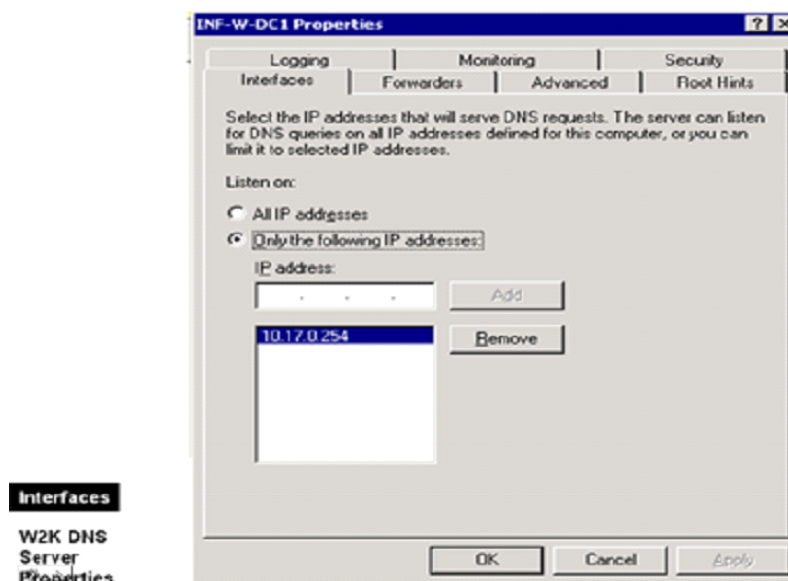
## Windows 2000 DNS configuration



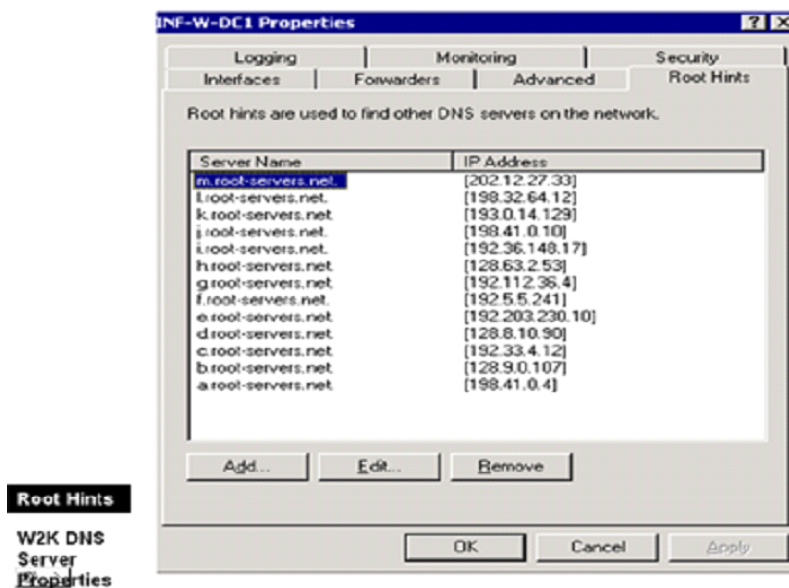
The preceding figure shows a sample of the forwarders tab on the Windows 2000 DNS Server properties window.

When working with Windows 2000 DNS, remember:

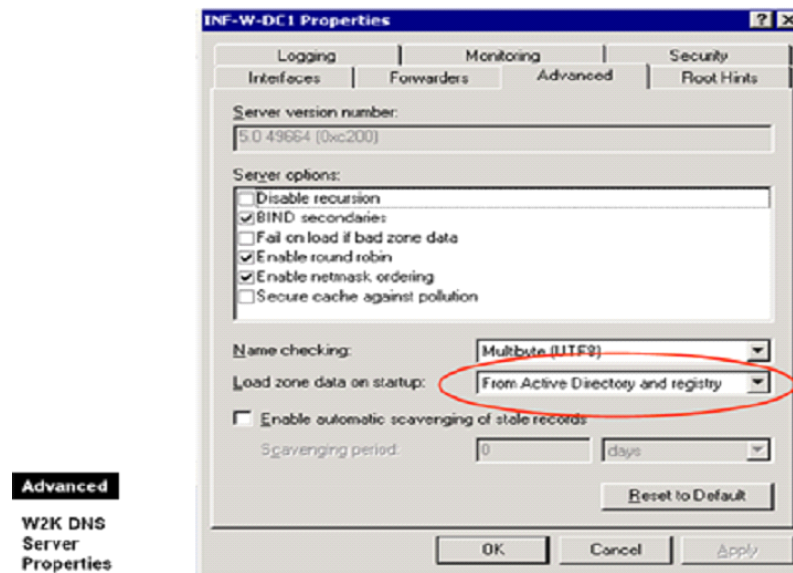
- Choose Active Directory Integrated (ADI).
- Recommend native mode.
- Use a minimum of two domain controllers.
- Configure DHCP servers.
- Support multiple interfaces.
- Configure forward and reverse zones as needed.
- Have subdomains.



For security reasons, you must name the interfaces, especially if the server has more than one NIC.



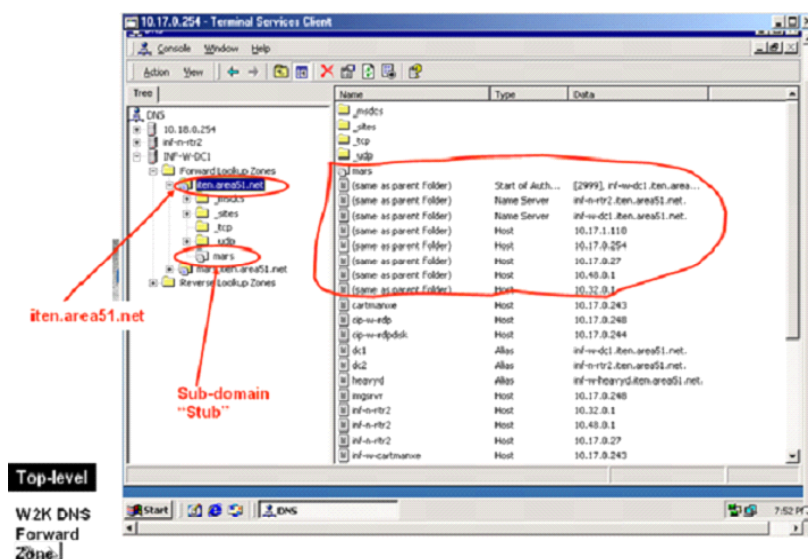
The preceding graphic shows a sample of what the root hints should look like.



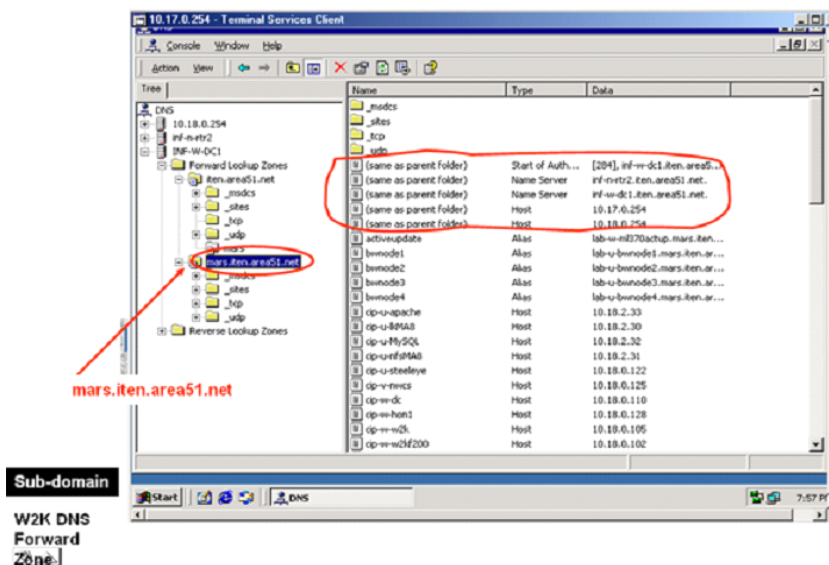
In the Advanced Settings window, select *From Active Directory and registry* in the Load zone data on startup field.

Other zone files are located at `c:\winnt\system32\dns`.

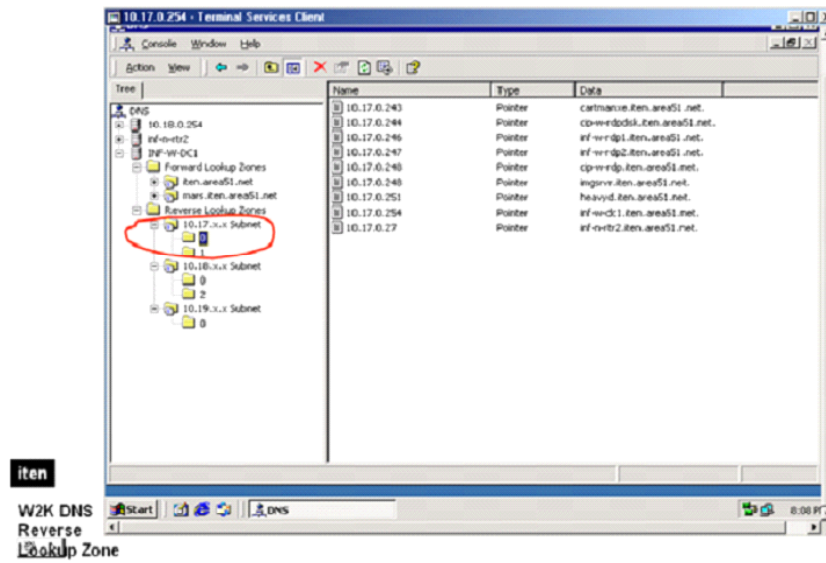
## Subdomains and DNS



Subdomains are created by right-clicking on the zone in the left menu and selecting *New Delegation*. This creates a DNS stub in the top-level domain. ADI DNS, all domain controllers are shown in designation.

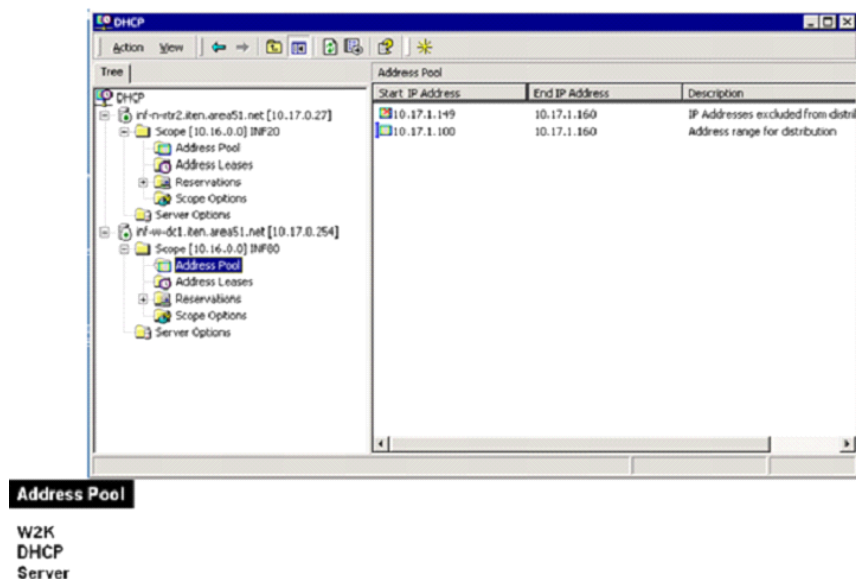


The subdomains are displayed in the preceding graphic.



The preceding figure shows the Windows 2000 DNS reverse lookup zone for iten. Servers and clients (DHCP) must be kept in different IP address ranges.

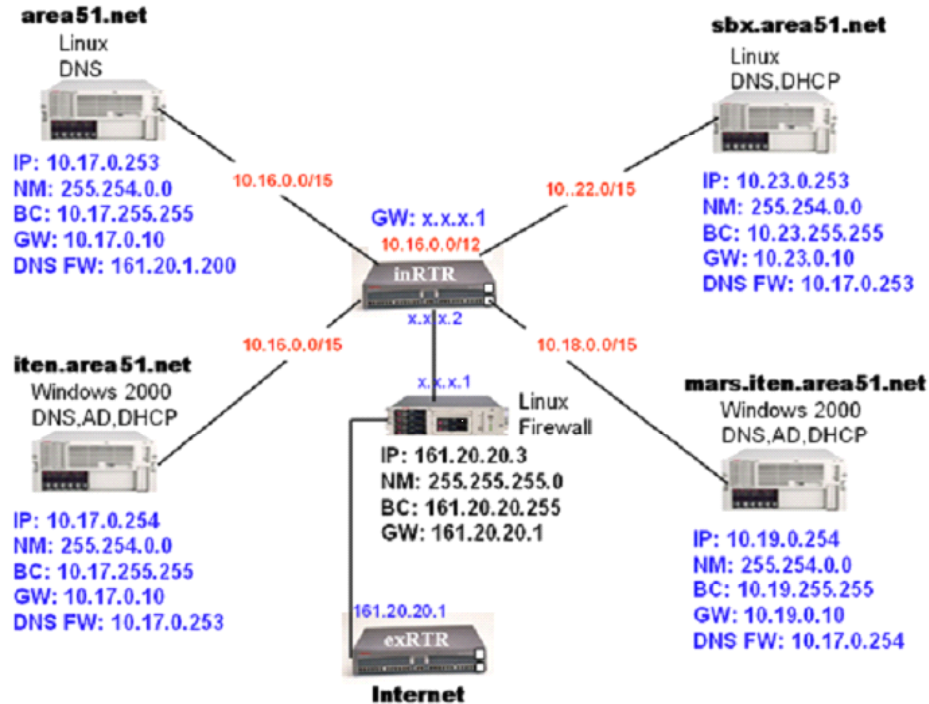
## DHCP address pools



HP recommends that two DHCP servers be configured to use the 80/20 – 20/80 rule.

In a mixed environment, many administrators believe that Windows 2000 DNS is easier to administer. However, if you are in a UNIX environment, use DHCP on the Linux server.

## Final network topology



The final network topology for the BIND 9.x and Windows 2000 Active Directory Integration is shown in the figure. The IP address, Net Mask, broadcast, gateway, and DNS forwarding addresses are listed for each server.

## Summary

Bob is impressed with server blade technology and decides that the return on investment for ProLiant BL20p G2 server blades will benefit RC Engineering in the long term. In addition to being able to take advantage of the powerful, scalable server blades, the company will be able to maximize the use of its available space and still have room for the growing staff.

You advise Bob to use RDP to perform a default installation of all the operating systems across the network. RDP is the easiest method and best suited to the environment because it reduces time and resources by enabling remote automated server deployment. Jackie is comfortable with the drag-and-drop method for deploying servers and the simplified management of server deployment through the remote console.

In addition, Jackie has worked with her staff to develop a timeline for implementing Samba across the network. She is grateful that you introduced her to the seven-step HP process for simplifying the Linux BIND 9.x/Windows 2000 Active Directory integration. As a result, Jackie has been able to configure the Linux server for file sharing. Because she has worked through the process in advance, she is confident that the document management system rollout will go smoothly.

## Learning check

1. List the configuration modifications you must implement to make your deployment server function optimally before you begin to use RDP.  
.....  
.....  
.....  
.....  
.....  
.....
2. Name the three main configuration and installation methods for RDP.  
.....  
.....  
.....
3. Samba allows connectivity between Linux operating systems and Windows computers by using \_\_\_\_\_.
4. Which ProLiant servers are well suited as firewalls? Select TWO.
  - a. ML330
  - b. ML370
  - c. DL380
  - d. DL360
  - e. BL server blades
  - f. DL580

5. Without \_\_\_\_\_, there is no DNS.
6. What is a service level agreement (SLA)?  
.....  
.....
7. List two of the four reasons it is important to have a consistent naming convention for all network devices.  
.....  
.....  
.....  
.....

### Objectives

After completing this module, you should be able to:

- Discuss how to use HP Systems Insight Manager to manage an enterprise
- Install Systems Insight Manager
- Configure Systems Insight Manager
- Use Systems Insight Manager to monitor managed systems
- Create reports in Systems Insight Manager
- Explain the requirements and procedures to develop a system software maintenance strategy using Systems Insight Manager
- Describe how to use the command line interface (CLI) to configure the Systems Insight Manager database
- Manage servers remotely using HP Lights-Out devices
- Describe the tasks that must be completed when a new server is introduced in a managed enterprise

## Introduction

Now that they have established their enterprise infrastructure, the RC Engineering staff needs to proactively manage all the systems on the network from a central location. Your job is to design a solution to meet this challenge.

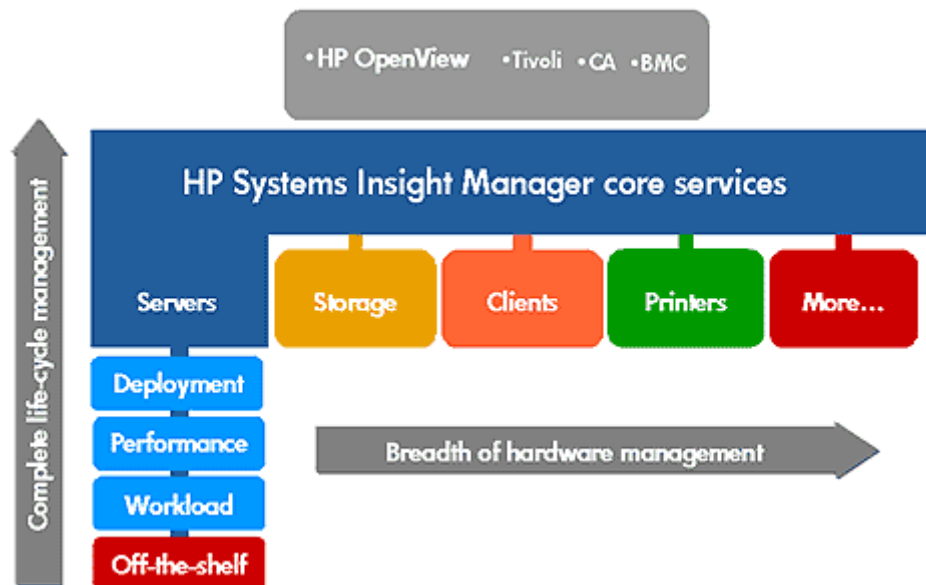
The main business goals of the company are to:

- Manage the network infrastructure from a centrally located management console
- Assign systems management responsibilities based on user roles
- Control servers at remote locations
- Implement a systems software maintenance strategy

After conferring with Bob, the CEO of RC Engineering, you recommend Systems Insight Manager because it helps maximize IT staff efficiency and hardware platform availability. Because Jackie's IT staff at RC Engineering is still small, Systems Insight Manager is ideal for the company because the staff does not have to be trained on multiple management tools. Its consistent user interface across Microsoft Windows and Linux platforms fits well into the RC Engineering environment. In addition, its modular architecture enables Jackie to plug in additional functionality as needed.

You must present your plan in a scope of work and document any changes you make as you work.

## Managing the enterprise



Successful management across the enterprise requires the ability to manage servers, storage, clients, printers, and other devices. The ideal solution would enable common and comprehensive management across hardware platforms and key operating systems and provide common fault, configuration, performance, and asset management across all assets in the network.

Systems Insight Manager enables effective enterprise management by offering capabilities in four key areas:

- Fault management
- Configuration management
- Asset management with automatic discovery
- Secure multisystem management












Systems Insight Manager delivers the essential capabilities required to manage all HP server platforms, including ProLiant, Integrity, and HP 9000 systems running Windows, Linux, and HP-UX. Systems Insight Manager also manages clusters, desktops, workstations, and portables, including third-party devices instrumented to the Simple Network Management Protocol (SNMP), Desktop Management Interface (DMI), and Web-Based Enterprise Management (WBEM).

Systems Insight Manager provides a choice of access modes—either through an intuitive web-based graphical user interface (GUI) or a CLI. It can also be customized with off-the-shelf or internally developed scripts and applications.

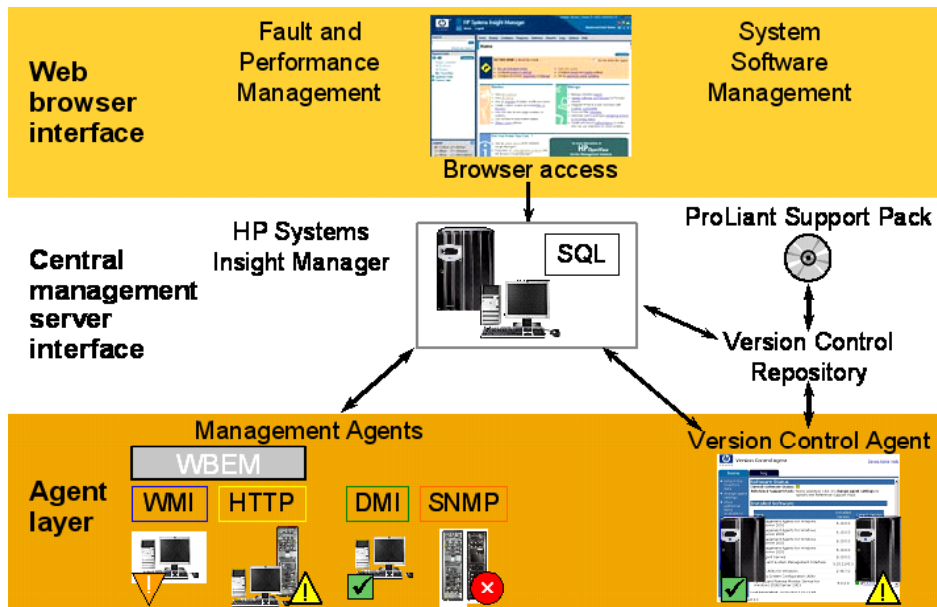
Systems Insight Manager can be easily extended to deliver enhanced server lifecycle management through plug-in components for specialized HP management tools, including:

- Partitioning
- Rapid deployment
- Performance management
- Workload management

Systems Insight Manager combines the best capabilities of HP management tools Insight Manager 7, Tootools, and Servicecontrol. The following table outlines the features of Insight Manager 7 and Servicecontrol Manager 3.0 that are now provided by Systems Insight Manager.

Feature	Insight Manager 7	Servicecontrol Manager
Full-function browser interface		
Role-based security	New!	
CLI	New!	
Automatic device discovery		New!
Inventory collection		New!
System health and fault monitoring		New!
Systems (device) list		New!
Extensibility through application plug-ins	New!	
Task execution	New!	
Reporting		New!

## Systems Insight Manager architecture



The Systems Insight Manager architecture is composed of three layers designed to provide fault and performance management and system software management:

- Web browser interface layer
- Central management server (CMS) layer
- Agent layer

Although each of these layers interoperates seamlessly to provide a consistent look, feel, and user experience, elements in each layer are built as independent units. Therefore, you can upgrade functionality within one layer of the management architecture without upgrading other layers of the architecture.

## **Web browser interface layer**

With Systems Insight Manager, a web browser serves as the primary means of accessing management functionality. Unlike traditional management systems that provide access to management capabilities through a Windows-based management console, you can access the Systems Insight Manager management server from anywhere on the corporate intranet or remotely through a secure Remote Access Server (RAS) or a virtual private network (VPN) connection.

Each element of the agent layer has a built-in HTTP server that enables browser-based access from any system equipped with a web browser and a network connection. The ability to browse directly to the agent layer is particularly useful when managing small server deployments that might not require a management server.

## **CMS layer**

The CMS is at the center of the systems management architecture and includes the Systems Insight Manager core, which aggregates fault, asset, performance, and configuration data from all discovered systems attached to the network. The CMS is also responsible for managing groups of systems through queries and tasks that control operations such as SNMP status polling, email and paging notification, and system software update.

Systems Insight Manager provides links to management applications that run at the agent layer. Examples of such applications are:

- HP Management Agents
- Version Control Agent (VCA)
- Version Control Repository Manager (VCRM)

## Agent layer

The HP Management Agents, VCA, VCRM, and Remote Insight Lights-Out Edition (RILOE) and integrated Lights-Out (iLO) management processors are designed to run at the agent layer. With the exception of VCRM, these applications typically run on each system within the managed environment.

Agent layer management applications provide fault, performance, and configuration management. Examples include:

- **HP Management Agents** — Provide detailed fault, asset, performance, and configuration information for individual HP servers, desktops, portables, and workstations. Management Agents are responsible for maintaining system status and reporting that status to Systems Insight Manager. The agents are also responsible for issuing alerts when they detect irregularities in the operation of the managed system.
- **Snapshot comparison** — Captures detailed configuration snapshots for individual HP servers. These snapshots can be compared to isolate changes in server configuration (including installation or removal of operating system patches) over time.
- **Disk thresholds** — Provide hardware-based graphical remote administration of HP servers both with and without a functioning operating system.

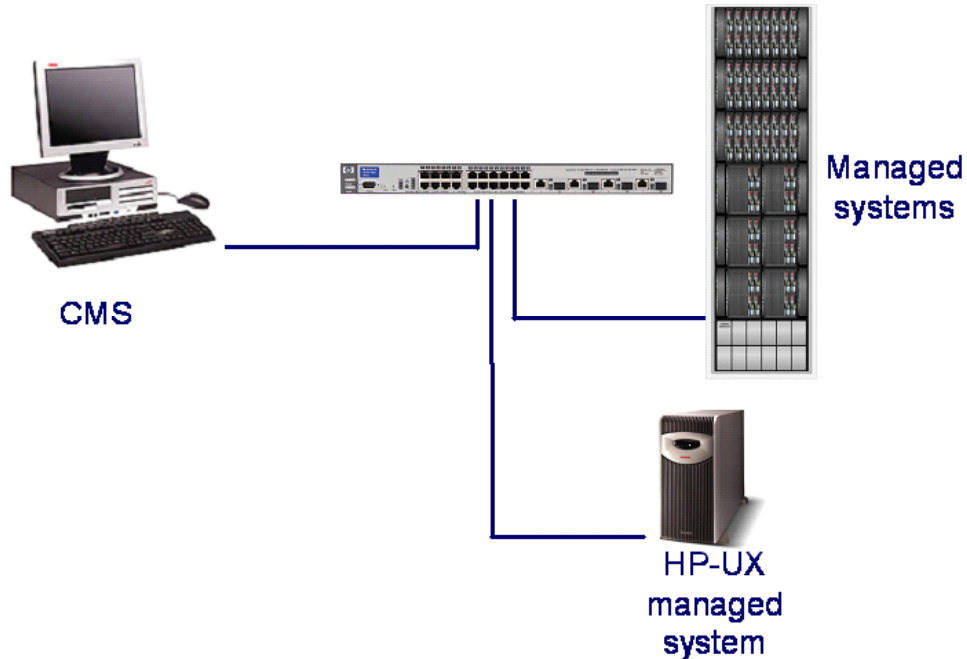
The agent layer also provides system software version and update capabilities such as VCRM and VCA.

## Systems Insight Manager usage scenarios

Systems Insight Manager can be installed in a:

- Branch office
- Data center
- Enterprise

### Branch office



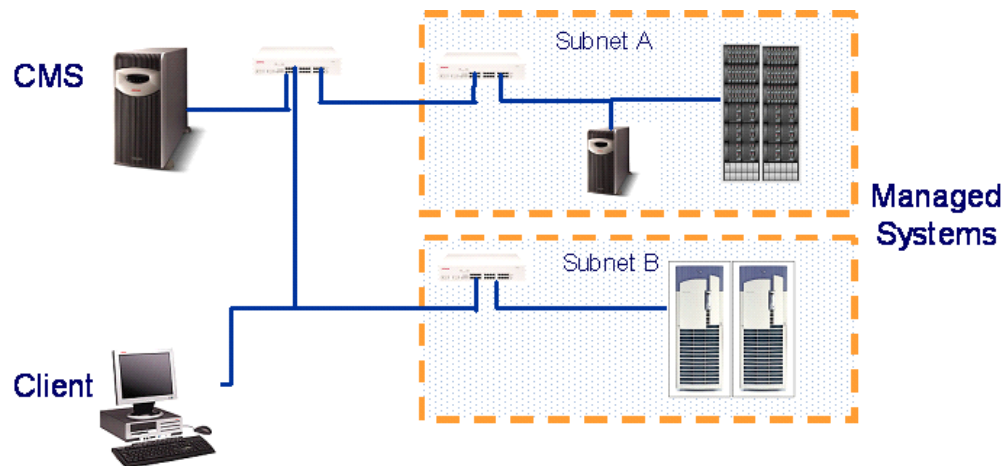
If necessary, branch offices can use a desktop or workstation for the CMS, which is where Systems Insight Manager would be installed. However, with this configuration, no more than 75 systems can be managed. Exceeding this amount might degrade the performance of Systems Insight Manager.

VCRM is also required on the network. In this environment, it is usually installed on the CMS. If a company has HP-UX managed systems, the HP-UX configuration utilities would also be required on the CMS.

The managed systems at each branch office require:

- Management agents
- OpenSSH
- VCA (Windows systems only)

## Data center

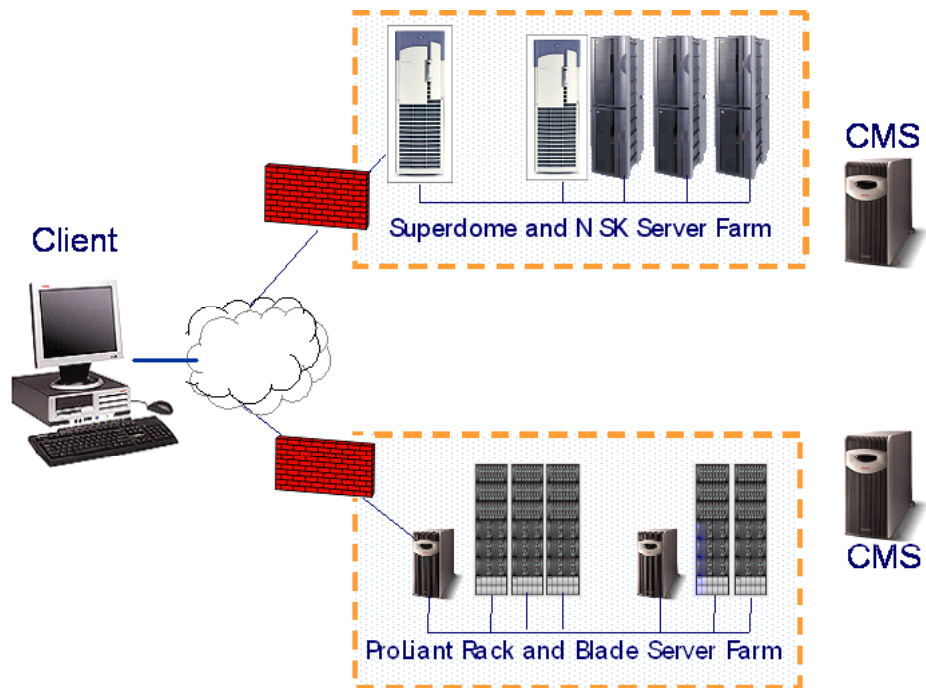


Data centers typically accommodate a variety of systems including servers, networking devices, printers, racks, and enclosures. To remove the limitation of the number of systems that can be managed, the CMS must be a dedicated server with Systems Insight Manager, a database, and VCRM installed. Although the network configuration includes two subnets, Systems Insight Manager can still discover and manage the systems.

The client, which can be used to access Systems Insight Manager on the CMS and manage software updates, requires only a web browser that has the Java JRE browser plug-in 1.4.1\_04 or later.

The requirements for managed systems are the same as for the branch office scenario. Managed systems running Windows must have Management Agents, OpenSSH, and the VCA installed. Linux and HP-UX managed systems do not require the VCA to be installed. However, the HP-UX configuration utilities are required on the CMS.

## Enterprise



When a business expands across geographic boundaries, managed systems must be behind firewalls with a CMS. In addition, VCRM must be installed on regional servers.

## Installing Systems Insight Manager

Before installing Systems Insight Manager, ensure that the system meets the installation requirements for the:

- CMS
- Systems Insight Manager database options
- Console browser

### CMS requirements for Windows

The CMS supports Windows, HP-UX, and Linux platforms. The following table lists the requirements for the Windows-based CMS.

Category	Minimum requirements
Hardware	ProLiant servers Compaq EVO desktops and workstations 1.5GHz Pentium processor on desktops or workstations
System memory	256MB RAM with VCRM and Microsoft SQL Server 2000 Desktop Engine (MSDE 2000) on the same server 1GB RAM with VCRM and SQL Server on the same server 2GB RAM or more recommended
Disk space	250MB Installation must be on an NT File System (NTFS) partition Local drive
Server operating system	Windows 2000 Server Service Pack (SP) 3 or later Windows 2000 Advanced Server SP3 or later Windows Server 2003 Standard Edition Windows Server 2003 Enterprise Edition Windows XP SP1 or later
Server software	TCP/IP SNMP services installed and active IPX for managing Novell NetWare servers that are not running IP Microsoft Internet Explorer 6.0
Database software	MSDE 2000 with SP3 or later SQL Server 2000 with SP3 or later SQL Enterprise Server 2000 with SP3 or later Windows authentication

## CMS requirements for Linux and HP-UX

The following table lists the requirements for the Linux- and HP-UX-based CMS.

Category	Minimum requirements
Hardware	ProLiant servers (Linux) HP PA-RISC 2.0 servers and Integrity servers (HP-UX)
System memory	1GB RAM with database on the same server 2GB RAM or more recommended
Disk space	110MB for Systems Insight Manager (/opt) 15MB for PostgreSQL (/var) 500MB minimum for data (/var/opt) 200MB incremental to previous swap space (500MB minimum)
Server operating system	Linux Red Hat Advanced Server 2.1 Linux Red Hat Advanced Server 3.0 SuSE Enterprise Server 8.0 (IPF) SuSE Enterprise Server 8.0 (UnitedLinux 1.0 kernel) HP-UX 11i v1.0 (11.11) HP-UX 11i v2.0 (11.23)
Server software	TCP/IP (Dynamic Host Configuration Protocol [DHCP] not supported) SNMP services installed and active DMI service installed and active Java SDK 1.41 or later X Server
Database software	PostgreSQL 7.2 or later

## Systems Insight Manager database options

Systems Insight Manager supports the following database applications:

- Windows
  - MSDE 2000 with SP3 or later
  - SQL Server 2000 Standard Edition with SP3 or later
  - SQL Server 2000 Enterprise Edition with SP3 or later
- Linux and HP-UX
  - PostgreSQL 7.2 or later

### MSDE 2000

MSDE 2000 is a free database engine based on and compatible with SQL Server technology. MSDE is also fully compliant with American National Standards Institute (ANSI) SQL and Transact SQL guidelines.

To install MSDE 2000 and Systems Insight Manager you need:

- 250MB of local hard drive space (not including the database)
- 256MB of system memory
- Windows 2000 Server or Advanced Server with SP3 or later, or Windows Server 2003 Standard or Enterprise Edition

Designed to run in the background, MSDE 2000 does not have a user interface; users interact with MSDE through Systems Insight Manager.

MSDE with SP3a is provided with Systems Insight Manager. After installing MSDE, check the Microsoft website for any new service packs or updates.

---

**INTERNET**

Microsoft service packs can be obtained at:

<http://www.microsoft.com/downloads>

---

### PostgreSQL 7.2

PostgreSQL is an open source, object-relational database management system. New versions that add significant features are released frequently. The PostgreSQL license is less constricting than commercial database applications.

You need 80MB of local hard drive space to install PostgreSQL.

## Console browser requirements

The following table lists the minimum requirements for installing the Systems Insight Manager console browser.

Category	Minimum requirements
Software	Windows NT with SP6a or later Windows XP SP1 or later Windows 2000 SP2 or later Windows Server 2003
Web browser	200MHz Pentium III or later Internet Explorer 6.0 or later Mozilla 1.4.1 for HP-UX or 1.4.2 for Linux
Java requirements	Java Plug-in 1.4
System memory	192MB RAM for Windows NT, Windows 2000, Windows XP, and Windows Server 2003
Monitor resolution	Minimum — 1024 x 768, 256 colors Recommended — 16-bit or better

The following browser options are usually enabled by default. You must ensure that they are selected for Systems Insight Manager to work properly:

- Enable Java
- Enable JavaScript
- Accept all cookies

## Systems Insight Manager installation tasks

Before installing Systems Insight Manager, complete these tasks:

- Review the system requirements.
- Review the release notes.
- Verify that an appropriate operating system is installed.
- Verify that TCP/IP and SNMP are installed on the server.
- Install a supported database.
- Create the database and tables on the database server.

After installing Systems Insight Manager, you must perform these tasks:

- Install the HTTP Web Agent (if needed).
- Install OpenSSH on managed systems that will be controlled remotely. It is not required to retrieve health and status information.
- Install VCRM to deploy software to managed servers. You can install it on the CMS or another Windows-based system on the network.

### WMI mapper

The WBEM Windows Management Instrumentation (WMI) mapper provides Windows support to WBEM applications. You need a Windows system to act as the WMI mapper proxy.

You can install WMI:

- On every managed node
- As a proxy on the Systems Insight Manager Windows CMS
- As a proxy on a different Windows server

If the CMS is running Windows, HP recommends using it as a proxy server. The WMI mapper proxy is automatically installed when a Windows server is configured as a CMS.

## Viewing the Systems Insight Manager Home page



The Introductory page, intended for novice users, displays when you first browse to Systems Insight Manager after installation. The term *Introductory page* is new to Systems Insight Manager, but its function is similar to the Quick Links section of the Insight Manager 7 Home page. It displays HP educational content and shortcuts and consists of areas that explain different specifics of the application.

## Customizing the Home page

### Home Page Settings

Description: Customize the "Home" link and the page displayed when HP Systems Insight Manager starts. Launch this page from the customize button on the introductory page or from the menu under Options->Home Page

Choose page to show when HP Systems Insight Manager starts and when I click on "Home".

☒ Introductory page
 

- ☒ Show Administrators tips for completing initial configuration
- ☒ Show the "Did You Know?" image

☐ Status Overview page

☐ This list:

OK

Systems Insight Manager provides Home page customization options that are accessible by clicking either:

- *Options → Home Page Settings* menu
- *Customize* link on the Introductory page

From the customization page, you can:

- **Select the desired destination for the Home link** — Choose from the Introductory page, Status Overview page, or a custom list as a target home page. By default, Systems Insight Manager loads the Introductory page after installation. You can subsequently access the Home page through the Home link in the banner area.
- **Disable the display of the Did You Know? image on the Introductory page** — By default, Systems Insight Manager displays this image on the Introductory page. It can be hidden through the Home Page Settings screen.

- **Disable the DO THIS NOW option to finish the install display on the Introductory page** — This option is only available to administrators. By default, Systems Insight Manager displays this section on the Introductory page until the initial configuration requirements have been met. These requirements include:
  - Setting up managed systems
  - Configuring protocol settings
  - Running discovery
  - Adding users
  - Configuring notifications (both email and paging settings)
  - Setting up automatic event handling
- **View system and event lists** — You can choose to have a list display instead of the Introductory page at startup and when the Home page link is clicked. Available lists are displayed by category. You cannot create a new list from this page, but only select preexisting lists.

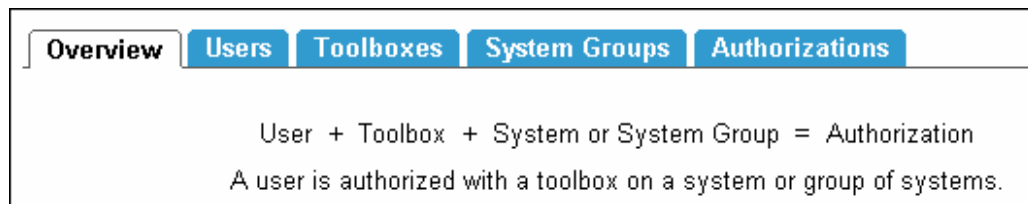
Home page preferences for each user are stored in the Systems Insight Manager database. When a user is removed from the database, the associated preferences are also deleted.

## Configuring Systems Insight Manager

After you install Systems Insight Manager, you must configure it to match your management needs. Regardless of the implementation, you must perform the following tasks when configuring Systems Insight Manager:

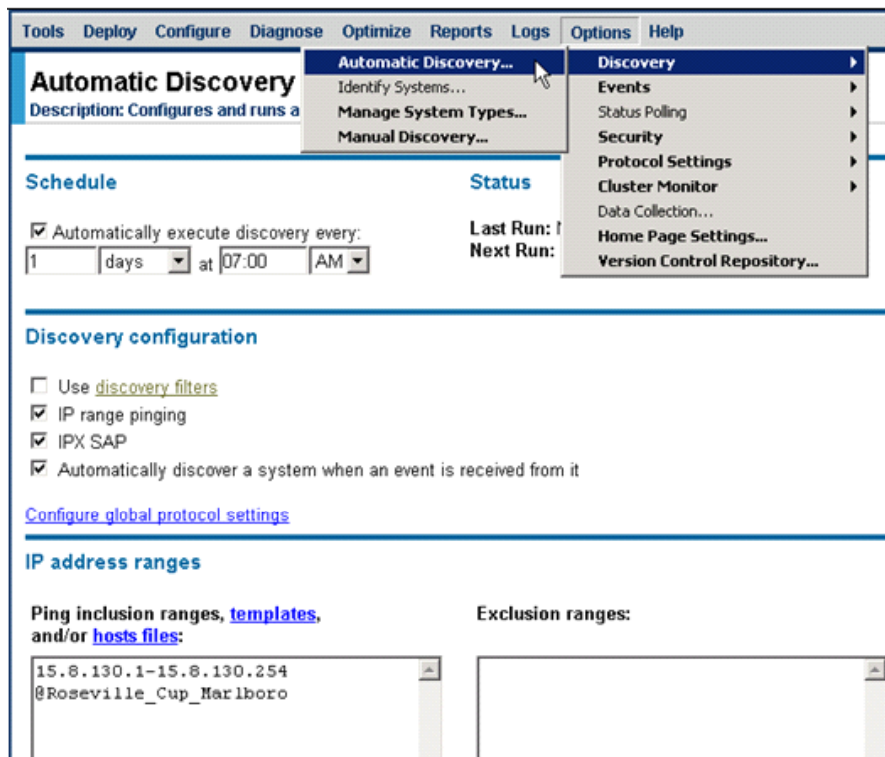
- Create and assign user rights
- Set up the next discovery
- Browse discovered devices
- Set up and customize status polling
- Receive notification of a problem
- Manage numerous devices
- Handle numerous events

### Creating and assigning user rights



Systems Insight Manager access is limited to authorized users. After a user is created, the role-based authorization is a combination of the user settings, the enabled toolboxes, and the system or system group the user is assigned to.

## Setting up the next discovery



To be managed by Systems Insight Manager, systems attached to the network must be discovered and identified. The information is then collected into a database for subsequent retrieval and activities such as lists, events, and tasks.

Discovery is the process of finding and identifying a device at a specific address on the network and collecting information about that device. Systems Insight Manager discovers and identifies devices on a network and maintains a database of information about them.

The discovery component is responsible for determining IP addressable systems on the network. It performs a simple ping sweep (Internet Control Message Protocol [ICMP] echo) of defined subnets. Any addresses that respond to the ping command are queued for system identification. You can also specify a list of systems to discover manually through a standard host formatted file or an eXtensible Markup Language (XML) file, or you can use a command line to add systems.

Systems also can identify themselves by sending events to the CMS. Examples include SNMP traps, WBEM indications, and HTTP events.

By default, Systems Insight Manager runs a discovery 24 hours after the previous discovery. You can run a discovery at any time from the Automatic Discovery option. You can also set up a schedule or change the scheduling to a more convenient time, add or subtract subnets, and choose the methods used to discover devices. If the network bandwidth is reaching its limits, you can optimize the protocol and ping settings to conserve resources.







## Browsing discovered devices

The System Page displays when you access the link for a system name. System links display as the result of queries such as events and the predefined and custom lists displayed in the left pane. The System Page displays the identity and status of a device, possible links the device may have, and its related events.

### Identity tab

**System Page**  
cmlab-2-1

Identity Links Events

General System Information	Status Information
<b>Address</b> 16.129.64.129	 <b>Hardware Status</b>
<b>System Name</b> cmlab-2-1	 <a href="#">from Insight Management Agents</a>
<b>DNS Name</b> cmlab-2-1.mro.cpqcorp.net	 <a href="#">from SNMP</a>
<b>Serial Number</b> D826BPV10043	 <a href="#">from Insight Management Agents</a>
<b>Management Protocols</b> DMI:2.0, SNMP:1.0	 <a href="#">from ping</a>
<b>Contact</b> Stanley Kou	 <a href="#">Software Status</a>
<b>Location</b> loc1-1127	
<b>System Type</b> Server	
<b>Product Model</b> ProLiant 2500	
<b>Hardware Description</b> x86 Family 6 Model 1 Stepping 9 AT/AT COMPATIBLE -	
<b>Software Description</b> Windows NT Version 4.0 (Build Number: 1381 Uniprocessor Free )	
<b>Server Role</b>	

The Identity page displays general information and the status of a managed system. A managed system can have a management processor and can be either in an enclosure in a rack or in a stand-alone unit. It can also be part of a cluster.

The status icon on the Identity tab indicates the hardware status stored in the database. The hardware status is collected from polling either the Insight Management Agent, DMI agent, WBEM agent, or with a ping command. The software status links to the software VCA of the device.

## Links tab

**System Page**  
owwpc029

Identity Links Events

**System Management Pages: owwpc029**

- [System Management Homepage](#)
- [Version Control Agent](#)
- [Configuration History Reports \(Survey Utility\)](#)
- [Version Control Repository Manager](#)
- [Insight Management Agents](#)

**System Web Application Pages: owwpc029**

- [Systems Insight Manager](#)

**HP Systems Insight Manager pages**

- [Data Collection Report](#)
- [System Protocol Settings](#)
- [Properties](#)

The Links page displays a list of website links that are related to the device. The sections on this page include:

- **System Management Pages** — System management and status
- **System Web Application Pages** — Web applications that the system hosts
- **HP Systems Insight Manager pages** — Links that are generated by Systems Insight Manager

## Events tab

**System Page**  
owwpc029

Identity Links Events

To view event details, make sure 'Event Type' column is displayed and click on desired link.

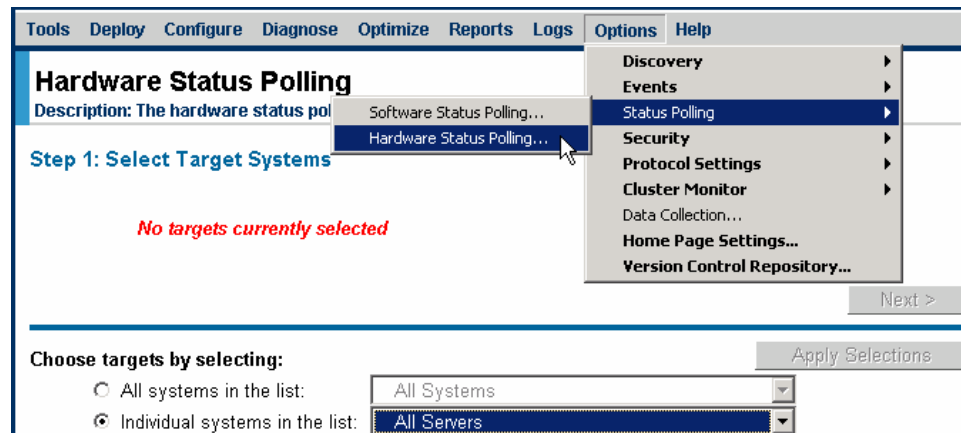
Events in table: 0 Critical 0 Major 0 Minor 1 Normal 113 Informational Total: 114

Se...	State	Severity	Event Type	System Name	Event Time	Assigned To	Comments
<input type="checkbox"/>	In Progress		<a href="#">Data Collection</a>	owwpc029	10/27/03 3:51 PM		
<input type="checkbox"/>	Not Cleared		<a href="#">Successful Login</a>	owwpc029	10/27/03 3:28 PM		
<input type="checkbox"/>	Not Cleared		<a href="#">Logout</a>	owwpc029	10/27/03 1:21 PM		
<input type="checkbox"/>	Not Cleared		<a href="#">Successful Login</a>	owwpc029	10/27/03 1:00 PM		
<input type="checkbox"/>	Not Cleared		<a href="#">Successful Login</a>	owwpc029	10/27/03 10:18 ...		
<input type="checkbox"/>	Not Cleared		<a href="#">Successful Login</a>	owwpc029	10/27/03 10:18 ...		
<input type="checkbox"/>	Not Cleared		<a href="#">Logout</a>	owwpc029	10/27/03 2:48 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Successful Login</a>	owwpc029	10/27/03 1:09 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Successful Login</a>	owwpc029	10/27/03 1:09 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Data Collection</a>	owwpc029	10/26/03 6:05 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Data Collection</a>	owwpc029	10/25/03 6:24 PM		
<input type="checkbox"/>	Not Cleared		<a href="#">Data Collection</a>	owwpc029	10/25/03 7:07 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Data Collection</a>	owwpc029	10/25/03 7:06 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Data Collection</a>	owwpc029	10/25/03 7:05 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Data Collection</a>	owwpc029	10/25/03 7:01 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Logout</a>	owwpc029	10/24/03 8:45 PM		

Clear Delete Assign To... Enter Comment... Print

The Events page lists all events associated with the system. Each underlined Event Type item is a link to more detailed information about that event.

## Setting up and customizing status polling



Status refers to the health of a device. Status polling returns health information from a device. Systems Insight Manager communicates with devices based on the protocol the device recognizes to retrieve status information.

Design your tasks so that the devices you are most interested in are polled most frequently and other devices are polled at a minimum. Always be aware that more traffic is generated when shorter polling frequencies are set.

You can also create polling tasks that monitor a specific group of devices based on a schedule.

### Example

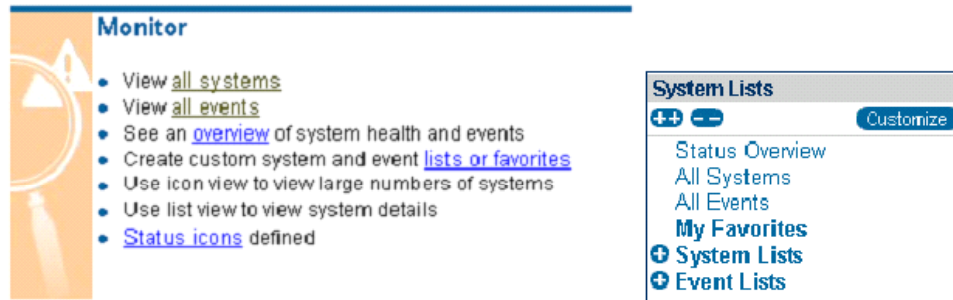
You might keep historical data on a certain type of device for later analysis and forecasting purposes.

## Receiving notification of a problem



If you are away from your console, you can use automatic event handling to set up email or paging notification or launching of a script. This script can be as simple as sounding an audible alert on management consoles.

## Managing numerous devices



An infrastructure with several subnets can have hundreds of devices to discover. Lists logically group the devices into collections based on information in the Systems Insight Manager database. Creating logical groups of systems reduces the number of systems viewed in a particular system list.

### Example

An enterprise has five system administrators who are responsible for 100 different systems in six locations. A list can be created for each administrator that includes only their systems or that includes only the systems located at a particular site.

Lists are shown by section and category. In addition to using the lists provided by Systems Insight Manager, you can also create, edit, or delete lists, or create categories of lists under each section to complement network management needs.

You can configure My Favorites to display your most frequently accessed lists. It can also contain categories, of systems, event lists, and cluster lists.

Public and private access levels do not apply in My Favorites because the Favorites folders are user-defined. Only the creator of the new Favorites folders can view or edit the hierarchy and contents.

Use the Customize button to tailor the System Lists tree. You can create, delete, or rename a section. You cannot use the following characters in category or list names: <, >, ", &, ', \_ , #, +, |, %, \, /, ;.

## Handling numerous events

An event is information sent to the administrator about something in the managed environment that has changed. Events are generated from SNMP traps or WBEM indications. Systems Insight Manager receives a trap when an important event occurs.

Systems Insight Manager reports the following severity levels for events:

- **Critical** — Indicates a failure and signals the need for immediate attention
- **Major** — Indicates an impending failure
- **Minor** — Indicates a warning condition that may escalate into a more serious problem
- **Normal** — Indicates proper system operation
- **Unknown** — Indicates status cannot be determined
- **Informational** — Provides useful information that does not require attention

## Monitoring managed systems

After systems are discovered on the network, they must be properly identified. Then they can be monitored using lists, tasks, events, and tools.

### System Type Manager

#### Manage System Types

Description: Create and manage rules to identify third party SNMP managed systems.

System type: All

Rules

Product model ↑	Product type	Subtype	Protocol
<input checked="" type="radio"/> 3Com Hub 10	Hub	None	SNMP
<input type="radio"/> 3Com Hub 40	Hub	None	SNMP
<input type="radio"/> 3Com Linkswitch 1000	Hub	None	SNMP
<input type="radio"/> AdvanceStack 12R with Manag...	Switch	None	SNMP
<input type="radio"/> AdvanceStack 16U with SNMP ...	Switch	None	SNMP
<input type="radio"/> AdvanceStack 24R with Manag...	Switch	None	SNMP
<input type="radio"/> AdvanceStack 24T with Manag...	Switch	None	SNMP
<input type="radio"/> AdvanceStack 8U with SNMP m...	Switch	None	SNMP
<input type="radio"/> AdvanceStack Switch 100	Switch	None	SNMP
<input type="radio"/> AdvanceStack Switch 200	Switch	None	SNMP

Total: 347

New...

Edit...

Delete

You can use the System Type Manager (STM) to customize the type and product name of systems using rules based on responses to SNMP and DMI (Windows only) queries from systems on the network. This procedure ultimately modifies the default behavior of the discovery and identification process.

Identification can be customized based on SNMP system object identifiers. Manufacturers assign unique system object identifiers to their SNMP-instrumented products. You can use the STM to create rules that map the system object identifiers to product categories and names of your choice. A DMI-based rule can be created using the CLI `mxstm` command.

DMI system information originates in Management Information Format (MIF) files. MIF files contain elements that have attributes and corresponding values. The information in the MIF files is the information used in the STM when the rule was created. The STM can request a value from a specific target for a specific attribute.

Only users with full rights can use the STM. This feature is integrated into Systems Insight Manager under the *Options* → *Discovery* → *Manage System Type* menu.

## Rules list

At a minimum, a rule consists of a system object identifier, device type, and product name. In addition, the community string of the target system must match the community string used to query the system. The default query community string is *public*.

Invoking the STM from the menu brings up the list of rules currently defined in the Systems Insight Manager database. The list can be filtered by system type by selecting a type from the drop-down list. SNMP rules can be added, edited, or deleted in the browser interface. However, DMI rules can only be added or deleted on Windows systems from the command line using the `mxstm` command.

A rule can be modified to edit the product type, subtype, priority, and URL to be displayed on the device page. When a rule is deleted, priorities of any remaining rules with the same system object identifiers (OIDs) are reordered appropriately.

## SNMP categories

Systems also supply information about themselves using variables described in Management Information Bases (MIBs). These variables are enumerated using an industry-standard structure.

MIBs are files provided by vendors for their systems and must be registered with Systems Insight Manager to be accessible and usable from the STM. HP preregisters all HP MIBs and many third-party MIBs. Remaining MIBs can be registered using the MIB compiler, if the related systems exist on the network.

A MIB contains modules or groups of variables. Some variables have multiple values. Each of these values has an object identifier as well. Systems Insight Manager queries these object identifiers to determine the system and its current behavior.

## DMI categories (Windows only)

DMI identification is based on how a system responds to a DMI request. Systems supply information in MIF files. The presence of an MIF file on a target system does not guarantee DMI identification.

MIF files cannot be registered the way MIBs are registered in Systems Insight Manager. An MIF file (for example, the generic Win32sl.MIF) consists of groups of attributes. The values returned in response to requests for MIF attributes can be used to determine which system you have and its current behavior.

### Example

The following extract is part of the Win32sl.MIF file. Notice that the group named *ComponentID* is followed by several attributes that identify one aspect of a DMI system, such as manufacturer, product, version, and serial number.

```
Start Group
Name = "ComponentID"
ID = 1 Class = "DMTF|ComponentID|001"
Description = "This group is required."
Start Attribute
Name = "Manufacturer"
ID = 1 Description = "Manufacturer of this system."
Access = Read-Only
Storage = Common
Type = String(64)
Value = "Intel Corporation"
End Attribute
Start Attribute
Name = "Product"
ID = 2
Access = Read-Only
Storage = Common
Type = String(64)
Value = "Win32 DMI Service Layer"
End Attribute
Start Attribute
Name = "Version" ID = 3
Description = "Version number of this component."
Access = Read-Only
Storage = Common
Type = String(64)
Value = "2.32"
End Attribute
Start Attribute
Name = "Serial Number" ID = 4 Access = Read-Only
Storage = Common Type = String(64)
Value = "unsupported"
End Attribute
```

Other MIF files have different groups and specify other aspects of a system. The information in the MIF files is the information provided to the STM when a rule is created. The STM can request a value from a specific target for a specific attribute.

## Managing hosts files

### Manual Discovery - Add System

Description: Add a single system to be managed by HP Systems Insight Manager

To add multiple systems, see the [Hosts Files](#) page. To add systems automatically, see the [Automatic Discovery](#) page.

Hosts files are created and imported to automate the process of adding devices or restoring device information.

You can bypass the need for an immediate discovery by importing a hosts file. For example, in the case of a catastrophic system failure, you can import a backup hosts file as the basis for reconfiguring your management environment and automatically repopulate a database.

Adding the devices using the hosts file utility does not replace devices in the database.

#### Example

If a device listed in the hosts file has the same IP address as an existing device, the duplicate is ignored. Any devices that previously existed in the database are not modified.

Add any IP address range to a hosts file using the same format as entries used in configuring a discovery.

You can import hosts files from the following sources:

- The Systems Insight Manager database that imports the device data, creates a hosts file, and sorts the data types according to your selection
- A client device that has an existing hosts file
- A client device that has a hosts file created by Systems Insight Manager

Hosts files can be created, edited, deleted, and executed on the Manual Hosts Files page of Systems Insight Manager. You can access this page from the Settings menu by selecting *Options* → *Discovery* → *Manual Discovery* → *Hosts Files*.

## MIB importing and compiling

Systems Insight Manager attempts to discover all the devices in the network. After it discovers a device, it attempts to identify it by asking a series of questions. The answers to the questions are compared to the expected answers stored in the Systems Insight Manager database. In this way, discovered devices are classified.

Using Systems Insight Manager, you can easily identify HP products. However, most networks contain devices from multiple vendors. MIB importing and compilation tools can be used to extend the identification capability of Systems Insight Manager to third-party devices.

The Systems Insight Manager MIB Importer includes components that orchestrate the management of MIB data so the data can be used to display event messages to the user in a meaningful manner. The parts of the MIB Importer are:

- MIB compiler
- MIB control layer
- MIB configuration

The clean separation of compiling MIBs compared to importing MIBs into the system solves two problems:

- Errors in a MIB compile are cleanly delineated from writing the data to the database.
- No rollback or commit scheme is necessary because imported MIBs are deemed to be correct.

### MIB importer CLI

The MxEvent command is used to perform a number of MIB utility functions such as importing new MIBs manually, processing a list of MIBs to import, viewing registered traps, and listing imported MIBs.

The following CLI switches are available with the MxEvent command:

- `MxEvent -i [-n "trapHandlerClass"] <mibfile.mib>` — Imports the configuration file for the MIB and then installs an associated trap handler for each trap associated with this MIB
- `MxEvent -f <filelist.txt>` — Imports all configuration files listed in a text file
- `MxEvent -d <mibfile.mib>` — Deletes all MIB-related data with respect to a specific MIB
- `MxEvent -l` — Lists all the MIBs installed in the system
- `MxEvent -t <mibfile.mib>` — Shows all the traps of a particular MIB installed in the system

## MIB compiler

The MIB compiler is V1/V2c compliant and performs all the semantic and syntactic checking of the MIB. If there are no errors, it creates a configuration file in the MIB directory.

---

### Note

The MIB compile is separate from Systems Insight Manager so it can be run at any time without adverse performance.

---

The MIB compiler is invoked through the `mcompile` command. On Windows platforms, the `mcompile` command runs from a DOS prompt, and on HP-UX and Linux platforms it runs from a shell. The compiler exports a configuration file that describes the pertinent event information.

### Example

`mcompile cpqrack.mib` — Compiles a MIB and exports the configuration file into the MIB directory

`mcompile -f somefile.list` — Compiles a list of MIBs and creates a configuration file for each MIB

If the MIB being compiled imports definitions from other MIBs, the imported MIBs must be located in the same directory as the MIB being compiled so that the `mcompile` command can resolve the imported definitions successfully.

You can direct the MIB compiler to look for the MIB file and place the configuration file in a specific location such as the Systems Insight Manager database. Files can be uploaded to the file directory using the `-d` command line switch. When a file is placed in the Upload file directory, the system automatically imports the configuration file at system startup or when directed through the `mxmib` command.

## MIB control layer

The MIB control layer searches the Upload file directory at system initialization for files to import. The files that exist in this directory will be read and parsed for keywords, and the data will be used to populate or replace MIB data in the database tables.

You can invoke actions in the MIB control layer either through a CLI using the `mxmib` command or through the browser.

The `mxmib` command enables MIB data to be listed, deleted, or imported into the system. The browser interface enables a user to view the traps registered in the system and modify certain trap-specific data that is stored in the database.

## MIB configuration

MIBs and the original configuration files reside in the MIBs directory. You can edit the configuration file to customize the event information using an ASCII editor. After you make the edits, the configuration file contains all the MIB data and the HP value-added messaging. When it is finished, you can use the `mxmib` command to import the configuration file into the system.

When you need to add new MIBs, they are compiled. When syntactically correct, a configuration file is generated. You can then edit the configuration file to create customized messaging for the user-defined fields. After you have edited the file, you can use the `mxmib` command to import it into Systems Insight Manager.

### Configuration file format and keywords

Rules for the configuration files include:

- There can be only one MIB per configuration file.
- The configuration file must have the same name as the MIB but with a configuration extension.
- Before any definitions, the configuration file lists the MIB file name or type such as WBEM indications.
- Following the MIB file name, a module name is listed.
- The module name is followed by object-type data and trap-type data.

You can modify the following keywords and their associated data in the configuration file:

- **DESCRIPTION** — Created from the original MIB. Modify this keyword to enhance the description data displayed for each event.
- **MSG\_FORMATTER** — Used to format a message that is used to format strings sent to a pager or email.
- **SEVERITY** — Created from the MIB but can be modified to change the severity of the event.
- **TYPE** — Used as a short description.
- **ENABLE** — Used to enable or disable a trap.
- **CATEGORY** — Used to group traps. Use this keyword as a selection mechanism to make it easier to find a type of trap that falls into a category.
- **RULE\_HANDLER** — Used to select a specific case rule handler such as rack traps.
- **NOTICE\_HANDLER** — Used to select a special trap handler such as rack traps.

Keywords display after trap type definitions and are preceded by `--#` characters.

## Tasks

Tasks are operations performed against groups of managed systems or events. All tasks are based on queries and are therefore self-updating. When you add a new system to the managed network, it is automatically added to the appropriate set of tasks. You can schedule tasks to happen immediately, periodically, or at some specified time in the future.

Status polling tasks track the status of the systems in System Lists. System status polling must occur continuously to determine when systems go offline or performance degrades. Systems Insight Manager provides the following default polling tasks:

- **Software Status Polling** — Collects status information for systems. This task:
  - Retrieves software and firmware inventory from systems
  - Determines the software and firmware update status
  - Sorts versions in the databaseBy default, this task is set to run every seven days, on Wednesday at midnight.
- **Hardware Status Polling** — Used to track system status. There are two types of Hardware Status Polling tasks:
  - **Hardware Status Polling for Non-Servers** — Used to collect status information for SNMP systems that are not servers, clusters, or management processors. By default, this task is configured to poll every 10 minutes and at startup. It does not send status change events.
  - **Hardware Status Polling for Servers** — Used to collect status information for SNMP servers, clusters, and management processors. By default, this task is configured to poll every five minutes and at startup.

## Task rules

The following rules govern the use of tasks in Systems Insight Manager:

- The user who initially creates the task is the editor and creator. The creator of the task never changes.
- The creator of a task and users with the task modification privilege can modify or delete that task. However, the task is always owned and run by its creator based on the original scope.
- If a user is removed from the Systems Insight Manager database, all tasks owned by that user must be removed or reassigned to another user.
- The last user to edit a task becomes the executor of that task when it is run according to schedule. If a user manually executes a task, then the user is the executor of that task. Because of enhanced security features, the executor of a task is shown in the Event Details page for tasks that track status.
- Only administrators can stop an executing task. The only tasks that can be stopped are time-scheduled tasks and manually executed tasks.
- The predefined SNMP Status Polling task is designed to work over a LAN connection. On a WAN connection, some customization is required.
- If you delete a task, the task is permanently deleted from the database and cannot be recreated.
- Default Systems Insight Manager tasks cannot be deleted.
- The list of task instances is based on user privileges and access levels. Users with full configuration rights can view all task instances known to the system.

## Predefined tasks

Systems Insight Manager includes several predefined tasks. Some tasks are associated with a plug-in application tool. Other tasks modify the configuration of target systems.

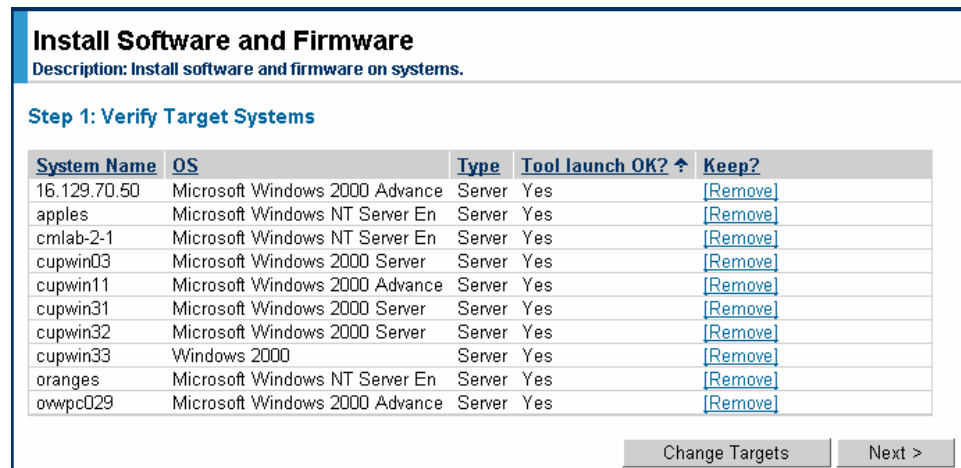
The following table lists the tasks that are associated with a plug-in application tool. Most of the tasks listed use HP OpenView Storage Area Manager applications that enable you to perform system administration tasks on HP-UX systems.

Menu option	Submenu	Task
Deploy	Software Distributor	CLI List Software
		CLI Preview Install
		CLI Verify Software
		Set SD Access
		View Depot Software
		View Installed Software
	Red Hat Package Manager (RPM)	Install
		Query
		Uninstall
		Verify
	License Manager	Collect Keys
		Deploy Keys
Configure	HP-UX Configuration	Accounts for Users and Groups
		Auditing
		Disks and File Systems
		Kernel Configuration
		Kernel Configuration (KCWeb)
		Peripheral Devices
		Printers and Plotters
		System Properties
		System Security Policies
		Verified Commands
	Partition Management	Partition Manager
		View and Manage Complex
		Create Partition
		Create nPartition
		Show Complex Details
	Diagnose	Event Monitoring Service
Optimize	Process Resource Manager	Process Resource Manager Console
		Display Resource Usage
		List Resource Availability

Systems Insight Manager also includes integrated tasks that enable you to configure or monitor target systems.

- **Install Software and Firmware** — Provides software update capabilities that use the VCA and VCRM.
- **Initial ProLiant Support Pack Install** — Installs a ProLiant Support Pack (PSP) on a Windows system that does not have Insight Management Agents installed and configures the system to use the trust certificate from the CMS and the setting for the desired VCRM repository.
- **Install OpenSSH** — Copies a Secure Shell (SSH)-generated public key from the CMS to target systems
- **Set Disk Thresholds** — Sets the threshold on all disk volumes on target systems
- **Remove All Disk Thresholds** — Removes disk thresholds set by Systems Insight Manager or through the Insight Management Agents
- **Replicate Agent Settings** — Enables Systems Insight Manager to retrieve and edit web agent configuration settings from a source system, and distribute that configuration remotely to target systems through web agents.
- **Ping** — Enables Systems Insight Manager to issue a ping command to an individual system or multiple systems

## Task Wizard



The Task Wizard enables you to create a task that can be run against one or more targets. You can launch a task by selecting the:

- Objects first and then the action to be performed
- Action first and then the objects on which to perform the action

With Systems Insight Manager, you can launch tools to run as a single execution or as multiple tasks. This process includes:

- Prompting for system or event selection
- Verifying selected systems or events
- Allowing changes to the system or event selection before running the task
- Requesting tool parameters
- Scheduling the task to run at a later time (periodic or specific)
- Scheduling the task based on a set of systems or events based on certain criteria
- Running a tool without a schedule

After you create a task, you can edit it to modify the target systems or the list currently configured for the task or owner.

### Note

Modifying the target systems of a task causes the task to run under the new owner's privilege and access level. The task is disabled automatically until the new owner edits the task to give it proper access to the tool, target systems, or target list.

## Scheduling a task

initial support pack install [set aside](#) [open in a new window](#) [HELP](#)

**Selected Node(s):** This is a comma separated list of the selected nodes...

---

**Step 6: Schedule Task**

**Name for this Task:**  **Refine schedule:**

**When would you like this Task to run?**

- ☐ Periodically
- ☐ Once
- ☒ Not scheduled

**In Addition:**

- ☐ Run when the Management Server is started
- ☐ Run now
- ☐ Disable this Task

Tasks can be scheduled:

- **Periodically** — Run on specific dates and times and execute a set number of times
- **Once** — Run on a specified date and time
- **At server startup**
- **Manually** — Only run when manually executed by a user with the appropriate privileges

A task can also be disabled prevent it from running regardless of the scheduling options.

If a System List is selected for a task, the task can be scheduled to only run when:

- New systems or events meet the list criteria
- Systems or events no longer meet the list criteria

## Time filters

initial support pack install

Selected Node(s): This is a comma separated list of the selected nodes...

Step 6: Schedule Task

Name for this Task:

When would you like this Task to run?

☐ Periodically

☐ Once

☒ When new systems or events meet the list criteria

☐ When systems or events no longer meet the list criteria

☐ Not scheduled

In Addition:

☐ Run when the Management Server is started

☐ Run now

Refine schedule:

☒ Use time filter:

Always

☐ Disable this Task

< Prev Done

Manage Time Filters

Filter Name:

New Edit Copy Delete

	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11
Mon																								
Tue																								
Wed																								
Thu																								
Fri																								
Sat																								
Sun																								

Apply OK Cancel

You can use time filters to limit a task to run only during certain hours. This feature is useful for notification tasks such as paging.

### Example

You can set two paging tasks to run by clicking *When new systems or events meet the list criteria*. One task can page the day-shift administrator and use a time filter of *Business Hours*. Another task can page the night-shift administrator and use a time filter of *Nights & Weekends*. Only one of them will be paged when an event occurs.

A time filter further defines the task schedule configuration. Systems Insight Manager supplies built-in time filters. In addition, you can create, copy, edit, and delete user-created time filters. The built-in time filters cannot be deleted or edited.

### Example

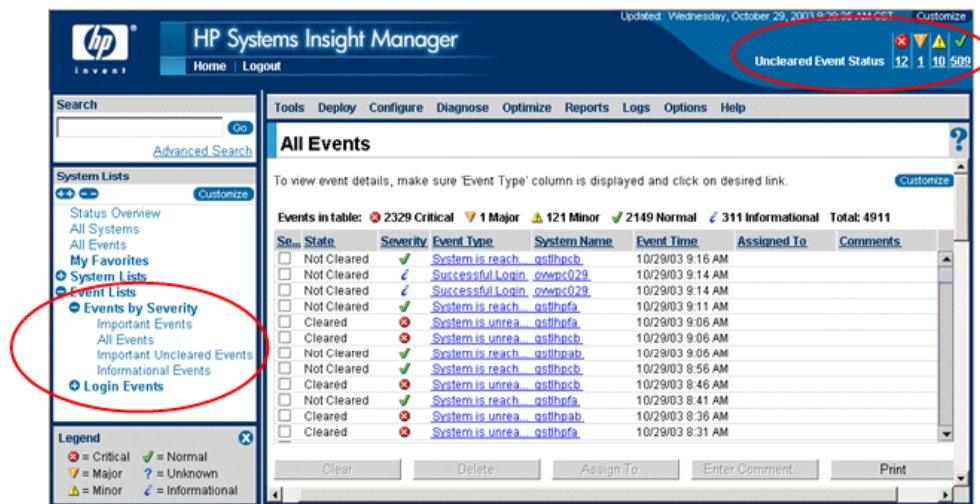
You can schedule a task to run every hour. However, you can apply a time filter to prevent the task from running during business hours.

## Events

You must configure managed systems to send appropriate alerts to the CMS. This process requires configuring the SNMP trap destination and community name. WMI does not send indications and currently, HP-UX operating system versions do not send WBEM indications.

There are several cases where the source of the event and system associated with the event is the CMS itself. Events generated on the CMS include all login, logout, and failed login events. Everyone can view these events in the event list, but only users who have the necessary authorizations can see the details of these events.

## Displaying events



The event list page is accessed from the following links:

- Predefined event lists or user-defined events from the System Lists pane
- Uncleared Event Status in the banner area
- All Events Associated with this System from the System Page of the system

When events are received in Systems Insight Manager, they are logged in the database automatically and displayed across the banner of the browser interface.

Event attributes include:

- Severity
- Node
- Event type
- Event state
- Time received

Event lists are filtered based on authorizations—only events for systems on which the user has an authorization display in the event list.

## Events list details

**All Events**

To view event details, make sure 'Event Type' column is displayed and click on desired link. [Customize](#)

Events in table: 2335 Critical 1 Major 121 Minor 2153 Normal 311 Informational Total: 4921

Se...	State	Severity	Event Type	System Name	Event Time	Assigned To	Comments
<input type="checkbox"/>	Not Cleared		<a href="#">System is reach... gsthpdpa</a>		10/29/03 10:21 ...		
<input type="checkbox"/>	Cleared		<a href="#">System is unrea... gsthpdpa</a>		10/29/03 10:17 ...		
<input type="checkbox"/>	Not Cleared		<a href="#">System is reach... gsthpfpa</a>		10/29/03 10:16 ...		
<input type="checkbox"/>	Cleared		<a href="#">System is unrea... gsthpfpa</a>		10/29/03 10:11 ...		
<input type="checkbox"/>	Not Cleared		<a href="#">System is reach... gsthpi1</a>		10/29/03 10:06 ...		
<input type="checkbox"/>	Not Cleared		<a href="#">System is unrea... gsthpi1</a>		10/29/03 10:06 ...		
<input type="checkbox"/>	Cleared		<a href="#">System is unrea... gsthpi1</a>		10/29/03 10:01 ...		
<input type="checkbox"/>	Not Cleared		<a href="#">System is unrea... 16.101.170.45</a>		10/29/03 9:56 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">System is reach... gsthpfpa</a>		10/29/03 9:46 AM		
<input type="checkbox"/>	Cleared		<a href="#">System is unrea... gsthpfpa</a>		10/29/03 9:36 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">System is reach... gsthpcb</a>		10/29/03 9:16 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Successful Login_owwpc029</a>		10/29/03 9:14 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">Successful Login_owwpc029</a>		10/29/03 9:14 AM		
<input type="checkbox"/>	Not Cleared		<a href="#">System is reach... gsthpfpa</a>		10/29/03 9:11 AM		
<input type="checkbox"/>	Cleared		<a href="#">System is unrea... gsthpfpa</a>		10/29/03 9:06 AM		
<input type="checkbox"/>	Cleared		<a href="#">System is unrea... gsthpcb</a>		10/29/03 9:06 AM		

Clear Delete Assign To... Enter Comment... Print

The events list includes several columns:

- **Selection** — Enables you to run context-sensitive tools on systems on which critical events occurred.
- **State** — Displays whether the event is in the Cleared or Not Cleared state. Events start in an Uncleared or In Progress state and transition to the Cleared state. The Cleared state indicates that the user is aware of the event and it is considered completed. A Pending event indicates that not all the data for the event has been logged. These events cannot be removed or cleared. A restart of the CMS moves any Pending states to an Uncleared state.
- **Severity** — Displays the event status icon to indicate the level of the problem associated with the event.
- **Event Type** — Displays event types such as SNMP traps, login failures, or when a managed system is unreachable.
- **System Name** — Displays the system name on which the event occurred. Clicking the link displays the System Page of the system.
- **Event Time** — Displays the time stamp when the CMS received the event.
- **Assigned To** — Displays a name to which the event is assigned.
- **Comments** — Displays comments entered by a Systems Insight Manager user for an event.

## Managing events

Event management procedures include:

- Clearing events
- Assigning events to users
- Entering comments on events
- Deleting events

You can access these events and begin to troubleshoot the problem by reading the details on the System Page.

### Example

A hard drive failure alert is received in Systems Insight Manager. The event is logged and the Critical Uncleared Events value in the upper-right corner of the Systems Insight Manager page increases. Click the link to see the event and drill down on the server to see what has occurred.

After the drive has been replaced, you can access the event again to clear it. The cleared events remain in the database until they are deleted. You can delete the event by selecting it and clicking *Delete* on the All Events page or with an automated task that deletes events based on certain criteria.

You can assign an owner of the event by selecting one user name from the list or by creating a new user name. The users listed are users to whom responsibility was assigned previously. Assigning events enables you to keep track of who is responsible for handling the associated problem.

You can enter a comment with up to 1,000 characters in the Comments field. You can also paste text in this area, provided that it does not exceed the maximum number of characters allowed.

---

### Note

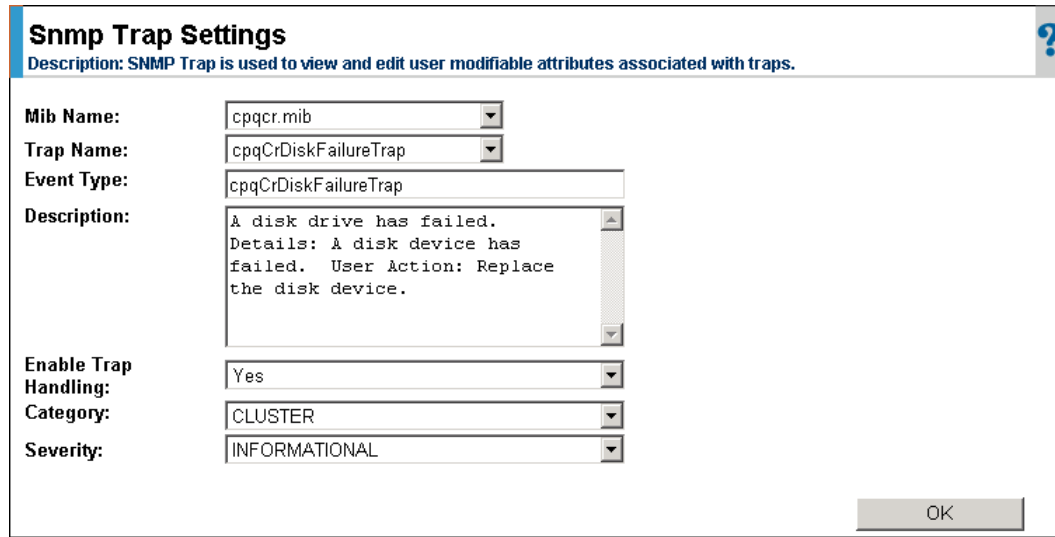
You cannot delete pending events, discovered system events, or service events.

---

When an event occurs that affects an entire rack or enclosure, several systems in that rack or enclosure might generate a trap for the event. These container traps are filtered such that only one event will be logged per rack or enclosure trap.

Although the source of the trap is a server blade or management processor, Systems Insight Manager sets the Event Source and Associated Device for the logged event to the rack or enclosure, as appropriate.

## SNMP Trap Settings



**Snmp Trap Settings**

Description: SNMP Trap is used to view and edit user modifiable attributes associated with traps.

Mib Name: cpqcr.mib

Trap Name: cpqCrDiskFailureTrap

Event Type: cpqCrDiskFailureTrap

Description: A disk drive has failed.  
Details: A disk device has failed. User Action: Replace the disk device.

Enable Trap Handling: Yes

Category: CLUSTER

Severity: INFORMATIONAL

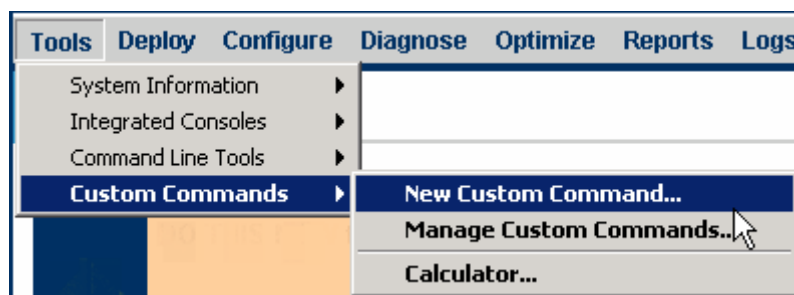
OK

SNMP traps are events generated by a managed system when an irregular condition exists. Trap messages can be cryptic, poorly written, and incomprehensible. The SNMP Trap Settings page enables you to customize trap messages by modifying the MIB information in the database representation.

The SNMP Trap Settings page displays by selecting *Options* → *Events* → *SNMP Trap Settings*. From this page you can modify the following trap data:

- Event Type
- Description
- Enable Trap Handling
- Severity

## Tools



The New Custom Command option enables Systems Insight Manager to run any application on the CMS. Users have the same permissions in the application as they do in Systems Insight Manager.



### Caution

If there is no need for Systems Insight Manager to launch external applications, disable this feature to protect the system from running potentially damaging applications.

Custom commands require the use of environment variables, which are parameters passed to the launched application to make it perform as expected. The command string can include system variables and user-defined variables for the application.



### Important

The application must be able to execute in the security context provided to Systems Insight Manager.

When a new custom command is created, it is added to the Systems Insight Manager menu list of tools on the Manage Custom Commands page. This page displays the name of the tool with the related description, command, and parameters. When you select a command, you can edit, delete, or run the application, or schedule the application to run at a specific time.

You must authorize custom commands in order to run them successfully. Add the application to the toolbox of each user to provide the necessary authorization to access the command.

## Application plug-ins

Systems Insight Manager provides the architecture for integrating applications into the menu structure. This feature offers several levels of compatibility and integration. At its most basic level, a menu entry can be added to launch an application. The next level allows some type of integration such as receipt of events, and the third level is where Systems Insight Manager takes action based on events from the application.

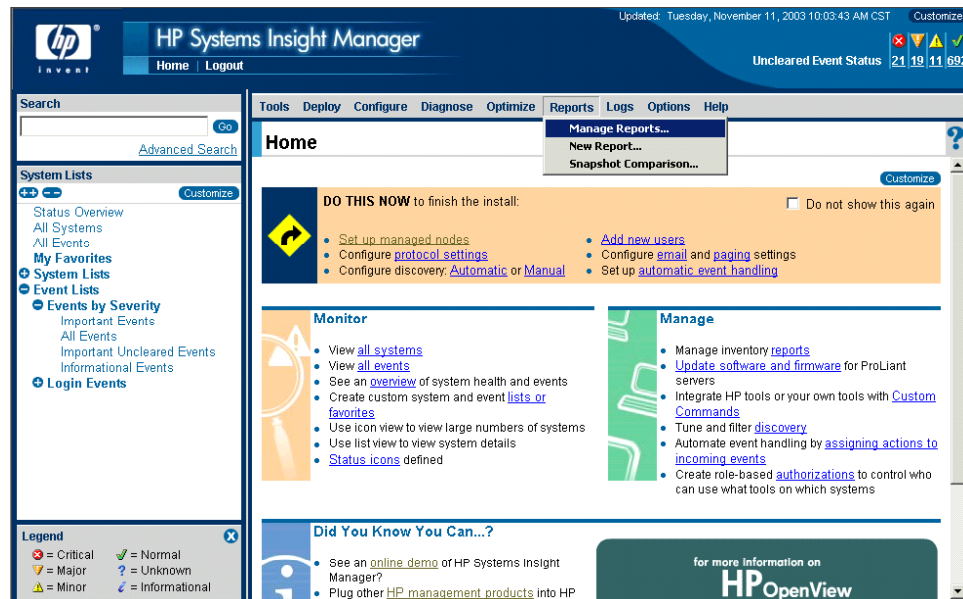
The `mxtool` command supports reading and writing tool definitions in the XML format. This capability can be used to share tool definitions between different Systems Insight Manager environments and to modify tool definitions within a Systems Insight Manager environment. Each file can contain the definitions of one or more tools.

Systems Insight Manager provides the following tools that you can run on target systems:

- **Single-system aware (SSA)** — SSA tools are executed on a target system and are only aware of the target system environment. When executing an SSA tool, the Distributed Task Facility (DTF) sends the tool information to each Systems Insight Manager agent to execute the tool. An example of an SSA tool would be a tool that wraps a common UNIX command, such as `ls`, `cat`, or `cp`.
- **Multiple-system aware** — multiple-system aware tools are executed on a central system, sometimes the CMS, and can interface with a list of target systems. An example of a multiple-system aware tool is the software distributor that installs software and PSPs.
- **Web launch aware (WLA)** — WLA tools are executed in a browser and are specified by a URL. For example, a WLA tool can access an embedded web agent such as Command View on a managed system or an embedded web server on a switch.

The DTF controls the execution of tools on target systems.

# Creating reports in Systems Insight Manager



Systems Insight Manager reporting provides an easy-to-use reporting wizard that enables you to select the set of system data you want to display and report on. You can access the reporting menus within any page in Systems Insight Manager.

You can select a system list or a set of devices in the System List and then create a new report or run a standard report based on the selected items.

When you select the New Report menu, the New Report reporting wizard displays and the system list or the set of devices selected is reflected in the Select Targets page of the reporting wizard.

## Managing reports

**Manage Reports**  
Description: Create, copy, delete, modify or run reports against central management server database

Report Name ↑	Target Systems
<input checked="" type="radio"/> Array Controllers - Servers	All Servers
<input type="radio"/> CPU - Clients	All Clients
<input type="radio"/> CPU - Servers	All Servers
<input type="radio"/> DIMM Slots - Servers	All Servers
<input type="radio"/> Installed Controllers - Clients	All Clients
<input type="radio"/> Installed Controllers - Servers	All Servers
<input type="radio"/> Inventory - Clients	All Clients
<input type="radio"/> Inventory - Servers	All Servers
<input type="radio"/> Logical Disk Drives - Servers	All Servers
<input type="radio"/> Network Interface - Servers	All Servers
<input type="radio"/> Operating System Information - Clients	All Clients
<input type="radio"/> Operating System Information - Servers	All Servers
<input type="radio"/> Physical Disk Drives - Servers	All Servers
<input type="radio"/> Power Supply - Servers	All Servers
<input type="radio"/> ProLiant BL Server Rack - Servers	All Servers
<input type="radio"/> System License Information - Systems	All Systems
<input type="radio"/> System Software - Servers	All Servers

The Manage Reports page displays all the predefined and user-defined reports. It enables you to select a specific report in the list to run, copy, edit, or delete. You can also create a new report.

All users can run a report. You must have full or limited access to copy, create, edit, or delete a report. The associated options are grayed out if the access level requirement is not met.

## Creating reports

The New Report wizard displays three pages:

- Select Targets
- Verify Targets
- New Report

Whether the New Report wizard is invoked from the New Report menu, Manage Reports menu, or Manage Reports page, the selected system list or set of devices from the System List page is carried over and reflected in the Select Targets page.

When the Select Targets page initially displays, no default report name is displayed in the report name box, and none of the categories or data items are selected in the tree view box. The Run Report button is grayed out initially because no report has been selected. The Save Report button is also grayed out until a report name is provided and the categories or data items are selected.

Use of this page is governed by the following rules:

- You must name a report to save it and run it. The maximum length for the report name is 150 characters.
- You can run a report after you enter a report name or make a selection in the tree view box. If you do not specify a report name, the title of the report displays *Unnamed Report*.

## Copying reports

Use the Copy Report options to make a copy of an existing report before making changes. Modifying a predefined layout is easier than creating an entirely new report.

## Editing reports

The primary purpose of editing a report is to make changes to the selection of data items and change the view option. Because the Edit Report wizard is invoked from the Manage Reports page, the selected system list or individual systems from Systems Insight Manager are not carried over to the Edit Report wizard. Instead, the original system list or list of devices associated with that report is reflected.

## Deleting reports

Users with full and limited configuration rights can delete any reports (including the predefined reports). However, a user with no configuration rights cannot delete reports. If a predefined report is deleted, it cannot be restored.

## Viewing reports

CPU - Servers			
<b>Associated system list: All Servers</b>			
Report date and time: Sun, 15-Feb-2004, 3:28 PM PST			
<a href="#">Export to file in CSV format</a>			
<a href="#">Show SQL queries</a>			
CPU			
System Name	CPU Type	CPU Speed (MHz)	Slot Number
16.101.170.1	Pentium III	1133	1
16.101.170.10	Pentium III Xeon	900	0
16.101.170.10	Pentium III Xeon	900	0
16.101.170.10	Pentium III Xeon	900	0
16.101.170.10	Pentium III Xeon	900	0
16.101.170.137	Pentium III	800	0
16.101.170.137	Pentium III	800	0
16.101.170.16			
16.101.170.16			
16.101.170.19	Pentium III Xeon	933	0

Report results display in a separate browser window in a tabular form and provide a time stamp indicating when the report was generated. The report page enables you to export the report results to a comma-separated value (CSV) file or to display the SQL queries used to generate the report.


Clicking a column heading link sorts the entries in that column. The default sort order is based on the system name.

### Export to file in CSV format

	A	B	C	D	E	F	G	H	I
1	DeviceName	CPUType	CPUSpeed	SlotNumber					
2	16.101.170.1	Pentium III	1133	1					
3	16.101.170.1	Pentium III	900	0					
4	16.101.170.1	Pentium III	900	0					
5	16.101.170.1	Pentium III	900	0					
6	16.101.170.1	Pentium III	900	0					
7	16.101.170.1	Pentium III	800	0					
8	16.101.170.1	Pentium III	800	0					
9	16.101.170.1	Pentium III	933	0					
10	16.101.170.1	Pentium III	933	0					

The Export to file in CSV format link enables you to export the report to a .csv format file. If Microsoft Excel is installed on the system in which Systems Insight Manager is accessed, the file can be opened to display in Excel.

## Show SQL queries

**CPU - Servers** 

SQL QUERY USED TO GENERATE THE DATA IN THE TABLE: CPU

```
select R_CPU.DeviceName, R_CPU.CPUType, R_CPU.CPUSpeed, R_CPU.SlotNumber from
R_CPU WHERE ((R_CPU.SnapShotID=-1) OR (R_CPU.SnapShotID is NULL)) AND DeviceKey
in (select devices.deviceKey from devices JOIN notices ON devices.deviceKey =
notices.deviceKey and notices.NoticeType = 1 JOIN consolidatedNodeAuths ON
devices.mxGuid = consolidatedNodeAuths.NodeId and consolidatedNodeAuths.userId =
'00000000864162de0000000040000000c' WHERE ( ( devices.productType = 1) ) ) order by
R_CPU.DeviceName ASC
```

Close

The Show SQL queries link displays the SQL Queries page in a separate browser window. Viewing SQL queries enables you to customize queries that can be used in SQL Server 2000 or PostgreSQL to export information to third-party applications.

## Developing a system software maintenance strategy

A carefully planned system software maintenance strategy maximizes server stability and availability. By developing well-regulated system software baselines for business servers, you can reduce the time required to update HP software and firmware on existing servers and ensure that new servers are set up with tested and stable software configurations.

Systems Insight Manager enables you to standardize software maintenance and update procedures on Windows NT 4.0, Windows 2000, and Windows Server 2003 systems. To set up a systems software maintenance strategy, complete the following tasks:

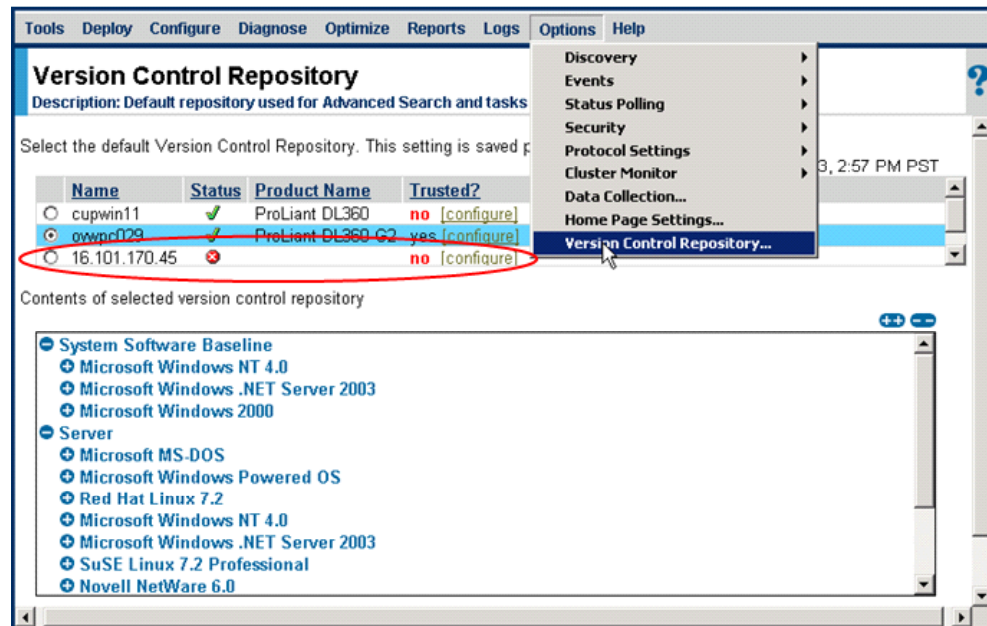
1. Install VCRM on a server that contains a repository.
2. Install the VCA on every managed system on the network.
3. Populate the repository with PSPs.
4. Configure the automatic update feature in VCRM.

### Version Control Repository Manager

Software updates must be applied on a timely basis. With multiple remote servers within an infrastructure, an efficient method is needed to keep track of software versions and ensure that software is current.

The Version Control feature in Systems Insight Manager uses the VCA to retrieve software status and perform software and firmware updates. Both the Version Control feature in Systems Insight Manager and the VCA on managed systems use the VCRM to retrieve software and firmware catalog information, PSPs, and Smart Components.

## Version Control Repository



When you install Systems Insight Manager on a server, you can create a default repository folder on the Systems Insight Manager server. However, Systems Insight Manager is not limited to using the repository on the Systems Insight Manager server. You can create software repositories on any system by using standard operating system commands to create a folder or directory.

Systems Insight Manager can use and deploy software from any repository, provided that:

- The repository is available to Systems Insight Manager on the network
- VCRM is trusted to Systems Insight Manager

### ! Important

For optimal integration with the software deployment features of Systems Insight Manager, ensure that the VCRM manages each repository.

The Systems Insight Manager Version Control Repository page on the Options menu has a setting to specify the default VCRM repository. A table lists the discovered systems with VCRM installed that the user is authorized to view. The selection box shows the current default repository or the first trusted repository alphabetically if no default has been chosen.

For maximum manageability and flexibility across operating system platforms, each repository that is created should be:

- Located on shared network drives managed by the Windows domain or Active Directory security
- Updated automatically
- Managed by the VCRM

## Maintaining a repository

After you create a repository, you must populate it with PSPs and Smart Components before deployment on the target servers. You can maintain the repository by using the HP website or the HP SmartStart for Servers CD.

---

<b>INTERNET</b>	You can download the latest VCA as well as PSPs, PSP deployment utilities, and Smart Components for Windows, NetWare, and Linux operating systems from: <b><a href="http://www.hp.com/support/files">http://www.hp.com/support/files</a></b>
-----------------	--

---

Perform the following tasks to maintain the repository:

- Monitor PSPs and Smart Components stored in the repository
- Delete multiple PSPs and Smart Components from the repository
- Copy multiple PSPs and Smart Components to another repository
- Create custom PSPs based on multiple Smart Components

The practice of updating PSPs and components from single or multiple repositories saves time and is key to standardizing software maintenance and update procedures on distributed systems. The system user account that updates the repository must have write access to the repository, which can be located on a local or shared network drive.

## Discovering software status

After the VCA and VCRM are installed and configured, Systems Insight Manager runs Version Status Polling tasks to determine when software on target managed servers requires an update. Systems Insight Manager has a predefined software polling task called *Software Version Status Polling* that runs weekly.

Data Collection tasks also perform the Software Version Status Polling functionality. There is also an Initial Data Collection task that runs when a device is newly discovered.

The Software Version Status Polling task is used in conjunction with the VCA installed on managed servers to retrieve the software status. The VCA determines the status based on whether the currently installed software matches the PSP baseline specified, or if there is no baseline, whether a newer version of the installed software is in the VCRM repository.

## Advanced Search

When you use the Advanced Search criteria to query for software versions, you see a list of only the software types found in a Version Control repository. The list includes version numbers found in the repository and versions found installed on systems (assuming any version of that software type is available in the repository).

### Example

If version 5.30 of the Foundation Agent for Windows is installed on a managed system, but there is no version of the Foundation Agent for Windows in the repository, then the Foundation Agent for Windows will not display in the Advanced Search list.

## Contacting the VCA

hp Version Control Agent [Device Home](#) [Help](#)

home log

→ refresh the inventory data  
→ change agent settings  
→ show additional items available in the repository

### Software Status

**Overall Software Status:**

**Reference Support Pack:** None selected. Click on **change agent settings** to specify the Reference Support Pack.

### Installed Software

Name	Installed Version	Latest Version
hp Management Agents for Windows 2000/Server 2003	6.10.0.0	6.40.0.0
hp Management Agents for Windows 2000/Server 2003	6.10.0.0	6.40.0.0
hp Management Agents for Windows 2000/Server 2003	6.10.0.0	6.40.0.0
hp Management Agents for Windows 2000/Server 2003	6.10.0.0	6.40.0.0
Web Agent Service	6.10.0.0	
hp ProLiant System Management Interface Driver	5.25.2195.0	
Survey Utility for Windows	2.48.7.0	2.53.0.0
Compaq System Configuration Utility		
hp ProLiant Remote Monitor Service for Windows 2000/Server 2003	5.0.2.0	5.0.2.0

**Last Generated:** 10/30/2003 7:53:19 AM

Version Control Agent 2.0

The VCA is an Insight Management Agent that is installed on a server to enable you to view the HP software and firmware that is installed on that server. The VCA can be configured to point to a repository being managed by the VCRM, enabling easy version comparison and software update from the repository to the server on which the VCA is installed.

When the VCA link is accessed for a device, the Version Control Agent screen displays if the VCA was detected on the server. The Software Status icon indicates whether software updates are available for the targeted system and how critical the software updates are.

The VCA determines server software status by comparing each component installed on the local device with the set of individual components or a specified PSP listed in the VCRM.

The four status levels for software are:

- **Current** (green) — All components on the device match the repository.
- **Minor** (yellow) — At least one of the components on the device can be updated, but none of the updates are critical.
- **Major** (orange) — At least one of the components on the device requires an update that is critical. There could be other components that are minor updates.
- **Unknown** (blue) — The device or repository could not be contacted and the status of that system is not known at this time.

---

**Note**

Discovered servers running HP-UX do not display a status.

---

If the Software and Firmware Version Status Polling task has not run or a VCA is not discovered for the target device, an Unknown (blue) icon displays.

If no VCA has been discovered for a server, a page displays indicating that the VCA was not detected on the target server along with instructions to install and configure the VCA and associated agents. However, even if a VCA was detected, the VCA page does not display if a trust relationship problem exists. Instead a page displays indicating that the problem exists and pointing to help on establishing a trust relationship.

The Version Control Agent permits users with full or limited configuration rights to perform the following tasks to maintain the software inventory of the server:

- View the currently installed software
- View whether any applicable updates are available in the VCRM
- Select a VCRM as a reference point for obtaining software updates
- Select whether to use a Custom Software Baseline or PSP as a managed baseline
- View the details associated with a Custom Software Baseline, PSP, or individual software component in the VCRM
- Install a Custom Software Baseline, PSP, or individual software component from the VCRM

You can also update individual components or entire PSPs by clicking the install icon located next to the system software status icon.

## Creating a task to update a managed system

To upgrade software or firmware on target systems, create a task using a repository. Administrator privileges to the devices in the repository are required. In addition, a trust relationship must be established with the selected repository.

The tree associated with a repository contains an inventory of the software sorted into categories by division (Server and System Software Baseline), operating system, and category.

Components can be forced to be downgraded even if the same or a later software version is already installed on target systems.

You can configure Systems Insight Manager to bring systems out of low power mode before initiating an update.

---

### **Note**

Only systems with network adapters that support magic packet technology can be brought to full power during a software update.

---

A managed system can also be rebooted after an update is successfully completed.

## Running the Update Software and Firmware task

When you run the Update Software and Firmware task, Systems Insight Manager communicates with the selected VCAs.

### Updating software and firmware status

Because the Update Software and Firmware tasks are launched asynchronously, the status should be tracked independently. Two tracking mechanisms are used:

- Event notification
- Deployment status polling, which includes:
  - Software Deployment Status Polling task
  - Software and Firmware Version Status Polling tool

#### Event notification

The VCA sends Systems Insight Manager an HTTP event when a deployment task completes. Systems Insight Manager retrieves the status from the event and stores the status in the database.

The VCA also sends an event when components have been downloaded but before beginning to install them. This provides dynamic job status information to determine which systems are downloading components, which are installing them, and which have completed installation.

#### Deployment status polling

A Software Deployment Status Polling task runs every 15 minutes by default, updating the status of each deployment task item. The task scans the notice log, looking for open software update notices. An open software deployment notice is one for which Systems Insight Manager has not received an HTTP event and has been open for at least 15 minutes.

Update tasks not completed in two hours are marked as failed. The notice is closed so that no further polling occurs. If an HTTP event is received from the system after the notice is closed, the notice will be updated accordingly.

---

#### Note

The Software Deployment Status Polling task is a background task that cannot be scheduled or modified by Systems Insight Manager users.

---

## Managing the Systems Insight Manager database

The Systems Insight Manager database contains all the information about the systems on the network and the lists and tasks that have been executed. This information includes everything you need to know to manage the network.

Systems Insight Manager uses Windows authentication to communicate with the database (MSDE or SQL Server) on Windows systems. PostgreSQL is used to create the database on Linux and HP-UX platforms. Systems Insight Manager creates its own database user (mxadmin) during database configuration.

### Database configuration

The installation application automatically configures PostgreSQL on HP-UX and Linux servers and MSDE on Windows servers. It also creates the Systems Insight Manager database and its associated database user. The user password is accessible only to root users on Linux and HP-UX and to the administrator of a Windows system.

### CLI configuration

- `mxconfigrepo -a [ -F]` — Configures the database. This option is used if users lose their passwords or if the database is corrupted. This option only applies to PostgreSQL on Linux and HP-UX platforms.



---

**Caution**

All data in the repository is lost with this option.

---

- `mxconfigrepo -s` — Sets up the database. This option assumes that the Systems Insight Manager database is created, the database is started and configured, and a user is created with permissions to set it up. This option is supported on all SQL servers on all supported platforms.



---

**Caution**

All data in the repository is lost with this option.

---

- `mxconfigrepo -k` — Keeps the data and updates the database. Data in the repository is updated with this option. This command is used if users lose their passwords or if the database is corrupted. This option assumes that the Systems Insight Manager database is created, the database is started and configured, and a user is created with permissions to set it up. This option is supported on all supported SQL servers on all platforms.
- `mxconfigrepo -u` — Removes the database. This option is supported on all SQL servers on all supported platforms.
- `mxconfigrepo -r` — Removes the database and the Systems Insight Manager. This option is only supported on the Linux and HP-UX platforms using PostgreSQL.

## Using Lights-Out devices to manage remote servers

A centralized network infrastructure must be able to manage a server from anywhere, at any time. HP Lights-Out technology enables you to manage servers and perform a variety of server administration tasks remotely.

Servers that do not include iLO can have RILOE II cards installed. These devices enable you to:

- Integrate Lights-Out devices into a directory service
- Control the power of the servers
- Use a local diskette or CD on a remote server to update the ROM
- View the console of the remote server

## Integrating management processors

Integrating management processors with a directory service enables you to:

- Authenticate users from a shared, consolidated, scalable user database including HP products
- Create management processor objects to control user privileges within the directory service
- Use roles to manage management processors and users

Even if you use a directory service, the local user database is retained. You can avoid using a directory service, use a combination of a directory service and local accounts, or use a directory service exclusively for authentication.

The four steps to integrate management processors with a directory service are:

1. Upgrade the management processor firmware.
2. Extend the directory services schema.
3. Install the management snap-in in the management application of the operating system.
4. Create and manage objects and roles in the directory service.

Directory services store information about objects on the network and make this information easy to find and use. Directory services use a structured data store as the basis for a logical, hierarchical organization of directory information.

The following table lists operating system support for the directory services in which management processors can be integrated. Active Directory is the directory service for Windows 2000 Server and Windows Server 2003. Novell eDirectory is a directory service that supports multiple platforms including Windows 2000 Server, NetWare, and Linux.

Operating system	Active Directory	Windows Server 2003 Active Directory	eDirectory 8.6.2	eDirectory 8.7
Windows 2000 Server	X		X	X
Windows 2000 Advanced Server	X		X	X
Windows Server 2003	X	X	X	X
NetWare 5.x			X	X
NetWare 6.x			X	X
Red Hat Enterprise Linux Advanced Server 2.1			X	X
Red Hat Linux 7.3			X	X
Red Hat Linux 8.0			X	X

## Authentication

Network authentication confirms the identification of the user to any network service that the user is attempting to access. An authentication mechanism such as Secure Sockets Layer (SSL) provides the security required to verify the identity of the user.

Directory information must be accessible to users and be managed to control access to resources. Integrating information for management processors with a directory service eases management and increases management processor access security.

## Directory Service objects

One of the keys to directory-based management is proper virtualization of the managed devices in the directory service. Virtualization enables you to associate the managed device with users or groups in the directory service. Administering a management processor in a directory service requires three basic objects:

- Management processor object
- Role object that contains users and management processor objects
- User objects

## Roles

Roles are categories of users used to determine access permissions to network resources, such as management processors. These role assignments are then used to determine which user categories have permission to access which elements within the device.

Ideally, roles should be identified according to the associated tasks, such as `Mgmt_Proc_User` or `Mgmt_Proc_Admin`.

## Users and groups

The preferred way to populate roles is with user accounts and groups. A user is assigned to the appropriate groups and then the groups are assigned to the appropriate roles. Using groups to populate roles makes it easier to manage large numbers of users.

In enterprise computing environments, it is often difficult to effectively track each user's place within the organization and determine how that information maps to the role-based security policy particular to each management processor. When the number of users, administrators, and devices increase this task becomes increasingly complicated.

## Upgrading management processor firmware

Management processors require a minimum firmware version that supports directory service integration. Upgrade RILOE II to at least version 1.10 and ensure that iLO has version 1.40 or later.

You can upgrade management processor firmware using one of these methods:

- Management processor browser interface
- HP Lights-Out Directories Migration utilities
- Group configuration using Remote Insight Board Command Language (RIBCL)

### Browser interface



Using a standard web browser, you can upgrade the management processor firmware remotely from any network client. This is accomplished through the Upgrade Firmware option on the Administration tab in the browser interface. To perform this task, you must have the Configure iLO Settings or Configure RILOE Access privilege.



#### Important

Do not attempt to upgrade the firmware from a ROMPaq diskette using the Virtual Floppy feature.

---

## Migration utilities

Two utilities automate some of the migration steps to management by directory services for previously installed management processors:

- **HP Lights-Out Migration (HPQLOMIG)** — An automated utility with a GUI and a wizard approach to implementing or upgrading many management processors. This utility, unlike HPQLOMGC, discovers management processors on the network.
- **HP Lights-Out Migration Command Line (HPQLOMGC)** — A CLI utility that enables migration of a single management processor. Used in conjunction with Systems Insight Manager, HPQLOMGC can be launched by itself or within an XML file.

Both utilities perform these functions:

- Upgrade the firmware on the management processors to the version that supports directory services
- Assign names to the management processors to identify them in the directory
- Create objects in the directory corresponding to each management processor and associate the objects to a role
- Configure the management processors to enable them to communicate with the directory

## Group configuration

The RIBCL enables you to write scripts to configure and administer management processors.

The UPDATE\_RIB\_FIRMWARE command is used in a script to update the firmware of management processors. This command copies the firmware upgrade file to the management processor, starts the upgrade process, and restarts the device after the image has been flashed successfully.

The management processor is reset after the firmware upgrade is complete. You must be logged in to the management processor with the Configure iLO Settings or Configure RILOE Access privilege.

### Example

```
<RIBCL VERSION="2.0">
  <LOGIN USER_LOGIN="adminname" PASSWORD="password">
    <RIB_INFO MODE="write">
      <UPDATE_RIB_FIRMWARE IMAGE_LOCATION="C:\ILO140.BIN"/>
    </RIB_INFO>
  </LOGIN>
</RIBCL>
```



### Important

The UPDATE\_RIB\_FIRMWARE command must be displayed within a RIB\_INFO block and RIB\_INFO must be in write mode.

---

## Administering management processors

Management processors require specific software that extend the schema and provide snap-ins to administer management processors on the network.

---

**INTERNET** An HP Lights-Out Directory Package containing the schema installer and management snap-in installer (as well as the migration utilities) is available for download from: <http://www.hp.com/servers/lights-out>

---

### Schema installer

The files containing the schema to be added to the directory (.xml files) are bundled with the schema installer. One of the files contains core schema that is common to all the supported directory services. Additional files contain only product-specific schemas.

The installer includes three stages:

- **Schema Preview** — The Schema Preview displays the proposed extensions to the schema in a tree view. This view is based on the contents of the selected schema files and parsed XML. It lists all the details of the attributes and classes that will be installed.
- **Setup** — The Setup screen contains a Directory Server section and Directory Login section. The type of directory service (Active Directory or eDirectory) is selected in the Directory Server section. This section also enables you to set the computer name and the port to be used for Lightweight Directory Access Protocol (LDAP) communications.

A login name and password might be required to complete the schema extension. The Directory Login section of the Setup phase enables you to enter this information and set the form of secure authentication. The options are Directory authentication using SSL, NT authentication, and clear text.
- **Results** — The information displayed on the Results screen includes whether the schema could be extended and what attributes were changed.

## Management snap-in installer

The management snap-in installer installs snap-ins for managing management processor users in Active Directory Users and Computers Microsoft Management Console (MMC) and Novell ConsoleOne. The management processor snap-ins are used to:

- Create management processor objects and role objects.
- Add or remove users to role objects.
- Set the rights and restrictions of role objects.

### Create objects

To add or remove the HP devices to be managed within a role:

- MMC uses the HP Devices tab.
- ConsoleOne uses the Role Managed Devices tab under the HP Management tab.

### Members

After user objects are created, the Members tab allows you to manage the users within the role. You can add a user or select from the list of valid members to remove a user.

### Role restrictions

The Role Restrictions tab allows you to set login restrictions for the role. These restrictions include:

- Time restrictions
- IP network address restrictions including IP/mask and IP range
- Domain Name Service (DNS) name

## Directory Service administration

After you create a role, you can assign rights for that role. You can make users and group objects members of the role, giving the users or group of users the rights granted by the role. The available rights are the same as in the management processor browser interface.

Rights are managed in the:

- MMC Lights Out Management tab in Active Directory.
- ConsoleOne Lights Out Management Device Rights tab under the HP Management tab in eDirectory.

Additionally, an Administer Local Device Settings option enables you to configure the options available on the Global Settings, Network Settings, SNMP Settings, and Directory Settings screens of the browser interface.

## Directory settings

The Directory Settings screen of the management processor browser interface includes the options shown in the following table.

Option	Description
Directory Authentication	Designates whether a directory service is used to authenticate a user login. By default, this setting is <i>Disabled</i> .
Local User Accounts	Enables a user to log in using a local user account instead of a directory account. By default, this setting is <i>Enabled</i> .
Directory Server Address	Designates the IP address or DNS name of the directory server or the name of the domain.
Directory Server LDAP Port	Designates the port used for LDAP communications. The default setting is the secure LDAP port 636.
LAN on motherboard (LOM) Object Distinguished Name	Specifies the full distinguished name of the management processor object in the directory service. Limited to 256 characters.
LOM Object Password	Specifies the password that the management processor object will use to log in to its corresponding object in the directory. Limited to 40 characters.
Directory User Context	Specifies search contexts when authenticating a user. These settings point to areas in the directory service where users are located so the user does not have to enter the complete tree structure when logging in. Limited to 128 characters each.

### Note

At this time, the LOM Object Password option is not used. This option provides forward compatibility with future firmware releases.

## Virtual power

If the operating system is not functioning, it can be bypassed to reset the server. After the server has been reset, you can observe the full video text of the reset process, from memory count through loading of the operating system.

The iLO Virtual Power button simulates pressing the physical power button on the server with the following options:

- **Momentary Press** — Used to power a server on or off.
- **Press and Hold** — Performs a cold start of the host server regardless of the condition of the operating system. This type of reset does not shut down the operating system gracefully.
- **Power Cycle System On, After Off** — Simulates powering the server off and back on.

RILOE II provides two virtual power buttons. The Shutdown Server and Power OFF button initiates a graceful shutdown. However, if the operating system is not responding, the Force Server Power OFF button forces the server to power off without shutting down. The Virtual Power interface for RILOE II also enables you to restart the server.

## Virtual Media

Virtual Media devices connect to the host server using USB technology. If the operating system on the host server supports USB, you can direct a host server to boot and use a CD, standard 1.44MB diskette, or an image file.

## Virtual Floppy

Virtual Floppy enables you to perform the following remote operations:

- Install an operating system from a network drive
- Apply a ROMPaq firmware update (iLO only)
- Perform user diagnostics

You can use a local diskette drive or an image file. You can create the image file on the management client using the Create Disk Image option.

## Remote console

The remote console is the video output from the server in a fully functional GUI.

The Lights-Out graphical remote console resides in the firmware of the board itself, enabling it to function independently from the operating system of the host and management PCs. An HTML server integrated on the board delivers the video output of the server in a form that can be reproduced as a virtual desktop with any standard web browser on a remote PC.

The remote console option presents a single mouse cursor during remote console. Synchronization of two cursors is no longer required, making navigation easy.

### iLO Remote Console (dual cursor)

The Remote Console dual cursor option uses two mouse cursors to represent the mouse cursor of the remote server and the mouse cursor of the local client. The local client cursor is displayed as a crosshair in the Remote Console window. If the two cursors drift apart, they can be synchronized and brought back together.

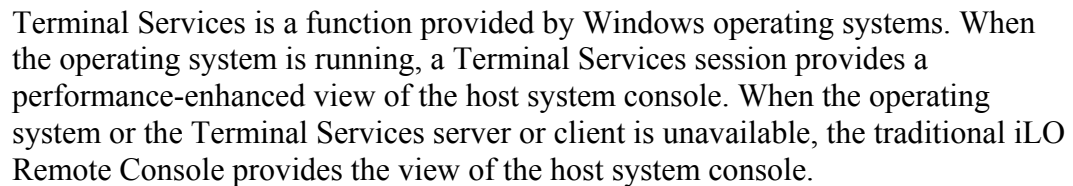
### Program Remote Console Hot Keys

Remote Console Hot Keys enable you to define Ctrl keys with key combinations that are not available in a remote console session.

#### **Example**

You could assign the key sequence *Alt+Ctrl+Del* to *Ctrl+X*. When you press *Ctrl+X* during a remote console session, *Alt+Ctrl+Del* will be transmitted in its place.

Each key combination can contain a maximum of five keys.



## Adding new servers to the managed enterprise

When you introduce new servers to a centrally managed enterprise, you must perform the following tasks:

1. After deploying the operating system and configuring network settings, install OpenSSH.
2. Ensure that a trust relationship exists between the new server and Systems Insight Manager.
3. Manually add the server to Systems Insight Manager.
4. If a server does **not** have the VCA installed, use the Initial ProLiant Support Pack Install feature of Systems Insight Manager to install the supported PSP.
5. If a server does have the VCA installed, run the Software Version Status Polling task on that server.
6. Add new Lights-Out devices to the directory service.
7. Assign new servers to an IT administrator in Systems Insight Manager.

## Summary

Bob agrees that Systems Insight Manager would be the perfect tool to enable the company to manage its growing network.

After you install Systems Insight Manager and configure it according to RC Engineering's requirements, Jackie used Lights-Out technology to poll the company's managed systems to determine which software and firmware needed to be updated, all from the Houston headquarters. The Version Control feature in Systems Insight Manager simplified these tasks.

Jackie also configured Systems Insight Manager to monitor all the RC Engineering servers, so that she will be warned of potential system faults that could cause failures and a technician will be paged if a hard drive or other component is about to fail.

Jackie is happy to know that she can manage the RC Engineering servers remotely and perform administration tasks from one location without traveling to each machine.

## Learning check

1. Which layer of the Systems Insight Manager architecture serves as the primary means of accessing management functionality?  
.....
2. Which Systems Insight Manager usage scenario described in this module requires the VCRM utility to be installed on regional servers?  
.....
3. The server designated for Systems Insight Manager is running HP-UX. Which database can be installed on this server to support Systems Insight Manager?
  - a. Oracle
  - b. SQL Server 2000
  - c. SQL Server 7
  - d. PostgreSQL
4. List the three options for customizing the Systems Insight Manager Home page.  
.....  
.....  
.....
5. An organization has five system administrators who are responsible for 100 different systems in six different buildings. Which feature of Systems Insight Manager will enable each administrator to view only those systems for which they are responsible?  
.....  
.....

6. Why add a rule to the System Type Manager?

.....

.....

.....

.....

7. You need to compile MIBs. However, you do not want to affect the performance of Systems Insight Manager. What must you do?

.....

.....

.....

8. When would it be useful to copy a report?

.....

.....

.....

.....

9. List the steps you should take to integrate management processors with a directory service.

.....

.....

.....

.....

10. List the tasks that you must complete when you add a new server to the network.

.....

.....

.....

.....

## Objectives

After completing this module, you should be able to:

- Describe the best practices for securing servers in an enterprise environment in regard to:
  - Physical security
  - Network security
  - Host security
  - Application security, including:
    - ◆ Website security
    - ◆ Email security
- Identify security vulnerabilities to common attacks and the measures you can take to secure against such attacks
- Develop an information security policy that outlines requirements, roles, and responsibilities to determine access to network devices and applications
- Configure security using HP Systems Insight Manager
- Develop an enterprise antivirus strategy

## Introduction

Bob, the CEO of RC Engineering, has discovered that the escalating cost of real estate in California makes it difficult to justify the cost of doing business there. Even though he got a great deal when he bought in during the dot-com bust, Bob is being forced to close the company's Silicon Valley location and ship the servers and other networking equipment to the corporate headquarters in Houston.

Bob has asked for your help determining the most secure location for this equipment. You will need to survey the Houston site with Jackie, the IT manager, to establish the physical parameters. Jackie realizes that this is a good time to identify the company's security vulnerabilities and has asked for your help in performing a risk analysis.

Jackie has also informed you that Bob recently received the company's financial statements in an email from the corporate accounting firm. Unknown to either Bob or the accountant who sent the email, the attachment was infected with a virus, which spread across the entire RC Engineering network. While troubleshooting the problem, an employee inadvertently erased files essential to the corporate website, bringing the website down for three days.

Although this incident did not result from malicious intentions, Bob and Jackie are now very conscious of the vulnerability of the company to attacks from intruders, both inside and outside. Jackie has implemented a firewall to protect the private network, but needs guidance to develop a complete enterprise security policy. Jackie would also like to update the encryption and decryption technology the company currently uses.

## Securing enterprise servers

Security is a balance between maintaining ease of use and controlling access. Designing a security policy that restricts users and potential attackers can be time consuming and costly. A security program that is too controlling can disgruntle users with policies that limit them from doing their work effectively.

Established in 1988, the CERT Coordination Center is a center of Internet security expertise located at the Software Engineering Institute, a federally funded research and development center operated by Carnegie Mellon University. According to CERT, all security best practices can be grouped into five top-level steps:

1. Secure your systems by establishing secure configurations.
2. Prepare for intrusions by laying the proper groundwork for detection and response.
3. Detect intrusions quickly.
4. Respond to intrusions so that damage is minimized.
5. Improve your security for the future.

To accomplish these steps, you need a security planning team. The team should include people involved in different aspects of IT from different areas of the enterprise. Knowledgeable people—savvy in business requirements, technology, and security—are necessary.

After a team is created, the first step is to analyze the business requirements. The hardest part of this step is distinguishing wants from needs. The planning phase involves developing a vulnerability/risk/threat assessment and includes an appraisal of existing countermeasures and their cost: benefit analysis. Whatever the security policy is, ongoing testing and continual process improvement is required to ensure continued good health.

The general categories for security are:

- Physical
- Network
- Host (operating system)
- Applications
  - Websites
  - Email

Remember to consider people as a source of vulnerability. One infamous hacker routinely used employees who answered his telephone calls to gain the information he needed to break into systems and networks.

## Physical

Although physical security does not receive as much press attention as network or operating system security, it is a maxim that without physical security there is no security.

Best practices include, but are not limited to:

- Do not allow piggy-backing through controlled access points.
- Control access to all areas containing servers, website hardware, firewalls, networking equipment, and so forth.
- Use human guards.
- Use a combination of keypads, smart cards, proximity badges (for example, coat tags), and biometric devices.
- Make it easy to identify who is allowed in more sensitive areas by use of special badges indicating proper access privilege.

Physical security does not stop at the door to the building. Users must be aware of their surroundings to ensure that passwords are not “shoulder surfed.” Cell phones and text pagers are inherently insecure methods over which to distribute passwords or other sensitive information. Another commonly employed tactic is “dumpster diving,” or going through a target’s trash looking for sensitive information. Be sure to dispose of all confidential documents securely.

## Network

In 2002 the CERT Coordination Center published a seven-step best practices procedure for network security methodology called *Security Knowledge in Practice* (SKiP). There is an associated set of security practices for each step, but in summary you should:

1. Select systems software from a vendor and customize it according to your unique requirements.
2. Secure the system against known vulnerabilities.
3. Prepare the system so that anomalies can be noticed and analyzed for potential problems.
4. Detect those anomalies and any other system changes that could indicate evidence of an intrusion.
5. Respond to intrusions when they occur.
6. Improve practices and procedures after updating the system.
7. Repeat the SKiP process as long as the organization needs to protect the system and its information assets.

In addition to the SKiP method, the following are also good network security best practices:

- Use Virtual Private Networks (VPNs), firewalls, proxies, and IP routing with Network Address Translation devices.
- Know your network and analyze your traffic.
- Enable logs and check them regularly.
- Know how to gather forensic data.
- Back up data regularly.
- Use intrusion detection systems.
- Perform periodic vulnerability scans.
- Know who to contact in an emergency.

## Host

Require two-factor authentication for your most secure servers. The two factors should be something the user knows, such as a password, and either something the user has (mobile phone, smart card, pager, or even a known PC) or something that would use biometrics (fingerprints or retina scans). Deciding to invest in this technology has a lot to do with how much security is required compared to the cost of implementation.

Other best practices include:

- Audit all user accounts for administrative privileges and make certain that only the needed level of permissions are granted to users.
- Grant only the file-level permissions needed to enable users to do their jobs.
- Include banner warnings and policy statements when a computer attaches to the network or attempts to log in.
- Configure computers to provide only the minimum selected network services.
- Distribute services across more than one system so that if a single system is compromised, not all of the services will be.
- Ensure that the software used to examine systems has not been compromised.
- Monitor and inspect system activities and logs for unexpected behavior.
- Disable unused or unneeded peripherals such as:
  - Printer, serial, and USB ports
  - Diskette drives
  - CD-ROM drives
- Disable system autorun functionality that automatically launches a program when a diskette, CD, or USB device is found.

## Applications

When testing new applications, install servers off the main network for the initial testing. Remember that many applications offer types of simple, complete, and custom installation methods. Most simple installations do not leverage all of the possible security settings and can leave the application vulnerable. Best practices dictate that you choose custom installations so you can properly customize the default settings and change insecure default passwords.

It is critical that you keep applications current by applying patches as soon as practical after they have been tested. If you use an open source application, inspect the code to evaluate its robustness and vulnerability to known exploits overflows.

## Websites

Websites are one of the most common objects of attacks; Microsoft Internet Information Services (IIS) is a favorite target. Website attacks can vary from simple defacement to theft of information such as credit card numbers. It is especially critical to keep the web server software current by applying patches as soon as practical after they have been tested.

Other best practices include:

- Isolate the web server from the public network and the internal networks of the organization.
- Configure the web server with appropriate object, device, and file access controls.
- Configure the web server to use authentication and encryption technologies, where required.
- Identify and enable logging mechanisms specific to the web server.
- Consider security implications before selecting programs, scripts, and plug-ins.
- If scripts and plug-ins are used, configure the web server to minimize their functionality.
- Maintain the authoritative copy of your website content on a second secure host.

## Email

Many companies consider email to be a mission-critical business tool. Best practices to ensure availability of company email include:

- Use a gateway or relay server to bring email from the Internet to your mail servers. Place the gateway server in the demilitarized zone (DMZ) or (preferably) on your side of the firewall.
- Use an antivirus scanner on email gateways.
- Disable all ports and services not directly related to processing email.
- Employ content filtering and blocking utilities to search for embedded malicious code such as scripts.
- Employ spam detectors and blockers.

For more information regarding security best practices, visit one of the following commonly used security sites:

- Computer Security Institute: **[www.gocsi.com](http://www.gocsi.com)**
- Federal Bureau of Investigation: **[www.fbi.gov](http://www.fbi.gov)**
- High Technology Crime Investigation Association: **[www.htcia.org](http://www.htcia.org)**
- Information System Security Organization: **[www.nsa.gov/isso](http://www.nsa.gov/isso)**
- National White Collar Crime Center: **[www.cybercrime.org](http://www.cybercrime.org)**
- National Security Agency: **[www.nsa.gov](http://www.nsa.gov)**
- National Security Institute: **[nsi.org](http://nsi.org)**
- RSA Data Security: **[www.rsa.com](http://www.rsa.com)**
- SysAdmin, Audit, Network, Security (SANS) Institute: **[www.sans.org](http://www.sans.org)**

## Security vulnerabilities

The majority of successful attacks on computer systems by means of the Internet can be attributed to exploitation of security flaws. Hackers are opportunistic, taking the easiest and most convenient route. They exploit the best-known flaws with the most effective and widely available attack tools. They count on organizations not fixing the problems, and they often attack indiscriminately, scanning the Internet for any vulnerable systems. Attackers do not generally find the original vulnerability, but instead find the code to exploit it. It does not take a security expert to exploit a vulnerable area and attack others.

System administrators typically do not correct many system flaws because they do not know which vulnerabilities are most dangerous and are too busy to correct them all. Most recent and well-publicized attacks on computers and networks exploited known flaws that already had patches released for them. It is critical that the operating system is kept current by applying patches as soon as practical after they have been tested.

Some vulnerability scanners search for up to 800 vulnerabilities, which broadens the focus needed to ensure that all systems are protected against the most common attacks.

The SANS Institute and the National Infrastructure Protection Center released a document summarizing the most critical internet security vulnerabilities. The following vulnerabilities account for the majority of successful attacks:

- Default installations of operating systems and applications
- Accounts with no passwords or weak passwords
- Nonexistent or incomplete backups
- Large numbers of open ports
- Not filtering packets for correct incoming and outgoing addresses
- Nonexistent or incomplete logging
- Vulnerable Common Gateway Interface (CGI) programs
- Uninformed employees
- Noncompliant employees
- Neglected operating system and server updates

The ability to scan a computer for open IP ports is an unavoidable vulnerability because of the behavior of the IP protocol. Additionally, unknown vulnerabilities, such as a buffer overflow attack on a Domain Name Service (DNS) server, exist that must be predicted. If an anomaly exists in the DNS server, sending a command with a long hostname could cause the DNS server to fail.

## Common attacks

There are several types of attacks used by intruders. Attacks can prey on weak passwords, application code, email, open ports, or known vulnerabilities that administrators have not yet patched.

### Weak passwords

Ensure that all computers require user authentication for access. Strong passwords of at least eight characters are advisable and should include letters, numbers, and symbols, which will help secure the system from dictionary attacks.

Man-in-the-middle attacks and hijacking are other methods used by intruders to obtain passwords or other critical enterprise network information.

- **Dictionary** — An intruder uses a program to test a security accounts database against a long list of names or passwords. These attacks rely on automated programs.
- **Man-in-the-middle** — An intruder detects passwords and obtains information from legitimate transactions.
- **Hijacking** — An intruder trespasses in a transaction between two parties, imitates one of the participants, and then continues the connection.

### Application code

Intruders use a variety of methods that result in enterprise downtime or data loss.

- **Back door** — An undocumented opening in an operating system or program designed with benign or malicious intentions.
- **Trap door misuse** — A piece of residual code, a trap door enables the original programmer to access the product easily to provide support. However, trap doors also make the system vulnerable to hacker exploitation.
- **Buffer overflow** — An action that allows the program of a target system to be modified spontaneously and remotely. If a user sends more data than the target system can receive at one time, the server will fail. The extra data overflows the program storage buffer and overwrites the actual program data.

- **Teardrop** — An attack that takes advantage of code that does not properly reassemble overlapping User Datagram Protocol (UDP) packets.
- **Trojans and worms** — Files that operate in an expected way, but also have a secret operation that subverts security. Trojans include files that can send sensitive information back to a hacker. Worms propagate from system to system.
- **Virus** — Programming code that, when run, attaches to other programs and executes each time the infected files run.
- **Default settings exploitation** — An attack on the default settings of software before it has been configured.
- **UDP bomb** — An attack from an illegal UDP packet containing incorrect values that can cause older operating systems to fail.

### Denial of service

A denial of service attack occurs when a hacker makes it impossible for a legitimate user to access or service a host. The goal is to flood the communication ports and memory buffers of the targeted site to prevent receipt of legitimate messages and the service of legitimate requests for connections.

- **Microsoft Windows out-of-band** — An out-of-band denial of service attack is used to cause the computer to fail or a loss of network connectivity.
- **Synchronous idle (SYN) flood** — A hacker creates multiple half-open TCP connections. When the SYN queue is flooded, no new connection can be opened.
- **Smurf attack** — A smurf attack involves manipulating the Internet Control Message Protocol (ICMP), the protocol invoked by the ping program. It is a form of IP spoofing that results in a denial of service attack.
- **Website defacement** — A hacker defaces a website by changing or destroying the content.
- **Cyber-extortion** — Cyber-extortion involves hackers blackmailing companies by threatening to turn over purloined strategic data to their competitors or expose it to the public. The more infamous cases resembled kidnapping in which credit card data was held hostage in exchange for money.

## Data interception

Information on a network is transmitted in packets. A hacker can intercept the packets and imitate the originator of the message or creator of the information for nefarious purposes.

- **Land attack** — A hacker sends a TCP synchronize/start (SYN) packet with a fake source IP address and port number that matches the destination IP address and port in an attempt to cause the networking to loop and the computer to fail.
- **Ping of death** — A hacker appends large amount of information to an ICMP echo request (ping) packet. This causes a kernel buffer overflow when the computer attempts to respond, which causes the computer to fail.
- **IP spoofing** — In an attempt to defeat authentication, a hacker uses a packet sniffer and a program to kill a TCP connection, generate a new TCP connection, and mimic an IP packet.
- **Blind and nonblind spoofing** — Blind spoofing occurs when a hacker manipulates a connection that exists on a separate physical line. Nonblind spoofing occurs when a hacker manipulates a connection on the same subnet or the same physical line.

## Open ports

- **Port scan** — A port scan is used to access the open IP ports on a target computer or network.
- **IP half scan** — An IP half scan is used to avoid detection. TCP communication uses a three-step process to establish a connection. The IP half scan completes only half the connection, avoiding detection.

---

### INTERNET

To learn more about packets and protocols, refer to these websites:

- [www.ieee.org](http://www.ieee.org)
  - [www.iana.org](http://www.iana.org)
  - [www.ietf.org](http://www.ietf.org)
  - [www.packet-level.com](http://www.packet-level.com)
-

## Information security policies

The purpose of an enterprise information security policy is to define the appropriate measures that must be employed to manage the operational risk associated with the information systems of a company.

Risk management is a complex issue involving a variety of potential activities. In an information security policy, risk management focuses on the risks arising from the generic internal business activities of an enterprise and the risks associated with external network interfaces related to the operation of an enterprise-wide information system.

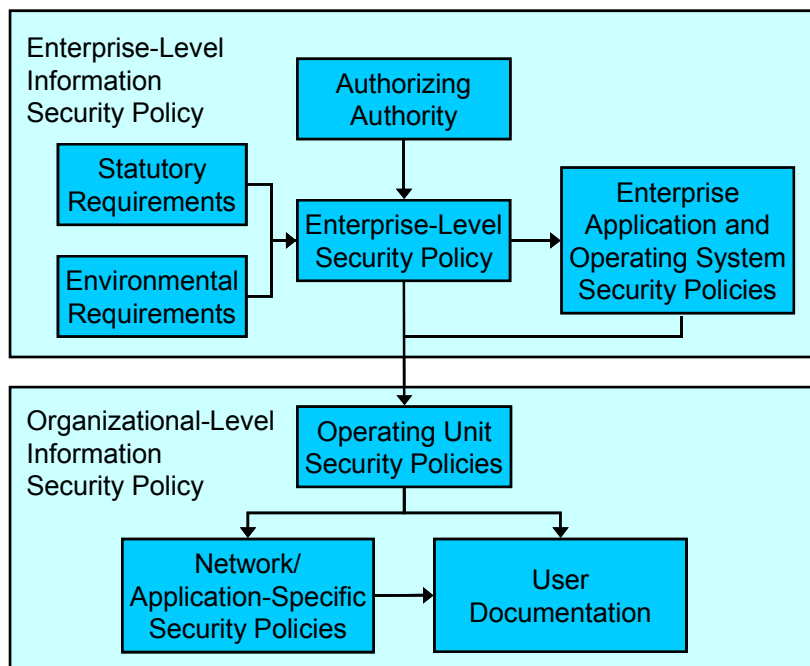
A security policy does not create a secure environment. The policy must be properly implemented using relevant procedures and technology to have an impact on risk.

An enterprise information security policy focuses on the following goals:

- To provide a level of system availability that is sufficient to ensure appropriate support for performance of the enterprise mission.
- To provide appropriate protection of the ownership and confidentiality of sensitive information, including proprietary, personal, and confidential information.
- To provide appropriate data integrity for information used in critical business decisions.

In addition, you can use audit trails to track system usage to determine the source of problems that can arise. These records can also be used for billing, capacity planning, and other purposes.

## Hierarchy of information security policies



A hierarchy of information security policies usually exists within the security profile of an enterprise. Policies begin at a high level and become increasingly detailed as they apply to operational environments.

In most cases, a company executive issues a document that establishes an information security program and delegates authority to an information security program manager. This manager then develops an enterprise-level security policy that applies to the entire company. The manager might also issue security policies for enterprise applications and operating systems.

Enterprise-level security policies should consider the operational environment and any applicable statutory requirements imposed by federal, state, and local governments. To support development and implementation of security policies, an enterprise can set up a security policy board to give organizations within the enterprise a voice in policy development and to build consensus to support implementation of the policy.

Enterprise-level policies are usually applied at the organizational level. When the policies apply to specific hardware or software components, detailed policy descriptions and explicit guidance are often provided through user documentation. A network-specific policy typically lists equipment and provides setup and maintenance instructions.

In small and medium companies, the enterprise-level and organizational-level policies are often the same document. Each enterprise modifies its policy model as the enterprise grows. Therefore, flexibility is a key consideration when developing and implementing an information security policy.

## **Application and network security policies**

When creating an information security policy, you might want to perform a vulnerability, risk, and exposure analysis. It will inform you about the trade-offs between security and usability that all computer systems are subject to in a networked environment. You must evaluate and document all such trade-offs in the enterprise security policy.

Account policies, auditing, and security patch management mitigate some security risks for the enterprise. However, you must follow up with detailed implementation and management plans.

### **Account policies**

Setting account policies ensures adequate password protection. Account policies provide a vehicle to set complexity and change schedules for passwords. Account policies also make it easier to track unsuccessful password logon attempts and initiate account lockouts if necessary.

Complex passwords that change regularly reduce the likelihood of a successful password attack. However creating strict requirements for password length and complexity do not necessarily translate into users using strong passwords. To educate users on choosing strong passwords, some companies display posters describing poor passwords in common areas, such as near the water fountain or copy machine.

Intruders are not limited to individuals external to an organization. By knowing the person who created a password, a person with malicious intentions might be able to guess a password based on the creator's favorite food, car, or movie.

## Security patch management

Developing a system security policy is extremely powerful because it allows an administrator to create a configuration management solution for the entire enterprise. One critical aspect of a system security policy is security patch management, which employs the tools, utilities, and processes for keeping systems current with new software updates that are developed after a software product is released.

Security patch management is best achieved when it is a consistent and integral part of an organization's standard operational processes. Without operational consistency, a separate process for security patch management can increase the overall cost of ownership and will introduce unnecessary ambiguity in the organization.

Security patch management is a necessary process on all platforms—every major software vendor committed to security releases security patches in response to newly identified vulnerabilities. There is no widely used operating system or application that is immune from attackers who spend their time trying to locate vulnerabilities to exploit.

Proactive security patch management is a requirement for keeping the enterprise environment secure and reliable. To maintain a secure environment, organizations must have a process for identifying security vulnerabilities and responding quickly. This process must involve applying software updates, configuration changes, and countermeasures to eliminate vulnerabilities from the environment and mitigate the risk of computers being attacked. Many attacks require only a single vulnerable computer on the network, so this process must be as comprehensive as possible.

## Auditing

An audit log records an entry whenever users perform certain actions that are being tracked. The audit entry shows the action performed, the associated user account, and the date and time of the action. You can audit both successful and failed attempts at actions. Failure logs are often more informative than success logs because failures typically indicate an error.

### Example

A user successfully logging on to the system would be considered normal. However, if a user unsuccessfully tries to log on to the system multiple times, this could indicate that someone is trying to break in to the system.

Because computer security is dynamic, security levels can be lowered temporarily to enable immediate resolution of an administration or network issue. However, changes such as this are often forgotten and never undone. As a result, a computer might no longer meet the requirements for enterprise security.

Regular analysis of the audit log enables an administrator to ensure an adequate level of security on each computer. This analysis enables an administrator to tune the security levels and, most importantly, to detect any security flaws that occur in the system over time.

Sometimes audit logs provide the only indication that a security breach has occurred. But if the breach is discovered another way, proper audit settings generate an audit log that contains important information about the breach.

HP Systems Insight Manager can configure an auditor-level user that only has rights to view, query, configure, and edit the audit log. The auditor also can provide and restrict access rights to other users, including the administrator or root super user. Only a trusted user can assign another user the read/write or read-only auditor role.

The audit log manager can:

- View raw audit log file content in a simple web page
- Limit audit log file growth with an automated roll-over feature
- Manually archive the log from the command line
- View the audit log

The common audit logging feature provides a secure way to record a history of operations and actions that occur within the Systems Insight Manager framework. The Systems Insight Manager audit log has the following features:

- A message is automatically entered in the audit log specifying that the audit log was archived by a particular user at a particular time.
- You can export selected rows to an ASCII file (comma-separated columns).
- Users who delete entries from the audit database are tracked through an entry that includes the user and the number of entries deleted.

## Configuring security with Systems Insight Manager

Because Systems Insight Manager is used with mission-critical and highly secure systems, it enables IT managers to track who logs on, when they logged on, and from where. It records all commands executed in a secure fashion.

Every Distributed Task Facility (DTF) tool launched by Systems Insight Manager is logged to the central management server (CMS) log file, which by default is readable by everyone but is writeable only by the Systems Insight Manager logger. In addition, the actions of web applications that make modifications to the CMS or a managed node are logged.

The log file records all login attempts, both successful and failed. This log file is protected by the CMS operating system and can be modified only by administrators or root users.



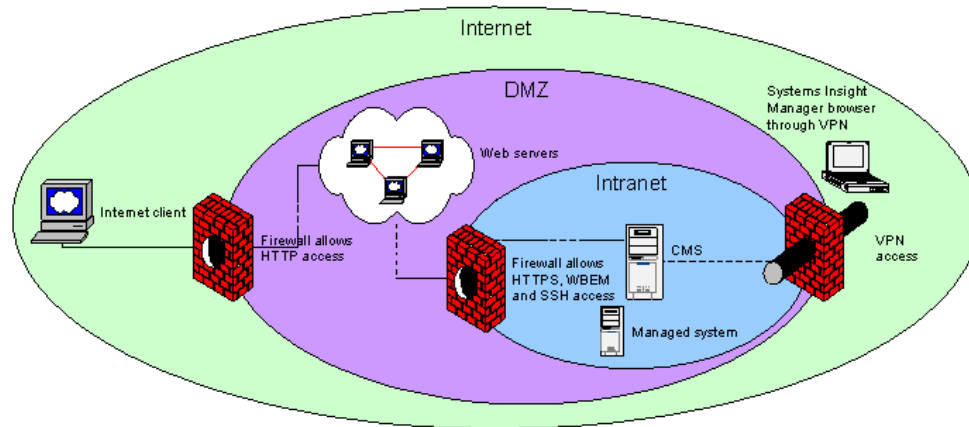
---

### Caution

A command line tool allows any Systems Insight Manager user to add entries to the log, enabling a malicious user to fill the log with junk. Although this breach will not allow existing records to be deleted, it is a form of denial of service because the CMS disk might become filled.

---

## Network security



HP management software such as Systems Insight Manager enables you to manage and control remote systems. However, the features that make Systems Insight Manager so powerful can also make it a security risk. If you do not take precautions, malicious users can compromise a management server and its managed devices.

The network should be secured through the use of firewalls between managed devices and the CMS and between the CMS and the client browser. The devices to be managed should be located within the corporate intranet, through firewalls separating different levels of trust (such as between an intranet and a DMZ), or accessed through a VPN.

Systems within a DMZ can have management agents (Web-Based Enterprise Management [WBEM] and DTF) that are accessed by a Systems Insight Manager server within the more secure corporate intranet. One firewall could prevent management traffic from passing through to the Internet, and a second firewall could restrict management traffic passing between the intranet and the DMZ. Although management requests can pass from the Systems Insight Manager server to the DMZ, only event notifications are initiated in the reverse direction. The Systems Insight Manager browser can access the server over the Internet using a VPN.

## Password security

Access to Systems Insight Manager requires one account in Systems Insight Manager and another account that can be authenticated against the operating system. Logging in requires the use of a user name and password that are authenticated against the operating system. The password should be sufficiently formatted for strength and kept properly secured.

Employees with access to the CMS should be encouraged to log off if they step away from the CMS. If a user logs in and then steps away, Systems Insight Manager will log off the user after 20 minutes.

---

### Note

Users who log in to a managed device can cause denial of service attacks at the CMS. These attacks, however, will not affect the data on the CMS, only its ability to deal with the data in a timely manner.

---

## Program access

You can assign specific program access to an individual user account to ensure that users only use those programs they need to do their jobs.

Organizations must understand that the administrator password allows full access to all Systems Insight Manager functionality. In other words, if you are an administrator, you can perform any action from any Systems Insight Manager tool. Organizations must carefully consider who should be given administrator status, and the administrator must be very cautious when issuing roles to other users.

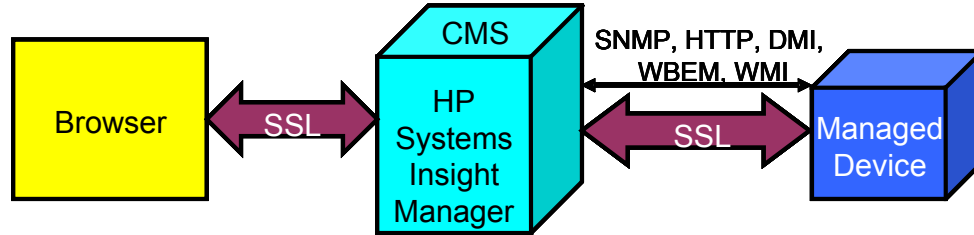
You can set different configuration rights for each user account on the CMS:

- Full rights for an administrator, who can control security for other users
- Limited rights for operators who need to manage reports for other users but do not need to alter security settings
- No rights for users who are not allowed to configure the CMS.

In addition to assigning program access to individual users, you can also control which Systems Insight Manager programs are allowed on a particular managed device.

Systems Insight Manager can restrict where users log in based on IP addresses. You can assign additional restrictions through the corresponding operating system account, such as account lock-out after invalid login attempts.

## Authentication



Systems Insight Manager uses Secure Sockets Layer (SSL) to provide secure communication between the CMS, the managed device, and a user's browser. An integral part of SSL is a certificate, which is a public document used to identify Systems Insight Manager.

### Authentication between the CMS and the managed device

The SSL certificate enables managed devices to authenticate a session with the CMS. To set up a connection between a managed device and the CMS, the managed device must be running a supported agent and the agent must be configured to trust the Systems Insight Manager server certificate.

### Trust configurations

Several levels of trust configurations are available for the supported agent.

- **Trust All** — The managed device will trust any Systems Insight Manager server, without verifying its digital signature. This level is the easiest to enable, but leaves the managed device open to any Systems Insight Manager server request, including requests from Systems Insight Manager servers installed in a test environment or bogus requests formatted to resemble valid requests.
- **Trust By Name** — The managed device will trust any of the Systems Insight Manager servers specifically named in a list, without verifying a digital signature. The Trust By Name configuration is still susceptible to bogus requests formatted as valid requests from the named servers.

- **Trust By Certificate** — The managed device will trust only Systems Insight Manager servers in the certificate list and will verify their digital signatures. This is the most secure option because it validates the digital signature of the request against a locally stored certificate, authenticating the sender and ensuring message integrity.

The browser might initially display a security alert when browsing to Systems Insight Manager, describing the certificate as *untrusted*. This occurs because the certificate is self-signed (signed by the Systems Insight Manager server) and the signer is not in the browser's list of Certification Authorities.

If this happens, an administrator has two options:

- Add the Systems Insight Manager server to its list of trusted certification authorities
- Obtain a certificate from a third-party certification authority and import it into Systems Insight Manager

### **Public key**

The SSL certificate contains a 2048-bit public key for the CMS. The managed device uses the public key to encrypt any communication it sends to the CMS. The public key supports a variety of cipher suites.

Any communication encrypted with the public key can only be decrypted with the private key of the CMS. If someone intercepts the encrypted data, it is useless without the private key.

### **Distributed Task Facility**

The DTF is an agent that provides a mechanism for launching a program on one or more managed devices and returning the results to the CMS. When the DTF receives a request from the CMS to execute a specific task on a set of managed devices, it communicates with each managed device, sending the request to run the specified task and consolidating the results.

The DTF communicates using the Secure Shell (SSH) program. Each managed device must be running an SSH server for the DTF to communicate with it. The SSH server on the managed device uses the public key to authenticate the CMS. After the managed device authenticates the CMS, it will completely trust that CMS without further authorization checks.

A log file tracks all the modification actions that have been taken.

## SSH security considerations

SSH is becoming an industry-standard program for remote execution and file copy, especially on Linux and UNIX platforms. It was created to address security issues with the “r” services: rexec, rcp, and rlogin. Risks that the SSH program mitigates include:

- Eavesdropping, by encrypting all communication between the client and the server
- Name service or IP spoofing, by using known “host key” lookups
- Connection hijacking, by checking that data has not been changed in transit using MD5 hashes
- Man-in-the-middle attacks, by using known host keys to identify the server and using public key authentication for user identification
- Insertion attacks, by using cryptographically strong integrity checks on all data sent between the client and the server

## WBEM

WBEM instrumentation uses HTTPS with a server certificate to identify managed devices. Systems Insight Manager verifies that the certificate is valid and that the name of the certificate matches the name of the managed device.

The managed device uses certificates to control whether it will accept a command from a particular CMS. Selecting *Trust By Certificate* on the managed device causes the system to verify the digital signature of the request. Other trust modes do not verify the digital signature.

Systems Insight Manager minimizes the number of ports used with managed devices, uses standard protocols over these ports, and enables the port numbers to be configured to meet the requirements of the enterprise security policy.

The HTTPS requests handled by WBEM are limited to WBEM requests, with no general-purpose web server access possible. Reserved ports 5988 and 5989 are used by default for this communication. HP recommends protecting these ports with firewalls.

Systems Insight Manager does not require other web servers on managed devices, although some functionality might be degraded if web-based management applications are not available. In general, management applications should be able to run on the Systems Insight Manager CMS and use WBEM to access the managed device.

## **Authentication between the CMS and applications**

Access to CMS applications is authenticated to the CMS by an application user name and password. Each application user account can have different authorizations and privileges. The Simple Object Access Protocol (SOAP) interface is designed to prevent unauthorized access by Systems Insight Manager applications.

## **Authentication between the CMS and the database**

Systems Insight Manager uses a database on the CMS as its primary storage facility. Access to the database varies according to the operating system:

- On a Windows CMS, access is controlled through Windows authentication using the account credentials attached to the Systems Insight Manager service.
- On a Linux or HP-UX CMS, access is controlled using a password, which is automatically created at installation and encrypted in an external file. The password is not stored or handled by Systems Insight Manager.

The Systems Insight Manager database can only be updated by the root CMS processes. The communication between the root CMS processes and the database is protected.

## **Authentication between the CMS and the browser**

You can communicate with the CMS through a browser. All communication between the browser and the CMS is encrypted using SSL.

When users log in to Systems Insight Manager, the browser does not display their passwords in clear text.

## Storing credentials

On the CMS, Systems Insight Manager uses the local file system to store certain configuration settings and parameters, including the SSL private key, WBEM passwords, and passwords used to access the database on Linux or HP-UX systems.

Windows 2000 and Windows 2003 provide advanced file system restrictions with NT File System (NTFS) and encryption capabilities to help protect sensitive files on the file system. NTFS is required on Windows systems running Systems Insight Manager.

On managed devices, Systems Insight Manager stores credentials to access management information, such as the DTF agent key, Simple Network Management Protocol (SNMP) community name, and WBEM access password, which should be protected with appropriate password policies.

## Installation considerations

Before installing Systems Insight Manager, consider adjusting security settings on your computer. With any system you are administering, log in as administrator or as root. The objective is to limit access to the administrator or root user.

Other considerations include:

- Configure firewalls to pass the appropriate traffic.
- Install the SSH server from the distributed software bundle on each system for DTF-related features. It is installed by using the management agent.
- Get access rights to any of the devices you will be managing. This includes:
  - SNMP — Read community name
  - WBEM — Appropriate user account on the managed device with domain administration rights

In addition, Systems Insight Manager provides several security-related options that must be configured during installation:

- Require trusted certificates instead of the default self-signed certificates.
- Configure SSH.
- Assign authorizations to user accounts.
- Configure each managed device to trust the CMS certificate.

## **Vulnerability assessment and patching**

Systems Insight Manager includes an HP ProLiant Essentials Patching and Vulnerability Assessment (PVA) Pack that integrates with the Harris STAT Scanner and Novadigm Patch Manager.

The PVA Pack prevents operating system disruptions by detecting vulnerabilities caused by inappropriate configuration settings or inadequate patch maintenance, and by recognizing malicious code that attempts to interfere with operating system or application operation.

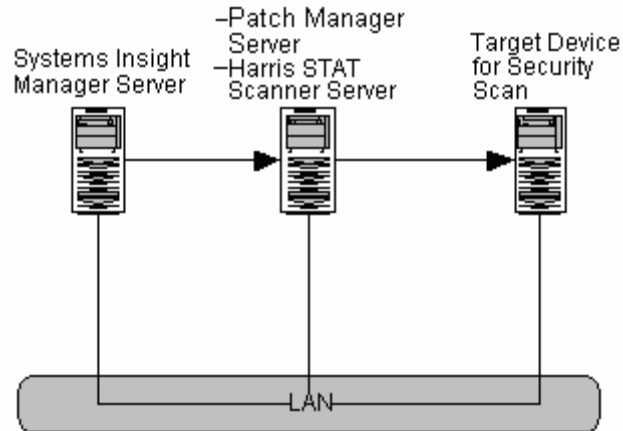
The STAT Scanner performs an accurate and comprehensive security vulnerability assessment of Windows systems, identifying operating system misconfigurations and missing patches. The STAT Scanner uses a unique database containing thousands of vulnerabilities. Many of the detected vulnerabilities can be automatically repaired.

The Novadigm Patch Manager can be configured with the Systems Insight Manager task manager to patch individual servers or groups of servers. It can be extended to include hot fixes and service packs.

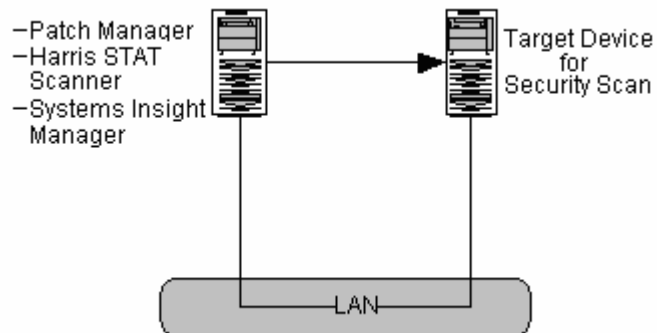
## Infrastructure

The PVA server environment consists of the following components:

- STAT Scanner server
- Patch Manager server
- Systems Insight Manager server
- Target servers

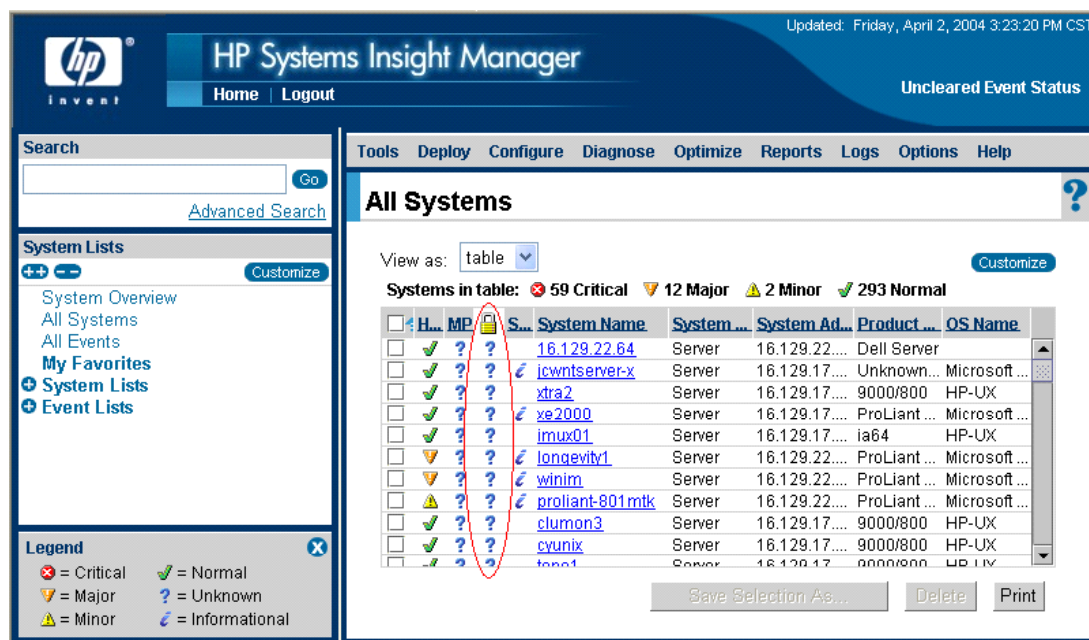


In the configuration shown in the preceding graphic, a shared server configuration is depicted with Systems Insight Manager on a dedicated server and the STAT Scanner and Patch Manager installed on a shared server.



In this configuration, all the components of the PVA infrastructure share a single server. The STAT Scanner and Patch Manager components can only reside on the Systems Insight Manager Windows console platform.

## PVA Interface



PVA vulnerability information is displayed in the Security column of the Systems Insight Manager console, shown circled in the preceding figure.

The server vulnerability is depicted using the color-coded icons in the following table.

Icon	Status
❓	The "unknown" icon might display because: <ul style="list-style-type: none"> <li>The server is not licensed.</li> <li>No seats are available.</li> <li>The server is licensed, but a scan has not yet been performed.</li> </ul>
ℹ	PVA can run on this server, but is not installed.
✓	No vulnerabilities are detected.
⚠	Minor vulnerabilities are detected.
⚠	Major vulnerabilities are detected.
✗	Critical vulnerabilities are detected.
No icon	PVA cannot run on this server.

## Antivirus strategy

Viruses spread from computer to computer when infected programs or data files are shared or exchanged (intentionally or unintentionally). Viruses can be designed to carry a payload or perform an undesirable action. Although some payloads are relatively benign, such as those displaying a derisive message, others can be damaging enough to corrupt all data on the hard drive of the infected server, steal data, or completely shut down a network.

Companies are increasingly overwhelmed by the propagation of viruses, Trojans, worms, and mixed-threat attacks. In 1988 there were six reported attacks; in 2003, from January through September, more than 114,000 incidents were reported.

These attacks take advantage of multiple entry points and vulnerabilities in enterprise networks. As a result of the increased complexity of these attacks, IT professionals are facing significant challenges in managing day-to-day operations and defining long-term strategies to deal with these threats.

Almost all attacks exploit known system vulnerabilities and poorly configured or misconfigured systems. A complete solution should begin with an assessment of all systems for these known issues.

---

**INTERNET**

CERT offers the most comprehensive information on current and past threats. To research current types of vulnerabilities, visit:  
<http://www.cert.org/>

---

The most common tool used to fight virus outbreaks is antivirus software. Most antivirus software comprises the following three components:

- A scanning application containing the user interface and configuration options
- A scanning engine that conducts the actual scan of the files
- Virus definitions or a signature file containing a database of known virus signatures

## Selecting antivirus software

Antivirus software can be installed almost anywhere in the enterprise: gateways, mail servers, production servers, client machines, and so forth. Different products offer different types of functionality depending on operating system and usage model. When helping a customer select a product, remember that antivirus software should:

- Be easy to install, manage, and update
- Run real-time scans continuously and unobtrusively in the background
- Provide on-demand scanning
- Scan for viruses when files are copied, opened, or downloaded, or when email attachments are received
- Scan programs when they execute

Three industry-recognized antivirus certification organizations can be useful in helping customers select a vendor:

- ICSA Labs Certification: [www.icsalabs.com/html/communities/antivirus/certifiedproducts.shtml](http://www.icsalabs.com/html/communities/antivirus/certifiedproducts.shtml)
- West Coast Labs Check-Mark Certification: [www.check-mark.com/cgi-bin/redirect.pl](http://www.check-mark.com/cgi-bin/redirect.pl)
- Virus Bulletin VB100 Certification: [www.virusbtn.com/100/](http://www.virusbtn.com/100/)

---

### INTERNET

Other antivirus and network security research organizations include:

- [www.icsa.net](http://www.icsa.net)
  - [www.eicar.org/](http://www.eicar.org/)
  - [www.wildlist.org/](http://www.wildlist.org/)
  - [www.cerias.purdue.edu/](http://www.cerias.purdue.edu/)
- 

In general, when a new threat is publicized, you should:

- Gather data about the possible attack
- Assess the vulnerability and deploy any temporary workarounds
- Notify management and users
- Test and deploy the update from your antivirus software vendor
- Scan and eliminate the virus from all systems
- Perform a post-mortem and implement improvements in your security process

## Summary

After learning more from you about best security practices, Jackie formed a security planning team to design the RC Engineering Security Policy. After signing off on the policy, Bob feels confident that his corporate network is secure from malicious and inadvertent attack.

The servers and other equipment brought from the company's location in Silicon Valley have been placed in a secure location at the corporate headquarters in Houston. Physical security measures have been implemented, including keypads and badges. In addition, Jackie has installed an intrusion detection system and has established a routine to perform periodic vulnerability scans. She has also updated the encryption and decryption technology currently used at RC Engineering.

Jackie and her security planning team identified the security vulnerabilities at RC Engineering and have taken steps to reduce their exposure to risks. The team has worked with all employees to strengthen user passwords and have hardened application code against actions that could result in enterprise downtime or data loss. The team also hung a poster in the break room that describes poor passwords and methods of data interception.

After the incident with the deleted web server files, Jackie audited all user accounts for administrative privileges and made certain that only the needed level of permissions have been granted to users.

When you installed Systems Insight Manager, you configured SSH and set each managed device to trust the CMS certificate. Additionally, you set up the program to require trusted certificates instead of the default self-signed certificates. These actions not only simplified the IT staff's job of configuring network security across all locations, but also gave them confidence that communication between the CMS and managed devices is secure.

After assessing all systems for known vulnerability issues, the RC Engineering enterprise security policy set forth written antivirus procedures. The security planning team installed antivirus software enterprise-wide—on gateways, mail servers, production servers, and client machines. Bob is so satisfied with the preventive measures you instructed his staff to implement that he has given Jackie a raise, out of money formerly budgeted to put out security fires after the fact.

## Learning check

1. Security is a balance between \_\_\_\_\_ and \_\_\_\_\_.
2. Two-factor authentication involves something the user knows, such as a password, and either: Select TWO.
  - a. Something that would be difficult for an intruder to guess (a user's birthday, maiden name, or favorite movie)
  - b. Something difficult to remember (a long string of numbers or a complex combination of symbols)
  - c. Something the user has (mobile phone, smart card, or pager)
  - d. Something that would use biometrics (fingerprints or retina scans)
  - e. Something that changes regularly (random codes or text against a patterned field)
3. What are the three levels of a trust configuration?  
.....  
.....  
.....
4. Which protocol is used to encrypt communication among the browser, the CMS, and the managed device?  
.....
5. When you browse to Systems Insight Manager, a message displays indicating a trust problem with the certificate. What are your two options to prevent this message from displaying again?  
.....  
.....  
.....
6. What should a complete antivirus solution begin with?  
.....  
.....

7. To prevent downtime and data loss, what steps should you take when you learn about a new virus threat? Select FOUR.
- a. Scan the subject line of all incoming email messages for the word *virus*.
  - b. Gather data about the possible attack.
  - c. Assess the vulnerability and deploy any temporary work-arounds.
  - d. Configure the supported agents on managed devices to trust the Systems Insight Manager server certificate.
  - e. Notify management and users.
  - f. Test and deploy the update from your antivirus software vendor.
  - g. Email employees about the nature of the threat, with the update procedure documentation in an attachment.
  - h. Move important files off the mail server.



### Objectives

After completing this module, you should be able to use the six-step HP troubleshooting methodology to:

- Collect performance management data
- Evaluate the data to detect and analyze performance bottlenecks
- Develop a plan of action to:
  - Scale up performance
  - Scale out performance
- Execute the plan of action
- Determine the effectiveness of the action
- Implement performance-based preventive measures

## Introduction

As a result of a recent acquisition of a widget design company, RC Engineering is in two locations and with more than 750 employees. The network spans both locations; it has grown to comprise a variety of server platforms and hundreds of peripherals.

In the past few months, users have been complaining of degrading network performance. In addition to the documentation management system being slow to load, the IT staff has been reporting that the network management functions they perform are now also taking substantially longer than expected.

Jackie, the IT administrator, has already implemented a management system based on HP Systems Insight Manager, Insight Manager 7, and Version Control Repository Manager (VCRM). HP ProLiant Essentials Performance Management Pack (PMP) has been installed as part of the Insight Manager 7 installation. However, Jackie and her staff have not yet licensed all the servers and have not yet configured the performance monitoring infrastructure.

RC Engineering also has successfully implemented Lights-Out technology across remote locations. Jackie has noticed that performance of the integrated Lights-Out (iLO) management processor is slow in the remote console interface. Although iLO is the tool of choice for controlling the power on the server and accessing the server when the operating system is not running, performance is less than optimal for servers when using the remote console feature on a Microsoft Windows platform.

After adding the IT staff as users, Jackie has detected a noticeable decline in server performance. You have arranged a meeting with Bob, Jackie, and the rest of the IT staff to discuss options to improve performance. During the meeting you are asked to assist with the following action items:

- Determine which servers, management agents, and HP programs must be updated, and perform the necessary updates.
- License selected servers to be monitored and configure them.
- Perform static analysis to determine if any configuration issues might impact performance.
- Show Jackie and her staff how to create reports based on logged performance statistics.
- Configure performance-based alerts.
- Detect, analyze, and resolve performance bottlenecks.
- Compare current performance with baseline performance.

## Using the HP troubleshooting methodology

The six steps in the HP troubleshooting methodology are:

1. Collect data.
2. Evaluate the data to determine the cause of the problem.
3. Develop a plan of action.
4. Execute the plan of action.
5. Determine the effectiveness of the action.
6. Implement preventive measures.

You can use the HP troubleshooting methodology in performance management to quickly and effectively restore server performance to optimal.

---

**Note**

For more information on the HP troubleshooting methodology, refer to the HP Accredited Information Specialist (AIS) Implementing ProLiant Servers course, Module 6 — HP troubleshooting methodology in an SMB environment.

---

## Step 1 — Collecting data

Collecting data is the first step of the HP troubleshooting methodology. Before analyzing performance and attempting to correct performance issues, you must understand what the performance expectations are and what behavior is normal. Determine:

- What performance issues is the user community experiencing and when are the users experiencing them?
- What is the performance baseline of the system?
- What is the maximum performance expectation for the system?
- What impact will the selected performance monitoring and data collection tools have on the environment?
- When it is best to monitor the system behavior?
- What performance metrics are being used?

### Selecting monitoring tools

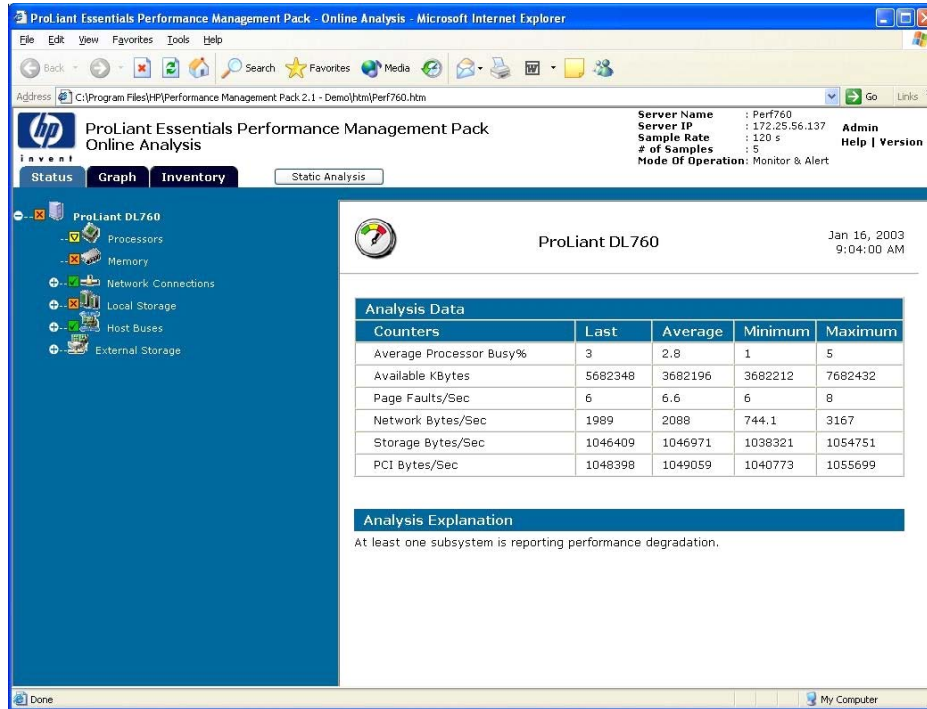
When monitoring system performance, remember that one isolated statistic will not provide the overall view that you need to resolve bottlenecks. Select monitoring tools from three broad categories—hardware, operating system, and application—to provide subsystem statistics. Viewed together, they can give you a larger understanding of system performance.

Several monitoring tools exist that can meet your objectives, but they differ in terms of complexity, price, capabilities, and supported platforms. Some of the more-sophisticated (and more expensive) tools can monitor more than one area.

To choose the most appropriate monitoring tool, complete these steps:

1. Define exactly what you need to monitor.
2. Determine your infrastructure requirements.
3. Choose the capabilities the monitoring tool must possess in addition to performance monitoring, such as:
  - Bottleneck detection and reporting.
  - Logging and offline analysis.
  - Alerting based on a performance change.
  - Performance analysis and suggestions for resolution.
  - Multiple platform (hardware and operating system) support.
4. Set budgetary constraints related to purchase, support, and staff training.
5. Compare each monitoring tool against your requirements and select the tool most qualified for the job.

## Performance Management Pack



Performance tuning and optimization is essential to good infrastructure management. Although availability technologies such as fault tolerance are reactive in nature, the HP ProLiant Essentials Performance Management Pack (PMP) is a proactive tool that helps an administrator identify bottlenecks on ProLiant servers before they become critical issues.

### Note

The current version of PMP is supported on Windows platforms only. The next release of PMP, version 3.0, will monitor performance of ProLiant servers running Linux.

The two distribution mechanisms for PMP are:

- The <http://www.hp.com/servers/proliantessentials/pmp> website
- The HP Management CD, which is part of the ProLiant Essentials Foundation Pack

### Demo

Your instructor might now demonstrate PMP 2.1.

## Features

Easy to use because you can access it locally or remotely, PMP offers these advantages:

- **Predetermined performance parameters** — Out of the hundreds or even thousands of performance parameters available to monitor, PMP monitors only those that directly or indirectly affect performance of your particular ProLiant server configuration.
- **Explanation of values and results** — In addition to monitoring selected performance parameters, PMP explains the meaning of its parameters and their values, and, when possible, provides potential solutions and recommendations.
- **Integration with Insight Manager 7 and HP Management Agents** — PMP installs silently as part of Insight Manager 7 service pack (SP) 2, uses the Insight Manager 7 alerting mechanism, and adds performance-specific indicators to its home page. They also share the same information repository. When released, PMP 3.0 will integrate with HP Systems Insight Manager.
- **Logging and reporting** — PMP logs the performance statistics in a relational database based on Microsoft Data Engine (MSDE) or Microsoft SQL Server. No additional setup of the database is necessary to accommodate the performance repository. Logged data can be extracted for analysis and reporting.
  - **SQL Server** — Insight Manager 7 and PMP support local and remote repositories when implemented with SQL Server, and can manage up to 5000 monitored devices. This configuration provides the highest performance and is recommended by HP. However, SQL Server must be licensed.
  - **MSDE** — MSDE is a freely distributed database engine from Microsoft that is included on the HP Management CD and can be used with Insight Manager 7 and PMP. However, compared to SQL Server, MSDE supports only a local repository, monitors up to 500 devices, and offers slower performance.

- **Browser-based user interface** — PMP uses a standard browser-based user interface, currently based on Microsoft Internet Explorer. This type of interface eliminates a need for specific software to be installed at the client machine and facilitates access from any machine running Windows, Internet Explorer, and Java.
- **Analysis of all major components** — The current version of PMP analyzes and detects bottlenecks on all major server components, including processors, memory, disk subsystem, network, and PCI buses.
- **Support for all major ProLiant servers and options** — The server to be monitored must be a ProLiant server for which PMP has been calibrated. Otherwise, PMP displays an unknown status for the server or for the unsupported options.

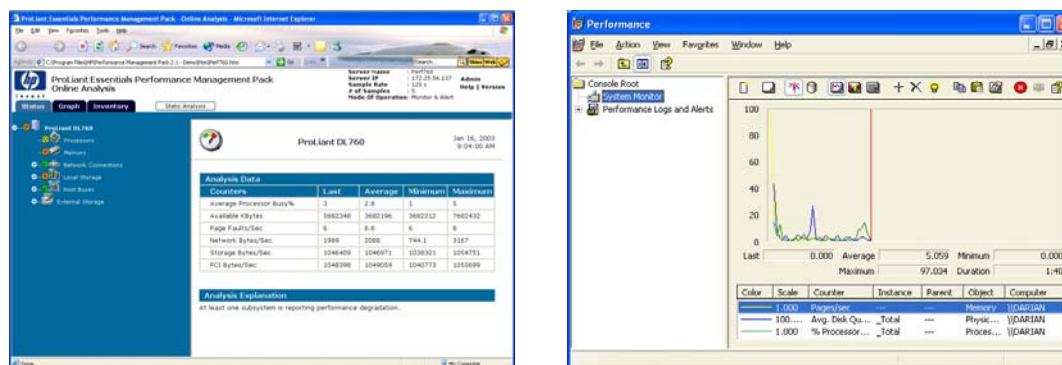
---

**INTERNET** All supported servers and options are available in the PMP Support Matrix available from: <http://h18004.www1.hp.com/products/servers/proliantessentials/valuepack/PMP/index.html>

---

- **Real-time and offline analysis** — PMP performs performance analysis in real time. You can log this information in a database for offline analysis.
- **Static analysis and hardware inventory** — Static analysis determines whether the server has any configuration issues that could result in future performance bottlenecks. Examples of such issues are half-duplex network mode, unassigned disk drives in an array, reduced SCSI speeds, and unbalanced PCI buses. PMP also gathers server inventory information.

## Competitive tool comparison



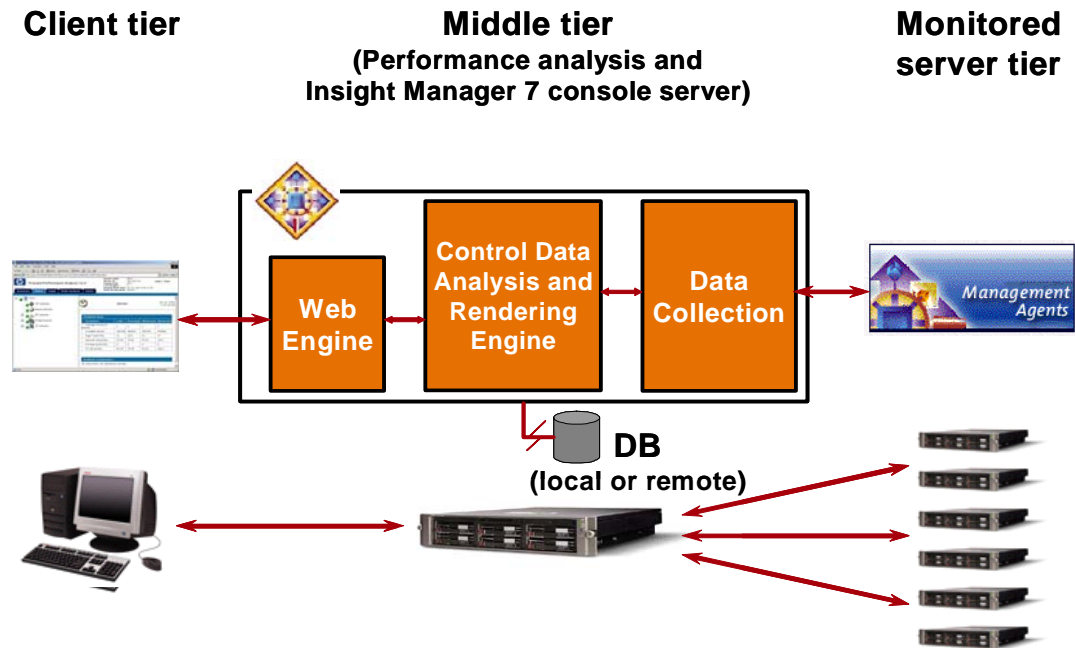
Many operating systems, including Windows NT, Windows 2000, and Windows Server 2003, include a comprehensive performance monitoring tool. These tools are often free and are included with the operating system. Key differentiators for PMP are:

- **Hardware-specific monitoring** — It is possible to collect and analyze many of the same performance statistics that PMP uses with operating system tools. Some of them, however, are either not available from the operating system or must be calculated by combining other monitored parameters. PMP monitors only those parameters that directly or indirectly impact server performance.

In addition, PMP is designed to monitor ProLiant servers and their options and often communicates directly with the hardware. It is aware of the performance capabilities of ProLiant hardware and uses this knowledge when determining performance levels.

- **Explanation and suggestions** — Operating system tools do not help you understand and interpret which parameters are important, what their relationship is, and how they apply to ProLiant hardware. PMP incorporates this intelligence and distills it to easily understood explanations and recommendations.
- **Integration with Insight Manager 7 and performance-based alerting** — Because PMP is integrated with Insight Manager 7, you have just one management tool to install, use, and maintain. If you already use Insight Manager 7 to alert you of impending hardware and software issues, you can add performance-based alerting as well.
- **Static analysis and inventory** — Hardware configuration analysis and inventory typically are not available with operating system tools.

## Architecture



The PMP architecture is based on three tiers:

- **Client tier** — Consists of Windows-based workstation and a supported web browser. The browser is used to access Insight Manager 7, and thus PMP.
- **Middle tier** — Hosts Insight Manager 7, PMP, and optionally the database repository. The database repository can be local (MSDE or SQL Server) or remote (SQL Server only). The PMP engine consists of three components:
  - **Web engine** — Communicates with the web browser, which is used to access the PMP console.
  - **Control data analysis and rendering engine** — Analyzes collected statistics, generates alerts, and renders the information.
  - **Data collection** — Collects data from the management agents running on the monitored servers. PMP uses Simple Network Management Protocol (SNMP) to collect data from the monitored servers.
- **Monitored server tier** — Consists of supported ProLiant servers monitored by PMP. These servers must have the minimum supported version of the management agents running; no additional software needs to be installed.

These tiers can be combined if necessary. The client tier and the middle tier can reside on the same machine; you can also run all three tiers on the same server.

## Client requirements



<b>Microsoft operating systems</b>	Windows NT 4.0: Workstation, Server, or Advanced Server
	Windows 2000: Professional, Server, Advanced, or Datacenter
	Windows Server 2003: Enterprise or Datacenter
	Windows XP
<b>Microsoft web browsers</b>	Internet Explorer 5.5 with SP1 or later (128 -bit)
	Internet Explorer 6.0

A supported web browser is used to connect to the PMP home page, which is accessible from the Insight Manager 7 home page.

PMP clients can be used to interactively monitor the real-time performance of a server or to replay information extracted from performance logs that analyze stored data. The server to be monitored must be a ProLiant server for which PMP has been calibrated. Only supported server types are displayed in the administration.

---

### Note

For additional information, refer to the HP ProLiant Essentials Performance Management Pack Support Matrix available from: <http://h18004.www1.hp.com/products/servers/proliantessentials/valuepack/PMP/index.html>

---

## Middle-tier requirements



<b>Microsoft operating systems</b>	Windows NT 4.0: Server or Advanced Server
	Windows 2000: Professional, Server, Advanced, or Datacenter
	Windows Server 2003: Enterprise or Datacenter
	Windows XP: Professional
<b>Microsoft databases</b>	SQL Server 7.0 with SP4
	SQL Server 2000 with SP2
	MSDE 1.0 with SQL Server 7.0 SP4 or MSDE 2000

The server on which the primary analysis is conducted is called the *performance analysis server*. PMP requires this server to run a supported version of the operating system and have access to a database repository.

The database repository can be local to the middle tier or remote on another server. The local repository can use MSDE or SQL Server; the remote repository must use SQL Server.

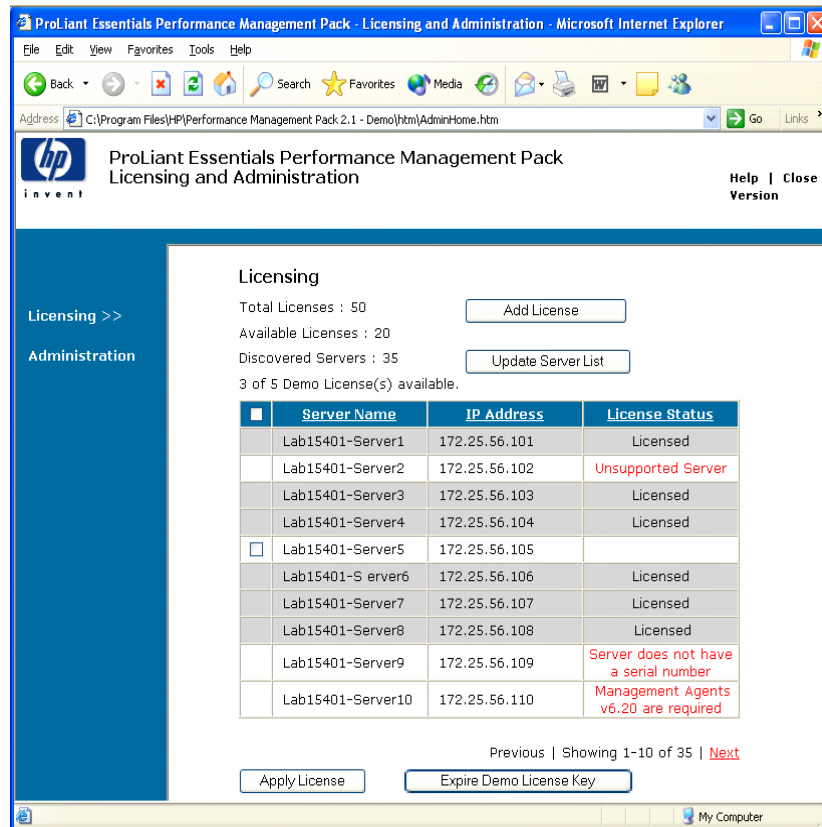
## Monitored server requirements



<b>Microsoft operating systems</b>	Windows NT 4.0: Server or Enterprise
	Windows 2000: Server, Advanced Server, or Datacenter
	Windows Server 2003: Enterprise or Datacenter
<b>HP management agents</b>	Windows NT 4.0: agents version 6.10.01
	Windows 2000: agents version 6.20
	Windows Server 2003: agents version 6.31
<b>Servers and options</b>	Refer to the support matrix document

The monitored server must be a ProLiant server for which PMP has been calibrated. PMP administration screens display only supported server types and options. If a server or an option is not supported, PMP continues operating normally, but shows such a server or an option as *unknown*.

## Licensing and license management



With every installation of the middle tier, the user is granted one free full-use license. This license is valid only for monitoring any single server on the network. Purchased licenses are good for the life of the hardware configuration on which they are used.

PMP licenses are tracked and priced by the number of servers to be analyzed.

### Example

If 25 servers are in the monitored environment, 25 licenses must be purchased (unless one is licensed with the free license).

When you use a license, it becomes “attached” to the monitored server and stays with that server until the server is retired. You cannot remove or transfer PMP licenses, nor can you “float” them among servers.

### Note

Exceptions to the no-transfer/no-removal license policy include licenses lost from replacement of in-warranty servers and licenses purchased for retirement or consolidation of servers. In these cases, contact your sales or channel representative to recover these licenses.

PMP licenses are stored on the middle-tier server. Multiple middle-tier servers can monitor the same target server; however, you must use multiple licenses. Multiple middle-tier installations require multiple license keys, each with their own number of licenses relative to the number of monitored servers.

HP offers flexible PMP licensing through ProLiant Essentials Value Packs, suitable for a single copy to large volumes. Three licensing choices are available:

- **Single License Kit** — A single copy of the software, documentation, and a single-server license.
- **Flexible Quantity License Kit** — A single copy of the software, documentation, and license key. The single license key can activate the exact number of licenses desired and purchased. Thus, a single license key can be purchased for 1-to-n servers and licenses can be applied to any monitored server. For example, a Flexible Quantity License Kit with a quantity of 25 enables you to monitor any 25 servers with a single instance of PMP.
- **Master License Agreement (MLA)** — Consists of a single license key account for incremental purchases that take place over time, with a negotiated price under such an agreement. Licenses are then ordered against this MLA and billed accordingly. The MLA requires the active involvement of an HP account manager. The license mechanism at the product level works the same for both MLA and non-MLA scenarios.

License keys are additive. Adding to your monitored server selection means purchasing only the additional number of licenses.

The Internet list price for a PMP license is \$99 (U.S.) per monitored server. Although there is no volume-tiered pricing, standard ProLiant Essentials internal and channel discounts apply.

---

**Note**

The price of \$99 US applies to United States only, and may vary in other countries.

---

To purchase PMP licenses:

1. Determine the type of license needed.
2. Establish an agreement with HP, if necessary.
3. Purchase the appropriate licensing option using the corresponding part number (through an HP account manager or directly, depending on the type of license).

**INTERNET**

---

For additional licensing information, refer to: <http://h18004.www1.hp.com/products/servers/proliantessentials/valuepack/licensing.html>

---

## Licensing monitored servers

PMP is a for-fee product. You receive one free license to use on any monitored server. However, you must purchase additional licenses to monitor the remaining servers. Demo licenses can be obtained from your HP account representative.

To license qualified servers for PMP monitoring, these servers must first be discovered by Insight Manager 7. If a specific server does not display on the Insight Manager 7 home page, execute the discovery and ensure that the server does display in the server list.

Before you can monitor a server, you must apply a PMP license to that server using the PMP Licensing screen. Using this screen, you can:

- Add a license key that contains one or more licenses
- Update the PMP server list to match the server list maintained by Insight Manager 7
- Apply a purchased, free, or demo license to a server
- Expire a demo license key

---

**Note**

When you apply a license to a server, you may not remove and reuse the license. Only demo licenses can be expired.

---

You can also determine the number of total licenses, available licenses, unreachable licenses, and demo licenses.

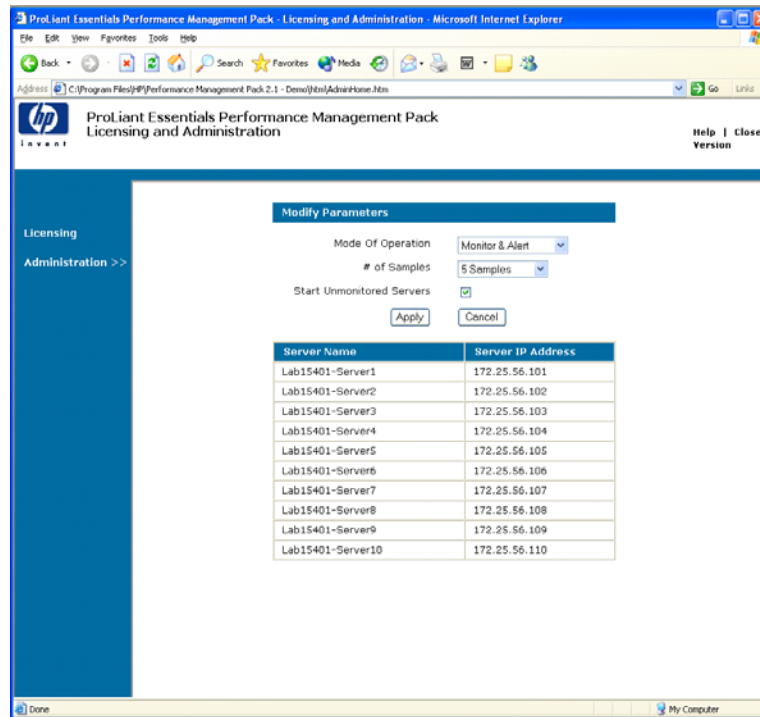
---

**Demo**

Your instructor will now demonstrate licensing and license management.

---

## Administration



You can administer PMP locally or remotely; you can access it through the Administration link.

The administration page displays a list of monitored servers and their configuration parameters. Select one or more entries and click *Modify Selected* to change the mode of operation, number of samples, and monitor status.

- **Mode of operation** — Determines whether PMP monitors server performance in real time (Monitor); logs the gathered statistics in a database (Log); alerts the administrator of a performance status change (Alert); or performs a combination of these operations.
- **Number of samples** — Controls the number of times PMP must collect information before it reports on a performance status. For example, if the number of samples is set to 5, PMP must collect five samples to determine the performance status of the monitored server.
- **Monitor status** — Can be started or blank (stopped).

You cannot use PMP to change the rate at which it gathers performance statistics. PMP determines the sample rate of the management agents at each server and gathers these statistics based on the same sample rate. Use Insight Manager 7 or the HP Management Agents applet in the Control Panel to change this sample rate.

---


### Demo

Your instructor will now demonstrate PMP administration.

---

## Cost of ownership and ROI calculator

**ProLiant Essentials Performance Management Pack**  
 return-on-investment calculator



performance resolution costs without PMP			performance resolution costs with PMP		
number of servers to analyze	<input type="text" value="25"/>	totals/yr	number of servers to analyze	<input type="text" value="25"/>	totals/yr
average number of performance issues experienced per month per server	<input type="text" value="1/4"/>	75	average number of performance issues experienced per month per server	<input type="text" value="1/4"/>	75
average hours spent analyzing each performance issue	<input type="text" value="3"/>	225	average hours spent analyzing each performance issue	<input type="text" value="1/2"/>	37.5
average hourly rates (loaded costs):			average hourly rates (loaded costs):		
internal personnel	<input type="text" value="\$85"/>		internal personnel	<input type="text" value="\$85"/>	
outside consultants	<input type="text" value="\$180"/>		outside consultants	<input type="text" value="\$180"/>	
%age of hours by outside consultants	<input type="text" value="0%"/>	\$19,125	%age of hours by outside consultants	<input type="text" value="0%"/>	\$3,188

acquisition and deployment costs of PMP			total savings with PMP		
net cost per PMP license	<input type="text" value="\$99"/>	\$2,475	performance resolution costs per year:		
Is Insight Manager already installed?	<input type="text" value="No"/>		without PMP	<input type="text" value="\$19,125"/>	
deployment time for IM7 console (hours)	<input type="text" value="2"/>	\$170	with PMP	<input type="text" value="\$3,188"/>	-/-
console server hardware	<input type="text" value="\$1,375"/>	\$1,375	cost of deployment of PMP	<input type="text" value="\$5,083"/>	-/-
management agent installation/upgrade time per server (hours)	<input type="text" value="1/2"/>	\$1,063	total first-year net savings	<input type="text" value="\$10,855"/>	
total acquisition & deployment costs	<input type="text" value="\$5,083"/>		net first-year savings per server	<input type="text" value="\$434"/>	
			time to roi (months)	<input type="text" value="3.8"/>	

The Return-on-Investment (ROI) calculator is a Microsoft Excel-based calculator that enables you to project the PMP implementation costs and determine the time that must elapse before you can realize a return on the investment.

**INTERNET**

The ROI calculator is available as a free download from:  
<http://www.hp.com/servers/proliantessentials/pmp>

### Demo

Your instructor will now demonstrate the Return-on-Investment calculator.

## Insight Manager 7 console integration

**Status indicators**

- Unknown
- Normal
- Approaching bottleneck
- Confirmed bottleneck
- Critical condition

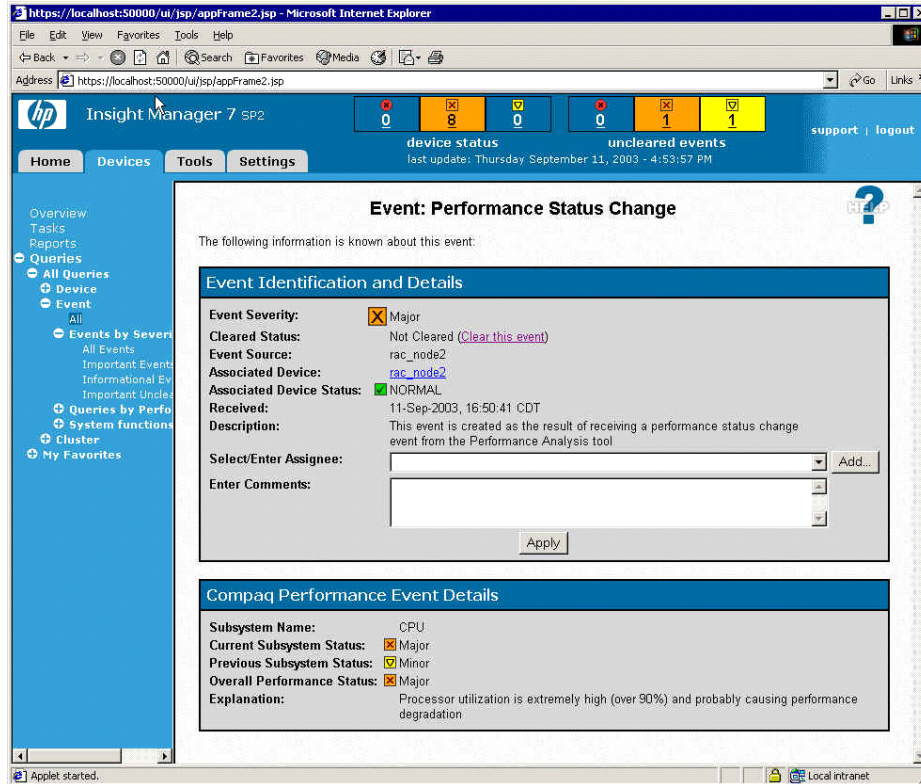
HW	MP	SW	PF	Device Name
✓		●	▲	oala
✓		●	✓	idesrv
✗		●	✓	abqwebstg
✗		●	✓	abqdev01

PMP adds a status indicator column to the Insight Manager 7 home page. This column is labeled *PF* (Performance) and displays one of five possible states:

- **Unknown (blue circle)** — Represents an unknown state of the server, which can result from:
  - The server not being licensed
  - The server being licensed but monitoring not being enabled
  - PMP not collecting enough performance information to determine the server state
  - The server not being supported or not meeting minimum requirements
- **Normal (green square with a checkmark)** — Represents a server with a normal (non-bottlenecked) performance
- **Approaching bottleneck (yellow square with an inverted triangle)** — Represents a server that has at least one subsystem approaching a bottleneck
- **Confirmed bottleneck (amber square with an 'X')** — Represents a server that has at least one subsystem in a bottleneck condition
- **Critical condition (red circle with an 'X')** — Indicates that a critical error occurred when PMP attempted to communicate with the monitored server

Each status indicator represents the overall state of the monitored server. When you click the *Unknown* or *Critical* indicators, a screen explains the potential causes. Clicking the other indicators displays the PMP home page.

## Proactive performance monitoring



PMP supports proactive notification of device status change using the Insight Manager 7 notification mechanism. Using any of the supported notification methods, you can set up rules to generate these notifications when a performance issue is detected. These notification methods include the following tasks:

- **Email** — Sends notifications through email
- **Pager** — Sends notifications to pagers
- **Application launch** — Launches an application on the system running Insight Manager 7

To set up performance notifications:

1. Create an event query (or use an existing one).
2. Set a schedule to execute this query.
3. Associate the query with an email or pager task, or launch the application.

An alert is generated when the performance status of a monitored server changes.

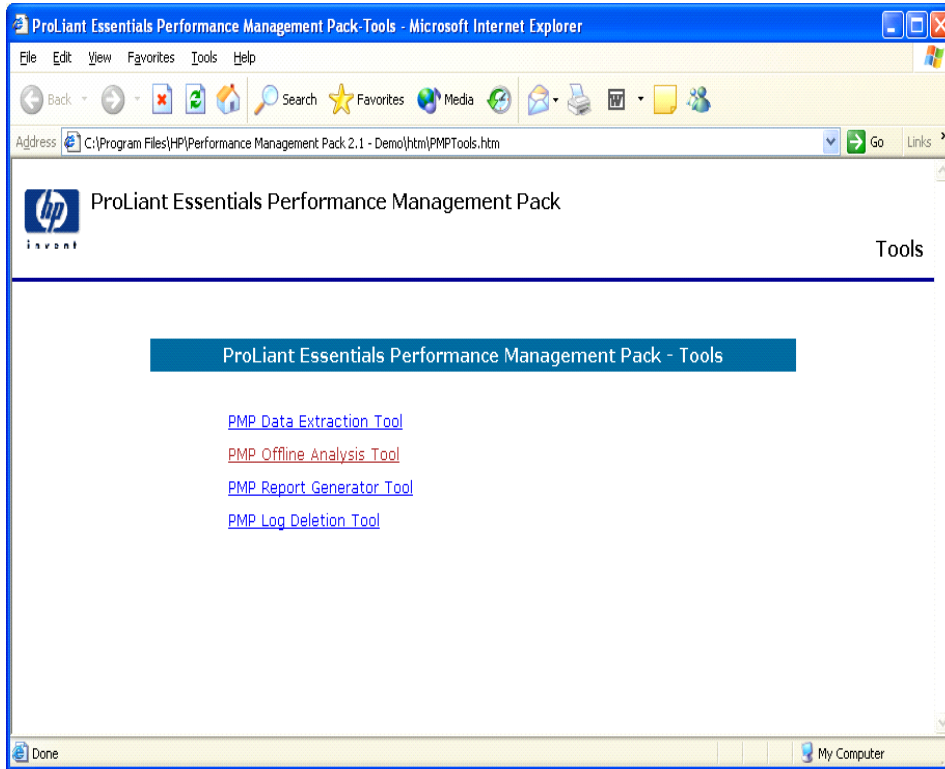
---

### Demo

Your instructor will now demonstrate how to configure email-based performance alerts.

---

## Logging and logged data manipulation



PMP enables you to log the performance statistics into a database and extract the data for playback (offline analysis) and reporting. To enable logging, use the *monitor and log* or the *monitor, log, and alert* settings for the mode of operation.

Depending on the configuration of your Insight Manager 7 tier, the performance database receiving the collected statistics can be local or remote. For the local database, you can use MSDE or SQL Server. A remote database (a database that is not located on the same server as Insight Manager 7) must use SQL Server.

---

### Note

Although logging can be enabled remotely, logged data manipulation such as data extraction, log deletion, offline analysis, and report generation must be performed at the performance analysis server (the middle tier).

---

To enable logged data manipulation, PMP provides these tools:

- Data Extraction
- Offline Analysis
- Report Generator
- Log Deletion

## Data Extraction

Before you can analyze data offline, you must extract it from the performance database. Data is extracted from the database into an XML-formatted file. To perform data extraction, use the Data Extraction tool.

To extract statistics from desired servers, specify the:

- Logged server name (the server from which the measurements were taken)
- Recorded session (a set of contiguous measurements)
- Date and time range within the recorded session
- Name and location of the output file

The statistical output is an XML file that is used for offline analysis.

## Offline Analysis

The Offline Analysis tool is functionally similar to the online version of PMP. The major difference between the two is how the data is retrieved and manipulated on screen. During offline analysis, you must specify an XML file instead of a live server. This XML file is created with the Data Extraction tool.

To manipulate displayed data, the Offline Analysis tool enables you to advance through the data stream automatically or manually using VCR-like control buttons. You can pause the playback, move back and forth by one sample or by 25 samples, and jump to the beginning of the session.

To use the Offline Analysis tool, specify the:

- XML file
- Sampling rate
- Date and time range

## Report Generator

Report generation uses data stored in and extracted from the database to provide a performance summary for the server or raw statistics to use in trending analysis.

The Data Reporting tool supports two report-generating options:

- **Generate a summary report in HTML format** — Generates an HTML report that shows the percentage of time the server was in a bottlenecked state and the overall performance utilization for a server, categorized by its subsystems.
- **Extract data to a CSV file** — Extracts every data point stored in the performance database and stores this data in a comma-separated value (CSV) file for import into desktop analysis or reporting tools, such as Microsoft Excel.

The Report Generator tool requires:

- Logged server name
- Recorded session
- Date and time range within the recorded session
- Choice of the summary report or the CSV report, along with the name and location of the output files

## Log Deletion

The Log Deletion tool enables you to delete recorded sessions from the performance database. Use this tool on a regular basis to remove unwanted recorded sessions and to reduce database size.

The Log Deletion tool requires these two parameters:

- Logged server name
- Recorded session

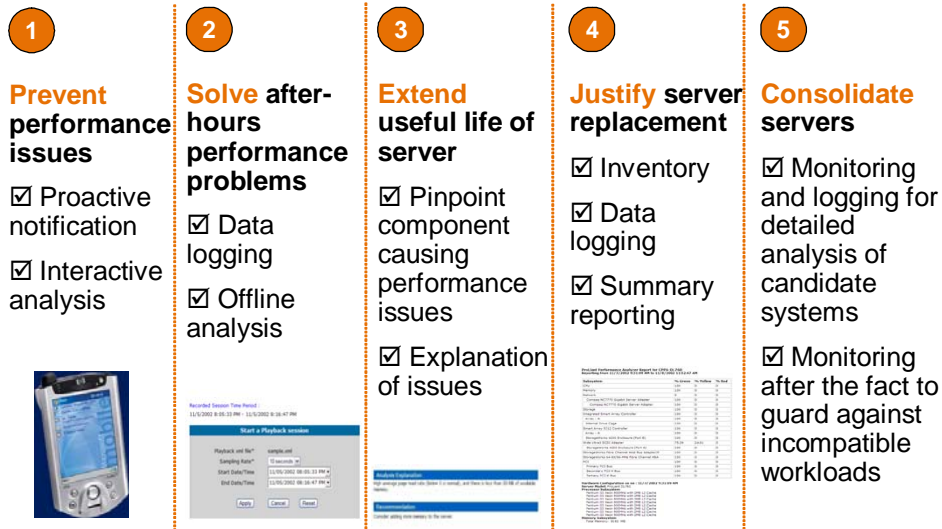
---

### Demo

Your instructor will now demonstrate how to use these PMP tools.

---

## PMP usage scenarios

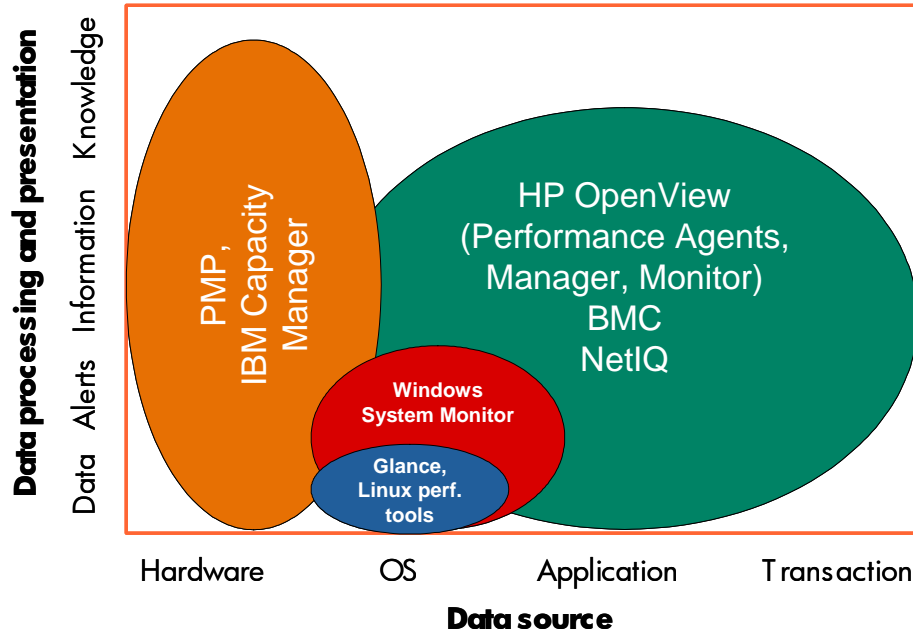


Typical scenarios include using PMP to:

- Prevent performance issues** — Customers who deploy ProLiant servers in production environments can use PMP to identify and diagnose performance problems that can lead to end-user dissatisfaction. In many cases, such diagnosis can be made or predicted before the performance issues occur. Active notification of impending problems can be set up to alert administrators for resolution.
- Solve after-hours performance problems** — Most customer sites do not have the staff or resources to monitor their network environment 24 hours a day. Performance problems can arise any time, day or night. Using a combination of Insight Manager 7 notification and remote access to the PMP tool, administrators can diagnose the problems at any time and determine the appropriate type of response. Using data logging and the offline analysis tool, the administrator can replay the after-hours activity and diagnose it as if working with a live system.

- **Extend the useful life of existing servers** — With shrinking IT budgets, many administrators must make their equipment last as long as possible. The interactive and logging capabilities of PMP enable you to determine which server components need to be upgraded and at the same time minimize downtime and money spent on upgrading or replacing the system. The static analysis enables you to configure server options to maximize their performance output.
- **Justify replacement of old servers** — When justifying system or component replacement, you often need detailed statistical reports. PMP provides system health reports and detailed performance data that can be imported into desktop analysis tools to create historical performance trends.
- **Identify server consolidation opportunities** — The performance data collected by PMP can be used to identify servers running workloads that are complementary to one another with respect to which server subsystems are taxed and which are not. PMP enables you to take a systematic approach to server consolidation. After consolidating workloads, you can use PMP to monitor the server to ensure that the assumptions you made during the consolidation process were correct.

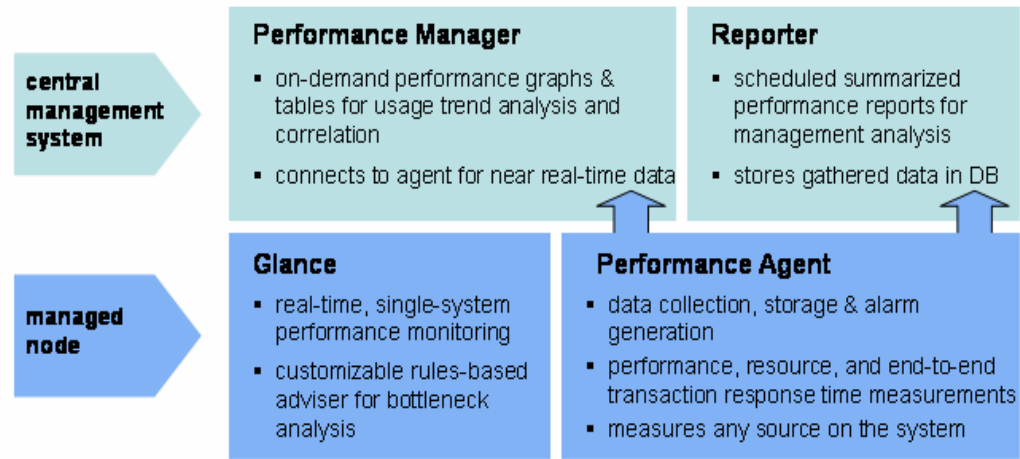
## Data collection, processing, and presentation



Performance management tools vary in the way they collect, process, and present data. The preceding graphic illustrates these dimensions and where the various performance tools discussed in this module fit in.

- **Data collection** — Refers to the source of the performance data. The source might be the physical hardware, which requires hardware-specific agents, operating system, applications, or even business-specific process transactions. The amount of data also varies from one tool to the next. Some tools monitor a few components or parameters but collect many details; others collect a few details about many components and parameters.
- **Data processing and presentation** — Refers to what the tools do with the collected data and how they inform the administrator of their findings. The methods include:
  - Simple data presentation
  - Alerting
  - Correlation and information
  - Intelligent processing and advise generation

## HP OpenView performance tools



HP OpenView system performance toolset

Segmenting the OpenView performance products enables customers to offload complex processing from the production servers being monitored, without inundating the network or analysis system with excessive data. Business requirements are matched by complimentary and integrated solutions, which include the following OpenView products:

- **Performance Agent** — Collects, summarizes, and logs resource and performance measurement data from applications, databases, networks, and operating systems. It is installed on each system to be monitored throughout the distributed environment.
- **Performance Manager** — Uses data collected from the Performance Agents and several other sources to isolate performance bottlenecks and maximize resource uptime.
- **Reporter** — Automatically transforms the data captured by OpenView agents running on supported platforms into management information, such as reports on application response times and service availability, at regularly scheduled intervals.
- **Glance** — Provides immediate system performance information, enabling you to examine system activities, identify and resolve performance bottlenecks, and tune the system for efficient operation.

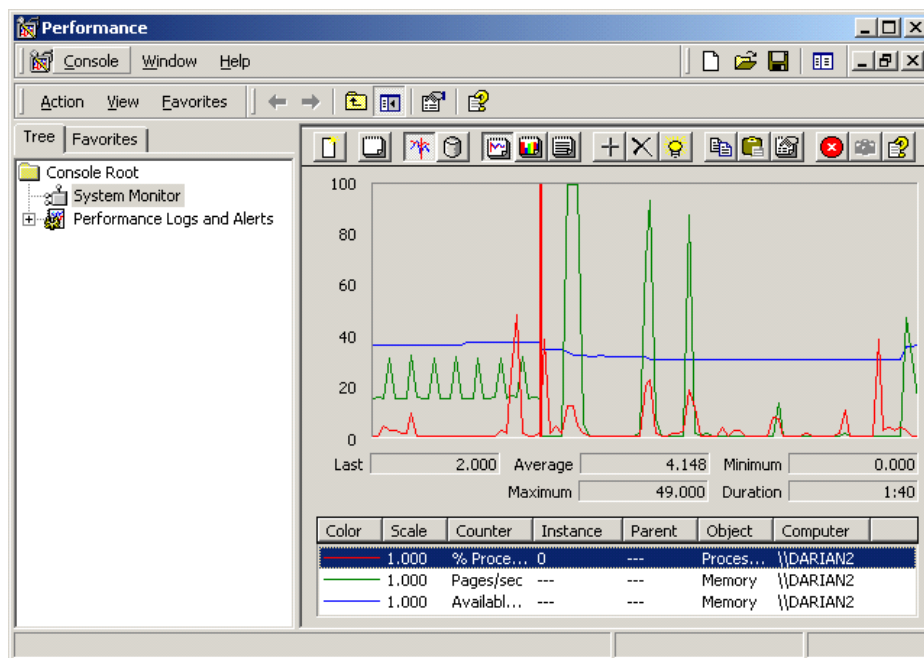
Operations for Windows is another performance product that provides event monitoring with embedded lightweight performance analysis. OpenView Operations performs Windows-based event management, proactive performance monitoring, automated alerting, reporting, and graphing for Windows, Linux, and UNIX systems, middleware, and applications.

## Operating system tools

Each operating system includes built-in or add-on monitoring tools that enable system administrators to observe the behavior of various operating system components. Most of these tools report on configuration, health, system activities, and performance of the system.

Because many operating systems run on a variety of hardware platforms, the associated monitoring tools must be generic in nature to be portable. Consequently, only some hardware and application monitoring capabilities are built into these tools, and they are generic in nature. To monitor specific hardware and applications, you must obtain specialized monitoring tools.

## Windows monitoring tools



Windows offers a number of monitoring and tuning tools that give you insight into resource utilization, health, and performance of Windows and hosted applications. These tools, which are included with the base operating system, are:

- **Windows System Monitor** — System Monitor is a graphical tool that enables you to monitor various aspects of the operating system and applications, set thresholds and alarms, analyze current and past data, and log the activity to disk-based logs for future reference. It also allows you to monitor both local and remote systems.

To access System Monitor, click *Start* → *Settings* → *Control Panel* → *Administrative Tools* → *Performance*.

- **Windows Task Manager** — Task Manager includes a subset of monitoring features offered by System Monitor. Task Manager provides information about computer performance and programs and processes running on the computer. Using this tool, you can close programs or processes, start programs, and view a dynamic display of the performance of your computer. Task Manager offers three tabs:
  - **Applications** — Lists running applications and their status
  - **Processes** — Lists running processes, their process ID, CPU statistics, and memory usage
  - **Performance** — Lists various system performance statistics such as processor and memory usage

To access the Task Manager, right-click the Windows taskbar and select *Task Manager* from the menu that displays.

---

**Note**

By default, security logging is disabled.

---

Additional monitoring tools from Microsoft for Windows operating systems include:

- **The Application Center Monitoring Service** — A web-based tool that helps you remotely monitor and determine the downtime for your servers. It also enables you to view performance and event-log data for one server or the entire cluster.
- **Ultrasound** — A monitoring and troubleshooting tool for the File Replication Service (FRS). FRS is used to replicate files and folders in the sysvol file share on domain controllers and files in Distributed File System (DFS) targets.
- **Terminal Services License Server Viewer (LSView.exe)** — A graphical user interface (GUI) that displays information about all available Terminal Services license servers in the current domain and current site of the computer. It is useful for monitoring and logging the status of license servers.

**INTERNET**

---

These tools are available for download from:

<http://www.microsoft.com/downloads/search.aspx?displaylang=en>

---

## Red Hat Linux monitoring tools

```
# /usr/ucb/iostat 1
```

tty		floppy0		dsk0		dsk1		cdrom0		cpu			
tin	tout	bps	tps	bps	tps	bps	tps	bps	tps	us	ni	sy	id
1	73	0	0	23	2	37	3	0	0	5	0	17	79
0	58	0	0	47	5	204	25	0	0	8	0	14	77
0	58	0	0	8	1	62	1	0	0	27	0	27	46

Understanding the monitoring tools available in Red Hat Linux Advanced Server 2.1 and the information they provide is important when detecting and solving hardware bottlenecks. The tools provide a snapshot of system resource utilization and supply subsystem information.

The monitoring tools available in Red Hat Advanced Server 2.1 are:

- The **sysstat** suite — Contains commands to collect I/O and processor statistics.
  - **iostat** — Measures processor statistics for devices and partitions; displays I/O statistics for one or more disk drives. The statistics returned can include read and write rates per second, average wait, service, and processor utilization.
  - **sadc** — The system activity data collector (sadc) collects system resource utilization information and writes it to a file.
  - **sar** — Measures system activity information by compiling cumulative statistics over a specified period of time. It also produces reports from the files created by sadc.
  - **mpstat** — Displays in-depth processor utilization statistics.
- **vmstat** — Reports virtual memory information; provides a concise view of system performance that enables you to determine if overall device activity is excessive. It can be configured to display resource utilization data at set intervals.
- **netstat** — Monitors network information; you can use netstat to query the network stack to verify which ports are listening.

Additional resource monitoring tools that ship with Red Hat Linux include:

- **free** — Displays memory utilization data one time, which is useful for short-term monitoring. You can display memory utilization figures repetitively by using the `-s` option, but the output scrolls, making it difficult to view the changes in memory utilization. Use the `watch` command to run `free` until you interrupt it by pressing `Ctrl+C`.
- **top** — Displays memory-related information, including processor utilization, process statistics, memory utilization. It runs continuously by default.
- **gnome-system-monitor** — Is a GUI that displays information regarding overall system status, process counts, process-level statistics, and memory and swap utilization. It also displays disk space utilization.

## NetWare monitoring tools

Several Novell NetWare monitoring tools can be used either separately or in conjunction with each other for viewing server statistics, health, and activities and for adjusting parameters to optimize a NetWare server. These tools are also an important diagnostics aid for troubleshooting and eliminating performance bottlenecks within the server.

The NetWare 6.5 server tools used to monitor processor, memory, disk, and network subsystem performance include:

- Monitor
- NetWare Remote Manager (NRM)
- Vtune

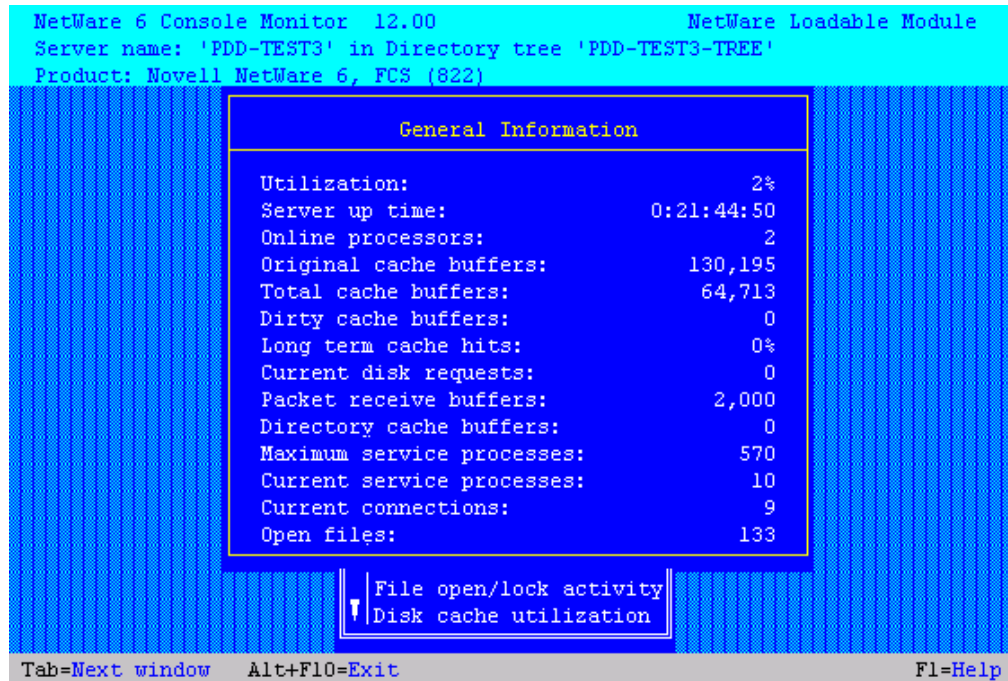
---

### **Note**

For additional information on these tools and general performance tuning guidelines for NetWare on ProLiant servers, refer to the integration note titled *Novell NetWare 6.5 Performance Tuning Guidelines for ProLiant Servers*, TC031102IN, 11/2003.

---

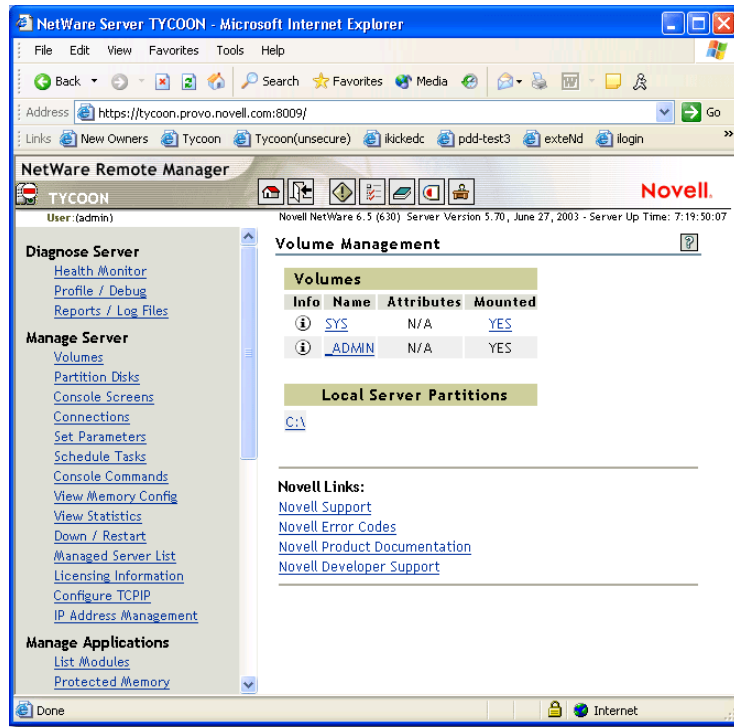
## Monitor



Monitor is typically used at the server console to display status information and statistics to help manage the server and to assess server memory and processor utilization. This tool also allows the user to set server parameters without returning to the console prompt to use the SET command.

To start this tool at the server console, enter *monitor*. After Monitor is running, select the monitor screen from the list of HTML-based screen pages.

## NetWare Remote Manager



NRM is a versatile, all-inclusive utility for displaying status information and statistics to help manage the server. This tool also enables you to set server parameters without returning to the console prompt to use the SET command. NRM performs the following main functions:

- Monitoring the general health of the server
- Changing the configuration of the server
- Diagnostics and troubleshooting of the server
- Viewing performance statistics and tuning the server

To invoke NRM directly from the local server through X Server Graphical Console, follow these steps:

1. Select *Novell* from the graphical console and then select the *Utilities* menu.
2. Click *NetWare Remote Manager* from the drop-down menu.
3. Enter the administrator user name and password and click *OK*.

To invoke NRM from a remote machine, follow these steps:

1. Log on to the server as an administrator.
2. At the browser, enter the Domain Name Service (DNS) name or IP address along with the server port number (8008) and press *Enter* to display the NRM window. For example:

`http://131.122.100.250:8008/`

## Vtune

The Vtune tool has two main components:

- The NetWare component (vtune.nlm) — Is loaded on a NetWare server to collect traces of information. For instance, vtune.nlm collects information such as where the processor is spending its time, memory read/write misalignments, or branch mispredictions.
- The Intel client component (vtune) — Is used to process the data collected by vtune.nlm. This component runs under Windows 9x, Windows 2000, and Windows XP.

To collect traces using this tool, follow these steps:

1. Load the vtune.nlm from the NetWare console.
2. Select the events of interest to profile and then select the sampling interval. Use the default trace file located on the root of the SYS volume or provide a name preference of your choice using the DOS 8.3 naming format. Save the file.
3. Import the trace file by using the Vtune client after you have collected all the traces needed. This allows you to view the graphical display of trace data.

---

### INTERNET

You can download the 30-day evaluation version of the Intel client Vtune at:  
**<http://developer.intel.com/software/products/global/eval.htm>**  
The NetWare component (vtune.nlm) can be downloaded from Novell at:  
**<http://developer.novell.com/support/sample/tids/topt2/topt2.htm>**

---

## Application monitoring tools

Network administrators monitor applications to ensure that they perform correctly and optimally. By measuring performance at the application level, administrators can align network performance with the business value of these applications.

It is critical for the administrators to establish guidelines and procedures for application monitoring and to communicate these procedures to developers. Together, administrators and developers can log and monitor information to help discover and diagnose problems. Throughout this ongoing process, the operations and development teams should strive to continually refine monitoring processes.

Service level agreements (SLAs) often dictate the application monitoring necessary to ensure that it is functioning correctly (health monitoring) and meeting the performance goals (performance monitoring).

- **Health monitoring** — Health monitoring enables you to identify the conditions that contribute to system or component failure and take corrective action. These failures can occur within the hardware, operating system, and applications and they result in a reduced or disrupted service level of the application. Use hardware, operating system, and application monitoring tools to observe health-related system events and performance thresholds, and design your system to withstand such failures with minimal impact on the application service level.
- **Performance monitoring** — Performance monitoring should also occur at the hardware, operating system, and application levels. Health-related issues, resource misalignment, application design, and other factors affect how the system and the application perform and meet the SLA.

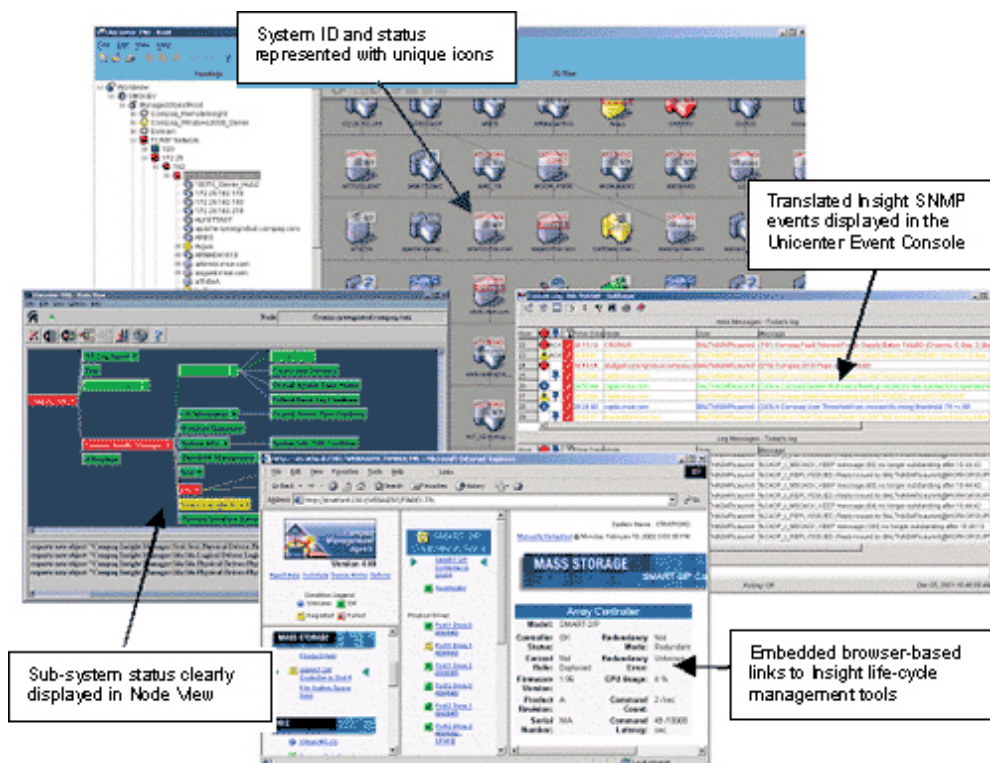
---

### Note

Many enterprise-level applications ship with performance tools that go beyond the capabilities of operating system tools. In addition, there are third-party tools that enable specific application performance monitoring.

---

## HP integration with third-party tools



CA Unicenter integration screens

Insight Manager 7 automatically integrates with certain third-party tools to detect all HP systems that have Insight Management Agents installed.

Insight Manager 7 creates a graphical representation of HP systems on a network segment map and reports information about these systems to the database. HP systems and subsystems are distinguished by unique icons; changes in health status are identified by icon color changes.

Because Insight Manager 7 and the Insight Management Agents support industry-standard management protocols and infrastructures, you can use available enterprise management platforms to gain access to the management information exposed by ProLiant servers and attached storage running the Insight Management Agents. These platforms include some of the OpenView products and several third-party tools:

- OpenView Network Node Manager
- OpenView Operations
- CA Unicenter
- Microsoft Operations Manager
- Microsoft Systems Management Server
- Tivoli Enterprise

Support for managing ProLiant servers from these enterprise management platforms is provided by linking to the web-based agents and through a snap-in module that provides knowledge about status events sent from the Insight Management Agents.

Network environments that implement these tools as the preferred enterprise management platform can integrate with Insight Manager 7 to streamline administration and increase system availability.

---

<b>INTERNET</b>	HP partners with NetIQ to offer products to optimize the performance of Windows 2000-based systems and applications running on any HP platform. You can find more information at: <b><a href="http://activeanswers.compaq.com">http://activeanswers.compaq.com</a></b>
-----------------	--

---

## Creating a performance baseline

Creating a performance baseline is one of the most crucial steps in performance planning and management. In this process, you capture the current state of your server or network performance and use this information as the baseline for comparing future performance levels.

It is important to capture and document the baseline performance properly, distribute the information to key stakeholders, and maintain accurate records over time. Typical documentation includes:

- System architecture
- Network architecture
- Application architecture
- Capacity plan
- Change management plan
- Test plan and procedures
- Maintenance plan

Ideally, a performance baseline for the network and all critical servers is created when performance is satisfactory. Having a baseline and comparing it to the current levels of performance can quickly pinpoint the problem areas. However, the performance baseline is not always created or available.

If the performance baseline exists, it will assist in analyzing network and server performance. This topic is covered in the *Comparing current performance with baseline performance* section later in this module.

If the performance baseline does not exist, you have one less tool to use during your performance analysis. However, you have an opportunity to create one after the performance returns to the satisfactory levels.

You can use a variety of performance management and monitoring tools to create a performance baseline. You can use PMP to collect performance levels from ProLiant servers and store the statistics in a database repository. Other performance tools might be necessary to do the same for applications and third-party systems.

## Creating a performance baseline using PMP

Creating a performance baseline using PMP involves configuring PMP to collect data at the desired collection intervals and logging it into the performance database. Because the data is permanently stored in the database, it can be extracted and manipulated at any time. Using the PMP tools, you can extract the desired data, create a detailed report, and import this report into a desktop analysis tool such as Excel.

To learn about collecting and manipulating performance statistics from a target server, refer to the *Configuring monitoring behavior* and *Logging and manipulating logged data* sections later in this module.

## Creating a performance baseline using other tools

PMP monitors hardware and operating system performance levels in ProLiant servers. PMP only monitors these levels in ProLiant servers. It does not monitor third-party components, network infrastructures, or applications. To monitor those areas and to create performance baselines for them, you must use other appropriate performance monitoring tools. When appropriate, select those tools that enable you to permanently store the collected statistics and manipulate them to make comparisons.

VeriTest offers two benchmarking tools for Windows platforms that enable you to measure server performance:

- WebBench — Measures web server software performance; supported on Windows platforms only
- NetBench — Measures how well a file server handles file I/O requests from 32-bit Windows clients

Httpperf is another tool for measuring web server performance. In Linux environments, you can use the ab tool to benchmark Apache HTTP servers.

## Step 2 — Evaluating the data

The purpose of this step is to analyze the collected information to determine the real (as opposed to the reported) source of the performance problem. The difference between a real source and a reported source of the performance issue is critical to resolving the bottleneck correctly. Often, adequate knowledge of the entire system—hardware, operating system, and application—is necessary. Otherwise, you can end up spending resources tuning the secondary source of the bottleneck with little or no results.

### Example

Users of a heavily accessed database server are reporting long response times. The number of users has doubled in the past month, but the server configuration has not changed. After you collect performance statistics from the server in question, you analyze the key server subsystems for utilization. You notice that the disk queue length is higher than recommended and determine that the disk subsystem is the performance bottleneck. The database queries are impacted by the high disk subsystem response times and the impact increases with more users. The disk subsystem is the reported source of the bottleneck.

Your first reaction may be to determine what can be done to speed up the disk subsystem. The real question, however, is whether the disk subsystem is also the real source of the bottleneck. If the other hardware components are working at their highest efficiency, you can be assured that your best course of action is speeding up the disk subsystem. However, if this is not the case, consider another course of action.

Many enterprise applications, databases included, require an adequate amount of memory set aside for use as their own cache. With few users, a small cache is sufficient. When the number of concurrent users increases, you might need to increase the cache size to maintain adequate performance. When the application cache is too small, the application might be generating an excessive volume of disk I/O requests, creating a disk bottleneck. Additional memory or an application cache reconfiguration would increase performance.

In this example, the disk bottleneck is the reported bottleneck, but the insufficient application cache is the real source of the performance issue.

## Using PMP to detect and analyze performance bottlenecks

Several tasks are associated with configuring the PMP infrastructure, such as:

- Determining which servers and management agents, if any, must be updated, and performing the necessary management agent updates.
- Determining whether PMP must be updated to the latest version and performing the necessary update.
- Configuring the monitoring of selected servers.
- Monitoring overhead as data is sent across the network.
- Performing static analysis to determine if any configuration issues might impact performance.
- Detecting and analyzing performance bottlenecks.
- Gathering performance statistics of selected servers over several days, creating reports, and analyzing the findings.
- Comparing current performance with baseline performance.

### Updating HP Management Agents

PMP 2.1 requires a minimum version of the HP Management Agents to be running on the supported ProLiant servers. This minimum version is dependent on the host operating system.

Operating system	Minimum agent version
Windows NT 4.0	6.10.01
Windows 2000	6.20
Windows Server 2003	6.31

To determine the version of the HP Management Agents running on the monitored servers, use either the HP Insight Manager 7.0 or the VCRM, or access the System Management Homepage directly.

---

**INTERNET** To access the system management homepage, point your browser to **https://localhost:2381** and log in with administrative privileges.

---

To update the HP Management Agents, use the HP Management CD to install the desired agent version. The installation process detects the previous agent installation, removes it, and installs the new version of the agents.

## Updating PMP

PMP 2.1 can be updated using two methods:

- **Using SmartStart 7.00** — If you deploy a ProLiant server using this version of SmartStart and install Insight Manager 7, you will automatically have PMP 2.1 installed. You must complete product licensing and activation.
- **Using the SoftPaq SP24900** — Download the SoftPaq from **<http://h18004.www1.hp.com/products/servers/proliantessentials/valuepack/PMP/index.html>** for a complete installation of PMP 2.1 that installs over any previous version. However, you must have Insight Manager 7 SP2 installed before running the update.

## Configuring monitoring behavior

You can configure PMP monitoring behavior locally at the middle-tier server or remotely. On the administration page, you can:

- Start and stop monitoring of selected servers.
- Check the monitoring status of servers.
- Modify PMP monitoring settings such as the mode of operation (monitor, log, and alert), sample rate, and number of samples.

Change the collection interval of the HP Management Agents at the monitored servers, using the HP Management Agents applet in the Control Panel.

## **Monitoring overhead**

### **Network overhead**

The amount of data sent over the network as a result of PMP polling is a function of the size and complexity of the server. An eight-way server with 80 disk drives generates a larger data sample than a two-way server with five disk drives. For most production servers, the data sample ranges between 3KB and 4KB. The frequency at which this data travels across the network is configurable using the SNMP sampling rate.

### **Processing overhead on the monitored servers**

PMP uses the same management agents as Insight Manager 7 and requires no additional software to be installed or running at the target servers. The polling interval is configurable and determines the frequency of retrieving performance data from the monitored servers.

Because the management agents are already running on the monitored server and collecting data, the overhead associated with PMP polling these agents is negligible.

### **Processing overhead on the middle tier**

PMP induces a greater workload on the middle tier server than Insight Manager 7. This additional overhead is attributed to the real-time calculations, comparisons, and other analysis the middle tier must perform. Therefore, PMP requires a more powerful server with more resources than an Insight Manager 7 server would require by itself.

### **Database load**

The size of the database depends on the number of servers monitored and logged, the size of these servers, and the sampling frequency. The database structures add approximately 15% to the size of the actual data stored. Unless performing offline analysis, the physical database I/O consists of write operations that record newly retrieved performance statistics in the database.

## Performing static analysis

*Static analysis* refers to configuration validation of the target server. PMP ensures that the hardware configuration of the server does not create a performance problem at a later time.

Perform static analysis by using the Static Analysis button on the PMP home page.

## Detecting and analyzing performance bottlenecks

Performance monitoring begins as soon as you license the target server and change the PMP mode of operation to any of the monitoring options. You might have to wait a few minutes for PMP to collect sufficient data points.

The overall server performance is displayed in the PF column of the Insight Manager 7 home page. If the status indicator is green, amber, or red, you can activate the PMP interface by clicking the status indicator corresponding to the monitored server.

Using the intuitive PMP interface, you can quickly determine which server subsystem is causing the bottleneck and which hardware component is performing at its capacity. PMP also displays a list of monitored parameters and their values, explanation of the analysis, and recommendations. Using the Inventory tab, you can determine the hardware configuration as well as software versions.

## Logging and manipulating logged data

When logging is enabled, PMP stores performance data gathered from monitored servers in a database repository. PMP provides powerful performance analysis tools to help manipulate the stored performance data.

- **Data Extractor** — Exports performance statistics from the database repository into an XML file. This XML file is then used by the Offline Analysis tool for playback.
- **Offline Analysis** — Enables playback of any monitored sessions recorded in the performance database. The Offline Analysis tool requires an XML file generated by the Data Extractor tool.
- **Report Generator** — Generates a summary report in HTML format or extracts all collected performance values into a CSV file for import into desktop analysis and reporting tools.
- **Log Deletion** — Provides an easy method of deleting the recorded sessions for a logged server. Logged sessions use database space and must be deleted manually when no longer needed.

## Comparing current performance with baseline performance

Comparing current and baseline performance is one method of analyzing performance trends and determining performance bottlenecks. PMP enables you to create a performance baseline and subsequent performance measurements, store this information in permanent tables in the database repository, and extract the data for import into a desktop analysis tool such as Excel.

To achieve this goal:

1. Log the performance statistics using the Logging feature of PMP.
2. Use the Report Generator tool to extract the statistics for a logged session from the database repository and to save them in a CSV file.
3. Import the desired CSV files into an Excel spreadsheet for comparison and analysis.

You can also analyze performance bottlenecks and trends using the playback feature of PMP. With the Offline Analysis tool, you can play back any performance session that is recorded in the database repository.

Many competitive tools allow you to generate performance statistics to support trending and performance comparisons. Only PMP, however, enables you to generate performance information specific to ProLiant hardware.

---

### Lab

Your instructor might now direct you to perform Module 4, Lab Exercise 1 — Using PMP.

---

## Step 3 — Developing an action plan

If performed correctly, the previous step—evaluating the data—will produce a clear indication of where within the system the performance bottleneck exists. The next step is to determine how to resolve this bottleneck.

For every bottleneck, several courses of action exist. Each action is defined by its cost, degree of difficulty, and level of effectiveness.

When tuning a system, you do not resolve a performance bottleneck in the absolute sense, but move the bottleneck from one location where it is most problematic to another location that is more acceptable. A properly tuned system is not a system with no bottlenecks. It is a system where:

- All components are performing most efficiently.
- The performance meets the requirements.
- The performance-limiting factor is the one you choose.

When developing an action plan to address a performance bottleneck, you must decide whether to focus your efforts on making the system run more efficiently or increasing performance through distributed computing or clustering. The first effort deals with performance scale-up; the other with performance scale-out.

### Scaling up performance

Scaling up refers to making the system perform faster by tuning resources within the system itself. The assumption is that you can make the system run faster, up to its performance limitations. This effort includes:

- Adding more hardware components.
- Upgrading existing hardware components with faster and more efficient ones.
- Upgrading the software to a faster and more efficient version.
- Tuning both hardware and software components to perform more efficiently.

System components that can be scaled up include:

- Processors
- Memory
- Disk subsystems
- Network bandwidth
- PCI buses
- Operating system
- Applications

## Processors

To solve a processor bottleneck, determine the type of task processing occurring. Then, depending on the predominant task execution and server configuration, upgrade to faster processors or add more processors. Task processing types are:

- **Serial** — Each task is executed serially, statement by statement, until the task is accomplished. Serially executing tasks are more complex than parallel tasks and run best on fast processors, where the raw processor speed contributes to fast execution.

### Example

A decision-support application analyzes sales data over the past quarter. In this case, performance will improve if you upgrade to faster processors.

- **Parallel** — Several tasks execute concurrently. These can be independent tasks or subtasks resulting from a division of a serially executing task. Parallel tasks are less complex than serially executing tasks. Parallel tasks run best on multiple processors, where each processor executes a subset of the tasks one at a time.

### Example

An online transaction processing database executes many concurrent transactions, or a serially executing database query is divided into multiple subtasks that execute concurrently. In this case, performance will improve if you add more processors.

Other methods of tuning processor performance include:

- **Tuning the thread execution priority** — Some operating systems allow you to assign processing priority to either foreground applications or background services. To make the server respond best to network applications, set the priority to background services.
- **Tuning the operating system** — If the operating system reports high kernel time (time spent executing interrupts or I/O requests), tune the disk subsystem. Also examine cache size and cache efficiency. A high user time (time spent executing applications) can indicate an inefficient application.
- **Improving efficiency of the application** — If the user time is high but the application is not processing requests as expected, tune the application. Some applications have advanced tuning options that, for example, reduce the number of context switches and allow the application threads to execute on a given processor for the entire time slice, regardless of the pre-emptive conditions. Others enable you to assign processor affinity, which results in the operating system scheduler assigning threads to execute on processors where they executed before. Such returning threads typically find their data already cached in the processor cache, thus running faster.
- **Upgrading the firmware, device drivers, operating system, or application.**

## Memory

If the system contains insufficient memory, adding memory will certainly improve performance. However, before spending money on additional hardware, consider the following tuning options:

- **Prevent paging** — Two types of paging activity occur:
  - **Soft paging** — Occurs within memory when the application does not find the required memory page within its working set because the required page is somewhere else in memory. Soft paging does not have a significantly negative impact on performance.
  - **Hard paging** — Occurs when the required page is not in memory, but is stored in the page file on disk. The operating system must generate physical I/O to retrieve it and often writes another memory page to disk to make room for this page. Hard paging occurs as a result of memory overallocation, when not enough memory is available to satisfy the operating system and application requirements. Hard paging should be prevented whenever possible by adding or reallocating memory.
- **Tune operating system and application cache** — Allocate sufficient operating system and application cache to meet the caching requirements of the environment. A file server, for example, relies heavily on the accessed data being stored in the operating system cache. Therefore, you must allocate a large operating system cache to meet the performance requirements. On the other hand, many enterprise applications, such as databases, allocate and manage their own cache. The application cache is optimized by the application manufacturer for the primary application activity and provides best performance. Consequently, the operating system cache becomes ineffective and should be minimized.
- **Configure sufficient page file (swap space) per operating system manufacturer recommendations** — These recommendations are typically 1.5 to 2 times the size of physical memory. Ensure that the paging activity is minimal. If paging cannot be avoided, place the page file away from other heavily accessed files.

- **Use Very Large Memory (VLM) support** — When applicable, VLM support enables applications to access memory beyond the upper memory limits placed on the applications by the operating system. For example, Windows 2000 restricts the application memory address space to 4GB. With VLM support, certain applications, such as Oracle9i, can address up to 64GB of memory. These memory addressing extensions are available with the operating system or certain applications.
  - VLM support in Windows is referred to as *Address Windowing Extensions* (AWE). It is available using the /pae (Physical Address Extension) switch in the boot.ini file.
  - Most enterprise distributions of Linux support VLM without any changes to the kernel. The application, however, must be designed to take advantage of memory beyond 4GB.
  - After the server.exe file is loaded in NetWare 6.5, remaining memory under the 4GB restriction is assigned to cache memory and virtual memory, which is available for NetWare Loadable Module (NLM) programs and other processes to use. The memory allocator cannot allocate memory beyond 4GB and the traditional file system does not use memory beyond 4GB. If you install more than 4GB, this memory is used by the virtual memory system and applications such as eDirectory. The applications that run in the memory beyond 4GB leave more memory below 4GB for the memory allocator to use. Although you cannot configure the memory beyond 4GB, adding more memory has the following performance advantages:
    - ◆ More memory is available to those systems and applications that run in protected memory or in the space beyond 4GB.
    - ◆ If the memory allocator has more available memory below 4GB to use, the server performs faster.
- **Use special memory tuning options available with the operating system** — The 4GB RAM Tuning (4GT) feature of Windows 2000 and Windows XP enables single-process applications such as Oracle9i to use up to 3GB of memory without AWE. Because the standard configuration enables Windows applications to use only up to 2GB of memory, the 4GT feature provides up to 50% more memory that is available to the application. You can enable the 4GT feature by adding the /3GB switch to the boot.ini file.

## Disk subsystems

To solve disk subsystem bottlenecks, first analyze the available memory resources and their allocation. Incorrect memory allocation, insufficient memory resources, and excessive paging activity often look like disk bottlenecks.

If the bottleneck persists after analyzing the memory subsystem and implementing appropriate adjustments, determine whether your application is generating a random or sequential I/O profile.

- **Random I/O profile** — Typically generated by online transaction processing applications and applications servicing many concurrent users, a random I/O profile generates a large volume of small random I/O requests. Consequently, it benefits from many disk drives capable of concurrently processing those requests. To increase performance with this type of profile:
  1. Add multiple drives first.
  2. Then add multiple controllers.
  3. Configure disk caching for write-back.
  4. Consider other tuning approaches.
- **Sequential I/O profile** — Typically generated by decision support, batch, and analytical applications, the sequential I/O profile transfers large volumes of data. Therefore, it benefits from disk controllers and drives with wide data paths and fast data transfers. To increase performance with this type of profile:
  1. Add multiple controllers first.
  2. Then add multiple drives.
  3. Configure disk caching for read-ahead.
  4. Consider other tuning approaches.



### Important

You should always back up your data before performing a major change to a disk subsystem.

---

PMP analyzes the storage subsystem utilization from the hardware and operating system perspective and determines whether additional controllers or disk drives should be added. PMP detects saturated disk drives and controllers and recommends additional or faster components when necessary.

- **Adding or upgrading disk drives** — Consider adding disk drives or upgrading to faster ones only if the current drives are continually saturated. This type of action is typically necessary in environments with a predominantly random I/O profile.
- **Adding or upgrading disk controllers** — Consider adding another disk controller or upgrading to a faster one only if the current disk controller is continually saturated. This is typically necessary in environments with a predominantly sequential I/O profile.

### **Migrating to another RAID level**

RAID migration is an option for improving performance of I/O-bound systems, especially when the added benefit is improved fault tolerance.

Depending on the environment, certain RAID levels perform better than others. This is especially true in write-intensive environments. However, tuning of other areas can mask the negative impact of a RAID level on performance. For example, write caching at the array controller level can significantly improve performance of RAID 5 and RAID Advanced Data Guarding (ADG).

In general:

- RAID 0 does not have any impact on performance because it uses all disk drives for the read and write requests.
- RAID 1+0 has a slightly negative impact on performance. Read operations use half of the spindles; the other half contains identical data. Split seeks are used for concurrent reads and often compensate for this performance loss. Write operations cause the array controller to generate two writes—one to the primary side of the mirror, one to the secondary side of the mirror.
- RAID 5 results in high read performance, because it can use all disk drives to service read requests. For write operations, the array controller must generate on average four additional I/O requests to maintain the data and parity information.
- RAID ADG maintains two distinct parity schemes to protect data and requires on average six I/O operations for every write request the array controller receives. Read operations are only slightly affected because of the two drives that carry parity information.

Advanced technologies built into modern array controllers address some of these issues. Examples of these technologies include:

- Deferral of redundant write requests
- Battery-backed caching of write requests
- Batching of individual write requests
- Split seeks
- Deferral of parity block generation until the entire stripe is filled

### Reconfiguring disk controller cache

Disk caching improves performance by:

- Reducing disk access (with cache hits)
- Reducing the negative effects of RAID overhead
- Assisting with disk I/O request sorting and queuing

Many disk controllers have configurable cache memory and support the following modes of caching:

- Read-only cache
- Write cache
- Combination of read-only and write cache

Read caching is used in two ways:

- **Provides read-ahead buffers** — Called *read-ahead cache*. These buffers are useful during sequential read access operations. When the disk controller detects sequential read patterns, it prereads anticipated information before the application requests it.
- **Holds reusable information in memory** — Called *most-recently used read cache*. The cache contains valid data that is reused for new read requests.

Read cache is beneficial in high-bandwidth application environments such as data warehouses, backup servers, and analytical/computational systems. It improves performance during sequential read operations. However, if the application maintains its own cache in system memory, the effects of the array controller cache are reduced because the data is more likely to be available from the application cache instead of the disk controller cache.

Write caching enables the application to post write requests to the disk controller cache and then receive an immediate completion status response. The disk controller writes the data to the physical disk later. Write cache is beneficial in random I/O environments. The write requests can be posted in cache, increasing the overall system performance. In high-bandwidth write environments, write cache becomes easily saturated and loses its effectiveness.

### **Increasing array controller cache size**

Most array controllers support cache memory upgrades to increase cache size. Increasing the cache size improves performance only in certain situations. More cache is better only when the current cache size is inadequate (100% used and saturated). When cache is only partially used (not saturated), adding more cache memory is a waste of resources. Usually, performance does not improve.

### **Lowering the rebuild and expand priority**

If the storage-related bottleneck occurs during a failed drive rebuild or logical drive expansion, consider lowering the rebuild and expand priority. You can configure both options using the Controller Settings button in the Array Configuration Utility (ACU).

### **Redistributing data**

Hot spots within the disk subsystem refer to data elements that face an unusually heavy I/O compared to the rest of the storage. For example, a single, heavily accessed data file within an Oracle database can be considered a hot spot.

To minimize disk contention caused by hot spots, redistribute the data and store the most heavily accessed portions among several array controllers and disk drives. Also balance the I/O among various hardware components, such as array controllers and disk drives.

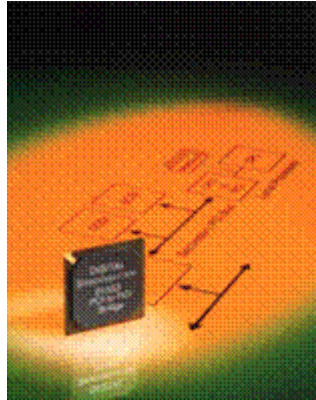
### **Separating sequential and random I/O**

If the system performance depends on the completion speed of sequential writes, isolate the sequentially accessed files to their own sets of disks. Combining sequentially accessed files with other sequentially or randomly accessed files results in random I/O for all of them and consequently high response times.

### **Optimizing paging**

Because excessive paging activity is undesirable and detrimental to performance, it is a good practice to place the page file where its I/O does not affect performance.

## Host buses



Host bus bottlenecks are almost nonexistent on servers released in the past few years. The advances in hardware speed and efficiency, the ratio of expansion slots on a single bus compared to the number of available buses, and wide adoption of chipset buffering all contribute to negligible performance differences between unbalanced and balanced host bus configurations.

The common practice for older servers, which you can still follow even on newer servers, to prevent host bus overloading is:

- Evenly distribute expansion boards among available host buses
- Use the same expansion board technology as the host bus

PMP detects unbalanced host bus configurations with the static analysis and host bus saturation with real-time performance monitoring. In rare situations where the host bus saturation occurs, PMP recommends that you move the installed expansion boards to better balance the host buses.

## Network

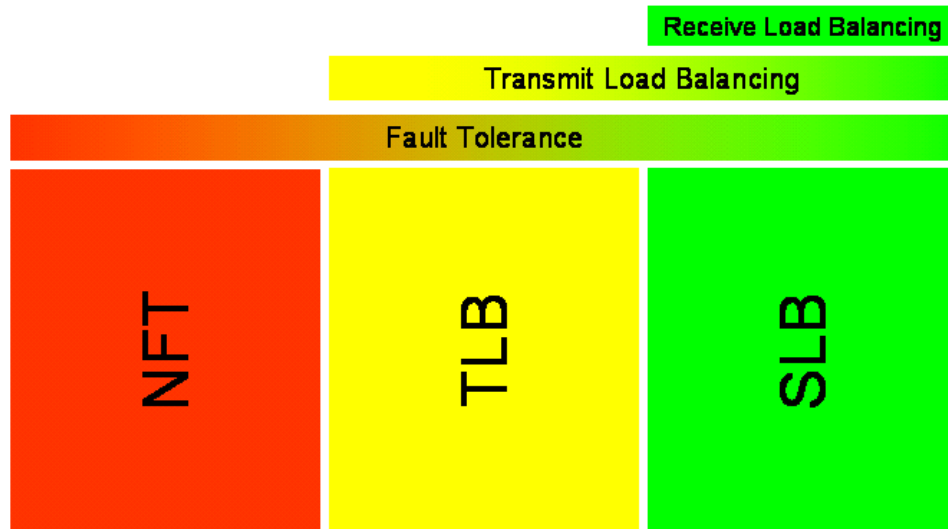
Optimizing processor speeds, memory, and hard drives can improve system performance, but users can still be left with networks too slow to connect them efficiently to the applications and information they need. The requirement for more data throughput as well as escalating data management requirements put additional demands on network scalability.

A practical approach to resolving network bottlenecks begins with the NIC. There are several cost-effective ways to use ProLiant NICs to increase network bandwidth without making cabling changes:

- **Install additional NICs** — Adding one or more secondary NICs in ProLiant servers can scale the network by building a bigger pipe.
- **Add upgrade modules** — ProLiant NIC upgrade modules plug into specific base modules to help scale the network. Installing upgrade modules to the base NICs adds ports and protect the customer's investment in existing hardware.
- **Use dual-port NICs** — ProLiant dual-port NICs allow more ports per slot. Therefore, replacing a single-port NIC with a dual-port NIC is a simple way to scale the network.
- **Optimize the queue length** — The NIC monitors the send queue (for best performance, the send queue should be on the NIC card). User process monitors the receive queue (for best performance, the receive queue should be in user memory). Other considerations include:
  - **Coalesce buffers** — Reduce the number of direct memory access (DMA) operations per transfer by combining small fragments into a single buffer.
  - **Increase number of receive descriptors** — In high network load situations, increase receive descriptors to increase performance. However, this also increases the amount of system memory the driver uses. If too few receive descriptors are used, performance suffers. If too many receive descriptors are used, the driver unnecessarily consumes memory resources.
  - **Adjust number of transmit descriptors** — Use to determine how many resources are allocated to transmit packets.
  - **Offload checksum** — Offload the task of computing the checksum for incoming/outgoing packets to the NIC hardware, thereby improving performance.

- **Configure NIC teaming** — By grouping multiple physical NICs into a single virtual interface, you can achieve greater throughput and performance from a single network connection.
- **Use Gigabit Ethernet** — Gigabit Ethernet is the current networking standard for high-performance networking. It is capable of transferring data at rates approaching 900Mb/s, almost nine times the rate of 100Mb/s Ethernet. To significantly increase throughput to and from the servers using the existing infrastructure, use multiple Gigabit NICs within each server to load balance traffic (also known as *port bonding* or *link aggregation*). This strategy can increase effective server bandwidth from 1GB up to 16GB (theoretically) using full duplex transmission with eight NICs.
- **Add virtual LANs (VLANs)** — A VLAN enables you to see and access only specified network segments. This increases network performance and improves network security.
- **Add switches** — On multiclient networks with unswitched hubs and repeaters, each client shares the bandwidth resource with all other clients on the segment. To overcome this limitation, intelligent network switches enable point-to-point communication between nodes.
- **Add network segments** — No matter which approach is chosen to scale a network, NICs are inexpensive options that can reduce network traffic bottlenecks by adding more segments to the network. Additional network segments allow increased traffic without affecting the traffic on existing segments.

## NIC teaming



Teaming types and teaming functionality

The three teaming modes for HP NICs are:

- Network fault tolerance
- Transmit load balancing
- Switch-assisted Load Balancing

Each mode gains in features and incorporates most features from the previous teaming mode.

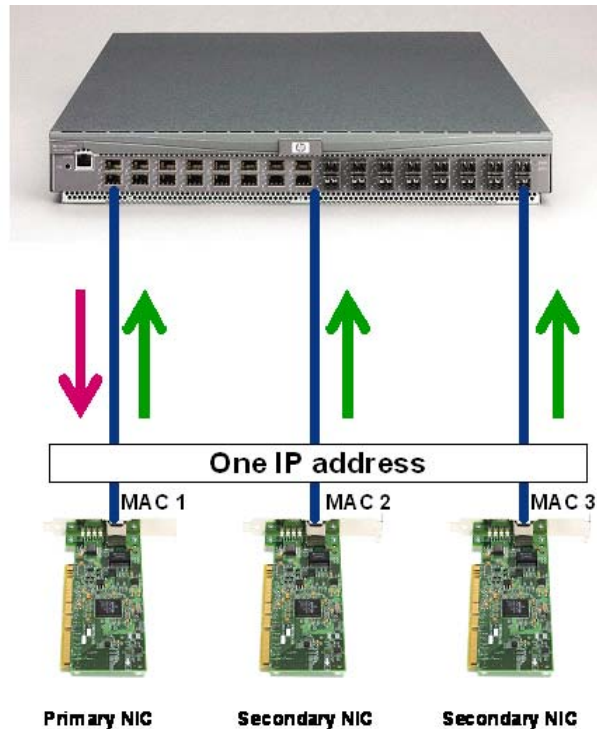
### Network fault tolerance

Network fault tolerance is a simple, effective, and fail-safe approach to increase the reliability of server connections. It enables you to set up link recovery to the server NIC in case of a cable, port, or NIC failure. By creating a team, the network fault tolerance approach enables you to maintain uninterrupted network connectivity.

Network fault tolerance provides simple redundancy with two to eight NICs in a fault-tolerant team. Each server can support up to eight teams where one NIC per team is defined as the primary NIC. All other NICs are secondary. Network fault tolerance teaming functions at any speed, on any media.

During normal operation, the backup NIC has transmit disabled. If the link to the primary NIC fails, the link to the backup NIC automatically takes over.

## Transmit load balancing



Transmit load balancing, also known as *adaptive load balancing*, incorporates all the features of network fault tolerance and adds load balancing of all transmit IP traffic. It is switch-independent and can be split across Layer 2 switches; however, all members must be in the same Layer 2 network. Transmit load balancing increases the transmission load potential of a server.

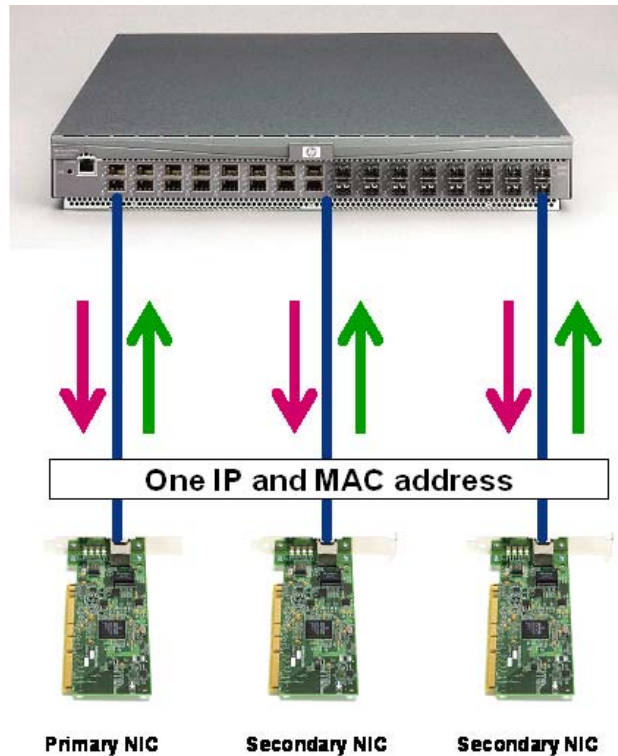
### Example

A transmit load balancing team containing four HP Fast Ethernet NICs configured for full-duplex operation can provide an aggregate maximum transmit rate of 400Mb/s and a 100Mb/s receive rate. In this example, total bandwidth is 500Mb/s.

With transmit load balancing, traffic received by the server is not load balanced. One NIC is primary and as many as seven other NICs are secondary. The primary NIC shares the transmit traffic and is responsible for receiving all traffic destined for the server. The secondary NICs only transmit data. Transmit load balancing distributes the data flow based on the destination IP or memory access controller (MAC) address. An algorithm that uses the last one or two bits of the source and destination MAC or IP addresses determines which port is used for a particular server-to-client communication flow.

Under transmit load balancing, all NICs in the team operate under the same IP address, but different MAC addresses. If one of the secondary connections fails, the server driver redirects the information flow from the failed connection to the remaining NIC team members. If the primary NIC in the team fails, then one of the secondary NICs assumes the MAC address and the duties of the primary NIC.

## Switch-assisted Load Balancing



Switch-assisted Load Balancing incorporates all the features of network fault tolerance and transmit load balancing and adds receive load balancing. If any of the NICs fail, the remaining NICs share the data load.

The operating system sees the multiple NICs as the same IP and MAC address and therefore as one NIC.

The differences between transmit load balancing and Switch-assisted Load Balancing center around three design issues:

- All network connections transmit and receive data simultaneously, enabling a theoretical maximum transfer rate of up to 16Gb/s.
- Switch-assisted Load Balancing is supported on Institute of Electrical and Electronic Engineers (IEEE) 802.3ad-capable intelligent switches and similar switches, such as those that support EtherChannel and multi-link trunking protocol (MLT).
- All ports must be connected to the same switch.

---

### Note

For more information on the IEEE 802.x standards, refer to this website:  
<http://www.ieee.org>

---

## NIC teaming comparison

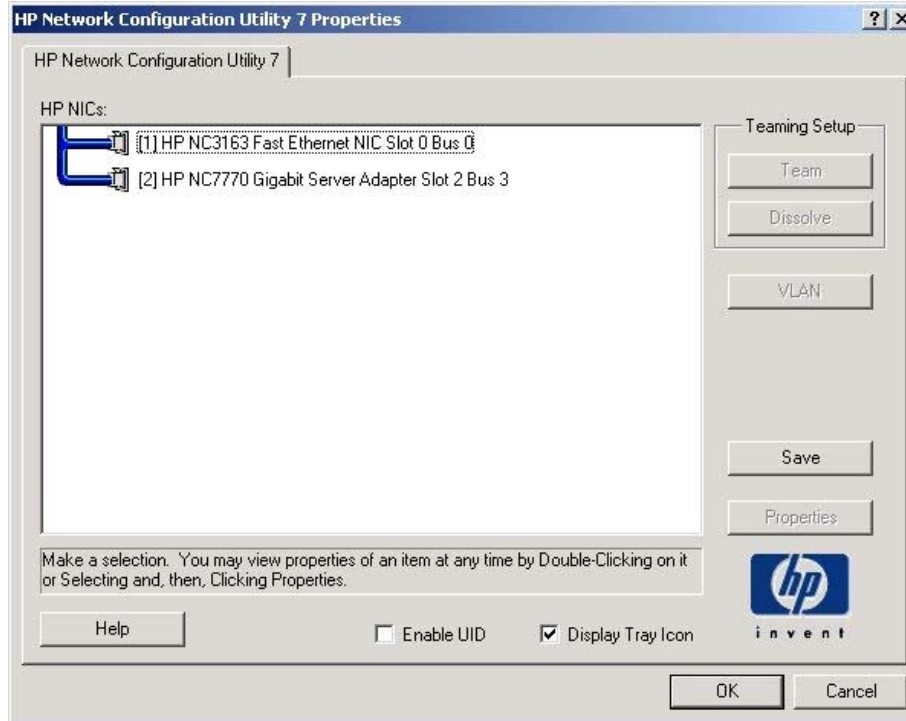
The differences among the NIC teaming types are detailed in the following table.

Teaming capability	Network fault tolerance	Transmit load balancing	Switch-assisted Load Balancing
Number of adapters supported per team	2 to 8	2 to 8	2 to 8
Fault tolerance	Yes	Yes	Yes
Transmit load balancing	No	Yes	Yes
Receive load balancing	No	No	Yes
Requires 802.3ad-compliant switch	No	No	Yes
Can connect a single team to more than one switch for switch redundancy (must be same broadcast domain)	Yes	Yes	Switch-dependent
Uses heartbeat for network integrity checks	Yes	Yes	No
Can team NICs that do not support a common speed	Yes	No	No
Can team NICs operating at different speeds as long as the NICs support a common speed	Yes	Yes	Yes
Maximum theoretical transmit/receive throughput in Mb/s with maximum number of 100Mb/s NICs	100/100	800/100	800/800
Maximum theoretical transmit/receive throughput in Mb/s with maximum number of 1000Mb/s NICs	1000/1000	8000/1000	8000/8000
Load balances TCP/IP	No	Yes	Yes
Load balances protocols other than TCP/IP	No	No	Yes
PCI Hot Plug support	No	No	Yes

### Note

Network fault tolerance and Switch-assisted Load Balancing are protocol-independent. Transmit load balancing load balances IP traffic only.

## HP recommendations for teaming



HP Network Configuration Utility Properties window

HP recommends procedures to optimize performance specific to the type of network teaming environment. The following recommendations are common to network fault tolerance, Transmit load balancing, and Switch-assisted Load Balancing environments:

- Heartbeats should be enabled by default (transmit validation heartbeats should be enabled by default in a Switch-assisted Load Balancing environment.)
- MAC addresses should not be set manually to a locally administered address (LAA) through the Microsoft user interface. You should not implement LAAs on individual NICs that are members of a team or teaming might not function correctly. You can set an LAA for the team through the HP Network Configuration Utility GUI.
- The spanning tree blocking, listening, and learning stages should be disabled, or bypassed, on all switch ports to which an HP NIC team port is attached. These stages are not needed when a non-switch networking device (for example, a server) is attached to the switch port. HP ProCurve switches have a feature called *STP Fast Mode* that is used to disable these spanning tree stages on a port-by-port basis. Cisco switches have an equivalent feature called *PortFast*.

The following recommendations are environment-specific:

- Network fault tolerance and transmit load balancing
  - Team members can be split across more than one switch to achieve switch redundancy. However, all switch ports that are attached to members of the same team must comprise a single broadcast domain (that is, the same VLAN). Additionally, if problems exist after deploying a team across more than one switch, reattach all team members to the same switch. If the problems disappear, then the cause of the problem resides in the configuration of the switches and not in the configuration of the team. If switch redundancy is required (team members are attached to two different switches), then HP recommends that the switches be deployed with redundant links between them and spanning tree protocol (or other Layer 2 redundancy mechanisms) be enabled on the ports that connect the switches. This helps prevent switch uplink failure scenarios that leave team members in separate broadcast domains.
- Transmit load balancing and Switch-assisted Load Balancing
  - Teams that communicate with TCP/IP network devices through a router should use the IP address-based load balancing algorithm (configured using the Network Configuration Utility).
- Switch-assisted Load Balancing
  - You should thoroughly understand the configuration guidelines set by the switch vendor because Switch-assisted Load Balancing is dependent on the switch being configured in a compatible mode. HP designed its Switch-assisted Load Balancing technology to provide flexibility. Therefore, the Network Configuration Utility can enable configuration of a Switch-assisted Load Balancing team that will not work correctly with a particular vendor's switch.
  - The load balancing algorithm for the switch should be set to XOR or SOURCE-BASED but not DESTINATION-BASED.
  - The load balancing algorithm of the switch should be set to balance by IP address if most traffic destined for the server originates on a different network and must traverse a router.

## Channel bonding

Channel bonding is a method of striping data across multiple NICs installed in Linux servers. As with NIC teaming, you can bond two or more NICs so that they function as a single NIC, providing either NIC with failover or load balancing capabilities.

Channel bonding requirements on a Linux server include:

- Use `ifconfig` to initialize the first NIC in the bond and `ifenslave` to initialize the remaining NICs.
- Configure a minimum of two physical subnets.
- Both NICs must have the same MAC address.

For bonded NICs in failover mode, only one of the NICs is active at any one time. Failover occurs when one of the NICs in the bond fails. Upon NIC failure, the standby NIC within the bond takes over for the failing NIC, with no loss of data. Failover generates an error in the log. There is no automatic failback between bonded NICs; the NIC currently handling traffic continues to do so unless it fails.

In load balancing mode, the bond handles packets in a round-robin manner, resulting in a higher aggregate throughput than is possible with a single NIC port. The amount of throughput gain is dependent on the bandwidth of the network. For load balancing mode, the bond must also be configured on the network switch to which the bonded NICs are attached. See the switch user manual for instructions on configuring a bond.

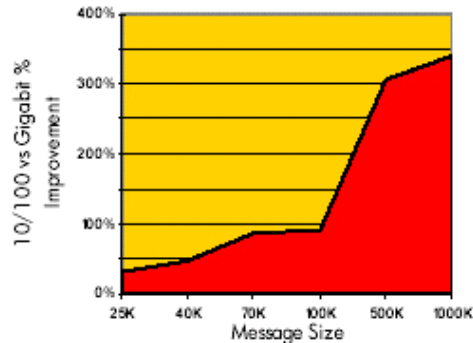
You need two switches or hubs as after you have successfully completed the bond. You must have one NIC connected to one switch or hub and the other NIC connected to a second switch or hub. After it is bonded, one NIC sends packets and the other receives packets.

## Teaming and bonding utilities

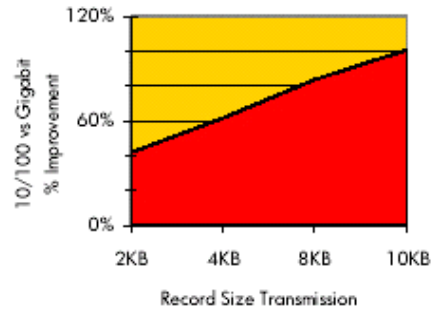
Depending on the operating system used in a ProLiant server, you must use different utilities to configure teams of ProLiant NC-series Ethernet NICs.

Operating system	Where to configure
Windows 2000 and Windows Server 2003	HP Network in the Control Panel
Linux	Execute <i>ifenslave</i> from the command line or see <i>bonding.txt</i> for more advanced options
Novell NetWare 4-6	Command line statements added to <i>autoexec.ncf</i>
Windows NT 4.0	Compaq Teaming in the Control Panel
SCO OpenUnix 8	Execute <i>netcfg</i> from the command line
SCO Open Server 5	Execute <i>cpqnim</i> from the command line

## Using Gigabit Ethernet



Microsoft Outlook 2000



Microsoft SQL Server 2000

Migrating to Gigabit Ethernet boosts network performance at the workstation and the server. Gigabit Ethernet provides a 10X speed increase over Fast Ethernet and a 100X speed increase over standard Ethernet 10Base-T. The latest standard, 10 Gigabit Ethernet, operates at 10Gb/s. However, it only functions over optical fiber and only operates in full-duplex mode, so collision detection protocols are unnecessary.

Because Gigabit Ethernet uses the same format as standard Ethernet and Fast Ethernet (standard Category-5 cable), a company can maintain its existing Ethernet infrastructure. This easy migration path allows seamless communication among the Ethernet family of protocols.

To upgrade to a faster network, you must:

- Upgrade all hubs to switches.
- Upgrade the wiring.
  - Use Category-5 unshielded twisted pair (UTP)/shielded twisted pair (STP) wiring when upgrading from standard Ethernet to Fast Ethernet.
  - Choose either copper-based Category-5 or fiber cabling when upgrading to Gigabit Ethernet.
  - Use optical fiber cabling for 10 Gigabit Ethernet.
- Upgrade all NICs.

ProLiant servers support both fiber-based and copper-based Gigabit NICs. These HP NICs offer the latest networking technology with high performance, reliability, fault resilience, and load balancing capability. They have been designed, developed, tested, certified, and validated specifically for ProLiant servers. Additionally, these NICs are integrated with ProLiant management utilities for easy, seamless deployment and management.

---

**INTERNET**

You can find more information on HP Gigabit NICs and download the latest software and drivers from: [http://h18007.www1.hp.com/support/files/networking/nics/Compaq\\_Gigabit\\_Ethernet\\_NICs.html](http://h18007.www1.hp.com/support/files/networking/nics/Compaq_Gigabit_Ethernet_NICs.html)

---

A key decision when planning to migrate to Gigabit Ethernet is whether to deploy Gigabit Ethernet over fiber or copper wiring. Generally, fiber is best used when:

- Distance is important.
- You know that you will eventually move to the 10 Gigabit standard.
- You need a generally more secure wiring type (EMI/RFI).

On the other hand, copper Gigabit NICs are best used when:

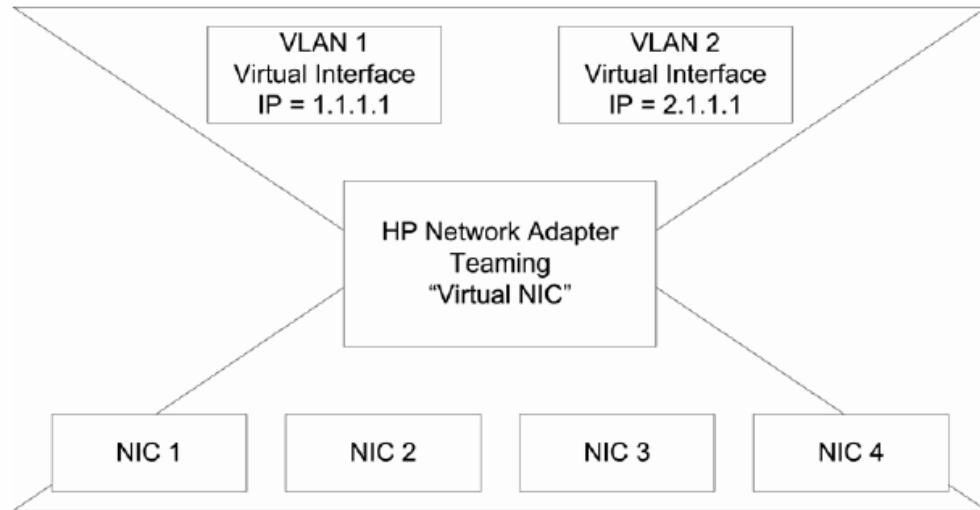
- You have budget constraints.
- You do not want to perform a major infrastructure upgrade (because most of the existing Category-5 wiring is already in place and ready for use).
- You want to perform a gradual (seamless) migration within a Category-5 infrastructure.

The following factors also influence the choice of media:

- Category-5 cable is the most common medium for horizontal cabling in ceilings and floors. It supports distances of up to 100m.
- Fiber cable is the most common choice for connecting buildings in campus settings and has been approved for 550m to 5km lengths.
- Either Category-5 or fiber cable can be used in the vertical risers that connect different floors within a building.
- By using vendor-specific, long-haul Gigabit Interface Converters (LH GBICs) in switches, Gigabit connections can be established up to 70km.

The migration plan should include prioritizing critical network segments, understanding bottleneck areas, standardizing on a platform, and testing that platform. Category-5 copper cabling is likely to already be in place within the data center, but fiber typically is used to connect buildings, to link segment switches to the data center, and to connect servers outside the enterprise.

## Adding VLANs



VLAN tagging used with HP NIC teaming

VLANs enable you to configure group users and stations together in logical workgroups. This configuration can simplify network administration when connecting clients to servers that are geographically dispersed across the building, campus, or enterprise network.

The Network Configuration Utility supports the configuration of VLANs on standalone NICS and on NIC teams, which enables a NIC or a NIC team to belong to more than one VLAN at the same time.

---

**Note**

The Network Configuration Utility installs during a SmartStart installation. To ensure that the utility is properly updated, always select *Install* when updating the network teaming configuration. The Support Pack and Smart Component will then update the software package if network teaming is already installed.

---

VLAN tagging uses the IEEE 802.1Q protocol to divide a physical network into one or more virtual networks, which reduces broadcast traffic within a network segment and adds security. When multiple VLANs are configured, 802.1Q VLAN tagging marks every transmitted frame with a VLAN identifier. VLAN tagging requires the support of the network infrastructure and/or the receiving network device.

The use of VLANs has only one effect on the operation of the NIC team, which is that heartbeats are transmitted only on the numerically lowest configured VLAN on the team. This means that if four VLANs are configured on the team and the numerically lowest VLAN is 20, the teaming driver will use VLAN 20 to transmit heartbeats between team members.

Much like deploying Switch-assisted Load Balancing, the use of VLANs requires the switch or switches to be configured properly. You must configure the switch port of every team member with the same VLAN configuration. This means that if a team is to operate on four different VLANs, every team member must have all four VLANs configured on their respective switch port.

The maximum configurable VLANs per team is 64. The valid VLAN number range per LAN is 1 to 4094.

VLAN tagging used with HP NIC teaming provides the same functionality to the operating system as having two NICs installed, but enables fault tolerance and load balancing across four NICs. The configuration is as follows:

- Four NICs are teamed together as a single virtual NIC using HP NIC teaming
- Two VLANs are configured on top of the virtual NIC to create two virtual interfaces

---

**Note**

For more information on VLANs in NetWare, see the nwteam.txt file on the HP NetServer 10/100TX PCI LAN Adapter Drivers CD.

---

## **Adding switches**

When you use network switches, each port on the switch is seen as the only client on that network segment. In this case, your network maximizes throughput because collisions are eliminated.

Switches allow greater utilization of bandwidth, but the added cost of intelligent switches might outweigh the performance gains in smaller, less stressed network segments.

When planning or implementing your network, be aware of utilization limitations. Either distribute network traffic between separate network segments or use intelligent switches to eliminate packet collisions and maximize throughput.

When migrating to Gigabit Ethernet, replace the switches that do not support Gigabit speeds with those that support auto-negotiation of 10Mb, 100Mb, and Gigabit technology. Auto-negotiation allows both Gigabit devices and legacy devices on the network to operate seamlessly in the faster infrastructure. Clients are connected directly into the switches. As the network expands, you can add Gigabit switches to maximize the investment in new hardware.

## **Adding network segments**

Another way to increase bandwidth is to break large network segments into smaller segments. Adding segments also helps span the network over greater distances, and is especially useful in Ethernet 10-Base T topology, which has a transmission limitation of 100 meters.

If a NIC senses a transmission on the network, it waits to transmit its frame. After the wait period elapses, it attempts to transmit the frame again. The NIC will wait 15 times before deciding the network is too busy. It will then clean the frame from its memory.

A packet collision occurs when two NICs transmit simultaneously. When a NIC detects a collision, it generates an alarm. If collisions continue to occur, the NIC dumps the frame and does not attempt to transmit. This happens if too many devices are on one network segment.

Each segment and its users have full use of the bandwidth available. For example, if there are 100 users on a 100Mb/s segment, each user has an average of 1Mb/s of available bandwidth. However, if this same segment were further divided into 10 segments with 10 users on each segment, each user would have an average of 10Mb/s of available bandwidth.

You can segment a network using a bridge, router, or switch. In most cases you must assign a contiguous range of IP addresses to each segment.

## Tuning the operating system and applications

You can tune the operating system and the applications to offload some of the I/O processing from the storage subsystem. This depends, however, on the type and version of the software. For example:

- **Tune the operating system cache to better coincide with the application requirements** — For example, configure a large operating system cache for file servers. Configure minimal or no operating system cache for applications that allocate and manage their own cache.
- **Tune the memory to avoid hard paging** — Hard paging results from over-allocation of physical memory and results in physical I/O to and from the page file.
- **If applicable, use high-performance drivers** — For example, you can replace the Windows miniport device driver with the Enhanced Command Interface (ECI) driver to lower I/O-related processor utilization.

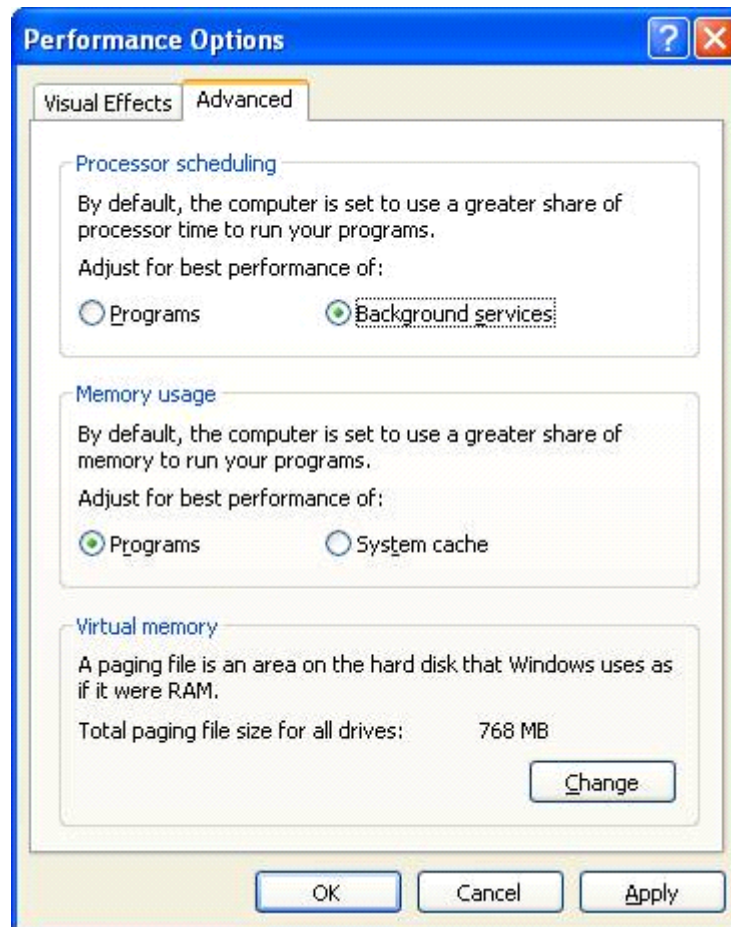
---

**INTERNET** Refer to the HP website (<http://www.hp.com>) for more details about the ECI driver.

---

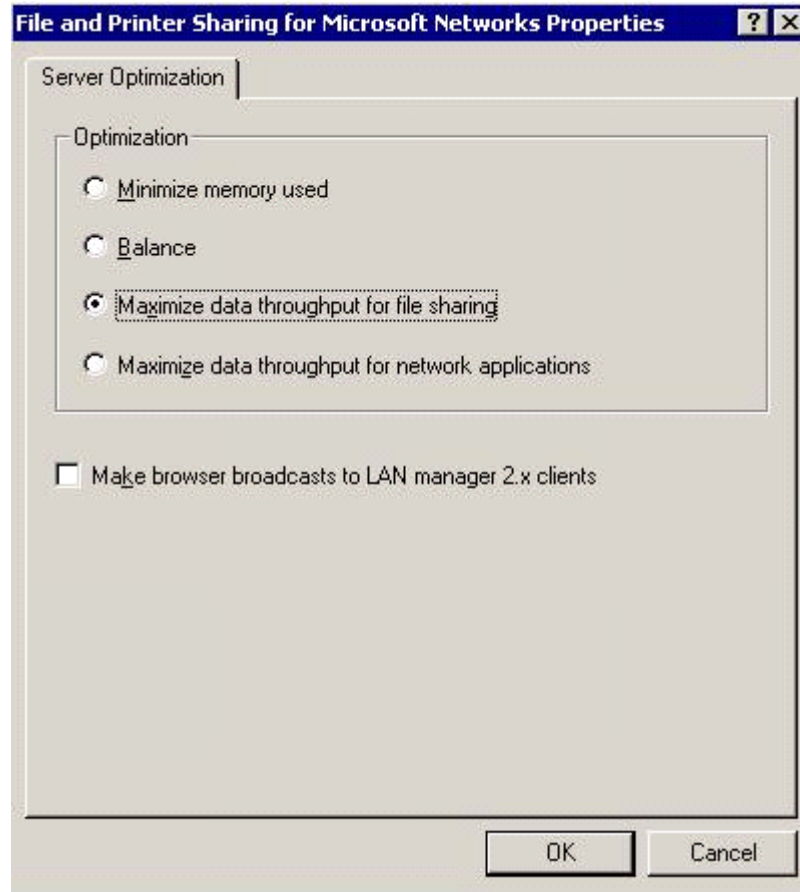
- **Remove any monitoring tools that interfere with the storage subsystem performance** — Performance monitoring tools are inherently intrusive; the actual impact on performance varies from one tool to another.
- **Use asynchronous I/O whenever possible** — Asynchronous I/O enables the operating system to issue I/O requests without waiting for their completion. The operating system can execute another task until it is interrupted by the completed I/O request. Asynchronous I/O is more efficient and faster than synchronous I/O, allowing multiple I/O requests to be issued in a single call, thus completing more work in the same amount of time.
- **Deploy a file system consistent with your performance expectations** — Certain file systems are designed to perform better under certain conditions. Many operating systems support raw partitions, which are unformatted disk devices that support I/O without a file system. Performance gains of raw partitions vary and are achieved by losing manageability.
- **Defragment the file system** — Many file systems are designed to keep the files contiguous and close to each other, thus avoiding excessive disk head movement (seeks). Determine whether the maker of your particular file system recommends defragmentation and which tools to use.
- **Optimize the application I/O** — If the application generates a large number of I/O requests but uses only a small portion of the data it retrieves, the application might need to be tuned.

## Configuring processor scheduling



Some operating systems, such as Windows, enable you to assign processing priority to either foreground applications or background services. On application servers, you should set the processing priority to background services to improve the application network response times.

## Optimizing the server



Server optimization refers to matching the operating system resources such as memory resources with the primary server role. Under Windows, for example, you can set server optimization to one of the following options:

- **Minimize memory used** — Optimizes the server for a small number of clients.
- **Balance** — Optimizes the server for a mix of file and print services, network applications, and console interactions.
- **Maximize data throughput for file sharing** — Improves performance for servers with high network usage, such as web and application servers. This setting assumes that the application does not cache its data and takes maximum advantage of the operating system file cache.
- **Maximize data throughput for network applications** — Improves performance for servers with low network usage, such as database servers. This setting minimizes use of the operating system file cache and allows the application to use its own caching methods.

## Scaling out performance

Software scaling increases the capacity of an application by adding servers. Although hardware scaling requires specialized servers, software scaling can be achieved using standard off-the-shelf servers.

Tools developed for Windows, Linux, and NetWare can be used to reduce the time and resources needed to perform software scaling. These tools can:

- Accelerate cluster deployment
- Scale with the growth of an enterprise
- Require no new application programming interfaces (APIs)

## Distributed applications

In a distributed application model, a client communicates with a middle layer that consists of the application server and an application containing the business logic. The application, in turn, communicates with a database that supplies and stores data.

In a distributed architecture, a middle layer is used to process requests from the client to the server. The middle layer alleviates the processing of rules and decision logic from both the client and the server. This permits the construction of thin clients and the removal of processing logic from the server layer. The server layer then can behave as a source of raw information. Processes can be distributed to any one of the layers.

Distributed architecture is based on a network model in which processes can be distributed on any processor and any two individual nodes of the network are in a client/server relationship with any number of intervening middle layers.

HP partners with TIBCO Software, Inc. to provide software that enables application developers to build scalable distributed applications on ProLiant platforms. TIBCO Rendezvous software is a messaging program that enables diverse applications to share data across LANs and WANs. Programs on heterogeneous platforms communicate transparently with self-describing data messages and subject-based addressing.

## Clustering



When the performance requirements outgrow the server capabilities, you have two basic choices:

- Replace the existing server with one with more capabilities.
- Add another server to the existing one and form a performance cluster.

Clustering technologies enable you to combine multiple independent servers into a single logical processing unit. The cluster then executes multiple copies (instances) of the application and accesses a common set of data. When a cluster node or a component within the node fails, the application and data accessibility are not affected. Thus, clusters are designed to meet two primary goals:

- Increase availability
- Increase performance

Not all clustering designs allow you to increase performance beyond a single server, and not all applications are designed to run in a cluster environment. Some address the availability issue only, some address the performance issue only, and others address both.

To cluster an application for performance gains, the application must be designed to execute within the cluster environment. The application must be able to execute multiple instances, typically with each instance running on an individual node within the cluster, and have a coordinated access to the shared information. This coordinated access can be achieved by partitioning the data and allowing each node to access only the assigned partition. Alternatively, a data access manager must be in place to solve the challenges of concurrent access to the same data.

Examples of enterprise applications supporting clusters include SQL Server 2000 and Oracle9i Real Application Clusters.

The adoption of clustering implementations using industry-standard servers, storage solutions, and cluster interconnects has increased in the recent years, driving down the purchase and implementation costs. The expertise with implementing and administering clusters has also increased, making clustering a viable option for companies of any size.

---

**Note**

Clustering technologies are discussed in more detail in Module 5 — High availability and clusters.

---

## Optimizing remote management performance

As an enterprise network grows, some configuration changes must be made to enhance the performance of Systems Insight Manager and Lights-Out devices. Most of these changes are on the client side.

### Systems Insight Manager client configuration

If the system list panels do not display, the problem could be related to the current set of Java Plug-in Applet Cache files stored on the system that might need to be deleted. Follow these steps to solve the problem.

1. From the client Windows machine, click *Start* → *Settings* → *Control Panel*.
2. Double-click the Java plug-in icon.
3. Select the *Cache* tab.
4. Click the *Clear* button.
5. A confirmation prompt displays. Click *Yes*.
6. Close the Java Plug-in Control Panel window.

### Systems Insight Manager database

Problems with Systems Insight Manager performance could be attributed to the use of the MSDE database. For improved performance, Microsoft recommends fewer than five concurrent users of MSDE. If an application requires more than five concurrent users, SQL Server is the best option. SQL Server is capable of supporting hundreds or thousands of concurrent users at the highest levels of enterprise-class performance.

## Lights-Out graphical remote console

The settings on the remote server determine the performance of the graphical remote console for Lights-Out devices. When using the remote console for Lights-Out devices, the resolution of the remote host server operating system display should be the same or lower than the settings on the client system. Higher resolutions transmit more information, which might slow the overall performance of accessing the management processor.

Use the following client and browser settings to optimize performance for the remote console:

- Use Microsoft Internet Explorer 5.5 (or later), Netscape Navigator 7, or Mozilla.

---

**Note**

Disable the Enable SSL version 3 and Enable TLS options in Netscape 6.2.2.

---

- Set the display properties on the client to greater than 256 colors.
- Use a 700MHz or faster single-processor client with at least 128MB of memory.
- Set the mouse pointer speed to the middle setting.
- Select a screen area greater than the resolution of the host server operating system.
- Set mouse pointer acceleration to *Low* or *None* to disable the pointer acceleration.

Memorizing the mouse settings can be a challenge. Instead, use the HP Lights-Out Mouse OPTimization (HPLOMOPT) utility, which is a Windows application that automates the configuration of the optimum mouse settings.

---

**INTERNET** This file can be downloaded from: [http:// www.hp.com/servers/manage](http://www.hp.com/servers/manage)

---

## Remote console hot keys

Using the remote console interface can also be enhanced with remote console hot keys. Ctrl keys are used often in the Windows interface for tasks such as locking the computer and logging in. Ctrl key combinations such as *Alt+Ctrl+Del* are not available in a remote console session. Remote console hot keys enable you to define such Ctrl key combinations.

### Example

Key sequence *Alt+Ctrl+Del* could be assigned to *Ctrl+X*. When *Ctrl+X* is pressed during a remote console session, *Alt+Ctrl+Del* is transmitted in its place. Each key combination can contain a maximum of five keys.

## Lights-Out device group administration

You can manage multiple Lights-Out devices by using the Remote Insight Board Command Language (RIBCL).

Systems Insight Manager uses the Lights-Out Configuration Utility to send an RIBCL file to a group of Lights-Out devices. The devices perform the action designated by the RIBCL file and send a response to a log file.

---

**INTERNET** This file can be downloaded from: [http:// www.hp.com/servers/manage](http://www.hp.com/servers/manage)

---

The Lights-Out Configuration Utility must reside on the same server as Systems Insight Manager; however, it can also be used through a batching process.

The following statement illustrates the cpqlocfg.exe command line and switches:

```
cpqlocfg.exe -s server_name -f c:/ribclfile.txt -l  
c:/logfile.txt -v
```

- -s denotes the RILOE II board to be updated.
- -f provides the location and name of the RIBCL file.
- -l defines the path and file name of the log file to be generated. When this switch is omitted, the file is stored in the directory where cpqlocfg.exe is launched and the log file name is the DNS name or IP address.
- -v enables the verbose messaging system.
- -c checks the XML syntax but does not open a connection to the board.

As with any other scripting task, using the Lights-Out Configuration Utility might require some trial and error before a script runs successfully. The Lights-Out Configuration Utility generates two types of error messages:

- **Runtime** — Occurs when an invalid action is requested
- **Syntax** — Occurs when an invalid XML tag is encountered; this interrupts the utility and the runtime script error is logged in the output log file

**Example**

Syntax error:

expected USER\_LOGIN=userlogin but found USER\_NAME=username

The supported Lights-Out Configuration Utility functions include:

- |                                       |  |
|---------------------------------------|--|
| ■ Add, modify, or delete a user       | ■ Obtain and set virtual floppy status           |
| ■ View user configuration information | ■ Insert, copy, and eject a virtual floppy image |
| ■ Modify network settings             | ■ Configure Remote Console hotkey settings       |
| ■ Modify global settings              | ■ Obtain and set Virtual Power status            |
| ■ Clear the RILOE II event log        | ■ Obtain the server power status                 |
| ■ View firmware version               | ■ Reset the server                               |
| ■ Update firmware                     |  |

## Step 4 — Executing the action plan

The next step is to execute your action plan. In particular:

- Ensure that you have a valid data backup and a tested recovery strategy.
- Follow your written action plan step by step, documenting each action you take and its result.
- Apply only one solution or modify only one variable at a time. If multiple solutions must be implemented or multiple variables changed at one time, do so in the smallest increments possible.
- Ensure that your action does not have an adverse effect in another area.

After the bottleneck devices have been identified, the next step to improve performance is to take measures to reduce response times. Two ways to reduce response time are:

- Reduce the service time of the devices.
- Reduce usage of the devices, which reduces the service time.

### Reducing service time

Reducing service time often requires replacing the existing drives with newer technology. Drives speeds of 15,000 rpm provide the greatest performance; Fibre Channel Arbitrated Loop (FC-AL) drives do not provide major performance gains over SCSI drives.

Another way to improve performance is to increase the number of queuing centers (increase the number of spindles).

### Reducing usage

Reduce the device request rate or the device service time to lower usage. You can accomplish by:

- Distributing the workload over more drives by using a drive array and RAID technology
- Shifting the workload to another device if the application permits
- Bypassing the device with cache

## Step 5 — Determining the effectiveness of your actions

After performing each action in your action plan, measure its impact on performance and compare the outcome with the desired result. Perform the measurement in a controlled environment, preferably with a predetermined set of tools and procedures. Whenever possible, use the same tools used in the initial data collection.

If the action you have taken does not solve the performance issue, reverse its effect and perform another action in your action plan. When you exhaust all possibilities without achieving the desired result, try an alternate solution or re-evaluate the order of the tasks in your action plan.

If your action plan still did not identify an alternate cause of the issue or resolve it, diagnose the mode of failure again. Retracing your steps in step 2, isolate the faults to a subsystem or software component level. If you still are unable to resolve the issue, the data with which you began your troubleshooting approach is inadequate. Start again to collect more information and step through the six-step troubleshooting methodology again.

---

### Lab

Your instructor might now direct you to perform Module 4, Lab Exercise 2 — Monitoring and tuning a database server.

---

## Step 6 — Implementing preventive measures

Performance management does not stop with solving the performance issue at hand. You must be proactive and implement measures that prevent the same problem from reoccurring in the future. Preventive measures vary from one site to another, but generally include:

- Keeping the system software current
- Maintaining the performance baseline
- Configuring performance-based alerting
- Implementing fault tolerance and failure recovery
- Partitioning resources

### Keeping system software current

A carefully planned system software maintenance strategy not only maximizes server stability and availability, but also ensures that the server performs efficiently by running the latest qualified system software.

Use any applicable system software maintenance strategy, such as VCRM, and ensure that the software patches and upgrades have been tested thoroughly from both the functional and performance perspective. Always document any system-level and application-level changes to the system and have a tested fall-back plan.

### Maintaining the performance baseline

The performance baseline should be created when a system is first placed into production and its performance levels meet the requirements. The performance baseline must be properly documented and distributed to key stakeholders.

Create a new performance baseline every time you change a hardware or software configuration that affects server functionality or performance. Be sure to document completely and accurately what has changed. If the performance requirements change, document those as well. Treat the new performance baseline as a starting point in addressing all newly arising performance issues. Save the previous baselines for historical trend analysis.

## Configuring performance-based alerting

Performance-based alerts warn you of changes in a monitored object performance. You can then tailor the level of response to the type of alert received.

### Using PMP

Using PMP, performance-based alerts warn of changes in the ProLiant server performance. An alert is triggered every time the server or subsystem performance changes between normal performance, approaching bottleneck, and confirmed bottleneck.

PMP generates these alerts based on the activity of the monitored servers and relies on the Insight Manager 7 alerting mechanism to process and propagate these alerts. To configure PMP alerting, complete these three tasks in Insight Manager 7:

1. Create an event query or use an existing one.
2. Set a schedule for the query execution.
3. Associate this query with an email or pager task, or an application launch.

### Using other tools

You can configure many operating system management and monitoring tools, such as Windows System Monitor, to generate administrative alerts when preset thresholds are exceeded. Many of these tools, however, have limitations in the areas of what they can monitor, how they interpret the information they collect, and what types of alerts they support.

Specialized management tools, such as HP OpenView and Systems Insight Manager, are designed to monitor health and performance levels of the target system and to generate administrative alerts when a degraded condition is encountered. They have built-in intelligence to determine hardware and software thresholds, detect pre-failure conditions, and use specialized management solutions for alerting and condition response.

## Implementing fault tolerance and failure recovery

Many performance-related problems occur as a result of a hardware or software component degradation or failure. For example, if a disk drive in a RAID 5 volume fails, the associated array operates in a data recovery mode at lower performance. Recovering from this failure (by replacing the failed drive and recovering the data) also restores the performance.

Proactive monitoring of server health, fault-tolerant components and configurations, and a sound fault recovery plan are all part of a preventive measures strategy that should be part of every enterprise.

## Consolidating workloads

Demands to reduce overhead cost and increase resource efficiency make it desirable to consolidate multiple applications onto larger servers. By consolidating applications from two or more under-utilized servers onto a single server, an IT administrator reduces cost and complexity, and uses the computing resources more efficiently.

There are several types of server consolidation. Workload consolidation combines applications from multiple servers onto fewer, more powerful servers. Workload consolidation increases performance when newer, more powerful hardware and software are deployed.

### Models for workload consolidation

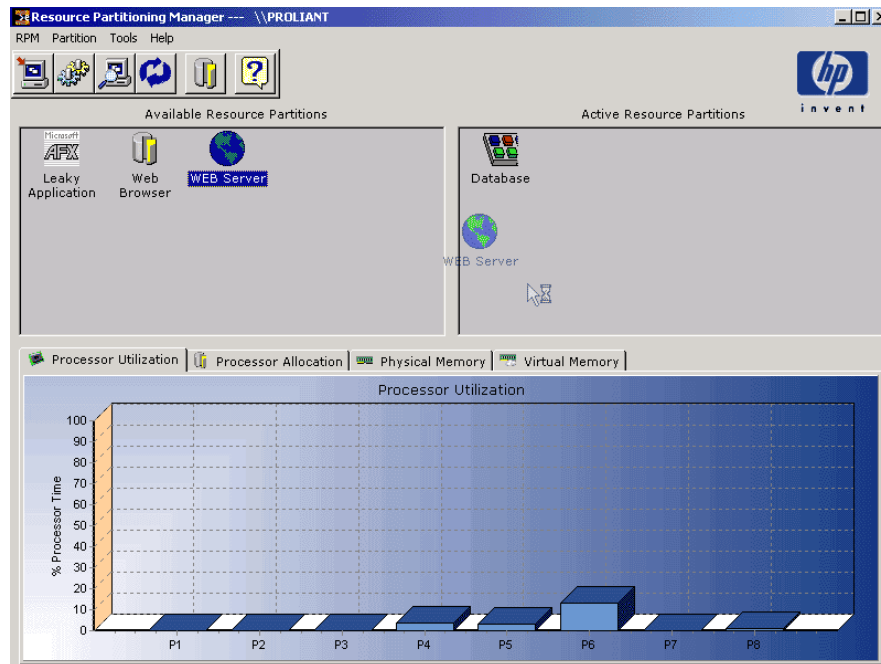
There are several ways to increase the number of applications that can run on a single server:

- **No partitioning** — Loading and running multiple applications on the server.
- **Hardware partitioning** — Running multiple operating system images on completely isolated hardware.
- **Logical partitioning** — Running multiple operating system images through a software layer.
- **Resource management** — Isolating specific resources for use by an application.

Partitioning and resource management can be accomplished statically or dynamically. Static partitions can be changed only at a reset of the system; dynamic partitions can be changed at runtime. Dynamic partitioning requires more sophisticated implementations of the operating system, middleware, and applications, so the software can properly de-allocate resources before giving these resources to a new entity.

Resource management brings many of the benefits associated with a logical partition. It provides the stability needed to run multiple applications on a single server without the added capital cost of a server using hardware partitioning. Resource management also reduces the complexity of a logical partition by managing only specific resources such as memory and processors, allocating these resources to specific applications.

## Resource Partitioning Manager



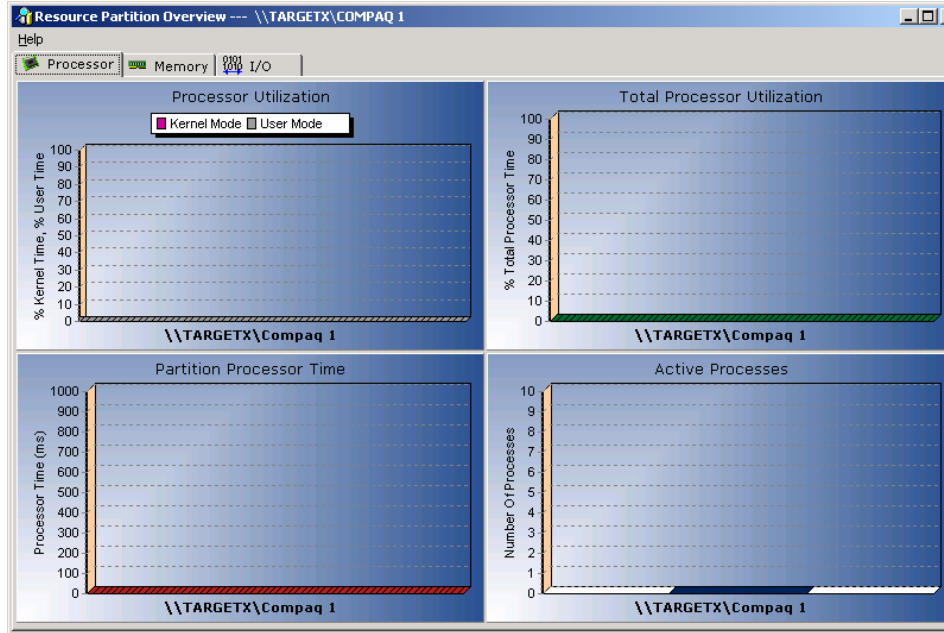
HP ProLiant Essentials Workload Management Pack provides the tools you need to safely consolidate applications and take advantage of any underutilized server assets. Workload Management Pack features Resource Partitioning Manager (RPM), which increases the stability and availability of applications, thereby enabling customers to deploy multiple applications on a single server.

RPM controls and dynamically allocates system resources to enable application consolidation and performance optimization on Windows 2000 server platforms. RPM helps protect your environment from memory-grabbing or disruptive applications but allows applications to share a server.

With RPM, you can configure resource partitions, which are application boundaries defined by their allocated quantity of processor and memory resources. Partitioning enables you to limit applications to the resources within their respective partitions.

Each partition can access specific, limited processor and memory resources. Limited access means that resources are available for your applications when they need them. It also means your system will no longer suffer decreased availability created by memory-leaking applications. You can restart just the partition that houses a troubled application—without rebooting the entire server.

## RPM functions



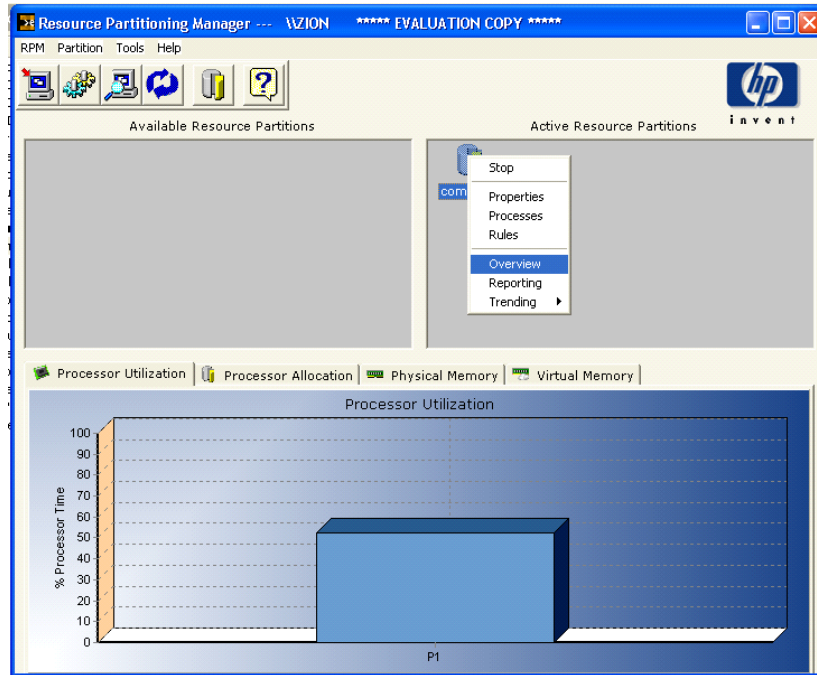
Resource Partition Overview Screen

Although you can limit the amount of resources available to processes, you can also add resources on demand wherever they are needed. With the RPM dynamic rules engine, you can assign rules to each partition to automatically adjust partition size based on utilization or time of day. This capability enables you to provide the service levels and application performance users need during peak load periods without having to keep idle resources standing by at other times.

RPM performs four basic functions:

- **Create resource partitions** — You can create the resource partitions on a target server by establishing a name and defining which processors and what amount of memory the partition will include.
- **Assign processes** — You can choose which executables, Dynamic Link Libraries (DLLs), and services run within the resource partition. This ensures that these processes use only the assigned partition resources.
- **Define rules** — A Dynamic Rules Engine creates flexibility in the amount of resources available to each partition. You can use the engine to establish rules that allow objects to dynamically increase or decrease processor and memory resources as needed.
- **Evaluate configuration and performance** — Allocation, utilization, and performance graphs give you quick performance information and feedback.

## Configuring an available resource partition



From the Resource Partitioning Manager window, you can change the configuration of an available resource partition by modifying its properties, processes, and rules.

In the Available Resource Partitions panel, right-click a resource partition to display the options drop-down menu. From this menu, you can select:

- **Start/Stop** — Initiates or terminates the resource partition
- **Properties** — Displays the Resource Partition Properties window, which allows you to modify the properties of a resource partition
- **Processes** — Displays the Resource Partition Processes window, which allows you to modify a process list
- **Rules** — Displays the Resource Partition Rules window, which allows you to modify existing rules
- **Overview** — Displays the Resource Partition Overview screen, which provides processor, memory, and I/O information for the selected resource partition

- **Reporting** — Displays the Reporting Preferences window, which enables you to create reports
- **Trending** — Displays options to enable or view trending data
- **Send to** — Displays options that allow you to export partitions to a file or to a target machine
- **Copy** — Creates a new resource partition with the same properties, processes, and rules as the selected resource partition
- **Delete** — Deletes the selected resource partition
- **Rename** — Renames a resource partition

RPM is installed on the managed server and can be accessed both directly from that server and from other networked servers running RPM.

Launching RPM displays the user interface, which serves as the control point for all RPM functions. However, because RPM controls resources using the RPM service, the user interface does not need to be active at all times. You can close the user interface window at any time without affecting the configurations, rules, or execution of any resource partition.

---

### **Lab**

Your instructor might now direct you to perform Module 4, Lab Exercise 3 — Installing and using HP Resource Partitioning Manager 2.0.

---

## Summary

Performance management encompasses performance planning, monitoring, and tuning. It entails:

- Designing the network infrastructure and the systems within it to meet the performance requirements
- Examining the behavior of critical components
- Determining whether they perform optimally
- Reacting to performance issues as they arise
- Preventing the issues from repeating themselves

Because the IT staff at RC Engineering had already installed Insight Manager 7, PMP, and VCRM, you were able to show Jackie how to license the servers and configure the performance monitoring infrastructure. You also performed the necessary management agent updates.

After they complained that their network management functions were taking longer than expected, you applied the HP Troubleshooting Methodology not only to solve system health-related problems, but also to solve performance issues. You performed static analysis and showed Jackie and her staff how to create reports based on logged performance statistics. You also configured performance-based alerts to prevent similar performance issues in the future.

## Learning check

1. To use PMP, you must license the servers of interest and enable monitoring.  
☐ True  
☐ False
2. When you use PMP to configure performance-based alerts, when are alerts triggered?  
.....  
.....
3. How do you design an application to execute within a cluster environment? (Select **two**.)
  - a. The application containing the business logic must communicate with a database that supplies and stores data.
  - b. Partition the data and allowing each node to access only the assigned partition.
  - c. Reconfigure the application cache to increase performance.
  - d. Put a data access manager in place to solve the challenges of concurrent access to the same data.
4. Write cache is beneficial in \_\_\_\_\_ environments; read cache improves performance in \_\_\_\_\_ environments.
5. Asynchronous I/O is more efficient and faster than synchronous I/O.  
☐ True  
☐ False
6. What should you do if multiple solutions must be implemented or multiple variables changed at a time?  
.....
7. Name three steps you should take if you have exhausted all possibilities in your action plan without achieving the desired result.  
.....  
.....  
.....

8. What does an amber status indicator alert you to in PMP?
- a. Normal performance
  - b. Unknown device
  - c. Approaching bottleneck
  - d. Critical condition
  - e. Confirmed bottleneck
9. When using RPM to restart a partition that houses a troubled application, you must reboot the server.
- ☐ True
- ☐ False
10. Match the types of NIC teaming with their characteristics.
- |                                   |       |  |
|-----------------------------------|-------|--|
| a. Network fault tolerance        | ..... | Is switch-independent and can be split across Layer 2 switches |
| b. Transmit load balancing        | ..... | Is supported on IEEE 802.3ad-capable switches                  |
| c. Switch-assisted Load Balancing | ..... | Provides simple redundancy with two to eight NICs              |



### Objectives

After completing this module, you should be able to:

- Describe ENSAextended concepts and architecture
- Discuss the technology involved with moving from direct attached storage (DAS) to a storage area network (SAN), a process called DAS to SAN (DtS)
- Identify the benefits and components of an HP SAN.
- Plan and design a SAN solution
- Describe a cluster
- Identify HP cluster solutions
- Discuss how to deploy and manage clusters
- Describe the technologies available to monitor clusters
- Explain the tools and methods used to troubleshoot SANs and clusters

## Introduction

As the result of a new widget they have developed to accelerate genetic sequencing, RC Engineering has caught the attention of the Greater Environment for the Expansion of Knowledge (GEEK), a nonprofit knowledge management organization. GEEK uses clusters attached to a SAN to warehouse the data resulting from its work in genetic mapping. The organization has agreed to merge with RC Engineering to leverage the RC Engineering product base and thus further GEEK's ongoing research into the genetic causes of learning disabilities.

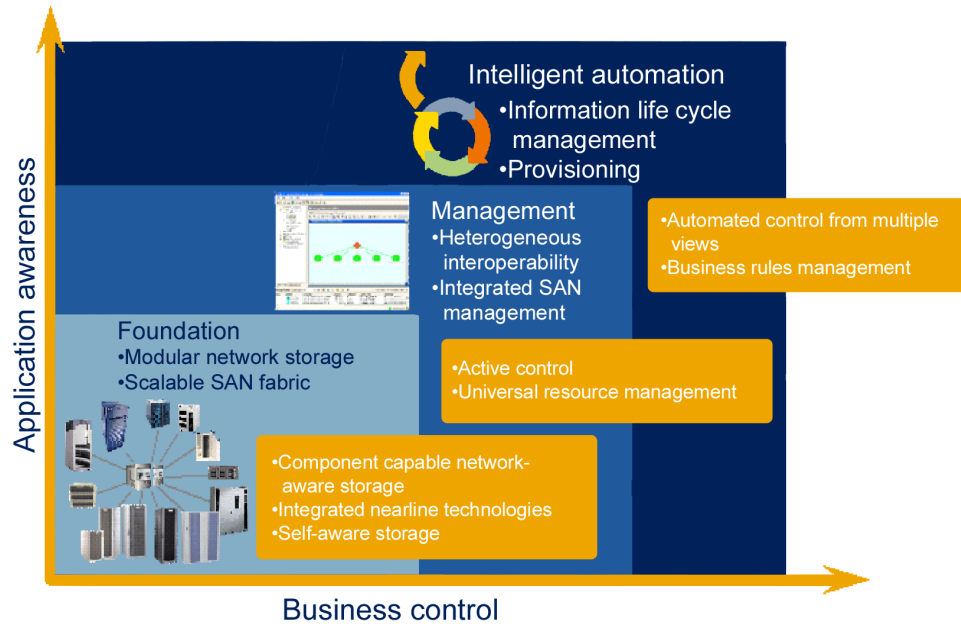
Carla is the managing director of GEEK. Her firm generates simulations that are compute-intensive, but with modest communications requirements. GEEK expects to receive a grant from the National Science Foundation to further its research, which will require expandable disk space, SAN switching, and the ability to split tasks among disks. These are all resources of an HP SAN. GEEK would have little flexibility to use disk space optimally without the architecture of a SAN.

Because RC Engineering had been using the DAS solution that you set up for them previously, you will need to perform a DtS conversion.

GEEK and RC Engineering also need higher availability. Carla and Bob, the CEO of RC Engineering, decide to build a set of cluster services using distributed event and naming services and a variety of application programming interfaces (APIs).

Because GEEK already has an environment built on HP ProLiant BL p-Class server blades, the RC Engineering equipment must be migrated from SCSI to Fibre Channel. This configuration will allow Carla to start the newly funded projects quickly with room for future growth.

# ENSA



The primary objective of the HP Enterprise Network Storage Architecture (ENSA) is to create a portfolio of modular, scalable, and highly available products. ENSA places physical disks into a large consolidated pool and creates individual user disks from the pool, transforming storage into a utility service that users access reliably and transparently.

ENSAextended builds on existing foundations of network storage and centralized management to deliver more automation, making it easier to control a storage environment. It enables customers to adapt and optimize storage and service capabilities through:

- **Universal networked storage** — Expands the total solution of all storage elements
- **Virtualization** — Adds greater simplicity and utilization of heterogeneous storage assets
- **Data services** — Includes performance optimization, data archiving replication, and migration augmented by offerings from HP network storage services
- **Application integration** — Delivers an application-aware storage approach that allows greater responsiveness to changing requirements

ENSAextended comprises:

- Storage products
- Servers
- Network infrastructure
- Linking devices (such as bridges, switches, and hubs)
- Additional components that administer and manage the entire environment

These elements communicate and cooperate to solve storage and data management problems. Customers see only a reliable, available utility that dispenses storage on demand.

A SAN is a dedicated, high-speed network that establishes a direct connection between storage elements and servers. A SAN functions as a separate high-speed network similar to a LAN and establishes a direct connection to storage resources—typically large RAID systems, robotic tape libraries, servers, or workstations. The SAN acts as an extended storage bus and uses the same networking elements as a LAN—including routers, hubs, switches, and gateways. These elements eliminate the distance limitations of traditional bus interfaces such as SCSI.

SAN technology also enables storage resources to be externalized from a single server and shared among multiple servers without affecting performance on the primary data network.

ENSAextended incorporates SAN technologies, but extends further than conventional SAN implementations in its intelligence, geographic distribution, and management. The ENSA concept of a single storage pool simplifies storage acquisition and management because it allows the storage devices to be physically near the applications they serve and retains the desirable attributes of centralized management.

## ENSAextended architecture

The ENSAextended architecture integrates open standards and industry-standard approaches toward managing storage by taking a portal approach. Web browsers provide access to applications, enabling execution from convenient locations.

The key elements of ENSAextended are:

- **Data path** — Is directly involved in moving data from storage to applications
- **Management path** — Includes services that manage data delivery and the policies and rules that govern automated management functions

The features of an ENSAextended architecture encompass these areas:

- **Storage interconnect technologies** — Include DAS and network attached storage (NAS), focusing on networking technologies such as Fibre Channel
- **Resilient technologies** — Include self-diagnosing, self-healing, and management capabilities for key components
- **Virtualization** — Encompasses the utility storage environment and its multiple applications, such as data replication and logical unit virtualization; applied at both the block and file levels
- **Data services** — Includes data replication, migration, and point-in-time copying
- **Life cycle data management** — Involves automated storage management processes that manage data placement and protection through the information life cycle
- **Active Intelligent Management (AIM)** — Includes storage provisioning technologies that enable a self-service model for delivering storage to users

## ENSAextended capabilities

ENSAextended features intelligent management capabilities that are broadly categorized as:

- **Integrated storage resource management** — Includes configuration and monitoring functionalities and storage provisioning
- **Availability and access management** — Securely present storage to authorized hosts, create and manage data paths through a SAN, and manage data recovery from point-in-time copies and other services

This architecture enables:

- **Transparent presentation of data to applications** — Uses block and file services and Fibre Channel and IP transport mechanisms
- **Rapid data protection** — Uses point-in-time copy services
- **Remote data protection and migration through a SAN** — Uses data movement and data placement services
- **Data access and sharing** — Uses distributed file systems to share data among clients
- **Tracking and management** — Uses attribution services to monitor information according to business-relevant policies

## DtS technology

DAS solutions consist of dedicated storage connected to a server that enables access and resource allocation. Storage access is bounded by the limits of the storage controller. Applications that use DAS usually do not take advantage of shared storage as provided by a SAN.

One of the most important factors contributing to the trend of moving from DAS to SAN (DtS) solutions is consolidation. Consolidation can take many forms:

- Centralization of an organization, department, or resource
- Physical consolidation of an organization, department, or resource
- Integration of data resulting from acquisitions, reorganizations, company initiatives, development of new capabilities, and so forth
- Integration of applications for many of the same reasons
- Consolidation of storage to save costs, make data more available, and so forth

Centralization can be either physical or logical:

- Physical centralization means all storage is located in the same place.
- Logical centralization means that storage seems to be centralized because dispersed storage resources are accessible through a SAN.

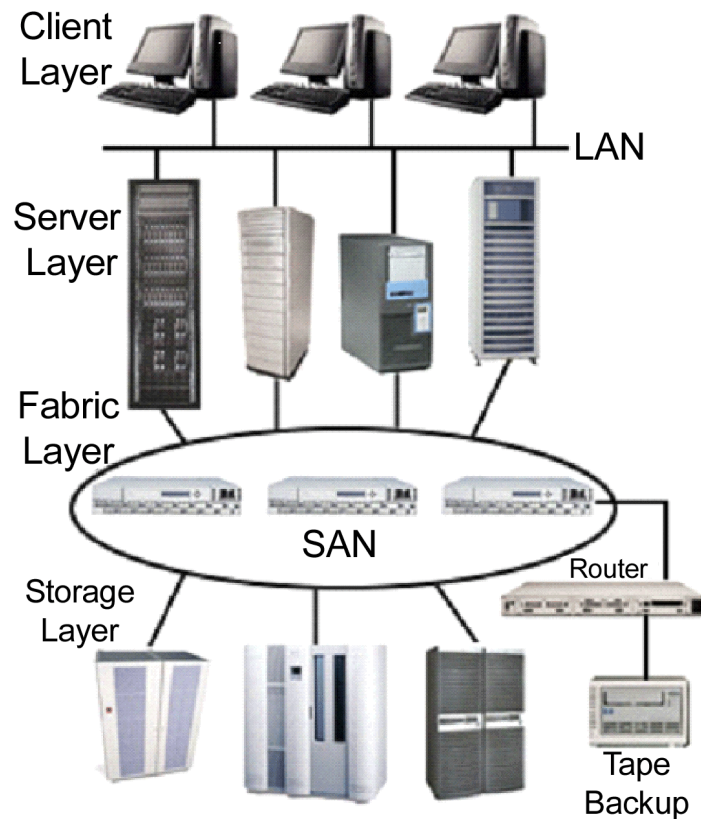
Customers who physically consolidate storage in a single enclosure, or logically consolidate with a SAN, can more efficiently configure, manage, and maintain storage resources.

Statistics show that in 2001, centralization, physical consolidation, and storage consolidation accounted for 85% of all server consolidation efforts and that the number is increasing.

Whatever form it takes, consolidation is driven by expected operational improvements and decreases in the cost of server and storage systems. Both of these expectations can be addressed most effectively by SAN solutions.

The HP StorageWorks Modular Smart Array (MSA) 1000 is the only storage system in the industry that protects a company's investment in disk drives and enables a cost-effective migration to SANs by offering DtS capability. This feature enables customers to easily migrate drives, including the associated data, directly out of an HP ProLiant DAS environment to the MSA1000.

## Storage area networks



SAN topography

A SAN is a high-speed (1Gb/s or 2Gb/s data transfer rates, with a future of 10Gb/s) network with heterogeneous servers accessing a common or shared pool of heterogeneous storage devices. SAN storage is external and can be shared without impacting system performance or the primary network.

SANs provide a secure implementation of storage I/O methods over transports through a scalable network that:

- Connects to servers throughout the enterprise
- Centralizes data
- Moves data automatically for availability and replication
- Provides accelerated data access
- Supports advanced storage management
- Enables highly available configurations
- Provides a natural platform for server clustering applications

---

### INTERNET

To learn more about SANs, take one of the HP storage and SAN courses. A list of curricula is available from: [www.education.hp.com](http://www.education.hp.com)

---

## SAN benefits

SANs provide high return on investment and reduce the total cost of ownership by increasing performance, manageability, and scalability. The benefits do not always justify the costs for small companies. As a general rule, if an enterprise has more than 16 servers, a SAN should be beneficial.

Some key benefits of SANs are:

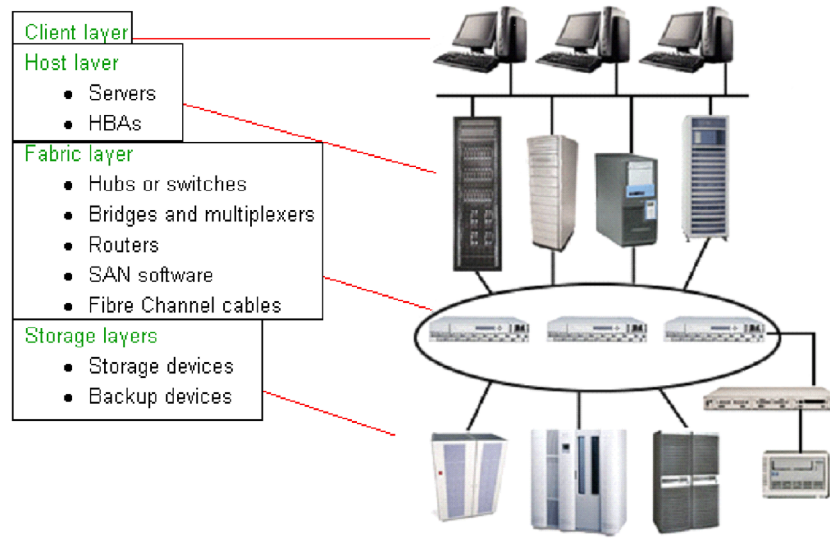
- **Reduced data center rack and floor space** — Because you do not need to buy big servers with room for many disks, you can buy fewer, smaller servers that take up less room in the data center.
- **Disaster recovery capabilities** — SAN devices can mirror the data on the disk to another location.
- **Increased I/O performance** — SANs operate faster than internal drives or devices attached to a LAN.

Although a SAN typically references Fibre Channel, it can be based on other technologies such as SCSI and iSCSI, which is a protocol that enables a server to access storage through an Ethernet network adapter.

Fibre Channel is the current industry-standard interconnect and high-performance serial I/O protocol that significantly improves the limitations of data transmission, storage capacity, and cabling distance associated with SCSI. Benefits that Fibre Channel technology provides include:

- **Distance** — Extended distances between the controller and the drives and between the host and storage system
- **Speed** — Gigabit speeds for high bandwidth applications
- **Efficiency** — Reliability and nondisruptive scalability
- **Manageability** — More devices supported by fewer people
- **Connectivity** — Any-to-any connections between servers and storage resources, including multiple paths
- **Cost-effectiveness** — Serverless backups and tape library sharing
- **Modular scalability** — Dynamic capacity
- **Consolidated storage** — Sharing of centralized storage

## SAN components



Components within the SAN topography

The components of a SAN infrastructure are:

- **Client layer** — The clients are the access point of a SAN.
- **Server layer** — The major components in this layer are the servers, the host bus adapters (HBAs) including the Gigabit Interface Converters (GBICs), and the software drivers that enable HBAs to communicate with the fabric layer.
- **Fabric layer** — This is the middle networking layer of a SAN, where hubs and switches bring all the cables together into a logical and physical network.
- **Storage layer** — This is where all the data resides on the disk drives.

The final component is the management layer, which organizes the connections, storage elements, and computer systems so that data transfer is secure and robust.

Fibre Channel SAN environments and components enable the development of solutions that provide high performance and high availability, which are the fundamental requirements of a storage network. Fibre Channel devices effectively combat the problems related to bandwidth, which generally occur during bulky operations, such as backup and restore operations.

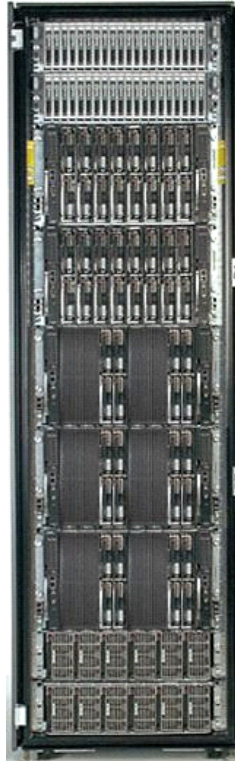
---

### INTERNET

For detailed information on designing a SAN, download the SAN Design Guide available from: <http://h18006.www1.hp.com/products/storageworks/san/documentation.html>

---

## Server layer



ProLiant BL server blades in a full 42U rack

After the client layer, the first SAN component that enhances performance is the SAN server. The SAN server:

- Acts as the access point for the clients
- Provides load balancing and data caching to improve performance
- Schedules backups

With a SAN, you can purchase fewer and smaller servers than with other topologies that rely on the servers to manage data. Because data is moved from the servers to storage devices, the servers can handle the server tasks more efficiently.

SANs support server clustering, which features a set of independent network servers working together to provide fault tolerance. The services, applications, and resources running on any node in the cluster are available to all connected network users. Clusters are invisible to users and interact as if each cluster were a single server.

HP offers a variety of hardware components with features to provide a range of solutions, from a small SAN to a high-speed, high-volume SAN data center. ProLiant servers supported in a SAN configuration include the ProLiant ML, DL, and BL lines.

## ProLiant BL server power requirements

The ProLiant BL20p G2 server blade delivers optional Fibre Channel support for SAN implementations and clustering capabilities. Through the implementation of a SAN with the ProLiant BL20p G2 server blade, customers can achieve improved data availability, scalability, and economy from consolidating disk resources. The ProLiant BL20p G2 server blade is optimized for HP StorageWorks and compatible with select EMC and Hitachi SANs.

When deploying ProLiant BL p-Class server blades in a high-density cluster, you must consider the proper power infrastructure guidelines carefully. The BL p-Class power infrastructure supports the following types of power:

- **Single-phase** — Can support half of a 42U rack or six ProLiant BL40p server blades.
- **Three-phase power** — Can support a full 42U rack of ProLiant BL server blades.
- **48V facility DC power** — Can support a full 42U rack of ProLiant BL server blades. The ProLiant BL p-Class power supplies are designed to convert AC input power to -48V DC power to the 6U server enclosures by using the bus bars. A customer can, however, connect the -48V ( $\pm 10\%$ ) facility DC power directly to the bus bars without using the power enclosures by purchasing a facility DC option kit to make the connection.

The ProLiant BL p-Class architecture supports dual power enclosures. Each 3U power enclosure is attached to two power circuits to enable redundant power feeds. Each 3U power enclosure must be able to support the full load of the rack in case of AC failure. Therefore, to ensure redundancy, the draw on each power feed must be limited to half its maximum power rating.

Graphical power calculators are available for a variety of ProLiant servers. Use the ProLiant power calculators to:

- Determine the power and cooling requirements for any server configuration
- Review the server loading and identify the number of power supplies required for redundancy
- Approximate the electrical and heat load per server to assist facilities planning

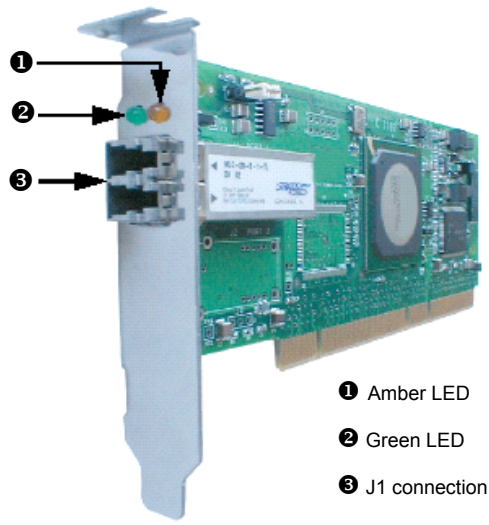
The ProLiant power calculators provide power, cooling, weight, and configuration information. The tools are designed to help you make decisions for single-phase and three-phase requirements based on your specific configurations.

---

**INTERNET** You can download the ProLiant power calculators from:  
<http://h18001.www1.hp.com/partners/microsoft/utilities/power.html>

---

## Host bus adapters



FCA2214 HBA

An HBA is a high-performance I/O solution for applications such as client/server configurations, database I/O environments, multimedia applications, and imaging technologies. Each of these environments requires high-level throughput with low-latency characteristics.

Each HP HBA ships with a unique address identifier that is stored in flash memory. Fibre Channel industry standards issue two unique identifiers:

- WorldWide Port Name (WWPN)
- Node Name (NN)

Each of these identifiers is derived from the Institute of Electrical and Electronic Engineers (IEEE) address assigned to the HBA. Combined, the WWPN and NN create the WorldWide Name (WWN), which is an 8-byte field that uniquely identifies an HBA on a Fibre Channel circuit.

---

**Note**

The WWN address and serial number are clearly marked on the HBA. HP recommends that you record the addresses for future reference.

---

## HP StorageWorks FCA2214/FCA2214DC HBAs

The HP StorageWorks FCA2214/FCA2214DC HBAs are 2Gb, high-performance Direct Memory Access (DMA) bus master host adapters designed for high-end systems. The performance is derived from ISP2312 chips.

The FCA2214 HBA is a 64-bit PCI device that functions in a 32-bit slot. It must be in a PCI bus master slot and it does not work in shared slots. It is designed and tested to operate at PCI bus speeds up to 133MHz.

The FCA2214DC is a dual-port PCI-X to Fibre Channel HBA for Microsoft Windows 2000, Windows Server 2003, and Linux that complements the single-channel FCA2214 product. It provides dual-channel connectivity for StorageWorks storage arrays.

---

### Note

The maximum number of HBAs supported per server is dependent on the server model.

---

These HBAs:

- Combine a Reduced Instruction Set Computing (RISC) processor, a Fibre Channel protocol manager with 2Gb Fibre Channel transceivers, and a PCI or PCI-X local bus interface in a single-chip solution.
- Use a bios utility to customize the configuration parameters on the adapter and attached drives.
- Support Fibre Channel SCSI, IP, and Virtual Interface protocols.
- Support point-to-point fabric connections.

---

### Note

This HBA does not support booting from a SAN on a Linux platform.

---

## Software drivers

After you install the FCA2214/FCA2214DC HBA in the system, you must install the driver.

---

**INTERNET** The software kit included with the HBA contains the latest version of the software files at the time of shipment. Software files are updated periodically and can be obtained from the HP website:  
**<http://www.hp.com/country/us/eng/prodserv/storage.html>**

---

The appropriate driver depends on the operating system:

- **Windows** — Windows 2000 and Windows Server 2003 require the hp2300.sys driver. The FCA2214/FCA2214DC HBAs are plug-and-play devices that are automatically detected by Windows operating systems.

---

### Note

For the FCA2214DC, the Windows Device Manager detects and displays two instances of the HBA. You must install the driver for each instance.

---

- **Linux** — To install the FCA2214 HBA drivers you must:
  - Patch the Linux kernel.
  - Install the qla2300 driver from the provided Red hat Package Manager (RPM) file. The driver must be loaded before the system can access the devices attached to the FCA2214/FCA2214DC HBA.

Because this process involves building a new kernel for the storage system, HP recommends performing this task during inactive periods. Depending on the kernel configuration, this process might take an hour or more. Some third-party drivers might require reinstallation on the new kernel.

- **NetWare** — Novell NetWare Cluster Service supports the FCA2210 2Gb Fibre Channel HBA. Review the installation documentation that shipped with the HBA or the Quick Install Guide (MSA1000) for detailed instructions.

## Fabric layer

SAN components in the fabric layer include:

- Hubs
- Switches
- Bridges

### Hubs

External hubs are ideal for data marts, web servers, and server farms that require scalability and bandwidth for growing storage needs. Use the following hubs depending on your application:

- **Fibre Channel Storage Hub 7** — Use this hub if low total solution cost and simplicity are key requirements.
- **Fibre Channel Storage Hub 12** — Use this hub if maximum capacity and manageability are required.

The MSA Hub 2/3 is an optional, hot-pluggable I/O device designed to replace the standard single-port I/O module on the MSA1000. Use this hub to provide inexpensive access to the storage controller without using an external hub or switch. Two small form-factor pluggable (SFP) ports are external and a third is an internal port that accesses the controller.

### Loop switches

Fibre Channel Arbitrated Loop (FC-AL) switches are low-port-count, low-bandwidth products. HP offers related products that act as loop-switching hubs and fabric-attached switches. These switches provide connectivity between attached FC-AL devices and between FC-AL devices and switched fabric elements.

This connectivity enables low-cost or low-bandwidth workgroup (edge) devices to communicate with fabric devices (servers, storage devices, or other peripherals), and ultimately to be incorporated into an enterprise SAN environment.

### Fabric switches

A set of interconnected switches is called a *Fibre Channel fabric*. Each fabric has ports into which several computer servers, storage systems, and related components can be integrated. Multiple fabrics can be included in a single SAN if needed to meet connectivity or availability requirements.

HP Fibre Channel switches have multiple ports and can be interconnected across long distances to achieve large network configurations.

## Brocade switches

HP has partnered with Brocade Communications Systems, Inc. to simplify the management of heterogeneous SAN environments. You can easily manage Brocade switches across multiple fabrics for common tasks such as:

- Configuration management
- Firmware download
- License management
- Sequenced reboot
- End-to-end performance monitoring
- Real-time health monitoring
- Zone management

Brocade offers an intelligent platform for a networking foundation for SAN solutions. The Brocade family of fabric switches and software is designed to optimize data availability and storage and server resources in the enterprise.

## Layer 3 switches

Layer 3 switches provide rapid IP routing, which dramatically improves network throughput. Working with existing routers, they enable enterprises to off-load IP routing from a router to an HP Enterprise Gigabit Switch and realize a significant performance increase in data rates through the network. Benefits of Layer 3 switches include:

- **Quality of service (QOS)** — Enables you to differentiate network traffic types and dedicate bandwidth to them, so you can shape network traffic and provide needed bandwidth to business-critical applications.
- **Hot-swappable GBIC ports** — Enable you to configure 1000Base-SX or LX ports as needed for a growing network.
- **Larger backplane** — Provides sufficient bandwidth to support non-blocking operation, even with all ports configured for full duplex. Up to 4MB of dynamically assigned buffer memory ensures that packet loss is prevented even during peak traffic conditions.

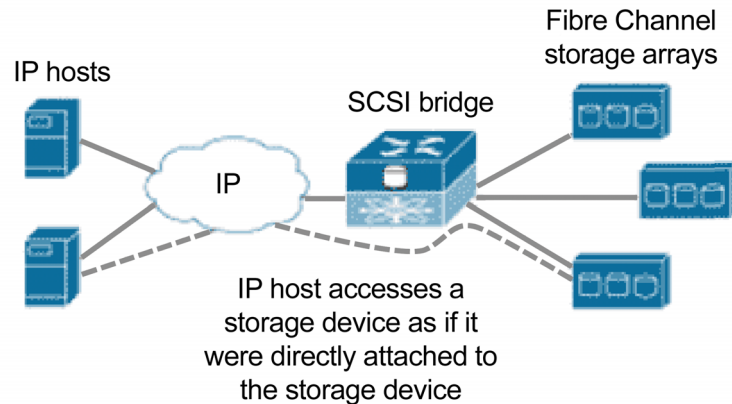
---

**INTERNET**

For more information regarding the switches HP supports in a SAN, visit:  
<http://h18006.www1.hp.com/storage/saninfrastructure.html>

---

## SCSI bridges



A SCSI bridge provides universal access to storage over IP networks. The storage bridge software controls the operation.

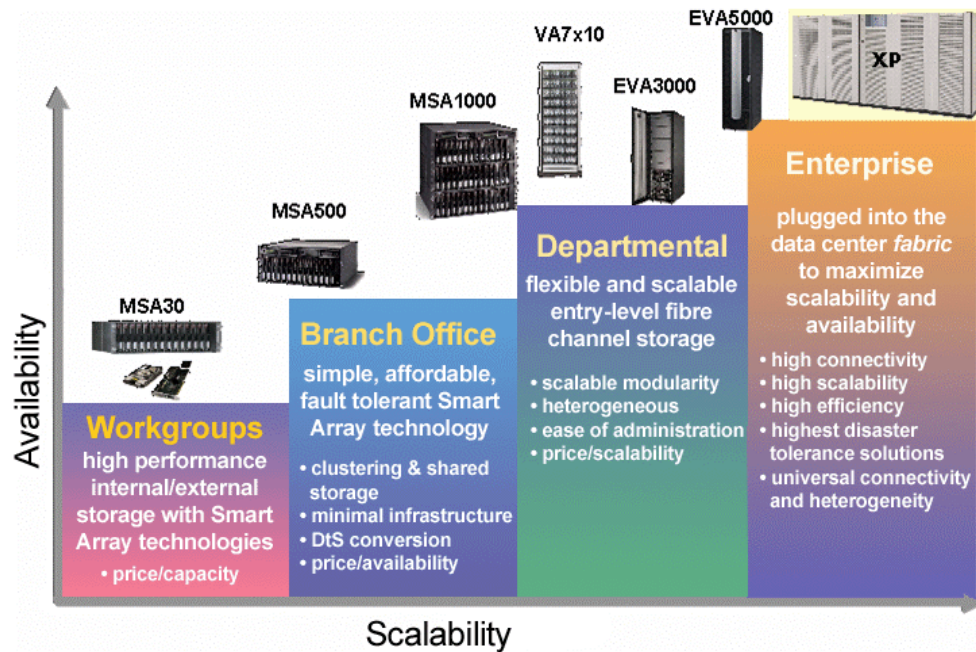
SCSI routing provides IP hosts with access to Fibre Channel storage devices as if the storage devices were directly attached to the hosts, with access to devices being managed primarily in the SCSI bridge. With SCSI routing, storage devices are not aware of each IP host—the storage devices are aware of the bridge and respond to it as if it were one Fibre Channel host.

An iSCSI target is an arbitrary name for a group of physical storage devices. The iSCSI targets are created and mapped to physical storage devices attached to the bridge. The bridge presents the iSCSI targets to IP hosts as if the physical storage devices were directly attached to the hosts.

Two types of access to storage over IP networks are available:

- **SCSI routing** — Provides IP hosts with access to Fibre Channel storage devices using the iSCSI protocol.
- **Transparent SCSI routing** — Provides IP hosts with transparent access to intelligent storage arrays using the iSCSI protocol. Each IP host is presented as a Fibre Channel host to an intelligent storage array. With transparent SCSI routing, availability of storage devices is managed primarily in the intelligent storage array.

## Storage layer



HP StorageWorks product family

HP StorageWorks products offer additional features that increase IT efficiency and stability, which contribute to overall business performance.

HP StorageWorks products provide:

- Stability using intelligent tape libraries
- Storage solutions for small, medium, and enterprise businesses
- Efficiency through the use of storage area management
- Choices in storage networking
- Business value with enhanced array functionality

**INTERNET**

For a complete list of enterprise StorageWorks products, visit:  
<http://h18006.www1.hp.com/storage/enterprisestorage.html>

StorageWorks subsystems and components include:

- **MSA500** — The MSA500 is an Ultra3 SCSI storage enclosure designed for single and multiple path two-node clustering and multiple node (up to four servers) shared storage. It provides data protection and increased uptime with redundant controllers, power supplies, and fans.

Because the MSA500 is SCSI-based, there are no Fibre Channel requirements such as hubs, switches, and cables, which results in a significantly lower initial investment when compared to Fibre Channel SANs. It also permits seamless conversion from SCSI to the Fibre Channel-based MSA1000.

When storage needs expand to require even greater capacity or attaching more than four servers, the MSA500 is easily convertible to a SAN because it is based on DtS technology. The MSA500 can be converted to a SAN by replacing the controllers and I/O module and installing a Fibre Channel hub or switch, providing the customer with extended hardware investment protection into the future.

- **MSA1000** — The MSA1000 provides a high-performance array controller in a 14-drive storage cabinet. These enclosures can be expanded to two additional storage cabinets for a total capacity of 42 Ultra 2 or Ultra3 SCSI drives. The MSA1000 is compatible with 1Gb/s and 2Gb/s Fibre Channel hub and fabric switch interconnects and includes options for embedded switches and hubs. Complete operating redundancy is supported, including redundancy when expanding disk and unit volumes.

The MSA1000 can operate in stand-alone, dual-node, or multinode cluster environments supporting multiple operating systems, including Windows, Novell NetWare, Linux, HP OpenVMS, or Tru64 UNIX.

The following software components are supported on the MSA1000:

- Array Configuration Utility (ACU)
- Command line interface (CLI)
- HP Insight Manager 7

- **SCSI adapters** — HP StorageWorks SCSI adapters provide support for up to 28 SCSI devices, with the adapter taking up one PCI slot. HP adapters feature one or two separate Wide Ultra3 SCSI channels that each provide up to 64-bit/66MHz bus-master capabilities, with a maximum data transfer rate of 160MB/s between storage subsystem and system memory. HP SCSI adapters can be used with all ProLiant servers and management software.
- **Smart Array controllers** — Smart Array controllers such as the HP Smart Array 4200 Ultra2 SCSI array controller offers ProLiant customers up to four times the processor performance of other array controllers. This increased performance is achieved with the introduction of a 64-bit architecture that provides more processing horsepower and bandwidth.
- **Universal hard drives** — Universal drive commonality across the full range of HP platforms provides customer benefits that include:
  - A maximum data transfer rate up to 320MB/s.
  - Compact one-inch design for maximum storage density.
  - Backward compatibility.
  - High data integrity and reliability in the enterprise environment.

---

**INTERNET**

For more information on StorageWorks products, visit:  
<http://h18006.www1.hp.com/storage/index.html>

---

## HP SAN solution service offerings

SAN installation, startup, and integration call for specialized expertise in planning and design, hardware configuration, network software installation, and operability verification. A SAN must be configured for high performance and availability and should be integrated with business applications, processes, and management environment. The HP SAN Solution Service encompasses all the activities required for a successful implementation.

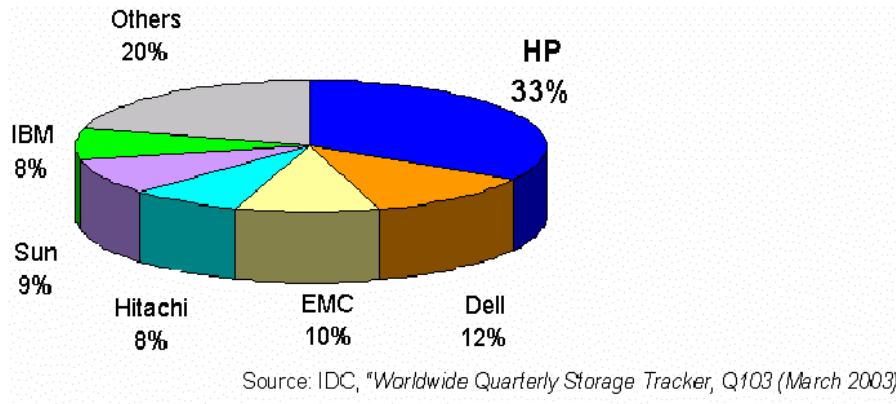
The SAN Solution Service provides three service levels:

- **Level 1: Basic SAN deployment** — Covers entry-level array solutions with:
  - Comprehensive SAN implementation strategy and planning
  - Hardware and software installation and configuration
  - Integration of the installed disk array in the SAN environment
  - Firmware verification
  - SAN testing
  - Knowledge transfer and solution handover
- **Level 2: Heterogeneous SAN integration** — Includes all Level 1 features, plus integration into a complex, heterogeneous storage environment:
  - Project management — A dedicated project manager determines any special requirements and customizes a statement of work.
  - Environment analysis — Current storage and related system configuration and operating system solutions are examined.
  - SAN architecture design — An HP consultant determines an optimal SAN design based on the system architecture and configurations, network and data center environment, and business requirements.
- **Level 3: Integration of applications** — Includes all Level 2 features, plus:
  - Smooth integration of key business applications such as SAP or Oracle into the SAN-based infrastructure
  - Review and testing of the configuration and functionality

HP Services also can provide ongoing support. Service options include:

- Access to HP Technical Support Engineers
- On-site services to resolve technical problems
- Advocacy within HP through a technical account manager
- Installation of software patches and bug fixes
- System maintenance of topologies and account profiles to expedite problem resolution

## HP position in the SAN marketplace



Worldwide open SAN market share (by units) in 2002

According to IDC, a leading international data consulting and research company:

- HP sold more production SANs in 2002 than its three nearest competitors combined, accounting for 33% of the market.
- HP was first in disk storage systems revenue for 2002 with 27% market share.
- In 2002, HP became the number-one vendor in external RAID storage revenues with 21% market share.
- HP continues to be the largest supplier of SAN-attached external RAID arrays in the world.

HP maintained the number one position in worldwide disk storage systems by factory revenues, and posted year-over-year gains in several storage segments for the second quarter of 2003.

In networked disk storage systems, StorageWorks led the open SAN market for the fourth consecutive quarter and gained three percentage points of market share. In addition, HP StorageWorks products held the number one position in market share by factory revenue in total external storage and external RAID, excluding NAS.

HP StorageWorks products also led in worldwide disk storage systems factory revenue with 26.7% revenue share and led the total market for external disk storage systems with a 21.5% revenue share.

## Market opportunities

SANs are the ideal solution for companies with the following characteristics:

- Three or more servers, or more than 100GB of storage
- Distributed storage systems
- Management or resource consolidation

SANs add business value by providing:

- Operational savings
  - Lower storage management costs
  - Better storage utilization
  - Lower backup costs
- Lower total cost of ownership
  - Management centralization and simplification
  - Increased data availability
  - Ease of growth
- Better business flexibility because storage or servers can be added easily to accommodate growth without disruption
- Service level improvements resulting from improved availability

Despite the slowdown in IT spending and decrease in IT budgets, customers continue to consolidate server and storage systems. This represents an important opportunity for the sale of HP SAN solutions.

## HP target markets

Customers are trying to control the volume of data created by new applications and are trying to lower IT management costs. They are also looking for new revenue opportunities and are trying to cope with the requirements of growing businesses. Those high-level needs can be translated into vertical and horizontal market segments.

### Vertical markets

- **Healthcare** — HP provides competitive solutions for medical image management and picture archival communication systems. In partnership with world-class IT associates, HP provides real-time diagnostic image systems and storage, retrieval, and communications systems for archived images, reports, and exam results.
- **Financial services** — The most pressing need in regulatory messaging compliance is email archival. Forecasts predict that worldwide corporate email traffic will more than double by 2006, requiring up to 8.6MB of storage per user per day. New government regulations are requiring brokers, dealers, and other financial firms to capture, index, archive, search, and retrieve their email communications.
- **Life sciences** — The 21st century promises to transform medicine and human life with life science innovations. HP infrastructure solutions powered the first major milestone in biosciences—mapping the first human genome.

Generating vast amounts of complex data is necessary when performing research and development. Fortunately, information measured in terabytes, not gigabytes, (with multi-petabytes soon to be a reality), can be stored, protected, secured, organized, distributed, and audited without interruption.

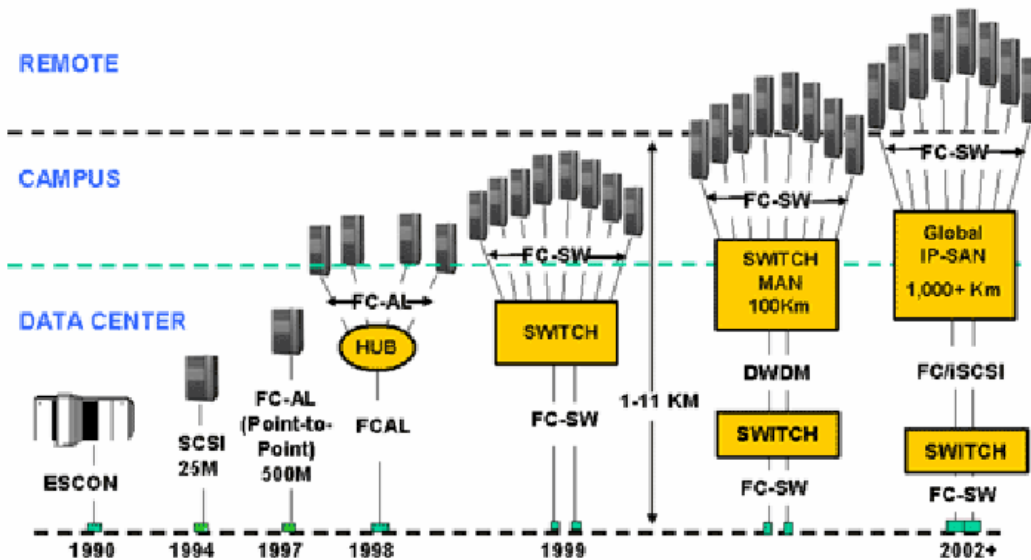
## Horizontal markets

Microsoft Exchange Server is designed to meet the messaging and collaboration needs from small organizations to large distributed enterprises. Exchange Server contains an enhanced platform of Microsoft Web Storage System and scalable, reliable, and easy-to-manage messaging and collaboration applications.

HP StorageWorks delivers stability through a variety of solutions:

- **HP Enterprise Tape Library Architecture** — HP adds reliability to the SAN with embedded automated intelligence in the Enterprise Storage Library (ESL) tape libraries. These libraries can perform backups quickly and efficiently.
- **Increased storage solutions** — HP provides improved IT efficiency and stability solutions for small and medium businesses and entry-level storage environments, including the NAS2000s, the DAT 24 and DAT 40 tape drives, and HP OpenView Storage Mirroring.
- **Greater efficiency with storage area management** — HP extended the HP OpenView Storage Area Manager to Oracle and Microsoft applications.
- **More storage networking options** — HP StorageWorks offers open heterogeneous storage with SAN fabric products:
  - HP StorageWorks M-series and B-series switch interoperability
  - C-Series products
  - Expanded HP OpenView Continuous Access Storage Appliance (CASA) support
- **Enhanced array functionality** — HP has enhanced functionality of the HP StorageWorks Enterprise Virtual Arrays (EVA3000, EVA5000, and EVA software), providing greater flexibility, improved availability, expanded business continuance, and enhanced management.

## Planning and designing a SAN solution



SANs pose challenges and raise questions about how the SAN will interoperate with an existing LAN and WAN. IT managers must decide which data and applications warrant the added cost and complexity of a SAN, and which storage should remain on server-attached RAID or Just a Bunch of Disks (JBOD) configuration.

Industry analysts target storage-intensive and mission-critical applications as the most obvious candidates for SAN migration. Because of advantages in performance, availability, and data protection, the following applications benefit most from migrating to storage hosted on a SAN:

- Transaction processing
- Email
- Groupware
- Enterprise resource planning
- Multimedia file serving
- Database acceleration

**INTERNET** For more information on planning a SAN, refer to the *SAN Design Guide* available from: <http://h18004.www1.hp.com/products/storageworks/san/documentation.html>

## Configuring a SAN

After deciding which applications to migrate, you must decide how to partition the storage arrays that are served by the SAN. User requirements should drive SAN partitioning, which can be determined by analyzing past and current storage consumption and use patterns.

Typical configuration planning questions include:

- Which servers (and workstations) will have access to which SAN partitions?
- What kinds of user activity can be expected?
- How much capacity should each SAN partition have?
- Which SAN partitions are shared by more than one server for failover?
- What is the ratio of application files to data files?

---

### Note

In a SAN with multiple servers, HP recommends designating one of the servers as a management server to centralize your management tasks.

---

## Documenting the SAN

HP recommends, at a minimum, documenting the following before beginning the actual SAN implementation:

- **Topology map** — Shows the logical SAN topology and fabric interconnect scheme. This map conveys the overall design from a strategic standpoint and can also convey how future growth and technological advances will be accommodated.
- **Configuration layout** — Shows the physical layout of the entire implementation. It is more detailed than the topology map. The layout is used during implementation to verify the correct connectivity. The layout is also helpful if troubleshooting is required in later phases.
- **Storage map** — Defines the storage system arrangement and configuration in the SAN as well as storageset settings such as Selective Storage Presentation (SSP) and RAID levels. This map effectively defines how all the storage is configured in the SAN.
- **Zoning map** — Defines the internode communication access within the SAN. This map defines which nodes or device ports are allowed to communicate with each other in the SAN.

---

### Note

A key decision in the zoning implementation process is determining whether hardware or software zoning will be implemented. Hardware zoning offers higher security than software zoning but is less flexible because device-to-switch cabling changes require zoning information to be updated.

---

## Designing for availability

To provide high device availability, critical servers, storage devices, or applications should be connected to more than one fabric element or to more than one fabric. To determine if dual-connection capability exists for a device, refer to the associated device documentation. To provide high fabric availability, consider the use of multiple fabric elements, multiple initial system loads, or redundant fabrics.

Many fabric-attached devices require highly available connectivity to support applications such as:

- Disk mirroring
- Server clustering
- Business continuance operations

High availability is accomplished by deploying a resilient fabric topology or redundant fabrics. A fabric topology that provides at least two internal routes between fabric elements is considered resilient. A single device failure does not affect the remaining elements and the overall fabric remains operational.

However, unforeseen events such as human error, software failure, or disaster can cause the failure of a single resilient fabric. Using redundant fabrics (with resiliency) mitigates these effects and significantly increases fabric availability.

## Designing SAN management

SAN topologies present new management challenges for IT organizations that are already spending a great deal of time, energy, and money managing server-attached storage. Automated tools that can deliver storage resource management features in a network context ensure the ongoing health and availability of a SAN.

Features to look for in SAN management tools are:

- Automated polling intervals
- Partition space thresholds
- Asset and configuration management
- Alarm/alert generation
- User consumption monitoring

## Configuration utilities

Use the following configuration utilities to implement SAN components:

- Fast!Util
- CLI

### Fast!Util

Use Fast!Util to configure the FCA2214/FCA2214DC HBA and connected devices. You can access Fast!UTIL, a BIOS-resident configuration tool, by pressing *ALT+Q* during the HBA bios initialization.

---

**INTERNET** For more information on StorageWorks products, visit:  
<http://www.hp.com/country/us/eng/prodserv/storage.html>

---

### Command line interface

Use the CLI to configure, manage, and monitor all aspects of the MSA1000, including array configuration. You can also use the CLI to display system setup information and status and to provide information on devices that are attached to the controller. You can access the CLI through a host computer connected to the customized RJ-45Z serial port on the front of the MSA1000.

---

#### Note

All supported operating systems can access the CLI.

---

Some CLI configuration and management tasks include:

- Configuring storage units (logical unit numbers [LUNs])
- Setting the addressing mode
- Limiting access to storage
- Viewing information on MSA1000 components
  - Controller
  - Unit
  - Cache

---

#### Note

If you plan to use the CLI to configure and manage your storage, HP recommends using the CLI exclusively and not the ACU.

---

---

**INTERNET** For more information on the MSA1000 CLI, visit:  
[http://h20015.www2.hp.com/hub\\_search/document.jhtml?lc=en&docName=c00032137#N1000B](http://h20015.www2.hp.com/hub_search/document.jhtml?lc=en&docName=c00032137#N1000B)

---

## Validating the design

StorageWorks blueprints detail pretested SAN configurations designed for particular environments and outline how to build the SAN (including a full bill of materials). Blueprint configurations have been developed for entry-level, middle-tier, and enterprise-wide applications.

---

**INTERNET**

For links to all the HP StorageWorks solution blueprints, visit:  
[www.hp.com/go/hpstorage\\_blueprints](http://www.hp.com/go/hpstorage_blueprints)

---

## Using HP sizing and planning tools

The HP Enterprise Configurator helps you configure a wide range of systems, including servers and storage systems. By answering a series of guided questions with recommended answers, you can configure a system quickly, complete with a bill of materials and rack diagrams.

Currently, the configurator provides configuration support for ProLiant servers. The features include:

- **Quoting Tool** — Export a consolidated parts list (with formulas) to a Microsoft Excel spreadsheet, which can be saved and used as a quoting tool by applying discounts on the line items or the total price.
- **Proposal Builder** — Export the configuration summary to a Microsoft Word document, which can be saved and used to build proposals.
- **Power Calculators** — Use to review the server loading to determine the number of power supplies required for the power supplies to be redundant. You can also use this tool to approximate the electrical and heat load per server for facilities planning.
- **Printer-Friendly Configuration Summary** — Export the configuration summary to a printer-friendly format with page breaks.

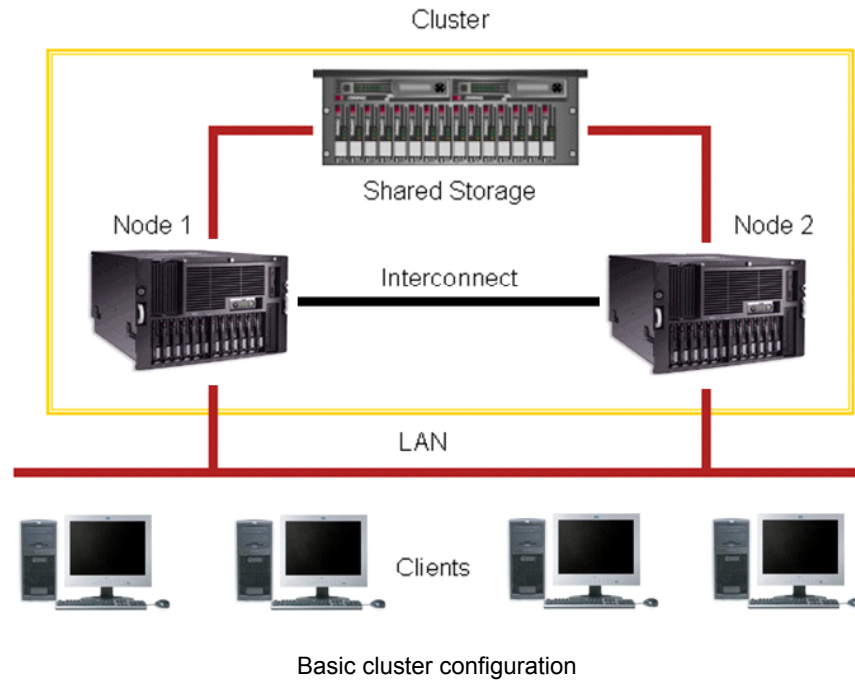
---

**INTERNET**

To use the HP Enterprise Configurator, visit:  
<http://h30099.www3.hp.com/configurator/>

---

## What is a cluster?



A cluster is a set of loosely coupled servers used as a single, unified computing resource. A cluster contains four basic elements:

- Servers, also known as *cluster nodes*
- Internode communication paths, called *interconnects* or the *private network*
- Shared physical storage
- Cluster-enabling software

A cluster offers a level of scalability and availability that exceeds that of stand-alone servers. To users, this can translate into increased performance and reliability. However, clustering should be implemented only after fault prevention and fault tolerance methods have been implemented.

Clusters keep server-based applications highly available, regardless of individual component failures. They also ensure minimal interruption of operations and can increase processing capacity and I/O bandwidth.

The services running on any system in the cluster are available to all connected network users. A client interacts with a cluster as though it were a single server and therefore might not be aware that it is a cluster.

Clustering traditionally has existed only in the realm of proprietary systems designed for mission-critical applications such as stock trading and aerospace mission control. HP clustering technology is mainstream, industry-standard computing designed to satisfy the increasing demand to keep business-critical applications available.

HP clustering solutions:

- Use industry-standard hardware and software
- Provide essential clustering features and benefits at a price lower than proprietary clustering systems
- Increase the usefulness and life span of software applications

Cluster solutions within the HP adaptive infrastructure enable organizations to:

- Adapt to changing market conditions
- Conserve valuable human and technology resources
- Respond rapidly to customers in a dynamic business environment

ProLiant clusters enable faster responses to customer demands with solutions such as embedded storage connectivity and no-single-point-of-failure configurations. These solutions ensure maximum reliability for enterprise customers, reducing the need for valuable, dedicated IT resources.

## Cluster-aware applications

A cluster-aware application recognizes that it is being installed on a cluster and creates the resources necessary for clustering. This makes clustering the application easier than if the application is non-cluster-aware. A non-cluster-aware application must be manually configured for failover.

In addition to being simple to install, a cluster-aware application is easier to manage and generally can recover from more faults than a non-cluster-aware application.

## Cluster models

There are three basic cluster models:

- **Shared-nothing** — Nodes in this model have access to a shared physical storage system, but the nodes cannot access the same logical drives at the same time.
- **Shared-disk** — A distributed lock manager is required to allow more than one node to access the same logical disks at the same time.
- **Shared-everything** — Memory, processors, and disks are shared by the cluster nodes.

---

### Note

A shared-everything cluster is also referred to as a *single system image* cluster. The administrator, clients, and operating system see the system as one unit. Examples of shared-everything clusters include HP NonStop Kernel clusters, HP OpenVMS clusters, and HP TruCluster systems.

---

## Advantages of clustering

Clustering provides the following advantages:

- High availability of resources
- Scalability for growth
- Centralized administration
- Load balancing

### High availability of resources

Clustering ensures the high availability of resources by reducing single points of failure in the cluster. A single point of failure is any one item whose failure would cause the service or application to become unavailable.

#### Example

One single point of failure is a direct connection from a device such as a printer or modem bank to one node of a cluster. If that node fails, the device will no longer be available.

If one node in a cluster either fails or the administrator takes it offline, the resources can fail over to a surviving node. During failover the remaining servers in the cluster automatically redistribute the tasks of the node that is down or failed. The process is transparent to clients in the network.

When a node that failed or was taken offline is ready, the resources can fail back to the original node. You can perform the failback manually or configure it to happen automatically.

## Scalability for growth

In addition to providing high availability, clusters can be highly scalable. Through clustering, the following resources can be expanded incrementally and efficiently:

- Processors
- I/O
- Storage
- Applications

Clustering provides reliable access to system resources and data as well as investment protection for all resources. By clustering existing hardware with new computers, you can protect your investment in both hardware and software. Instead of replacing a stand-alone computer with a new one of twice the capacity, you can add another computer of equal capacity.

### Example

The number of clients using an application on a server increases and performance degrades. The application is written so that the processing can be split across multiple nodes. Clustering this server with another server will improve performance and increase availability of the application.

## Centralized administration

In a typical server environment, various administrative tools identify the servers on the network and monitor their contents and activities. However, in the cluster environment, the administration of applications and services is centralized.

## Load balancing

Clusters not only provide high availability, scalability, and centralized administration, but also can provide load balancing.

### Example

A system administrator discovers that too many applications are running on one server and another server is barely being used. Ordinarily, the system administrator would have to shut down and reconfigure both systems. However, if the servers are part of a cluster group with application failover capabilities, the system administrator can manually fail over the applications to another server and balance the workload without shutting down the server.

During normal operation, most clustering software allows manual or automatic load balancing of resources. The Microsoft Cluster Service (MSCS) supports only manual load balancing.

## Determining the need for clusters

Research shows that the leading causes of downtime can be ranked in the following order:

1. Infrastructure problems (building, power, and network)
2. Software failures (operating system, application, tools, and drivers)
3. Operational and administrative activities (procedures, personnel, and maintenance activities)
4. Hardware failures (disk, power supply, and memory)

The least likely component to fail in an environment is the hardware. Downtime usually is related to events and activities outside the server. Issues such as poorly trained personnel, power or air conditioning outages, or flawed backup and restore procedures account for more downtime than events such as hard drive failures.

## Decision points

Each organization must determine its leading causes of downtime before deploying a cluster solution. The following decision points will help you determine whether clustering is right for an organization.

- **What are the availability requirements?** — Many organizations can tolerate hours or even days of downtime for their environment. Others need 99.999% availability.
- **Can availability requirements be achieved by investing in other areas?** — Availability of 99% and 99.9% can be attained without additional investment in clustering technology. HP provides many hardware and off-the-shelf technologies such as RAID arrays and redundant power supplies and fans that can increase availability without clustering. Personnel training or procedural redesign also can reduce downtime.
- **Does clustering solve the leading causes of downtime?** — After evaluating the leading causes of downtime in the environment, determine whether clustering addresses these issues. The leading causes of downtime might be operational or administrative issues, software failures, or infrastructure problems. Clustering increases availability and protects from hardware failures. If you choose to implement clustering but do not address the issues that are causing downtime, you probably will be disappointed in the results.
- **Is the increased complexity that clustering introduces worthwhile?** — To administer a cluster, you must be familiar with clustering concepts, administration tools, and procedures for failover and failback operations. Some applications are very complex. Clustering those applications drastically increases the total complexity.
- **Is investing in clustering justified by the return?** — If the environment is mission-critical, it might warrant investment in every tool available to increase availability.

## HP cluster solutions

ProLiant clusters combine industry-standard ProLiant servers with HP storage solutions. Early solutions increased data availability for multiple HP servers and allowed the servers to recover from hardware errors. ProLiant clusters are designed to recover not only from hardware errors, but also from operating system and application errors.

High-availability clusters from HP provide superior single-system manageability and easy deployment. The latest products are designed to meet customer requirements for industry-leading clustered solutions.

HP offers the following ProLiant cluster solutions:

- Windows Server 2003 and Windows 2000 Cluster Service
- HP Serviceguard for Linux
- SteelEye LifeKeeper for Linux
- Novell Cluster Services
- Oracle Parallel Server (OPS) for Oracle 8i
- Oracle 9i Real Application Clusters

In addition, HP offers the following high-end cluster solutions that run on AlphaServers:

- HP TruCluster Server software
- HP OpenVMS Clusters software

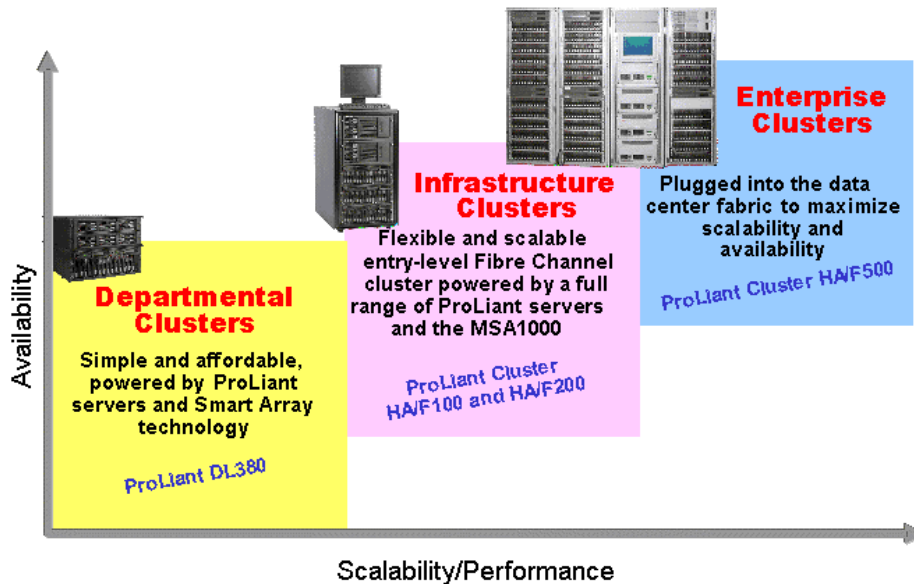
---

**INTERNET**

You can find more information on ProLiant high-availability solutions at:  
<http://h18004.www1.hp.com/solutions/enterprise/highavailability/index.html>

---

## ProLiant cluster family



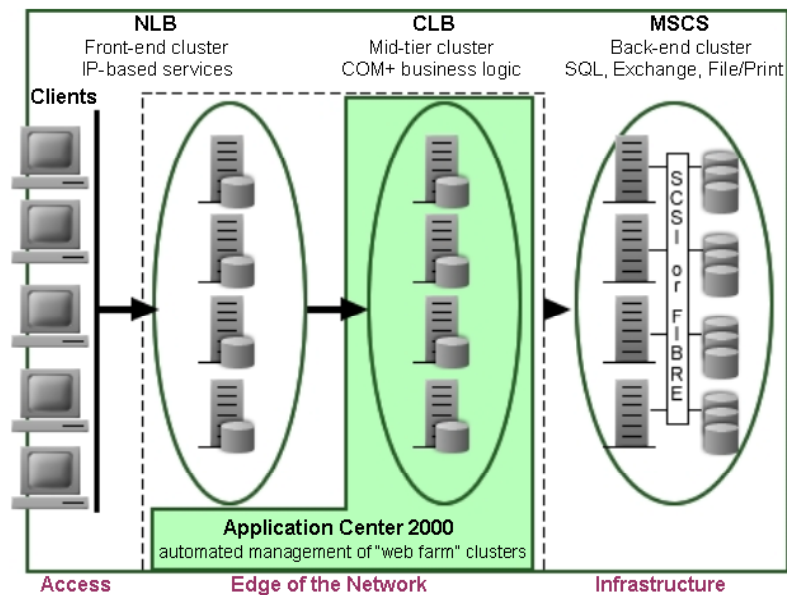
The ProLiant cluster family consists of three categories:

- **Departmental clusters** — Simple and affordable clusters powered by ProLiant servers and Smart Array technology, examples include the ProLiant DL380 G3 Packaged Clusters, which combine two ProLiant DL380 G3 servers with the MSA500.
- **Infrastructure clusters** — Flexible and scalable entry-level Fibre Channel clusters, these clusters are powered by ProLiant DL and ML servers combined with the MSA1000.
- **Enterprise clusters** — Plugged into the data center fabric to maximize scalability and availability, an example is the ProLiant Cluster HA/F500, which provides the highest levels of availability and performance.

### INTERNET

To obtain the most current list of ProLiant clustering solutions, visit the HP website at: <http://h18000.www1.hp.com/enterprise/highavailability.html>

## Windows clustering technologies



The Windows three-part cluster strategy

A highly available cluster solution enables clustered applications and related resources to move automatically from one node to another in the event of failure. Microsoft has introduced three cooperative clustering technologies that provide high availability solutions:

- **Network load balancing (NLB)** — Builds front-end clusters that distribute IP traffic. Network load balancing is built into all members of the Windows Server 2003 family.
- **Component load balancing (CLB)** — Builds middle-tier clusters to load balance Component Object Model Plus (COM+) components dynamically. Because component load balancing is not available in Windows Server 2003, detailed component load balancing information is not provided in this course.
- **MSCS** — Builds back-end clusters to provide high availability for applications and services such as databases and messaging platforms.



### Important

Do not install or activate network load balancing and MSCS on the same machine. This is not supported by Microsoft or HP. Network load balancing and MSCS must be configured on separate machines.

## Network load balancing

Environments that are not resource-intensive, such as web servers, are perfect for server farms, which are groups of multiple, independent servers. Each server can handle simultaneous connections. These servers distribute incoming requests using a load balancer, which can be either hardware- or software-based.

Hardware load balancers have the following benefits over software load balancers:

- Greater scalability
- Easier to manage
- No processor load on the application server farms

Microsoft software load balancing has these characteristics:

- All servers must be on the same network segment.
- Load balancing is performed on the same server as the application.
- No external hardware is needed, so the cost is lower than a hardware load balancer.

Network load balancing is ideal for enabling incremental scalability and availability for e-commerce web sites. It enhances scalability by distributing its client requests across multiple servers within the cluster. Network load balancing also provides high availability by automatically detecting the failure of a server and repartitioning client traffic among the remaining servers within 10 seconds, with no interruption of service.

Network load balancing enhances the availability and scalability of TCP/IP-based programs such as:

- Web servers
- FTP servers
- Streaming media servers
- Terminal Services

In addition, network load balancing seamlessly integrates into existing web server farm infrastructures.

Network load balancing is used to keep applications responsive under heavy client loads by allowing you to spread incoming requests across as many as 32 servers, thus expanding enterprise-wide services.

## **Network load balancing benefits**

The primary benefits of Windows 2003 network load balancing are:

- Ease of use
  - Allows clients to access the cluster with a logical Internet name and virtual IP address
  - Retains individual names for each computer
  - Allows preventive maintenance without disturbing cluster operations
  - Supports mixing Windows NT 4.0, Windows 2000, and Windows Server 2003 nodes in one cluster
- Fault tolerance
  - Automatically detects and recovers from a failed or offline computer
  - Automatically rebalances the network load when the cluster set changes
  - Recovers and redistributes the workload within 10 seconds
  - Handles inadvertent subnetting and rejoining of the cluster network
- Manageability
  - Specifies the load balancing for a single IP port or group of ports using port management rules that customize the workload for each computer and enable optional support for client sessions
  - Directs all client requests to a single host to further refine load balancing among different programs using optional single-host rules
  - Blocks undesired network access to certain IP ports
  - Remotely starts, stops, and controls network load balancing actions from any networked computer running Windows Server 2003 by using shell commands or scripts
- Scalable performance
  - Load balances requests for individual TCP/IP services across the cluster
  - Load balances multiple server requests from a single client
  - Ensures high performance and low overhead using fully pipelined implementation

## Microsoft Cluster Service

MSCS is a full-featured cluster management application that consists of a loosely coupled collection of two or more independent servers. It provides isolation, detection, and recovery from almost any failure in the hardware, operating system, or application software.

MSCS enables fully automated detection, failover, and failback at the hardware and application level in the event of hardware or application failure. It manages the components of the cluster as resources and maintains availability to the attached network.

MSCS is the Windows Server 2003 subsystem responsible for:

- Maintaining a record of the cluster configuration
- Maintaining intracenter communications
- Initiating failover of resources when informed of a resource failure by the Resource Monitor

---

**Note**

Microsoft Exchange 2000 is not supported for installation on the Windows Server 2003 operating system. Windows Server 2003 can be used as the domain controller for an Exchange 2000 server or cluster running on a Windows 2000 server with either the Exchange 2000 Service Pack (SP) 3 installed or the Exchange 2000 SP2 with the Q316463 directory fix installed.

---

MSCS capabilities are built on top of the existing foundation of the Windows Server 2003 operating system.

## MSCS communications

MSCS maintains two types of intracluster communication:

- The cluster heartbeat is a simple data packet that is sent periodically from one node to the other, similar to a TCP/IP ping. The heartbeat is used to verify that each node of the cluster is operational.
- Cluster and application configuration data is kept synchronized among all nodes. For example, data in the Windows Server 2003 registry is kept synchronized so that if an application or service is failed over, current registry settings on the first node will be configured correctly when the application starts up on the new node.

The nodes communicate over multiple separate physical links:

- **Private interconnect** — Heartbeat and configuration data is exchanged over the private interconnect.
- **Backup path** — Heartbeat and configuration data can be exchanged over a backup path if the private interconnect fails.
- **Quorum arbitration** — Periodically, each node writes to and reads from a special, shared logical volume known as the *quorum drive*. The Resource Monitor running on each node communicates through the quorum drive to verify that the other node still has access to the shared storage arrays. This type of heartbeat is known as *quorum arbitration*.

It is critical for preventing a false failover situation, where one node attempts to assume control of the logical volumes on the other node before the other node has relinquished control. Such an event would result in severe file system corruption.

## Choosing network load balancing or MSCS

The guidelines for choosing network load balancing or MSCS are as follows:

- Use network load balancing for data that is mostly static and can be easily replicated, which could include:
  - Web servers
  - Terminal servers
  - FTP services
- Use MSCS for data that changes often and is difficult to replicate, such as:
  - File and print servers
  - Database servers
  - Messaging servers

### Example

An e-commerce site usually consists of a database back-end and a website front-end. If only one web server and one database server are used, a single failure of either server would bring down the site. Increasing the reliability and availability of a site requires multiple servers.

You can create a web server farm with up to 32 machines. Each web server will have the same active web pages and network load balancing. Two to four servers on the back-end can be used with MSCS to run the database.

General recommendations for choosing the appropriate clustering technology in various scenarios are illustrated in the following table.

Scenario	MSCS	NLB
Web server farm		X
Terminal servers		X
File/print servers	X	
Database/messaging	X	

## Serviceguard for Linux

HP Serviceguard for Linux is specialized software used to protect critical applications from hardware and software failures. Serviceguard for Linux enables multiple nodes to be organized into clusters that deliver highly available application services to LAN-attached clients.

Serviceguard for Linux enables you to create high-availability clusters with HP ProLiant servers in a Linux environment. A high-availability computer system enables application services to continue during a hardware or software failure. Highly available systems protect users from software failures as well as failure of a system processing unit, disk drive, or LAN component. If one component fails, the redundant component takes over. Serviceguard and other high-availability subsystems coordinate the transfer between components.

A Serviceguard cluster is a networked group of ProLiant servers that have sufficient redundancy of software and hardware to prevent a single point of failure from a significantly disrupting service. Application services (individual Linux processes) are grouped together in packages. If a single service, node, network, or other resource fails, Serviceguard automatically transfers control of the package to another node within the cluster and allows services to remain available with minimal interruption.

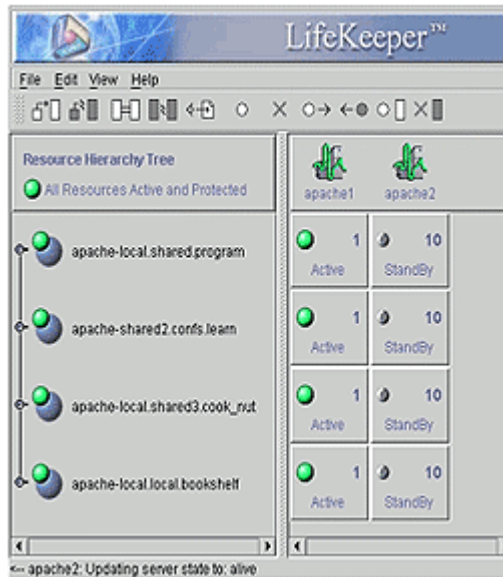
Serviceguard for Linux monitors the health of various components within each node and quickly responds to failures to eliminate or minimize application downtime. Components Serviceguard monitors include:

- System memory
- System and application processes
- LAN media and adapters
- Storage subsystem and adapters

Serviceguard for Linux:

- Supports Red Hat Enterprise Linux and SuSE Linux Enterprise Server 8 certified up to 16 nodes
- Includes MySQL and Oracle toolkits for rapid integration
- Provides increased HP ProLiant, Integrity, and StorageWorks portfolio coverage

## LifeKeeper



LifeKeeper for Linux offers a high-availability software solution for 2- to 16-node clusters. It supports mission-critical applications and provides data fault resilience on a range of ProLiant servers and StorageWorks solutions.

Hardware component and application faults at the resource level are detected in advance of a full system failure through multiple fault-detection mechanisms. LifeKeeper monitors Linux clusters using intelligent processes and multiple LAN heartbeats. By sending redundant signals between server nodes to determine system and application health, LifeKeeper confirms system status before taking action. This reduces the risk of a single point of failure and minimizes false failovers.

LifeKeeper also limits unnecessary failovers by recovering failed applications at the resource level, without a full failover to another server. If required, failover is transparent to clients and does not impact user productivity.

LifeKeeper migrates all applications and transfers connectivity in such a way that clients have continuous access to applications and data. This ensures that all clients are not affected by unanticipated system failures.

## ProLiant clusters for NetWare

NetWare 6 includes Novell clustering technologies, which provide hardware, software, data, and application redundancy for high fault tolerance.

Novell Clustering Services clusters feature shared disk pools and logical volumes managed by the logical servers. The application is installed and configured on each server that supports the application in the cluster. The application is installed using templates that are available for many applications and are easily created for new applications. These templates include load and unload scripts and volume mounts. They dictate how many of the nodes must be operational to consider the cluster alive.

All versions of NetWare 6 include Novell Clustering Services for two nodes; Novell Clustering Services 1.6 supports up to 32 cluster nodes.

The Cluster System Services is a set of APIs for cluster-aware applications to use to enable distributed shared memory and distributed locking.

- Distributed shared memory allows cluster-aware applications running across multiple servers to access the same data as if the data were located on physically shared RAM chips.
- The distributed locking service prevents one thread on one node from locking the same shared pool another thread has already locked. This ensures that pools are only active on one node at a time.

## Deploying and managing clusters

Throughout the life of the cluster, you will need to:

- Deploy new or existing nodes
- Improve performance
- Upgrade hardware components
- Upgrade software
- Increase storage capacity
- Restructure cluster groups
- Back up your cluster data
- Monitor ongoing activities

HP provides applications and utilities to deploy nodes remotely and manage cluster performance. The HP ProLiant Essentials Rapid Deployment Pack (RDP) simplifies cluster deployment by enabling you to drag and drop predefined images or scripts onto multiple target nodes. RDP includes unique features for ProLiant BL servers and supports all ProLiant ML and DL servers.

Advanced features detect and display server blades based on their physical rack, enclosure, and bay location. You can install or redeploy the configuration of a previous computer to a new blade automatically when it is replaced. You can also assign an event to be executed the first time a blade is inserted into the bay.

Performance can change when applications or resources transfer from one node to another. The extent of the performance change depends on how well the secondary node is equipped to handle the increase in workload. This change is especially obvious after an entire node failure, where all the cluster resources might move to the surviving node.

To determine the optimum time to manually move a cluster group to balance the load, consider which type of group needs to be moved and how many clients are using the group. File and print services generally are not business-critical and they do not sustain an extremely high usage rate. Therefore, file and print services are good candidates to move whenever load balancing needs to occur, even when the overall cluster usage is high.

However, applications such as databases should not be moved from one node to another during peak processing periods. When a database is moved from one node to another, the database application must be shut down and restarted. During the time it takes to restart, you cannot access the database. In most cases, moving a database group to another node should be performed during nonpeak hours.

## Cluster monitoring

HP Systems Insight Manager

Updated: Monday, October 27, 2003 4:50:23 PM CST

Home Logout

Uncleared Event Status 9 1 6 341

Search [ ] Go

Advanced Search

System Lists

Status Overview

All Systems

All Events

My Favorites

System Lists

Systems by Type

Systems by Status

Systems by Operating System

Clusters by Type

All Clusters

All MSCS Clusters

All TruClusters

All OpenVMS Clusters

Clusters by Status

System Functions

My Lists

Event Lists

Legend

✖ = Critical

✓ = Normal

⚠ = Major

❓ = Unknown

⚡ = Minor

ℹ = Informational

Tools Deploy Configure Diagnose Optimize Reports Logs Options Help

All Clusters

View as: table

Clusters in table: 0 Critical 0 Major 1 Minor 0 Normal 0 Unknown Total: 1

Se...	CS	Cluster Name	Clust...	Cluster Type	Cluster Description
<input type="checkbox"/>	juicyfruit	16.12...	MSCS	Microsoft Cluster Service	

Select All Save Selection As... Delete Print

The HP Systems Insight Manager cluster monitoring feature provides a view of the cluster environment, identifying clusters and their attributes. The Cluster Monitor is accessed when a cluster query is executed.

Cluster monitoring provides:

- The ability to identify clusters and their members using a cluster query
- The ability to set thresholds on disk capacity and processor utilization for cluster nodes
- Visual representation of clusters and their nodes
- A list of status messages signaling potential trouble such as a node failure or exceeding a threshold

Cluster Monitor supports the following attributes or cluster monitor resources:

- **Disk and processor** — The status for a disk or processor is based on the threshold assigned to the disks or processors. The threshold for a disk is based on the disk utilization. The threshold for a processor is based on the processor utilization.
- **Cluster resource** — The cluster resource displays cluster information such as whether a resource is online, what group it belongs to, and which node controls the group. The cluster types supported by the Cluster Monitor include MSCS, OpenVMS, HP Tru64 Unix, and Novell NetWare 6.0 Cluster Service. Currently, the Cluster Monitor supports a cluster resource for the MSCS clusters.
- **Cluster tree status** — The cluster status is independent of the hardware and software status shown in a device query result list. The cluster tree status is a propagated status.

The status of an attribute or resource is reflected in the color of the parent node and cluster icons. Changes in node or cluster resources are visible even when the cluster tree is not expanded. When more than one cluster or node resource is in an abnormal state, the parent cluster or node icon reflects the most severe state because the resource statuses are propagated upward. The node status in the cluster tree is not tied to any node status in tables provided by the MSCS resource, only to the resource status for that node.

In Systems Insight Manager, an All Clusters query is defined under the cluster list. The All Clusters query begins with the entire cluster list but then applies the user's authorizations to customize the list. In Systems Insight Manager, no matter what the query is, users will only see those devices that are part of their authorizations.

## Cluster Monitor

The screenshot shows the Cluster Monitor interface for a cluster named 'juicyfruit'. The interface is divided into several sections:

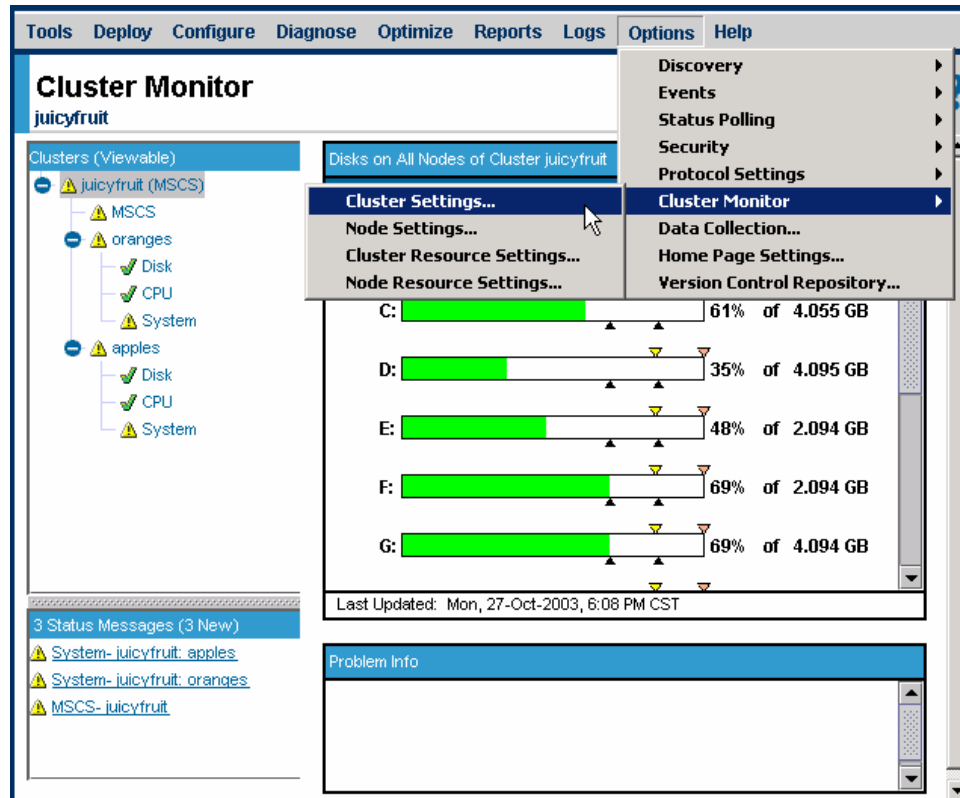
- Cluster Monitor juicyfruit**: The main title of the interface.
- Clusters (Viewable)**: A tree view showing the cluster hierarchy.
  - juicyfruit (MSCS)**: The root cluster node, marked with a warning icon.
    - MSCS**: A node under the root, marked with a warning icon.
      - oranges**: A node under MSCS, marked with a warning icon.
        - Disk**: A resource under oranges, marked with a checkmark.
        - CPU**: A resource under oranges, marked with a checkmark.
        - System**: A resource under oranges, marked with a warning icon.
      - apples**: A node under MSCS, marked with a warning icon.
        - Disk**: A resource under apples, marked with a checkmark.
        - CPU**: A resource under apples, marked with a checkmark.
        - System**: A resource under apples, marked with a warning icon.
- Aggregate Cluster Information**: A section providing information across the entire cluster. It includes links for [Disks](#) and [CPUs](#).
- Cluster Information**: A section for detailed cluster information, including:
  - Cluster Name: juicyfruit
  - Cluster IP Address: 16.129.64.132
  - Contact Person: (empty field)
  - Contact Phone: (empty field)
  - Contact Pager: (empty field)
  - Contact Email: (empty field)
  - Dial In Modem: (empty field)
- 3 Status Messages (3 New)**: A section listing status messages:
  - [System- juicyfruit: apples](#)
  - [System- juicyfruit: oranges](#)
  - [MSCS- juicyfruit](#)

The Cluster Monitor tree lists the cluster resources (such as MSCS), the cluster nodes, and all node resources (such as the disk and processor). Clicking an item in the tree displays information about that item.

Two classes of resources are distinguished by their scope:

- **Cluster-level resources** monitor attributes on a cluster-wide basis. Attributes monitored by cluster-level resources are listed with nodes at the second level of the cluster hierarchy tree.
- **Node-level resources** monitor attributes on individual nodes. Attributes monitored by node-level resources are listed with nodes at the lowest level of the cluster hierarchy tree.

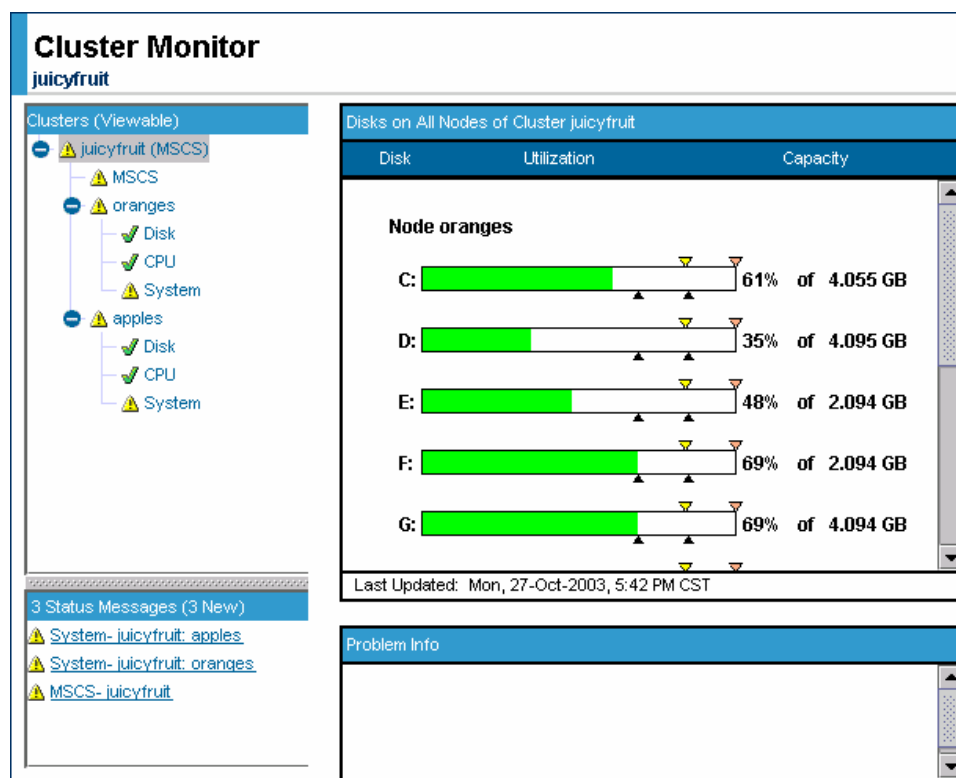
## Configuring Cluster Monitor



In Systems Insight Manager, there are four Cluster Monitor setting pages:

- **Cluster Settings** — Used to enter problem reporting and administrative information for the managed clusters. This information is displayed in:
  - The Cluster Monitor data area when the cluster is selected from the hierarchy tree
  - Events generated about the cluster
- **Node Settings** — Used to enter problem reporting and administrative information for the nodes. This information is displayed in:
  - The Cluster Monitor data area when the node is selected from the hierarchy tree
  - Events generated about the node
- **Cluster Resource Settings** — Used to set operational parameters such as polling rates for cluster-level resources.
- **Node Resource Settings** — Used to adjust operational parameters such as polling rates and thresholds for node-level resources.

## Cluster resources



The Cluster Monitor relies on resources to monitor specific cluster attributes. Standard resource types are Disk and CPU, which display the disk and processor capacity and utilization.

You can set thresholds (minor and major) on a particular disk or processor, and upon reaching that threshold, a Systems Insight Manager event and Cluster Monitor status message are created. By creating event thresholds, users can use the notification methods provided by Systems Insight Manager, such as email and paging.

## Smart Array Multipath software

HP Smart Array Multipath software provides dual-path functionality and supports failover capability for redundant HBAs in a server with redundant cabling to an MSA500 system. This fault-tolerance tool enhances availability in clustering and DAS solutions with up to two servers managing up to 2TB of stored data.

The software supports multiple I/O paths to the same logical volumes. When a path fails, the software moves the logical drives from the failed path to the path of the redundant HBA in the same server. Smart Array Multipath software works differently with Microsoft and Linux products to provide specific features that are compatible with each operating system.

### Microsoft feature support

Features supported under the Microsoft operating systems include:

- Static LUN balancing enables the administrator to optimize I/O by assigning specific volumes to specific paths in the storage system. If one path fails, the system moves volumes to the path for the redundant HBA.
- Server-based PCI Hot Plug functionality supports online HBA addition and replacement.
- Logical volume hot-add functionality supports online array expansion and extension.
- Configuration flexibility enables boot volumes for DAS solutions to reside on the storage system.

### Linux feature support

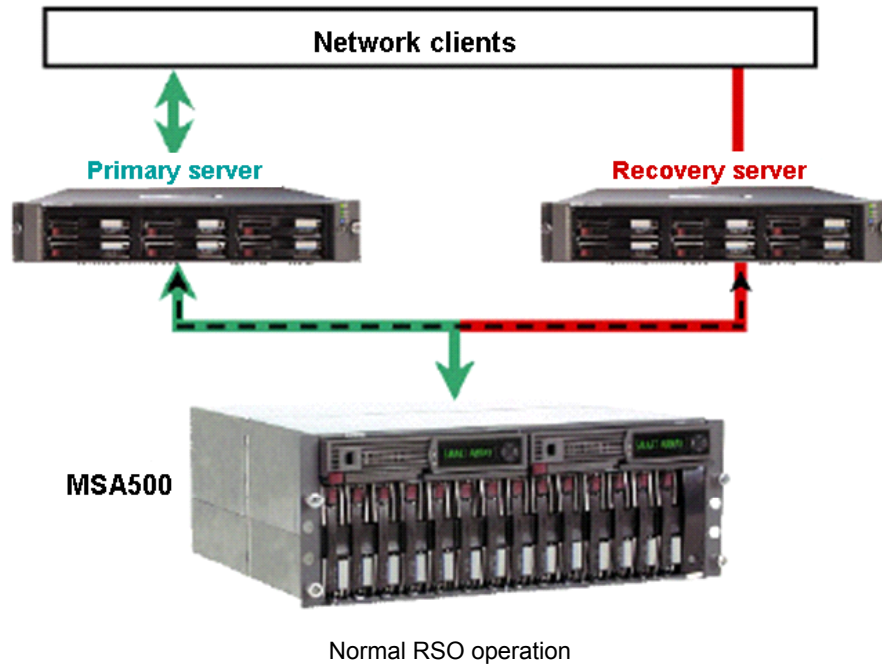
Features supported under the Linux operating systems include:

- Static load (read-only) balancing enables the HBAs to optimize I/O with dual paths to the same volumes in the storage system. If one HBA fails, the system moves volumes to the path for the redundant HBA for I/O management.
- Servers maintain boot volumes and enable greater storage capacity on the storage system.

### Multiple operating system support

For multipath configurations with multiple operating systems, use HP Selective Storage Presentation (SSP) to restrict HBAs from accessing logical volumes containing data from a different operating system.

## Recovery Server Option



ProLiant Essentials Recovery Server Option (RSO) enables active/standby configuration of two servers where:

- One server acts as the primary or active server.
- The second server is the recovery server acting in a passive or pre-initialization mode.

When a fatal fault is detected in either the hardware or operating system of the primary server, the recovery server initializes and takes over, making the applications available.

RSO maintains uptime as a result of fully automated server recovery in the event of a server failure. This eliminates the requirement for intervention from the systems administrator to recover from a server failure, providing an excellent solution for unattended operation.

Other benefits include that RSO:

- Provides a reliable entry-level high-availability solution that is simple to install and configure
- Is ideal for:
  - Non-cluster-aware applications, including custom-developed applications
  - Remote locations and distributed environments without on-site IT staff
- Uses a menu-based interface, which makes it easier to install and configure than a typical cluster configuration
- Works in the background and does not interfere with standard server operation
- Uses standard operating system editions, which keeps costs low

RSO operates independently of operating system and application software, but it requires the appropriate software drivers. RSO supports these operating systems:

- Windows 2000
- Windows NT 4.0
- Red Hat Linux
- SuSE Linux Enterprise Server 7
- NetWare

## Cluster technologies

### Distributed Lock Manager

The Distributed Lock Manager (DLM) defines a lock resource as the lockable entity. The lock manager creates a lock resource the first time an application requests a lock against it. A single lock resource can have one or many locks associated with it. A lock is always associated with one lock resource.

The lock manager provides a single, unified lock image shared among all nodes in the cluster. Each node runs a copy of the lock manager daemon. These lock manager daemons communicate with each other to maintain a cluster-wide database of lock resources and the locks held on these lock resources.

By allowing all nodes to maintain the master copy of lock resources, instead of having one primary lock manager in a cluster, the lock manager can reduce network traffic in cases when the lock request can be handled on the local node. Handling the requests on the local node also avoids the potential bottleneck resulting from having one primary lock manager and reduces the time required to reconstruct the lock database when a failover occurs.

### Non-Uniform Memory Access

*Non-Uniform Memory Access* (NUMA) refers to a hardware architectural feature in multiprocessor platforms that attempts to address the increasing disparity between requirements for processor speed and bandwidth capabilities of memory systems, including the interconnect between processors and memory.

NUMA systems group resources into building blocks that balance an appropriate number of processors and I/O buses with a local memory system that delivers the necessary bandwidth. The local building blocks are combined into a larger system by means of a system level interconnect with a platform-specific topology.

The local processor and I/O components on a particular building block can access their own “local” memory with the lowest possible latency for a particular system design. The local building block can in turn access the resources (processors, I/O, and memory) of remote building blocks at the cost of increased access latency and decreased global access bandwidth. The term *Non-Uniform Memory Access* refers to the difference in latency between “local” and “remote” memory accesses that can occur on a NUMA platform.

## **NUMA-CC**

Non-Uniform Memory Access cache-coherent (NUMA-CC) architecture systems are constructed out of multiple resource affinity domains (RADs) that group memory and processor resources. Access to memory that resides in a different RAD from the processor making the reference requires significant additional time (latency) as compared to accesses to memory that resides in the same RAD as the processor making the reference.

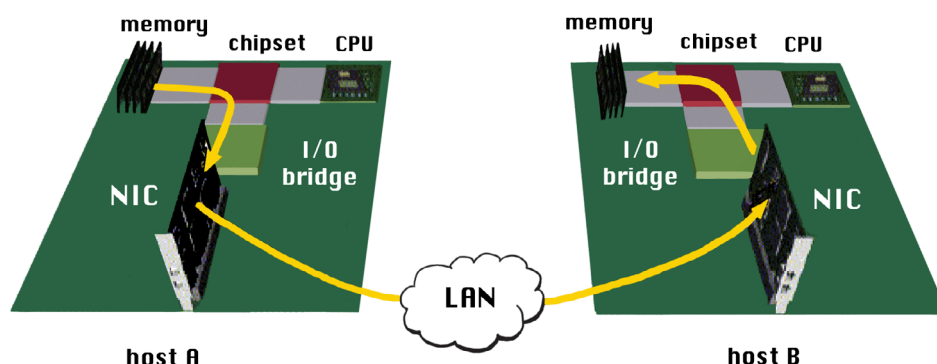
The impact of this access latency is often hidden by the data cache of the processor. However, only applications with poor cache behavior (that is, little access locality) or a high degree of tight synchronization of read-write data between multiple processors on multiple RADs are likely to experience noticeable performance reductions due to increased memory access latency.

## **Virtual Interface Architecture**

Virtual Interface Architecture (VIA) is a distributed messaging technology authored by Compaq, Intel, and Microsoft that specifies an open architecture. It provides high-speed and low latency communications between servers in a cluster and is used where server-to-server messages deal with application and data availability. It enables any server error or failure to cause an immediate transfer of the business-critical application to another server for uninterrupted processing.

The architecture also allows for parallel application processing and minimizes performance bottlenecks where more traditional communications protocols have caused application processing delays because of protocol stack handling and or network traffic overloads resulting from collision detection processing. These slowdowns cause application delays or inefficient use of the cluster.

## Remote Direct Memory Access



Remote direct memory access (RDMA) protocol provides a faster path for applications to transmit messages between servers by moving data from the memory of one computer directly into the memory of another computer with minimal involvement from their processors. It is used by systems, applications, and storage to communicate directly over a fabric infrastructure, such as Ethernet.

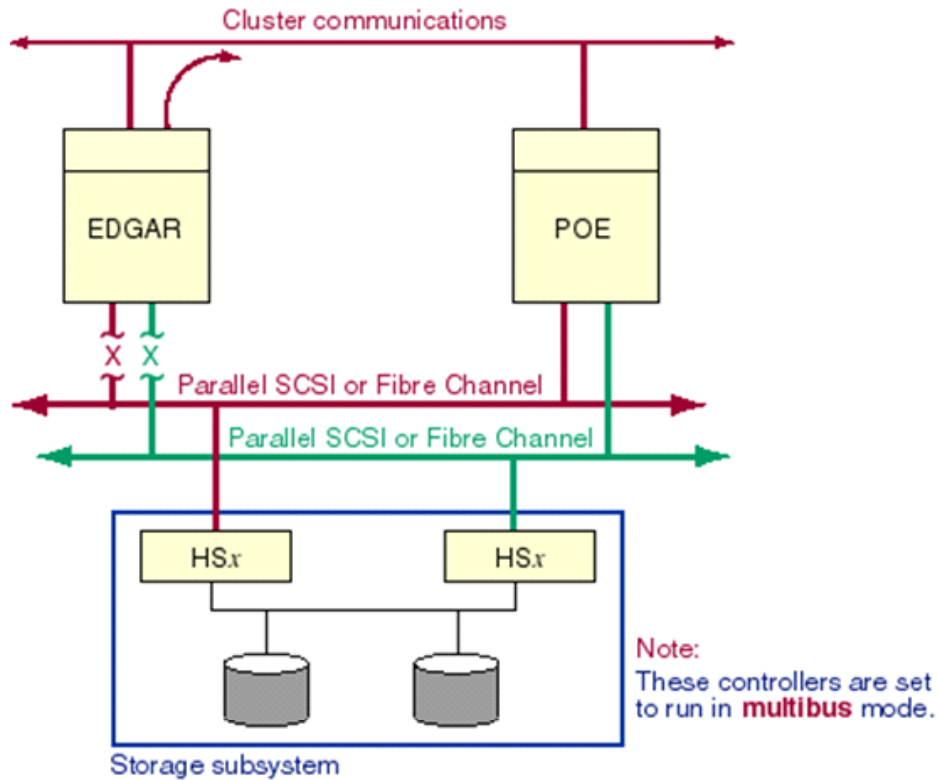
A typical data center uses a variety of interconnects to link servers and storage. The use of multiple system and peripheral bus interconnects decreases compatibility, interoperability, and management efficiency and drives up the cost of equipment, software, training, and the personnel needed to operate and maintain it. To increase efficiency and lower costs, the data center network infrastructure must be transformed into a unified, flexible, high-speed fabric.

A unified high-speed data center infrastructure requires a high-bandwidth, low-latency fabric that can move data efficiently and securely between servers, storage, and applications. Use of more efficient communication protocols, some of which run over existing infrastructures, frees processors for more useful work and improves infrastructure utilization. In addition, the ability of fabric interconnects to converge functions in the data center over fewer, or possibly even one, industry-standard interconnect presents significant benefits.

RDMA over TCP technology can diminish TCP/IP protocol overhead and constrain memory bandwidth, which are obstacles to faster Ethernet networks in the data center. Additional information included in the RDMA protocol allows a system to place the communicated data directly into its final memory destination without any additional or interim data copies. This capability provides the most efficient network communication possible between systems.

RDMA evolved from cluster interconnects such as ServerNet, VIA, and InfiniBand. HP is a founding member of the RDMA Consortium, an independent group formed to develop the architectural specifications necessary to implement products that provide RDMA capabilities over existing TCP/IP networks.

## HSx technology



Direct SCSI to MSCP served configuration with two interconnects

In a multihost SCSI cluster system, you can use HSx technology to increase disk storage availability by configuring the cluster for both types of multipath failover:

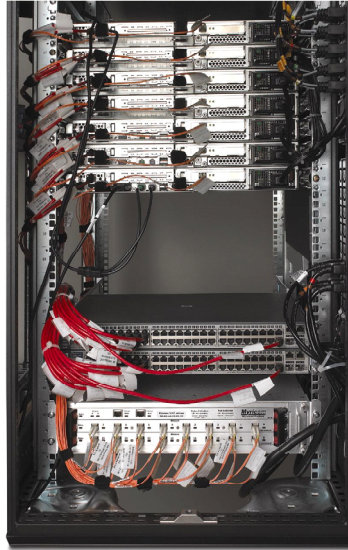
- Direct SCSI to direct SCSI
- Direct SCSI to Mass Storage Control Protocol (MSCP)-served SCSI.

In this configuration:

- Both nodes are directly connected to both storage interconnects.
- Both nodes are connected to a second interconnect for cluster communications.
- Each HSx storage controller is connected to only one interconnect.
- Both HSx storage controllers are in the same cabinet.

This configuration provides the advantages of both direct SCSI failover and direct to MSCP-served failover.

## Troubleshooting SAN and cluster problems



The cluster nodes must have enough network capacity to handle requests from the client machines and to handle failover and failback events gracefully.

Ensure that each server can handle the maximum number of clients that can attach to the cluster. If node A encounters a failure and its applications and services fail over to node B, then node B must handle access from its own network clients as well as those that normally connect to node A.

It is important to understand the effect of failover on network bandwidth. When the cluster encounters a server failover event, only one server is responding to network I/O requests. Ensure that the network speed and protocol of the surviving node will sufficiently handle the maximum number of network I/Os necessary to support critical services.

## Identifying points of failure

To have a highly available system, you must reduce single points of failure in both internal and external resources. When a cluster has a dependent, external resource, the resource must be highly available or the external resource becomes a potential single point of failure. In addition, the network path from the cluster to the external resource becomes a potential single point of failure.

A minimum of two networks are used in a cluster:

- The public LAN
- The private interconnect

Availability and performance of these networks are critical.

Several circumstances can contribute to failures in the network subsystem of a cluster:

- The domain controller is unreachable.
- The public LAN NIC fails and prevents clients from accessing services or applications that are currently running on that node.
- The interconnect communication between the cluster nodes fails and MSCS must determine the state of the cluster and take action. The node that currently has control of the quorum drive remains in control and MSCS on the other node is stopped.
- The switch in either the private or public networks fails and prevents network communication.

Each of these single points of failure can be resolved through:

- Domainlets
- NIC teaming on the public LAN
- Configuration of a backup interconnect path using MSCS
- Redundant hubs or switches



Assistance, service, and support are available for clusters by visiting:  
<http://www.hp.com/hps/mission/#services>

---

## Troubleshooting shared storage

This section addresses potential problems related to using HP Fibre Channel storage systems as shared storage devices. It does not address problems that are specific to the storage system components themselves or to using a Fibre Channel storage system in a stand-alone server configuration. For those issues, see the *HP Fibre Channel Storage System User Guide* or the *Recovery Server Option User Guide*.

Problem	Explanation and recommended action
Drives in the storage array are not recognized.	<ul style="list-style-type: none"><li>■ Microsoft Cluster Administrator takes a snapshot of the registry when it starts up. However, it can take up to a minute after the second node is restarted for the disk signatures to be written to both registries. You might not have waited long enough to view the most current data. Wait a minute and then click <i>Refresh</i>. Both cluster nodes should be restarted after installing MSCS. If you have not done so, restart the two nodes and verify that the drives are recognized.</li><li>■ Ensure that the Fibre Channel HBA driver or SMART-2 Array controller driver for Windows Server 2003 is installed and running on both nodes. If not, see the “Installing Drivers for Windows” chapter in the Fibre Channel Host Controller Installation Guide or in the ProLiant DL380 Packaged Cluster User Guide.</li><li>■ Use Event Viewer to ensure that there are no hardware errors or transport problems. Look in the event log for disk I/O error messages.</li><li>■ Ensure that the physical SCSI disks are supported by ProLiant clusters and the firmware revision level of all drives meets or exceeds HP recommendations. If the drives are still not recognized, start investigating the problem from the lowest level of drive configuration.</li><li>■ Run the ACU. You can run this utility online if at least one logical drive is configured and recognized. The online utility is installed as part of the HP Support Pack for Windows Server 2003. If you need to run this utility offline, shut down the nodes and restart with the SmartStart CD or with ACU Diskette 1.</li><li>■ If all drives are not recognized by the ACU, the problem is probably a physical connection.</li><li>■ If all drives are recognized by the ACU and are configured correctly, start Windows Server 2003 and run Disk Administrator. Verify that all drive volumes display in Disk Administrator and that they each have a permanent drive letter assigned. Verify that the drive letter assignments on the second cluster node match exactly the assignments on the first cluster node.</li><li>■ If all drives are recognized by Disk Administrator, verify that all drive volumes display as physical disk resources.</li><li>■ Verify that the short wave GBICs (GBIC-SWs) are properly seated in the HBAs, Fibre Channel array controller, and storage hub. Verify that the correct power-on sequence was used.</li><li>■ Verify that all Fibre Channel cables are properly connected to their GBIC-SWs. For details on how to connect the GBICs and Fibre Channel cables, see the <i>Fibre Channel Storage System User Guide</i>.</li></ul>

Problem	Explanation and recommended action
Data on the shared storage seems to be overwritten.	<ul style="list-style-type: none"> <li>■ Both cluster nodes have access to the Fibre Channel storage system and will write freely to the drive volumes unless MSCS is loaded and running. MSCS arbitrates the access to drive volumes in the shared storage.</li> <li>■ Launch the Service Control Panel applet on each node and verify that MSCS is running.</li> </ul>
The second node cannot connect to the shared drives.	<ul style="list-style-type: none"> <li>■ Ensure that a physical connection exists from the second node to the Fibre Channel storage hub or to the ProLiant storage system.</li> <li>■ Verify that the GBIC-SWs are properly seated in the Fibre Channel host controllers, Fibre Channel array controller, and storage hub.</li> <li>■ Verify that all HBAs are properly connected to their GBIC-SWs.</li> <li>■ Verify that the SCSI cables are securely fastened to the SMART-2 Array controllers and to the connectors on the ProLiant storage system.</li> <li>■ Ensure that the shared drives (located in the Fibre Channel array or ProLiant storage system) are assigned the same drive letters on both nodes. To do so, run Windows Server 2003 Disk Administrator on each node and ensure that all shared drives have been assigned identical drive letters. The drive letter assignments must be permanent.</li> <li>■ If the second node was powered on before either the Fibre Channel storage hub or the Fibre Channel array, shut down the second node, power it off, and then back on. The Fibre Channel storage system components must be powered on before the cluster nodes.</li> <li>■ If the second node was powered on before the shared ProLiant storage system, shut down the second node, power it off, and then back on. The ProLiant storage system components must be powered on before the cluster nodes.</li> <li>■ Verify any switch zoning conflicts.</li> <li>■ Verify SSP on the HSG controllers or Virtual Disk presentation on the EVA.</li> </ul>

## HP troubleshooting utilities

### Array Controller Utility

Use the ACU to verify connectivity on both servers. If the ACU or Disk Administrator fails to detect the Fibre Channel array, use the recommended actions listed in the following table.

Problem	Explanation and recommended action
The ACU fails to recognize the Fibre Channel array.	<ol style="list-style-type: none"><li>1. Ensure that all ROMs and drivers are the latest versions and that the ACU is version 2.50A or later.</li><li>2. Power off the nodes.</li><li>3. Power off the Fibre Channel storage system for one to two minutes.</li><li>4. Power on the Fibre Channel storage system and wait several minutes before restarting the nodes. (You can determine when the system is ready for the server to be powered on by observing the lights on the rear of the Fibre Channel array controller. The top light should be green and flashing steadily.)</li></ol>
The Windows Server 2003 Disk Administrator fails to recognize the Fibre Channel array.	Power off the storage system and the server as described in the preceding situation.
Windows Server 2003 reports multiple system or application errors.	Use Windows Server 2003 Disk Administrator to check whether Windows Server 2003 recognizes the Fibre Channel storage system. If it fails to recognize it, shut down Windows Server 2003 and power down the Fibre Channel storage system, as described in the first situation discussed in this table.
The Windows Server 2003 My Computer menu reports incorrect drive letter mapping.	Follow the suggestion described in the first situation discussed in this table.

## Array Diagnostic Utility

The Array Diagnostic Utility (ADU) is a Windows-based software tool designed to run on all HP servers that support HP array controllers.

The two main functions of the ADU are to:

- Collect all possible information about the array controllers in the system
- Generate a list of detected problems

The error messages and codes listed include all codes generated by HP products. The system generates only codes applicable to the configuration and options in the server. The ADU works by issuing multiple commands to the array controllers to determine if a problem exists. This data can then be saved to a file. In severe situations, this file can be sent to HP for analysis. In most cases, the ADU provides enough information to initiate problem resolution immediately.

---

**Note**

The ADU does not write to the drives, destroy data, or change or remove configuration information.

---

To start the ADU, follow these steps:

1. Insert the SmartStart CD into the CD-ROM drive.
2. Reboot the system from the SmartStart CD.
3. Select *Array Diagnostic Utility (ADU)* from the System Utilities menu.

The Please Wait panel displays, indicating that ADU is identifying the system parameters. ADU gathers information from all of the array controllers in the system. The time it takes to gather this information depends on the size of the system. When the information gathering process is complete, The utility displays the main screen or a panel indicating any problems detected.



---

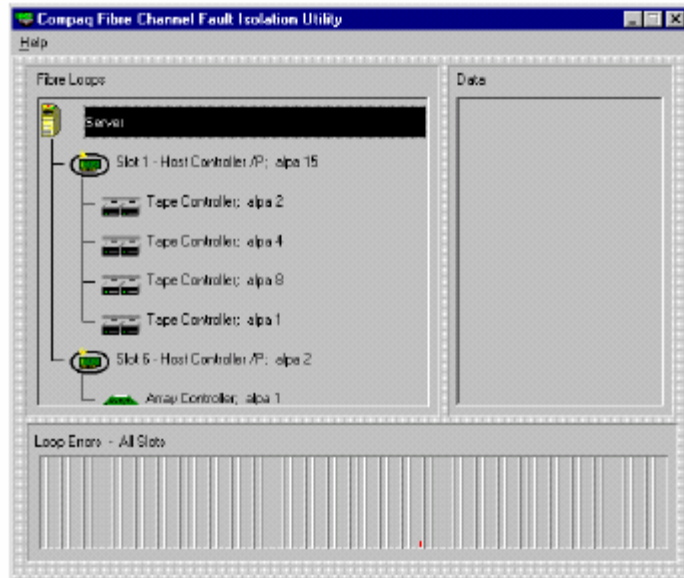
**Caution**

Do not cycle the power during this identification process. ADU must perform low-level operations that, if interrupted, could cause the controller to revert to a previous level of firmware if the firmware was soft-upgraded.

---

4. To generate an ADU report, select *File → Save Data* from the command menu.

## Fibre Channel Fault Isolation Utility



The Fibre Channel Fault Isolation Utility analyzes Fibre Channel components including HBAs, array controllers, and tape controllers. Fibre Channel storage hubs are logically transparent to the operations of the FC-AL, but even a failed hub can be detected when the Fibre Channel Fault Isolation Utility is used in combination with troubleshooting flow charts.

Use the Fibre Channel Fault Isolation Utility to:

- Verify controller firmware revisions
- Confirm active HBAs
- Determine if a loop segment is faulty
- Isolate a faulty loop segment

This utility is located on the HP SmartStart Diskette Builder. The SmartStart Diskette Builder is a utility that uses data stored on the SmartStart CD to create support diskettes. You can create support diskettes for specific configuration needs or for software that cannot be used directly from the SmartStart CD.

---

### Note

Detailed information on how to use this utility and troubleshooting flow charts can be found in the *Fibre Channel Troubleshooting Guide* that ships with HP storage hubs.

---

## Dumpcfg.exe: Dump Config

This command line tool simplifies the manual system recovery process associated with storage configuration. You must be a member of the Administrator group to run this tool.

Use Dump Config to:

- Display system, disk, and volume information
- Change your disk signature

Windows disk devices are partitioned. These partitions can be used directly as logical devices, subdivided further into smaller logical devices, or combined into larger, multidisk logical devices called volume sets. Normally, access to a logical device requires that a drive letter be assigned and the device be formatted with a file system, although there are techniques for accessing Windows devices as "raw" devices using either drive letters or a physical drive designation.

Windows maintains disk recognition based on a unique 4-byte disk signature that is recorded in the master boot record of each disk at off-set 0x1B8. DumpConfig can be used to modify this value. When the system initializes, and if the disk does not already have a non-zero value in this field, this signature is generated by the Disk Administrator.

This technique does not rely on knowledge of physical properties that might change such as the SCSI ID or which adapter is connected to a disk. Therefore, disk hardware connections can be changed without reducing existing software fault tolerance.

The advantage of using the signature to identify a disk is that even when the physical disk configuration is changed, Windows can still correctly identify volume sets and assign drive letters to their original devices. When a disk that already has a valid signature displays on a Windows system, Disk Administrator will not write a new signature, but will record the signature in the DISK key.

Duplicate signatures are not allowed because they can create problems when making copies of a disk at the track level. This occurs with some breakable mirror and disaster protection software, if the disk copy (which also has a copy of the signature) is visible from the same system as the original disk.

When Disk Administrator detects a duplicate signature, it will ask permission to write a new signature to the offending disk. If permission is denied, it will make the disk unavailable.

You can use DumpConfig to change the disk signature, making the disk readable again. The disk can still be accessed by programs using physical drive address semantics.

## Summary

RC Engineering and GEEK have successfully integrated their human and IT resources. The SAN enables the RC Engineers to access the GEEK database. In turn, GEEK scientists can view RC Engineering files remotely over the network.

The original office now has two file and print servers attached to an MSA1000 through an FC-AL connection. By clustering both servers from both companies, they have increased availability and reduced vulnerability to failure. Clustering also provided scalability and investment protection to both companies.

Now that Bob has received help troubleshooting his server, he must record the changes that were made along with the results of those changes. Any references to recommendations that were not performed, purchased, or could not be implemented should also be documented. This document should be a running history of all work performed on the system, which could prove invaluable the next time the system requires support.

## Learning check

1. What is the primary objective of the HP Enterprise Network Storage Architecture (ENSA)?
  - a. To integrate open standards and industry-standard approaches toward managing storage
  - b. To create a portfolio of modular, scalable, and highly available products
  - c. To establish a direct connection between storage resources
  - d. To automate the storage management processes that manage data placement and protection through the information life cycle
2. What is a SAN?  
.....  
.....
3. Which component is not a SAN layer?
  - a. Client
  - b. Server
  - c. Storage
  - d. Interconnect
  - e. Fabric
4. List three benefits of implementing a SAN.  
.....  
.....  
.....
5. The MSA1000 is the only storage system in the industry that enables a cost-effective migration to SANs by offering DtS capability.
  - ☐ True
  - ☐ False
6. Which Systems Insight Manager utility focuses on a computing environment from the perspective of clusters?  
.....

7. What is a cluster?  
.....  
.....
8. Which component provides initialization to Fibre Channel devices?
  - a. HBA
  - b. Disk array
  - c. SAN management software
  - d. SAN switch
9. Which applications benefit most from migrating to storage hosted on a SAN?  
.....  
.....  
.....
10. According to the latest research, what are the leading causes of downtime, from most common to least common?  
.....  
.....  
.....  
.....
11. List three utilities used to troubleshoot HP SANs and clusters.  
.....  
.....  
.....
12. What happens when more than one cluster or node resource is in an abnormal state?
  - a. A pager alert is automatically generated.
  - b. The parent cluster or node icon reflects the most severe state because the resource statuses are propagated upward.
  - c. The central management server administrator receives an email notification.
  - d. Systems Insight Manager restarts because the server in which it is installed is unreachable.

---

# **Business continuity planning and disaster recovery**

Module 6

## **Objectives**

After completing this module, you should be able to:

- Plan for business continuity.
- Identify HP solutions for fault tolerance.
- Design and implement an Enterprise Backup Solution strategy.

## Introduction

The Greater Environment for the Expansion of Knowledge (GEEK) offices on the island of Seahaven are in a small two-story building, in a bucolic valley not far from the coast. The servers and other network equipment are on the main floor and the research and support staff offices are in upstairs rooms.

Violent weather is common on the island in the summer and Fernando, the manager of the Seahaven office, monitored the storm system that was building in Truman Bay. Remembering the outcome of the previous tropical storms, Fernando took steps to prepare for the worst. Before leaving on Friday afternoon, the staff boarded the office windows and moved the desks to the middle of the room. They also unplugged all electrical equipment.

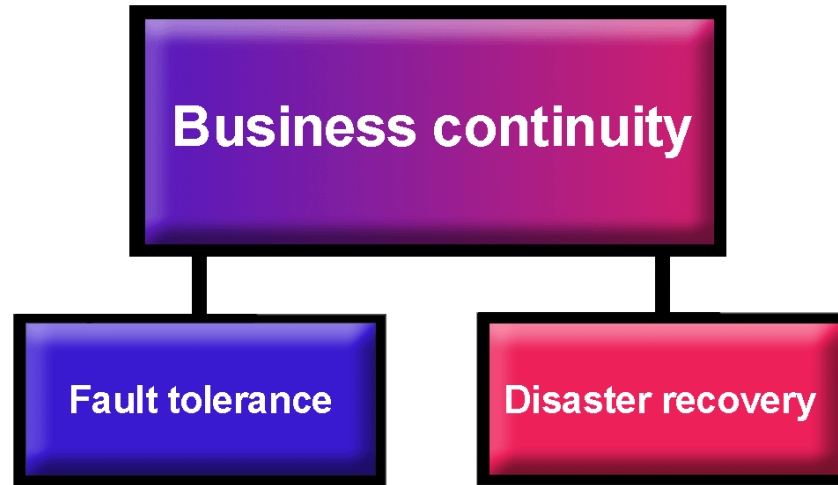
Overnight the storm developed into a Category 3 hurricane. Local law enforcement began evacuation procedures and closed the road to the offices. From his house a few miles away, Fernando watched the wind and rain bend the tops of palm trees to touch the ground.

The next morning Fernando waded through the parking lot to survey the damage. Hurricane Brent covered all the servers and UPS equipment with four feet of water; files and other business records of the company were also destroyed.

Luckily, the network administration staff had been diligent about keeping backups of company data. However, they never thought about backing up configuration and registry files. That means that after the company gets replacement equipment, the IT staff will have to reinstall and reconfigure the entire system before they can begin reloading the backed-up data.

Fernando and Carla, the managing director of GEEK, want to be sure that they will be better prepared if disaster should strike again. They need to send a report in to Bob, the CEO of RC Engineering, to show where the companies are most vulnerable to disasters and what their options are for minimizing those risks. This experience has taught Bob and Carla the difference between disaster tolerance—attempting to mitigate some of the causes of data loss—and disaster recovery—which involves picking up the pieces, much sadder but wiser.

## Business continuity

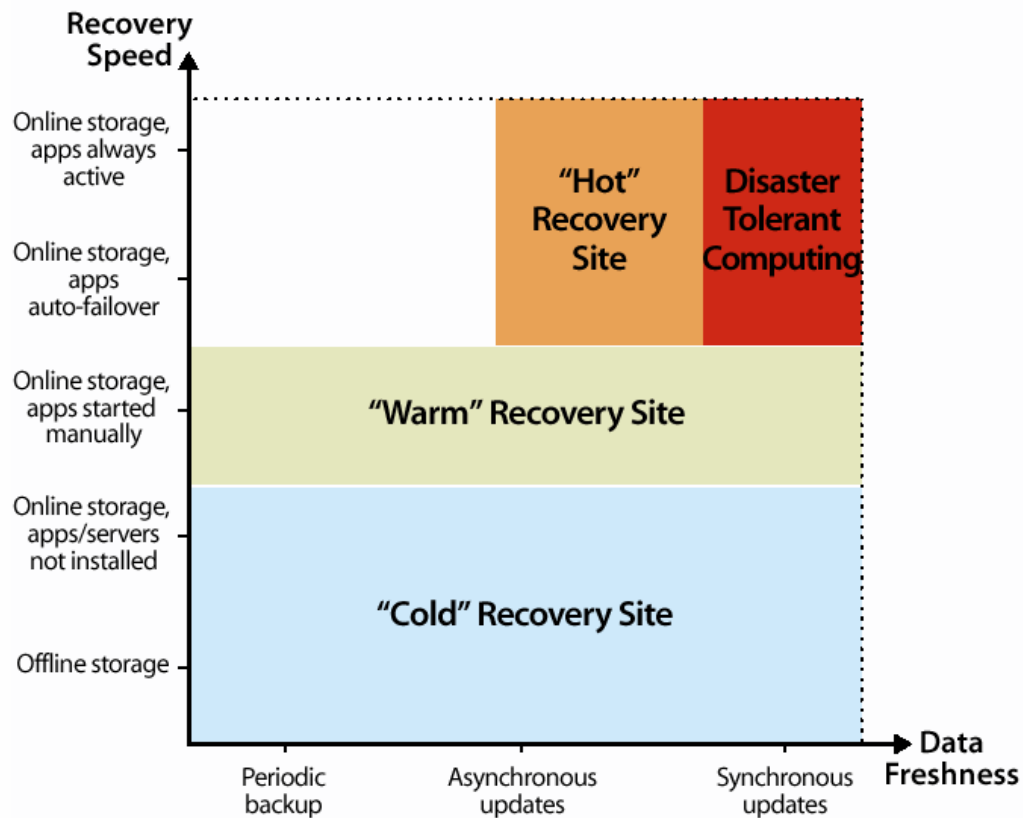


Business continuity is a broad term that encompasses all of the plans, resources, and actions used to ensure the continued operation of a business in the event of a site-wide outage, whether planned or unplanned. In the business world, computer disaster recovery planning is evolving toward the more proactive approach of business continuity planning. A thorough business continuity plan includes two main components:

- **Fault tolerance** — The ability of an organization's IT facilities to continue operations through an automatic hardware and software design that hardens the computing infrastructure to prevent a disaster. A fault-tolerant system must be able to detect the failure of a primary system, notify the people in charge, and (when authorized) proceed with a failover to redundant systems and sites without further manual intervention.
- **Disaster recovery** — A reactive plan that focuses on recovering the computing environment after the damage has occurred. A disaster recovery plan to return IT to an acceptable level of operation after a site-wide outage can include the use of servers, software, storage, networking, and staff duplicated at a remote site.

## Cont

## Disaster Recovery Continuum



Disaster recovery continuum

One way to approach a business continuity planning program is called the *Delphi method*. Experts in each business function identify their critical business processes and develop separate (but coordinated) continuity plans for each process. The benefits of this distributed approach are many:

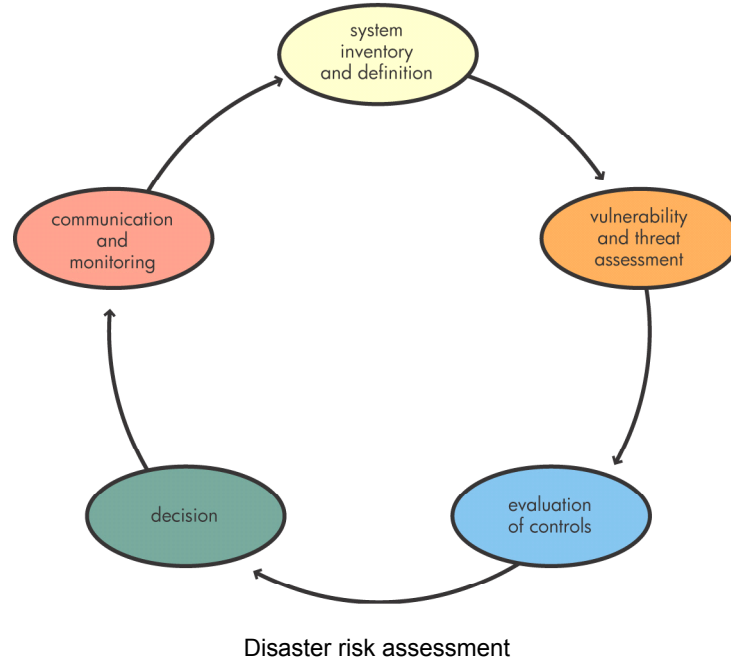
- Business processes that are not critical do not hinder those that are, so limited resources can be used effectively.
- An infrastructure that supports noncritical business processes does not get recovered.
- Multiple critical business processes or applications can be recovered in parallel.
- Applications that normally run on different systems can be recovered on the same system, if necessary.

## Beginning the continuity planning process

Developing a corporate continuity plan involves a number of steps as well as some basic concepts.

1. **Obtain management commitment** — The plan must be part of the strategic business plan and the company must budget appropriately and separately for the continuity planning program. A top-level policy statement should be issued that:
  - Affirms the value of business continuity
  - Acknowledges and accepts the associated costs
  - Documents management responsibilities
  - Includes the goals and expectations of the plan, as well as any organizational assumptions or parameters
2. **Identify critical business functions** — Identify how the company obtains its revenue in terms of business procedures. Communicate this information to management so they can confirm and rank the criticality of each business function.
3. **Build business process core teams** — Using the Delphi method, build teams consisting of IT operations management, end-user management, applications support staff for each critical business function, and the records management department. Through the Delphi teams, you develop a clearer view of the infrastructure (for example, processes, records, and IT applications) the teams believe is critical to performing their business functions.
4. **Build a corporate team** — The team should consist of members from core infrastructure support functions. The legal, public relations, and investor relations departments must also keep the public and stockholders informed of the company's operational status after an event has occurred.

## Risk assessment and business impact analysis



The purpose of a risk assessment and a business impact analysis is to answer the following questions:

- What am I trying to protect? (system inventory and definition)
- What am I trying to protect against? (vulnerability and threat assessment)
- What controls are currently in place or needed to prevent or minimize the effects of potential loss? (evaluation of controls)
- How much am I willing to spend on those controls? (decision)
- Is the money I am spending effective? (communication and monitoring)

The risk assessment involves identifying threats, vulnerabilities, risks, and the business impact of a disruption for each department or aspect of the business:

- **Threats** — Events that could disrupt an entity. Some examples are natural disasters (wildfire, flood, earthquake), disasters created by people (burst pipe, electrical fire), acts perpetrated by disgruntled employees, and mistakes.
- **Vulnerability** — Susceptibility to the threat (chances that an entity can be affected by a threat). For example, the closer a building is to an earthquake fault, the more vulnerable it is to an earthquake.
- **Business impact** — Loss of revenue, customers defecting to the competition, damaged reputation, or disgruntled employees if the company cannot pay them.

To start the risk assessment, rank all the departments and other components whose loss could negatively affect a business, gain consensus from each Delphi team, and then merge the results for presentation to upper management.

Remember that some threats have a time component. For example, a power failure that lasts a few minutes might not be a disaster, but one that lasts hours could be.

As part of the risk assessment, the Delphi teams estimate how long an entity can be unavailable, how old the information supplied by the entity can be, and how much of it might be lost. That is, they determine the:

- **Recovery time objective** — The time from when the event occurs until the business process (for example, the accounting department, the accounting application, or manual procedures used by accounting) must become active again (recovered).
- **Recovery point objective** — The point at which the data must be recovered—old or obsolete information no longer reflects the state of the company.

## **Application domains**

If a company is geographically dispersed and application uptime is imperative, you can create application domains at more than one site and distribute the load. With load balancing routers, redundant communication lines, or other methods, transactions can be split between multiple servers running in multiple sites.

When the load is being shared in this manner, you do not actually have primary and secondary systems or sites, but the beginnings of indestructible, scalable computing. New servers can be added at any time, and applications and database files can be migrated between the servers and sites as needed, so that any server or any site can be taken offline for testing, maintenance, or upgrades. In a properly designed application domain, a failure is undetectable, except perhaps for application slowdown.

## **Alternative plans and controls**

Because computer resources can be at a premium in a disaster, replacing computer-based processes with manual processes is an option that should be explored. Also consider such matters as alternative sites, temporary personnel, hotel and meal costs, off-site records and forms storage, and installation of new phone lines. The purpose of this step is to calculate the cost:benefit ratio of plans and controls for different recovery objectives.

When an actual crisis occurs, a crisis response team—composed of executive staff members and guided by one or more emergency response teams—determines if a disaster should be declared. The team considers the estimated time the affected business system will be unavailable, the recovery window, and the time it will take to execute the continuity plan. If, for example, the recovery window is two days and the system will be unavailable for only one day, it might not make sense to declare a disaster.

Declaration of a disaster needs to be reevaluated if the business system is down longer than initially expected, based on the time remaining before the it can be made available again and the time required to execute the continuity plan.

## Documentation and standards

Before developing a continuity plan identifying activities to be performed during a disaster scenario, you must understand how those same activities work every day. This means that you need access to some basic documentation for each business process. Determine if the following documentation exists:

- Change control process
- Standard operating procedures for users
- Run books
- Special forms requirements and special peripheral requirements for the operations staff
- Data flow diagrams and problem isolation procedures (for the business functions as well as the computer applications)
- Tape backup and rotation schedule used for records management

You should also gather more specific information about the business functions of the company. For example, find out if sufficient application downtime is scheduled to back up the databases used by the business functions, or, if the software allows it, if the databases are being backed up online. (If the company runs 24 x 7, you do not have time to take the system down for backups.)

Determine if there is an archival process to remove inactive records from hard-copy files and databases so that they are kept to a manageable size. Also, identify where critical records are being stored: on-site, off-site, or out of the region.

By having these documents and processes in place, the company will be more competitive in the long run as a result of:

- Faster isolation of application bugs
- Fewer operational mistakes
- Reduced support requirements
- Faster training of new staff
- Easier maintenance and enhancement
- Fulfillment of industry and legal regulations
- Easier auditing

## **Developing standards**

The continuity planner either develops or purchases a standard set of forms and procedures to be used by each function. Without standards, each business functions plan will look different, making coordination difficult.

At a minimum, this book of standards should contain:

- Step-by-step approach for each group to follow in writing the continuity plan
- Corporate team description, stating which corporate resources will be available to assist each business function in developing its plan
- Notification process
- Plan considerations
- Responsibility list

A responsibility list is a script or checklist, by job function, describing what each person will be required to do during the plan execution.

It is much easier to divide and assign the tasks necessary for recovery when the plan is being written, rather than in the middle of a disaster. The standards book should contain sample responsibility lists and layout hints. That tasks should be assigned by teams or by job titles and not to individuals.

## Writing the continuity plan

When developing the written continuity plan, ask employees within each business function to write the plan for that function. This achieves multiple goals:

- Employees know what the day-to-day activities are.
- Different functions can be generating plans at the same time.
- You gain buy-in of the business function employees.

Not only does each business function need to generate a plan, but each department represented on the corporate team needs to have a plan in case those departments are also affected by the disaster.

Although the plan can be written online, it must also be stored on paper to help ensure accessibility. Some continuity planning software vendors sell web-based services that make the plan available from any PC. However, if you do not have access to a PC or if the disaster also makes the vendor site inaccessible, you might not be able to access the plan.

Supply each business functions continuity planner and corporate team member with a binder containing a phone list with numbers for police, fire, ambulance, hospital, hazardous materials team, government authorities, and utilities. The binders should also contain the office, home, mobile, and pager numbers for each corporate team member, backup, and manager. Give each executive staff member a small packet for his or her briefcase, containing an emergency phone list, an executive staff phone list, and a “quick start” plan execution document.

## Exercising the plan

Continuity plans should not be tested—they should be exercised. Tests are passed or failed, whereas exercises are conducted for practice. Exercise the plans thoroughly to ensure that they work. During the exercise, note any problems that occur and encourage feedback from participants. The purpose of the exercise is to reveal any defective or missing components in the plans. It is counterproductive to reprimand someone for pointing out errors or omissions.

It is best to exercise and update the plans at least annually, or when major changes occur. Update call lists quarterly. An out-of-date plan provides a false sense of security and wastes time during an actual disaster.

---

**INTERNET**

For more information on business continuity planning, refer to the white paper *Developing a Business Continuity Plan* available from:  
<http://h71033.www7.hp.com/object/disrecsb.html>

---

## HP Services

With more than 50 global recovery facilities and experience helping clients recover from thousands of disasters, HP Services can set up business continuity solutions for a range of situations. HP Services offers a variety of products, solutions, and services for business continuity and disaster recovery, including:

- Fault-tolerant and highly available computer systems and storage modules
- Data replication software and hardware
- Outsourced (or internal) hot, warm, and cold-site planning, design, implementation, and operation
- Professional services for all phases of a planning program

HP Services can partner with you to:

- **Cost-effectively recover across distributed multivendor environments** — For HP and mixed-vendor platforms
- **Ensure nearly continuous availability** — Disaster-tolerant technologies and services to avoid data loss and minimize downtime
- **Maintain continuity amid IT and business change** — Ongoing protection with program management, business continuity audit, change management, and training services from HP
- **Comply with the latest regulations mandating business continuity** — For Basel II and HIPAA requirements, the Turnbull Report recommendations, and so forth
- **Safeguard enterprise application suites and strategic business processes** — Protection for applications such as SAP, BAAN, PeopleSoft, Oracle Financials, and Siebel, as well as processes in critical areas such as call centers, order entry departments, and trading floors

---

**INTERNET**

You can find out more about the offerings from HP Services at:

<http://www.hp.com/hps/continuity/>

---

## HP solutions for fault tolerance

Business continuity depends, in a large part, on technologies that keep a system in operation even if a fault occurs. Often these technologies involve adding redundant components that will take over if another component fails.

Data centers present special challenges for fault tolerance. Computer components grow more powerful, but this increased performance often requires tremendous amounts of electricity and generates more heat. Without the proper power, servers can fail and data can be lost. High temperatures can also damage data and components.

HP has determined best practices and offers targeted technologies to ensure fault tolerance in the data center, especially in the areas of:

- Thermal management
- Power protection
- Rack stability

In addition to making sure the hardware is protected, HP also offers data replication solutions so organizations can continue to operate if access to their data is interrupted.

### Thermal management

For HP ProLiant servers, managing thermal output is both an internal and an external process. Internally, ProLiant servers include fans to draw cool air over the heated components. HP engineers carefully consider air flow when they determine where to place components within a server. Many designs include baffles and heat sinks to help keep components cool.

Newer servers require more power to support high-performance components. More power generates more heat. In addition, servers are becoming more dense, which also affects the amount of heat generated.

Thermal management in the data center is expressed in tons of cooling. Many heating, ventilating, and air conditioning (HVAC) units can meet a particular cooling tonnage requirement, but the problem is getting the cool air where it is most needed in the data center.

---

**Note**

Whether designing a new data center or retrofitting an existing one, you should work with knowledgeable HVAC engineers to ensure adequate cooling.

---

## Determining the HVAC requirements

The heat load of the equipment in the data center determines the number and capacity of HVAC units required. Heat load is normally expressed in Btus/hour.

### INTERNET

You can find the heat load of the equipment in a data center by using the HP Site Preparation Utility at: [http://activeanswers.compaq.com/configurator/calc/Site Preparation Utility.xls](http://activeanswers.compaq.com/configurator/calc/Site%20Preparation%20Utility.xls)

After you know the data center heat load, you must determine the number of HVAC units you will need. The cooling capacity of all the HVAC units needs to exceed the data center heat load and is expressed in tons of cooling. Airflow is expressed in cubic feet per minute (cfm).

## Tons of cooling

One ton of cooling corresponds to a heat absorption rate of 12,000 Btu/hr. So if an HVAC unit is rated at five tons, it can absorb a heat load of 60,000 Btu/hr.

The tons capacity rating is measured at 80° F (27° C) and 0% relative humidity, but the recommended operating conditions for HVAC units are 70° to 72° F (20° to 22° C) and 50% relative humidity (RH). At 72° F (22° C), the cooling capacity of an HVAC unit is considerably reduced.

Furthermore, the tons rating is based on total cooling, which consists of “sensible cooling” and “latent cooling.” Computer equipment produces sensible heat only; therefore, the sensible cooling capacity of an HVAC unit is the most useful value. For this reason, HVAC unit manufacturers typically provide cooling capacities as “total Btu/hr” and “sensible Btu/hr” at various temperatures and RH values. To determine the required tons of cooling for a data center, use this formula:

Sensible cooling capacity (at the desired operating temperature and humidity)  
divided by 12,000 Btu/hr per ton

## Airflow

Airflow is measured in cubic feet per minute and is related to the moisture content of the air and the temperature difference between the supply air and return air. In general, a densely populated rack requires airflow of about 1,200 to 1,600 cfm.

## **Cooling footprint**

The cooling capacity of each HVAC unit determines its effective cooling area. The effective cooling area takes into account the cooling footprint of the equipment.

The floor area that each rack requires must include an unobstructed area to draw in and discharge air. Almost all HP equipment cools from front to rear so that it can be placed in racks positioned side-by-side and arranged in rows front-to-front and back-to-back to form alternating hot and cold aisles. The equipment draws in the cold supply air and exhausts warm air out the rear of the rack into hot aisles.

The amount of space needed between rows of racks is determined by the cooling footprint of the equipment. The cooling footprint includes width and depth of the rack plus the area in front for drawing in cool air and the area in back for exhausting hot air. Typically, a cold aisle should be least two floor tiles wide, or about 48 inches (122cm). A width of at least one unobstructed floor tile in the hot aisles is needed to facilitate cable routing. (Equipment that draws in air from the bottom or side or that exhausts air from the side or top will have a different cooling footprint.) The center of one cold aisle to the center of the next cold aisle should measure about 14 ft (4.3m), or seven full tiles.

## **Determining the placement of HVAC units**

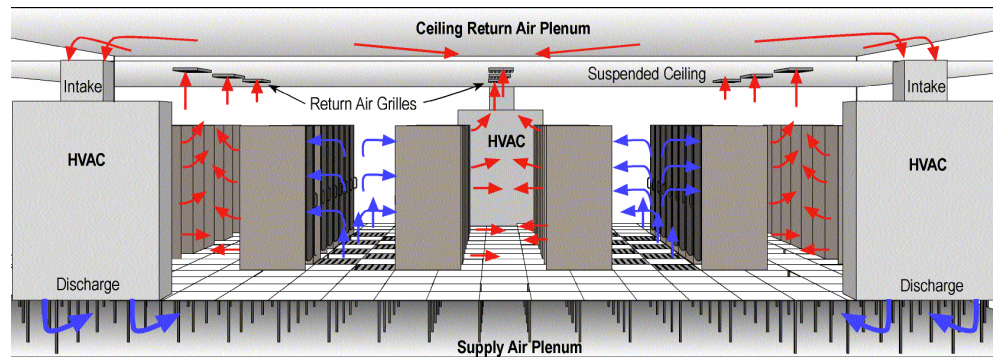
The geometry of the room and the heat load distribution of the equipment determine the best placement of the HVAC units. HVAC units can be placed inside or outside the data center walls. Customers should consider placing liquid-cooled units outside the data center to avoid damage to electrical equipment that could be caused by coolant leaks.

HVAC units should be placed perpendicular to the rows of equipment and aligned with the hot aisles. Rooms that are long and narrow can be cooled effectively by placing HVAC units around the perimeter. Large, square rooms might require HVAC units to be placed around the perimeter and through the center of the room.

## **Cooling range of HVAC units**

The effective cooling range of each HVAC unit is determined by its capacity and the heat load of the equipment in its airflow pattern. The most effective cooling begins about 8 ft (2.4m) from the HVAC unit. Although units with capacities greater than 20 tons are available, the increased heat density of newer servers limits the cooling range of these units to approximately 30 ft (10m).

## Airflow distribution



High-density data centers use a downward airflow pattern, which requires a raised floor configuration. The supply air plenum, or space between the building floor and the top of the tile, typically measures 18 inches (46cm) to 24 inches (61cm). Floor tiles are usually 24 inches (61cm) square and supported by a grounded grid structure. The plenum beneath the raised floor directs cool air to the racks.

In a downward airflow pattern, air currents are cooled and heated in a continuous convection cycle. The HVAC unit draws in warm air from the top, cools the air, and discharges it into the supply plenum beneath the floor. The static pressure in the supply plenum pushes the air up through perforated floor tiles in cold aisles. Ideally, the warm exhaust air rises to the ceiling and flows along the ceiling back to the top of the HVAC unit to repeat the cycle.

### Static pressure in the supply air plenum

For adequate airflow, the static pressure in the supply air plenum beneath the raised floor must be greater than the pressure above the raised floor. Typically, the plenum pressure should be at least 5% greater than the pressure above the floor.

The percentage and placement of perforated floor tiles are major factors in maintaining static pressure. Perforated tiles are classified by their airflow percentage. Airflow percentages vary from 25% (the most common) to 56% (for high airflow). A 25% perforated tile provides 548 cfm at a 5% static pressure drop and a 56% perforated tile provides 2,006 cfm.

Perforated tiles should be placed in front of at least every other rack. The high velocity discharge from the HVAC unit reduces the static pressure through perforated tiles nearest the unit causing inadequate airflow. The static pressure increases as the discharge moves away from the unit, thereby increasing the airflow through the perforated tiles.

To remedy this situation, use airfoils under the raised floor to divert air through the perforated tiles. Another option is to use a fan-assisted perforated tile to increase the supply air circulation to a particular rack or hot spot. Fan-assisted tiles can provide 200 to 1,500 cfm of supply air.

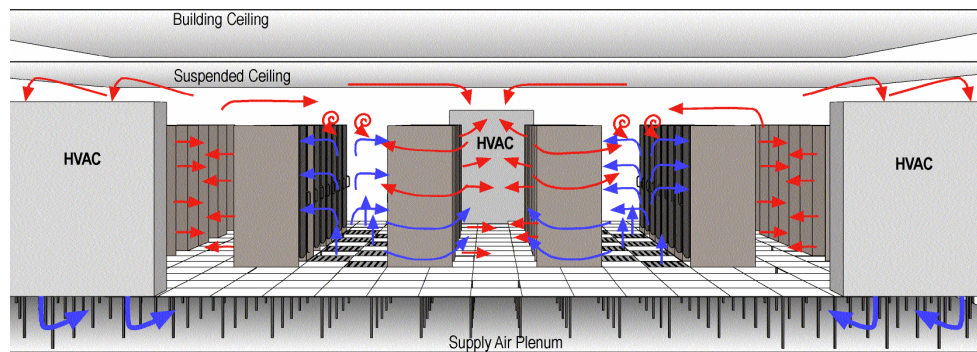
## Airflow blockages and leaks

The plenum is also used to route piping, conduit, and cables that bring power and network connections to the racks. In some data centers, cables are simply laid on the floor in the plenum where they can become badly tangled. This can result in cable dams that block airflow or cause turbulence that minimizes airflow and creates hot spots above the floor.

Use U-shaped “basket” cable trays or cable hangers to manage cable paths, prevent blockage of airflow, and provide a path for future cable additions. Another option is to use overhead cable baskets to route network and data cables so that only power cables remain in the floor plenum.

Electrical and network cables from devices in the racks pass through cutouts in the tile floor to wireways and cable trays beneath the floor. Oversized or unsealed cable cutouts allow supply air to escape from the plenum, thereby reducing the static pressure. Self-sealing cable cutouts are required to maintain the static pressure in the plenum.

## Airflow mixing



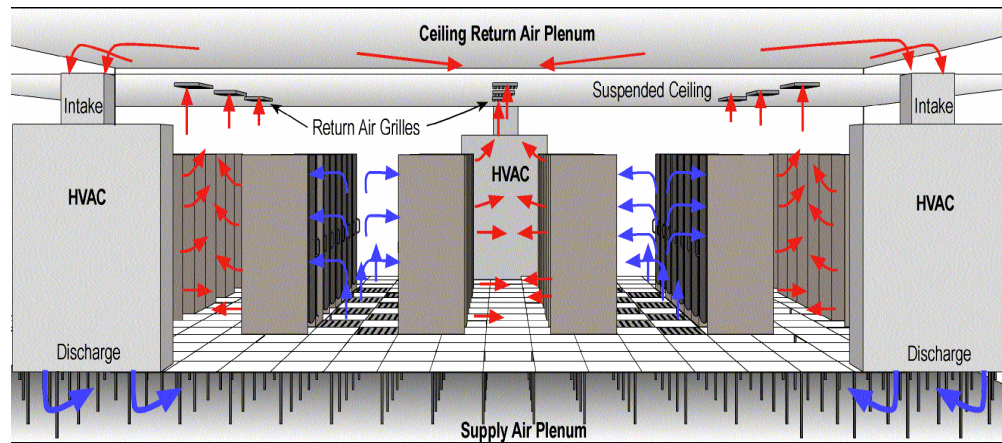
To achieve an optimum downward airflow pattern, warm exhaust air must be returned to the HVAC unit with minimal obstruction or redirection. Ideally, the warm exhaust air will rise to the ceiling and return to the intake of the HVAC unit. In reality, only the warm air closest to the intake can be captured; the rest can mix with the supply air. Mixing occurs if exhaust air goes into the cold aisles, if cold air goes into the hot aisles, or if there is insufficient ceiling height to allow for separation of the cold and warm air zones.

When warm exhaust air mixes with supply air, two things can happen:

- The temperature of the exhaust air decreases, thereby lowering the useable capacity of the HVAC unit.
- The temperature of the supply increases, which causes warmer air to be recirculated through computer equipment.

## Configurations for high-density data centers

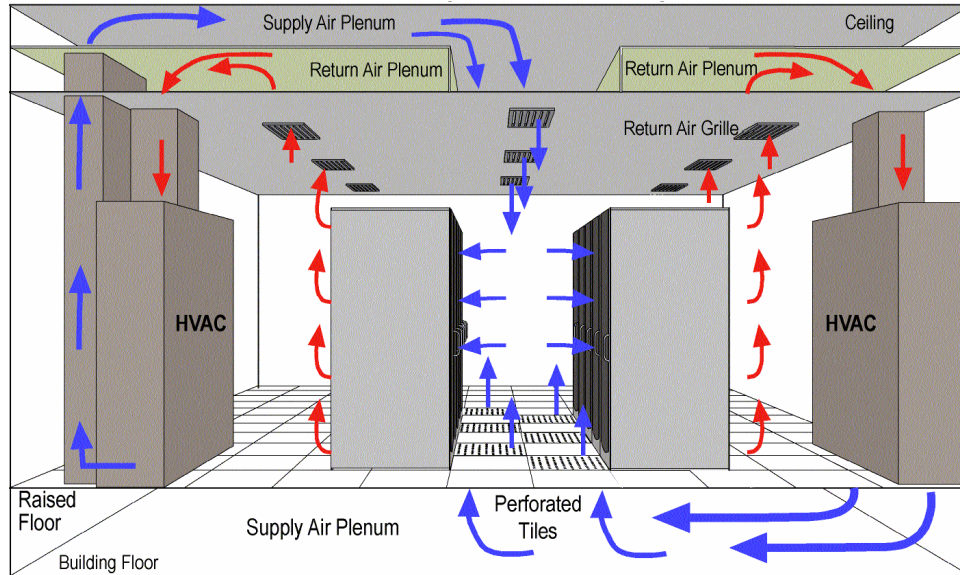
### Ceiling return air plenum



Raised floor computer rooms with very high heat density loads have begun to use a ceiling return air plenum to direct exhaust air back to the HVAC intake. The ceiling return air plenum removes heat and abates the mixing of cold air and exhaust air.

After the heated air is in the return air plenum, it can travel to the nearest HVAC unit intake. The return air grilles in the ceiling can be relocated depending on the layout of computer equipment.

## Dual supply air plenums



When power and heat densities climb, a single supply air plenum under the raised floor might be insufficient to remove the heat that will be generated. High-density solutions might require dual supply air plenums, one above and one below. In this configuration, additional supply air is forced downward in the cold aisle

### INTERNET

For more information on thermal management, refer to the *Power and Cooling Trends in the Data center* document available from:  
[http://wws1pro.compaq.com/support/reference\\_library/viewdocument.asp?source=tc030203tb.xml&dt=21](http://wws1pro.compaq.com/support/reference_library/viewdocument.asp?source=tc030203tb.xml&dt=21)

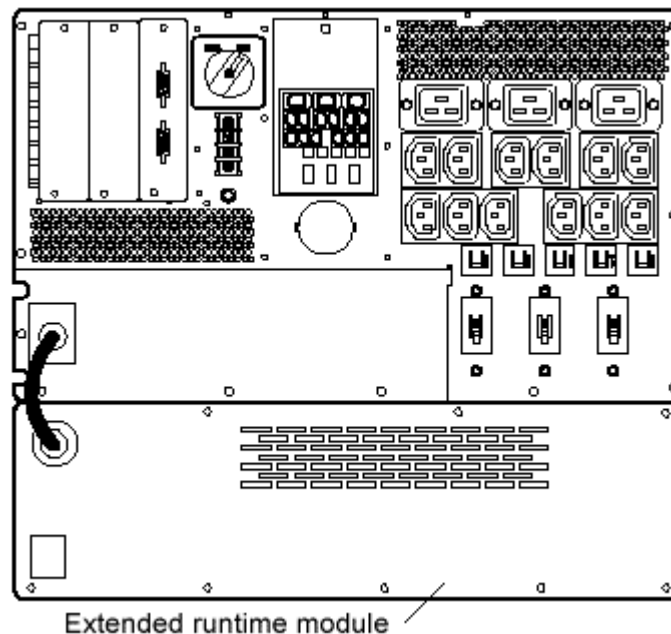
## Power protection

An HP UPS contains batteries that protect against power disturbances—natural or man-made—and circuitry that conditions (filters and enhances) utility power to provide more stable current.

If a UPS determines that the utility voltage is within the nominal operating range, the UPS supplies the utility power to the output receptacles.

If the utility voltage is outside the nominal operating range, the UPS supplies battery power to the output receptacles. In the event of a power failure, the UPS keeps equipment running for a few minutes to allow a graceful shutdown. The length of time a UPS can keep equipment running after a power failure is mainly determined by the size of its battery. The larger the battery, the longer the connected equipment can operate before shutting down.

### Extended runtime modules



HP makes modular UPSs that can be scaled to meet the demand for more power. Extended runtime modules (ERMs) are optional external battery modules that increase the available battery backup time in case of power failure.

## Enhanced battery management

HP UPSs employ enhanced battery management (EBM). EBM comprises three components:

- **Intelligent battery charging** — All UPS batteries need charging. Extended charging, however, significantly shortens battery life. The HP UPS uses a three-stage charging process that ultimately doubles battery service life:
  - a. First, the HP UPS rapid-charges the battery to 90%.
  - b. A constant voltage (float charge) continues until the battery reaches full capacity.
  - c. The charger is then powered down and the HP UPS goes into a rest mode, enabling the battery to be preserved for future power failures.

Most manufacturers use a trickle-charging method (a constant voltage feeding a low current to the battery) that dries the electrolyte and corrodes the plates. Batteries that are trickle-charged reach the end of their useful life in less than half the time of those charged using EBM technology.

- **Advance notification of battery replacement** — All batteries eventually fail. Because UPS batteries are valve-regulated, sealed, lead-acid cells, there has not been a practical way to provide users with advance notification of battery failure. The only way to determine that the batteries needed to be replaced was to wait until the power failed, taking the servers and computers down with it.

EBM is the only technology available that reliably provides advance notification of battery failure. The HP UPS microprocessor tracks the charge and discharge characteristics of the battery. These characteristics are compared to an ideal battery state. By monitoring the battery, the user receives prior notice when battery replacement is necessary.

- **Superior voltage regulation** — Most third-party UPSs correct input voltage variations as low as -25% but transfer to battery when a surge or a sag needs to be filtered in the system. This type of voltage regulation shortens the battery service life of the UPS.

Innovative buck/double boost voltage regulation ensures consistent input voltage to the load by automatically “bucking” it if it is too high, or “boosting” it if it is too low. Voltage variations as low as -35% or as high as +20% of nominal voltage are corrected—without transferring to battery. As a result, the number of charge/recharge cycles is reduced and the life of the HP UPS battery is extended.

## Load sharing

### N+x parallel redundancy

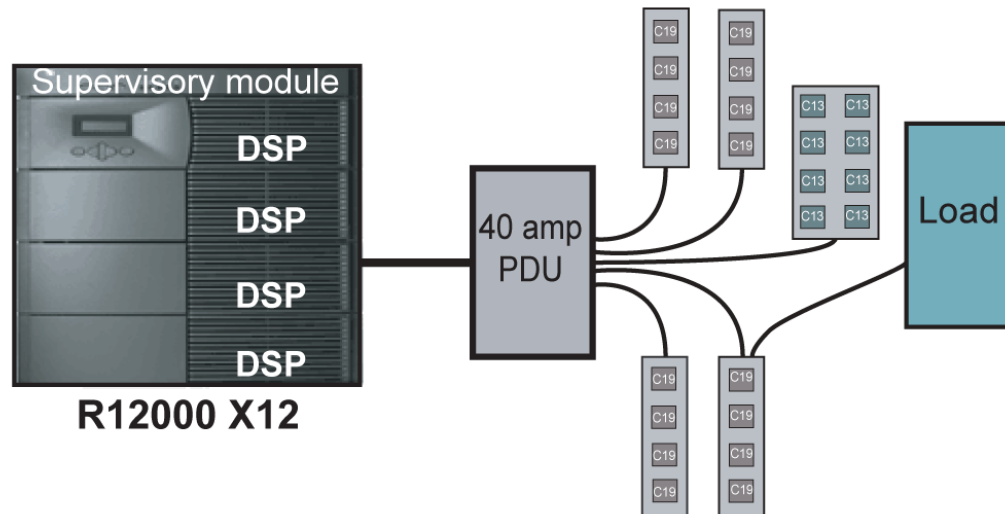
Load sharing is a process in which multiple battery modules share the connected electrical load. HP employs N+x redundant power-sharing technologies. If any module fails or is removed, the other modules take over its load. This ensures maximum uptime and continuous availability.

#### Example

The HP R12000 XR UPS supports four 3000VA modules. These modules can be configured four different ways.

Total voltage	Redundancy	Explanation
3000VA	N+3	One module carries the load. Three modules are standby in case the first one fails. This configuration can handle three separate UPS module failures.
6000VA	N+2	Two modules share the load. Two modules are standby in case both active modules fail. This configuration can handle two separate or simultaneous UPS failures.
9000VA	N+1	Three modules share the load. One module is standby in case any of the active modules fail. This configuration can handle a single UPS failure.
12000VA	N+0	All four modules share the load. There is no redundancy, so this configuration cannot handle any UPS failure.

## Wireless paralleling



When battery modules load share, they constantly monitor each other to stay in parallel. If one UPS detects that it is not functioning properly, it disconnects itself from carrying its share of the load, which transfers the entire load to the remaining UPS modules. Traditionally, communication between the modules occurs through a series of intricate wiring; however, these wires can be a single point of failure.

Wireless paralleling, available on some HP UPSs, enables redundancy without inter-module communication. Using a patented technology, a sensor in each UPS monitors the output power waveform of the other UPSs and an algorithm in the firmware keeps the UPSs in parallel.

## Unity power rating

When the watt rating of the UPS is equal to its amp rating, it is said to have a *unity power rating*. Unity-rated UPSs have improved power flow, thermal management, and increased efficiency.

### Example

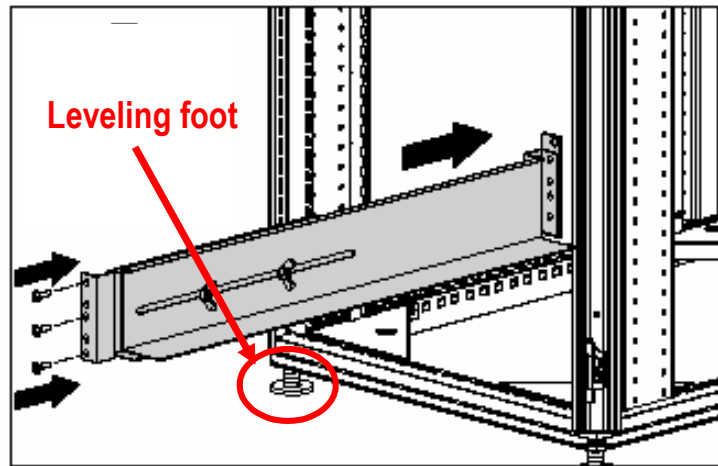
The R12000 XR UPS is a unity-rated, 12000 VA/12000W UPS. The R12000 XR has a unity power rating of 12kVA and 12kW.

Running at a unity power factor means that all of the electrical energy is producing real work and none of it is being wasted. A load with a 1.0 power factor results in the most efficient loading of the supply; however, a load with a 0.5 power factor results in much higher losses in the supply system.

## Rack stability

Rack stability is of special concern when equipment is routinely installed, removed, or accessed within the rack. Racks that are unstable can tip over, injuring people and damaging data. Stability is addressed through the use of leveling jacks (also called *leveling feet*), fixed stabilizers, and deployable stabilizers or ballast.

### Leveling jacks



Leveling jacks, located beside each caster on the rack, unscrew and extend to the floor, resting in leveling-foot bases provided with a rack. These feet support the rack and help compensate for uneven surfaces.

After positioning the rack in its final location, use an adjustable wrench to extend the leveling jacks into the bases until the weight of the rack is fully on the jacks and feet bases, not the casters. This stabilizes the rack for installation of components.

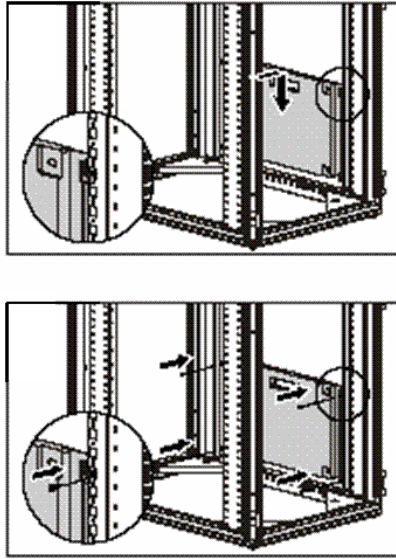


#### **WARNING**

To reduce the risk of personal safety or damage to the equipment, be sure that:

- The leveling jacks are extended to the floor.
- The full weight of the rack rests on the leveling jacks and not on the casters. The casters are designed **only** as an aid in moving the rack into position. They are not designed to support the weight of the rack and can become damaged if relied on to support the rack.
- In single rack installations, the stabilizing feet are attached to the rack.
- In multiple rack installations, the racks are coupled.
- Only one rack component is extended at any time. A rack can become unstable if more than one rack component is extended for any reason.

## Ballast kits



You can add ballast kits to rack cabinets to increase side-to-side and front-to-back mechanical stability.

Lightly loaded 9000 and 10000 series cabinets can require ballast to keep them from tipping over when a force is applied to the side of the cabinet. To ensure rack stability, the minimum weight of the installed equipment should be 210 lb (95kg). Ballast kits should be added in the event a system has less than 210 lb minimum weight.

### Example

If you have only 50 lb (23kg) of equipment, you would add two ballast kits. Each kit contains two 40-lb (18kg) ballasts. Two ballast kits (a total of four ballasts) equal 160 lb (72kg), bringing the total system weight to 210 lb.

In addition, if any single piece of equipment weighs more than 100 lb (46kg), there must be at least 200 lb (91kg) of additional equipment installed to maintain stability when the piece of equipment is extended on its rails. Add ballast kits if the additional equipment is less than 200 lb (91kg).



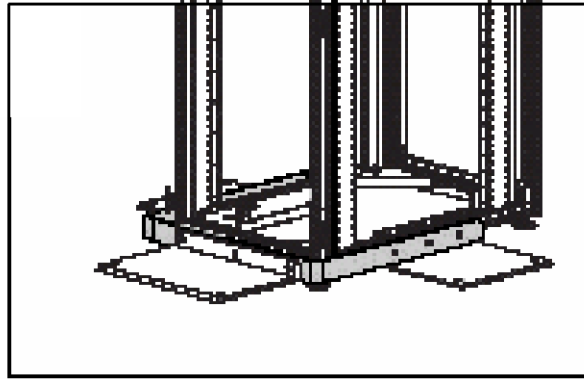
---

### WARNING

To reduce the risk of personal injury or damage to the equipment, extend only one component at a time. The rack may become unstable if more than one component is extended.

---

## Fixed stabilizer



An anti-tip stabilizer provides stability and support when equipment is installed, removed, or accessed within the rack. HP recommends that you use a stabilizer option kit, available in both a 600mm and 800mm version, with a stand-alone rack.

## Additional guidelines

- Use the HP Customer Builder to help determine where to install components.

---

**INTERNET** The Customer Builder utility is available from:  
<http://h30099.www3.hp.com/configurator/eco-cb/custombuilder.asp>

---

- Load the heavier components first and load the rack from the bottom up. When coupling racks, be sure to balance the weight load between the racks, placing the heaviest components at the bottom. For example, if you have several UPS units and several servers, do not put all of the UPS units into one rack—distribute them evenly in the bottom positions of each rack.
- Before working on the rack or extending a component on rails, be sure that the leveling feet extend to the floor, the full weight of the rack rests on the feet, and the rack is level and stable. Also, install stabilizing brackets or extend deployable stabilizers on a single rack or bay multiple racks together before starting work.
- Extend only one component at a time. The rack can become unstable if more than one component is extended. If a deployable stabilizer is installed, extend this part before sliding any component forward for access. When work is complete, replace the deployable stabilizer.
- Allow a minimum clearance of 30 inches (762mm) at the back of the rack.
- Allow a minimum of 48 inches (1219.2mm) clearance beyond the front of the rack to permit server installation and removal. This applies to both individual rack installations and when aligning rack rows so that the front doors are facing each other.

## HP solutions for disaster recovery — Enterprise Backup Solutions

HP offers HP Enterprise Backup Solutions (EBS) to facilitate the challenge of backing up the enormous volume of data stored on a corporate network. EBS combines a SCSI-based automated tape library with Fibre Channel technology and storage management software to create a consolidated data protection solution. This combination enables all backup data movement to be taken off the network backbone and placed on a separate Fibre Channel loop or fabric.

HP delivers EBS software for the following environments:

- Microsoft Windows NT/Windows 2000
- Novell NetWare
- HP Tru64 UNIX
- HP-UX
- Sun Solaris
- IBM AIX

### Recognizing a suitable environment for EBS

EBS is well suited in environments that employ five or more servers with more than 20GB of storage per server. These servers might be used for email and groupware, with databases or as data marts, as clustered servers, or as web servers.

#### **Example**

Microsoft Exchange Server databases historically have been limited to 16GB per server. Although the latest version of Exchange supports large databases, many Exchange users still use multiple servers in the 25GB to 30GB range rather than one large server. Managing the backup and restore on these servers can be cumbersome when performing local backups of each server.

Using EBS, customers can have a large tape library connected through a Fibre Channel storage network to up to 20 servers running Exchange Server. The result is a centrally managed backup of all their servers with performance equal to a local backup.

To determine if EBS is suited for a particular customer environment, begin by asking the following questions. Customers who answer “Yes” to at least one of these questions are candidates for an EBS.

- Is the company moving toward automation?
- Is data growing rapidly?
- Is there a shrinking backup window?
- Is there talk about centralizing storage and data management?
- Is scalability an important criterion?
- Does the company currently have business-critical applications running on NetWare, Windows NT/Windows 2000, Tru64 UNIX, Sun Solaris, AIX, or HP-UX servers?
- Are data administration costs rising?
- Are there LAN servers into the data center?
- Is the company implementing data marts or decision support applications?

### **Performing a needs analysis**

To gain a better understanding of the customer’s network topology, ask questions about the servers, storage systems, and network settings. Categorizing the data by its importance will help you make decisions about how and when backups are performed.

Understanding how important it is to retrieve the data quickly will also help determine if near-line storage — which uses sequential access tape media as well as CDs, DVDs, and random access magneto-optical devices — offers a practical solution. Near-line storage is an economical, scalable method of storing large volumes of data, although the data can be slower to retrieve than direct attached or network attached storage.

## Servers

- Which operating system is installed on each server?
- What type of host bus adapter (HBA) is being used?
- Which disk controller is being used?

---

### Note

The rate at which a controller can read data from the drives and transfer it to the tape library will determine how many and what type of controllers are required.

---

- Is the server part of a cluster?

## Network

- How fast is the LAN?
- Backup speed is generally limited to the capabilities of the LAN as well as the I/O capabilities of the server. This speed has a direct correlation with the backup window. The following table indicates maximum and typical LAN speeds, which can be regarded as upper limits to backup performance.

LAN type	Maximum speed	Typical speed
10Base-T	3.6GB/hr	2GB/hr
100Base-T	36GB/hr	15–20GB/hr
Gigabit Ethernet	360GB/hr	36–100GB/hr
FDDI	Similar to 100Base-T	Similar to 100Base-T
Fibre Channel	720GB/hr	560GB/hr

- What kind of interconnects will you be using?

## Volumes

- What is the feed speed of the hard disk system? This speed will help determine how quickly a volume can be backed up.
- What is the data compression ratio for the volume? The type of data contained on the volume greatly impacts compressibility, backup speed, and the number of tapes required to hold the data. Typical compression ratios are shown in the following table.

Data type	Compression ratio
Typical file/print server	2:1
CAD	3.8:1
MPEG/JPEG	1:1
Spreadsheet/word processing	2.5:1
Oracle/SAP databases	1.2:1
Microsoft Exchange/SQL databases	1.4:1
Lotus Notes database	1.6:1

---

### Note

These figures are typical and were not tested by HP. Your results might vary, depending on the actual data being compressed.

---

When compressed data is passed through a compression process a second time, it typically occupies more space rather than less because of the overhead of the compression algorithm. This fact can be important when calculating the number of tapes needed.

- How large are the files? As file size decreases, so does the tape backup rate because of the overhead of cataloging a higher number of files.
- How much data does each volume hold? How much do you expect the volume to grow in a year (in percentage)? You must consider the present data volume and anticipate the growth rate to arrive at the best storage solution.

## Data

- Which data is business-critical? Business-critical data, such as sales transactions, have precedence over user data. Volumes containing this information should receive preference during the backup window.
- Should differential backups be retained between full backups? A differential backup records all the changes since the last full backup, so there is redundancy in the data backed up during consecutive partial backups. The value of having this redundancy must be weighed against the cost of the greater storage capacity required.
- What data domains will your backup software require? If necessary, you will create backup domains based on the backup window, data importance, amount of data, and capacity of backup devices.

### Example

If a customer is using VERITAS NetBackup, the maximum number of servers allowed in the data zone is 32. If the customer wants to include 50 servers in the EBS, you will need at least two data zones.

- How much data will you need to back up? Calculate the entire amount of data to be backed up for the network, each server, and each volume.
- Which data will be in which backup jobs? It is important to group data into different jobs to provide more data security in case of a tape drive or other failure during the backup job. Grouping data also provides higher backup performance. With one job per tape drive, several drives working at the same time can complete the backup much faster.

Load balancing is also enabled when the data is grouped into jobs. Using drive pools, the backup software selects the drives for the individual backup jobs, ensuring that no one drive is handling the majority of the backup.

## Tapes

- How much will you be available to spend? Evaluate drive performance and drive capacity in view of the current drive and media costs and the company budget.
- How reliable do the tapes and devices need to be? The reliability of a backup device is related directly to its duty cycle (the number of hours per day that the device is in use). For example, if a tape drive designed for 1GB backups is being used to back up a 10GB server, that drive will be subject to premature aging and reliability problems.

The best method for building hardware reliability into a backup strategy is to ensure that the backup hardware is matched to the servers. The following table relates various servers with the appropriate tape drives.

Server capacity	Recommended drive format	Recommended drive
<20GB	Digital audio tape (DAT)	20/40GB
<50GB	Advanced Intelligent Tape (AIT)	50GB
>50GB	Super digital linear tape (SDLT)	110/220GB or 160/320
	Ultrium	Ultrium 230 or Ultrium 460

Another factor to consider is media life. The following table lists the typical life for the various media discussed in this module.

Media type	Rated passes	Rated backups
DAT	2,000	300
AIT (50GB)	20,000	10,000
SDLT I	1,000,000	17,850
Ultrium 200GB	1,000,000	20,000
Ultrium 400GB	1,000,000	20,000

### Note

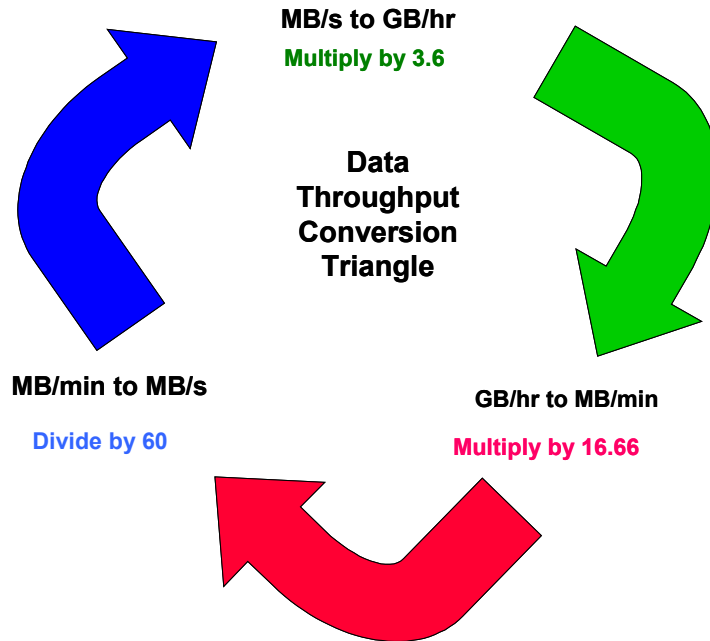
The backup rating was calculated under optimum environmental conditions of 50% relative humidity and 72° F (22° C). The rating also assumes that drive is streaming and that the entire tape capacity is used for each backup.

- How much capacity will you need? Unattended backups can significantly reduce the administrative costs. Deploying sufficient backup capacity and automation, if needed, to make Lights-Out backups possible is often a significant cost saver for the long term.

For unattended backups, the tape drive and media must meet the backup requirements in terms of overall storage capacity and backup performance.

- What type of tape technology will you use? Performance and capacity vary with the type of drive technology selected. The newer technologies have the highest capacities and fastest transfer rates.
  - DAT drives are usually the appropriate choice for servers with 4GB to 20GB capacity. An autoloader for DAT tapes is available from HP. Autoloaders reduce administrative costs by using a robotic mechanism to load and unload tapes.
  - DLT drives use simultaneous multichannel and multihead read/write technology to achieve capacities up to 40GB without compression. A DLT drive is the appropriate high-end backup solution for larger servers, especially those with storage of 40GB or more, owing to a wide range of tape array, mini-library, and large library options.
  - SDLT drives offer backward-compatibility with DLT tapes and increase storage capacity and transfer rate by an order of magnitude. Storage capacity starts at 110GB without compression. The SDLT drive is appropriate for enterprise-class servers, owing to its capacity and transfer rates, as well as improved reliability over DLT.
  - Ultrium drives offer outstanding reliability, excellent capacity, and data rate matching technology to optimize performance. Cartridge memory helps improve media management and reduce wear. HP Ultrium 2 drives are based on the Linear Tape-Open (LTO) standard to minimize network interruption and store up to 400GB on a single piece of media.

## MB/s to GB/hr conversion formulas



Tape drive capacities and throughput figures are often represented in gigabytes. Primary storage throughput figures are typically represented in megabytes. In the world of enterprise storage, it is important to learn the formulas for converting megabytes gigabytes. The preceding graphic demonstrates how to calculate this equation.

### Example

$$5\text{MB/s} \times 3.6 = 18\text{GB/hr}$$

$$18\text{GB/hr} \times 16.66 = 299.88\text{MB/min}$$

$$299.88\text{MB/min} \div 60 = 4.998\text{MB/s}$$

## Determining the backup strategy

After you understand the network topology, you must decide how to back up the data. Consider asking the following questions:

- What types of backups will you want?
  - Full backups only — All data will be backed up. The backups take longer, but the recoveries are faster.
  - Full and incremental backups — Incremental backups record the changes since the last backup (full or incremental). The backups are faster, but the recoveries take longer because all incremental backups must be restored.
  - Full and differential backups — Differential backups record all the changes since the last full backup. The backups take longer than incremental backups, but the recoveries are faster (and safer) because all data is restored from only the most recent backup.

The importance of the data helps determine the types of backups to perform. The type of backup you choose affects not only the speed of the backup and restore process, but also the amount of data being retained on tape.

- How many backup sets will you keep?
- When will backups be performed?
- What is the tape retention schedule?
- If performing partial backups, where will they be retained?
  - On the same tape set as the corresponding full backup.
  - On a tape set separate from the full backup.
  - On its own tape set (each partial backup)
- Will partial backups be differential or incremental?
- How long will partial backups be retained? If partial backups are retained on the same tape set as full backups, they must be retained as long as the full backup.

## Backup window

After identifying backup requirements, determine the backup window, which is the block of time spent performing the actual backup. If the company is performing an offline backup, the window is determined by the amount of time the company can afford to take servers offline each night.

With a Fibre Channel system, users can access information even during backups. Network bandwidth is not reduced during backups because there is a direct path between the server and the tape backup solution. Depending on the backup software being used, however, they might not be able to access files that are currently being backed up.

If immediate and constant access to files is critical, customers might need to plan for a dedicated backup window. In this case, the configuration of the solution will be based partly on the acceptable backup window.

In the following table, four HP servers are chosen as examples to illustrate how the required transfer rate formula can be used to choose a backup tape device.

Server	Amount of data to back up	Backup window	Suggested backup device
ProLiant ML350	25GB	7 hours	20/40GB DLT
ProLiant ML370	30GB	2 hours	Ultrium 230
ProLiant DL740	80GB	2.5 hours	MSL5060L1
ProLiant DL760	160GB	4.3 hours	MSL5060L1

When identifying the backup window, ask these questions:

- Can the volume to be backed up be taken offline?  
Some volumes contain data that must remain available at all times. For example, the database for an online store cannot be taken offline during backup or customers will be unable to make purchases during that time.  
Plans need to be made to keep this data available during backups. Possibilities include using tools such as the StorageWorks Virtual Replicator or Enterprise Volume Manager (EVM).
- How large is the backup window?

## Refining your backup and restore process

Many companies spend hundreds of hours refining their backup and restoration procedures. Some methods of decreasing your recovery time are as follows:

- Separate and back up files in the order that they need to be restored. One such grouping could be operating system, application, database files, and transaction logs.
- Pull and ship tapes from your off-site storage in “waves” so that the most critical files arrive sooner. When the next set of tapes is being pulled, the first set can be on its way to the recovery site. If you have duplicate tapes, ship them separately so that a transportation mishap does not derail your recovery.
- Determine how many tape drives are required at the backup site to meet your recovery time objective. A typical large company might use 64 drives in parallel.
- Tape is inexpensive. Filling tapes to save money could prevent you from recovering in the allotted time because multiple restore threads might need different files on the same tape. A company can cut its file restoration time by 50% by optimizing tape usage.
- If at all possible, automate your recovery tasks so that mistakes are minimized. A typical company can have 800 batch jobs that restore the application and database from tape, roll the database forward with the logs, and check consistency.
- Follow industry best practices to ensure data integrity by restoring the backed-up data on a test server and comparing it to the source data.

## Selecting backup devices

Selecting a backup device is a four-step process:

1. Calculate the compressed transfer rate — Begin with a 1:1 compression ratio. Then apply the compression formula.
  - a. For DLT drives — For a 1:1 compression ratio, the compressed rate equals the base rate times compression ratio. For compression ratios greater than 1:1, the compressed rate = (Base rate x compression ratio) – (compression ratio)<sup>2</sup>

### Example

$$(15 \times 2) - 22 = 26\text{GB/hr}$$

---

### Note

The maximum compressed rate for a 35/70 DLT drive is 43GB/hr. If the calculated compressed rate comes out to more than 43GB/hr, you must back this number down to 43GB/hr.

---

- b. For LTO and AIT drives — Base rate x compression ratio = compressed rate.

### Example

$$20 \times 1 = 20\text{GB/hr}$$

---

### Note

The maximum compressed rate for an AIT 50 drive is 45GB/hr. If the calculated compressed rate is more than 45GB/hr, you must back this number down to 45GB/hr.

---

2. Adjust the compressed rate for feed speed and lights-out operations.
  - An internal SCSI controller can feed approximately 60GB/hr.
  - HP Smart Array controllers can feed approximately 60GB/hr to 200GB/hr.
  - The Modular Data Router can feed approximately 100GB/hr (up to 16 drives). The Network Storage Router M2402 can feed approximately 600GB/hr.
  - Tape drives can handle a 3:1 feed-to-write ratio; a DLT drive can write information at 15GB/hr, which requires a 45GB/hr feed rate for maximum performance. A 3:1 feed ratio uses 100% of the write capability of the drive; a 2:1 ratio reduces write speed to 80%.
    - ♦ If the feed speed is 3:1 over the compressed rate, leave the compressed rate as is.
    - ♦ If the feed speed is more than 2:1 but less than 3:1, reduce the compressed rate by 20%.
    - ♦ If the feed speed is less than 2:1, reduce the compressed rate by 50%.
  - If full backups are more than one tape capacity, reduce the compressed rate by 3%.
  - If performing incremental or differential backups, reduce the compressed rate by 14%.
3. Determine the number of drives needed to meet the backup window.
  - The total data divided by the adjusted compressed rate equals the total hours of backup for one drive.
  - Hours of backup for one drive equals the total data divided by the adjusted compressed rate.
  - The total hours for one drive divided by the backup window equals the number of drives needed.
  - Number of drives needed equals the hours of backup for one drive divided by the backup window.
  - If performing full backups, the mechanical process of changing tapes could reduce the base backup rate by up to 7%.
  - Incremental backups have slower backup rates than full backups because all the data on the system must be read to determine which files are new or have been modified since the last backup.

4. Determine the quantity of retained media.
  - a. Calculate the number of full and incremental or differential data sets to be retained.
    - 1) Each tape holds up to 110GB (SDLT), 40GB (DLT), or 50GB (AIT) of data. As the data compression ratio increases, effective tape storage capacity increases.

**Example**

At 1:1 compression, a tape can store 40GB of data; at 2:1 compression, it can store 80GB.

- 2) Total storage capacity is affected by the data compression ratio, type of backups, and so forth.
  - 3) If your retention parameters demand more backups, you will require additional tapes, which could increase the number of libraries required.
- b. Assign percentage values to each backup set. Use a full backup as the benchmark and assign it a value of 100%. Partial backup sets will have a smaller value, based on the amount of data that has changed or been added since previous backups.

If the company estimates that 10% of its data files are changed or added to in some way each day, an incremental backup would have a value, on average, of 10%. If the company uses differential backups, the value for each set will be 10% times the number of days since the last full backup.

- c. Calculate the total retained percentage by calculating the value of the backup sets produced during the retention period.

**Example**

If the company performs weekly full backups (F) and no partial backups, and retains backup sets for three weeks before overwriting them, the total retained percentage would be:

$$F \times 3 = 100\% \times 3 = 300\%$$

If the company performs weekly full backups and incremental backups (I) on the other four days of the week, and retains them for three weeks, the total retained percentage would be:

$$(F + I + I + I + I) \times 3 = (100 + 10 + 10 + 10 + 10) \times 3 = 420\%$$

If the company performs weekly full backups, differential backups (D) on the other four days, and retains sets for three weeks, the total retained percentage would be:

$$(F + D + D + D + D) \times 3 = (100 + 10 + 20 + 30 + 40) \times 3 = 600\%$$

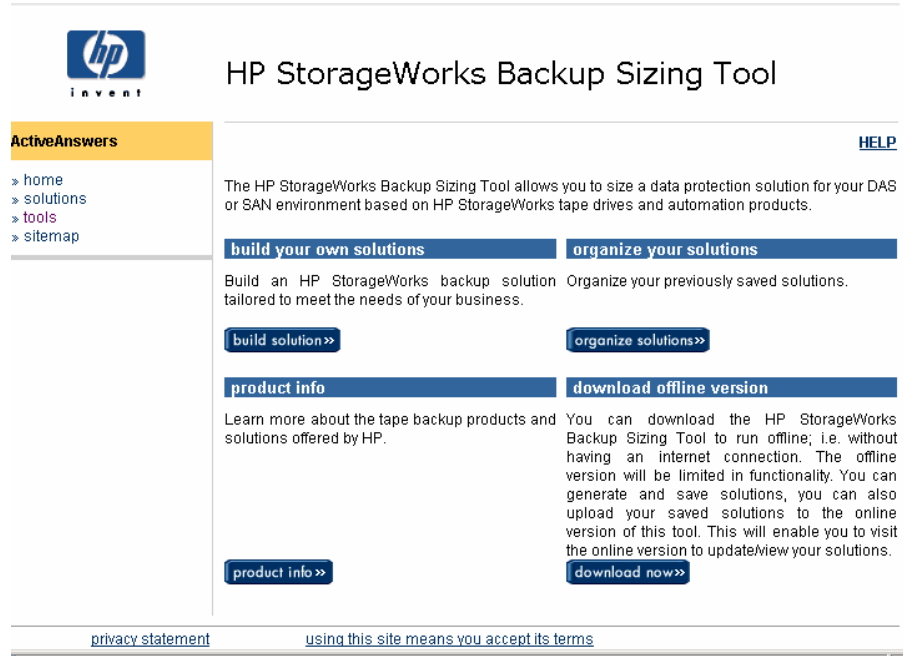
- d. Calculate the storage capacity of each tape.

- ♦ Tape capacity (SDLT) = 110GB x compression ratio
- ♦ Tape capacity (DLT) = 40GB x compression ratio
- ♦ Tape capacity (AIT) = 50GB x compression ratio

- e. Calculate the number of tapes required.

Number of tapes required = (total data x retained percentage) ÷ tape capacity

## Designing an Enterprise Backup Solution



After you have gathered information regarding the business processes and network environment of the company, use the HP StorageWorks Backup Sizing Tool to design a solution.

The Backup Sizing Tool takes basic parameters about the customer's tape backup environment—such as backup windows, amount of data to be backed up, tape rotation scheme, and number of servers—and produces a bill of materials for the recommended configuration of the EBS.

You are guided through the sizing process by a series of predetermined questions regarding such things as the applications being used by the customer, the system configuration, and the customer's requirements relating to price, performance, and capacity. Many rules, trade-off decisions, and assumptions must be considered to ensure successful results.

### INTERNET

The HP StorageWorks Backup Sizing Tool is available online at:  
<http://www23.compaq.com/SB/ETF/ETFSAUTH.ASP>

A stand-alone version for download is available at:  
<http://activeanswers.compaq.com/ActiveAnswers/Render/1,1027,6537-6-100-225-1,00.htm>

To use the Backup Sizing Tool, follow these steps:

1. Click the *Build Solutions* button.
2. Choose *Fibre Channel SAN Backup Solution*.
3. Configure the servers.
4. Configure the volume for each server.
5. Set up the backup schedule for the volume.
6. Repeat steps 3 through 5 until all the servers and volumes are configured.
7. Choose the interconnect that you will use.

The Backup Sizing Tool will display three recommended solutions. Recommended Solutions are divided into three categories:

- Good — This solution meets your needs and is cost-effective.
- Better — Growth is factored in and cost is secondary.
- Best — This solution offers the best performance.

Click the *View Solution* button next to any solution to see a list of the configured hardware, a suggested backup schedule, recommended StorageWorks products, and audit information, including usage data, on the interconnect and the tape devices.

## **EBS drivers**

Drivers are software that allows hardware components to run correctly. EBS requires drivers for the:

- HBA
- Tape drives
- Fibre Channel filter

### **HBA drivers**

Each HBA in the server is run by driver software. These drivers are located in the solution kit supplied with each solution. When you install an HBA, you must install the appropriate drivers for that HBA.

### **Tape drivers**

The tape drives in the library are run by tape driver software. If applicable, these drivers are located in the solution kit supplied with each solution.

### **Fibre Channel filter drivers**

Some SCSI commands to tape devices do not fit into the Fibre Channel protocol. The Fibre Channel filter driver allows those SCSI commands to be executed transparently over Fibre Channel. The filter driver is automatically installed when the HBA driver is installed.

## Supported backup software

A crucial component of HP EBS is the backup management software that enables all the hardware components to work together.

HP has partnered with several vendors to deliver an integrated, centralized, Fibre Channel storage solution. When configured with partner software, EBS enables multiple servers to share libraries. Supported backup software includes:

- HP OpenView Storage Data Protector
- Atempo Time Navigator
- Bakbone NetVault
- Computer Associates
- CommVault Galaxy
- Legato NetWorker
- Syncsort Backup Express
- Tivoli Storage Manager
- VERITAS Backup Exec
- VERITAS NetBackup

## Maximum supported configuration

In general, to configure backup software so that multiple servers are sharing one or more tape libraries, you must install a primary server. This primary server will manage the database that tracks devices, media, ownership, and other indices. The group that comprises the primary server, the tape libraries, and the other servers that share the tape libraries is often called a *data zone* or *backup domain*.

The members of this data zone can be members of different switch zones in the storage area network (SAN). However, for all servers in the data zone to access the tape library, the tape library must exist in each of these switch zones in the SAN or you must create a separate switch zone that contains the tape library and all of the servers that are part of the data zone.

In addition, depending on the specific ISV software, members of the data zone can also be running different operating systems.

Data zones have limited supported configurations. For example, with VERITAS Backup Exec, the maximum supported configuration is 32 servers sharing up to 32 tape drives in one data zone. If one were to exceed the maximum configuration, severe performance degradation or loss of communication could occur. A configuration involving fewer servers and tape drives might be necessary to achieve optimal performance.

The following table lists the maximum configuration supported by EBS.

<b>Backup software</b>	<b>Maximum configuration</b>
OpenView Storage Data Protector	16 servers and 16 tape drives
Atempo Time Navigator	16 servers and 16 tape drives
BakBone NetVault	16 servers and 16 tape drives
CommVault Galaxy	16 servers and 16 tape drives
Computer Associates Brightstor ARCserve	20 servers and 20 tape drives for NetWare 16 servers and 16 tape drives for Windows NT/Windows 2000
Computer Associates Brightstor Enterprise Backup	16 servers and 16 tape drives
Legato NetWorker	16 servers and 32 tape drives for Windows 32 servers and 32 tape drives for UNIX
Tivoli Storage Manager	16 servers and 16 tape drives
VERITAS Backup Exec	32 servers and 27 tape drives for NetWare 32 servers and 32 tape drives for Windows NT/Windows 2000
VERITAS NetBackup	32 servers and 32 tape drives

Although multiple servers can share multiple drives, only one server may control the robot at a given time and only one server may access a particular tape drive at a given time. With some backup management software, a primary server controls the robot while other backup servers send tape mount and dismount requests to this primary server.

In general, servers in a switch zone should be homogeneous in operating systems if possible. NetWare and Sun Solaris servers are incompatible in the same switch zone. Linux, HP-UX, and IBM AIX servers are each incompatible in zones with all other servers. The following table summarizes switch zone compatibility between different operating systems in a SAN. Platforms in the same column can coexist in the same switch zone.

<b>Zone 1</b>	<b>Zone 2</b>	<b>Zone 3</b>	<b>Zone 4</b>	<b>Zone 5</b>
NetWare	Sun	Linux	HP-UX	IBM AIX
OpenVMS	OpenVMS			
Tru64 UNIX	Tru64 UNIX			
Windows NT	Windows NT			
Windows 2000 Server	Windows 2000 Server			
Windows 2000 Advanced Server	Windows 2000 Advanced Server			

## OpenView Storage Data Protector

The OpenView Storage Data Protector utility provides reliable data protection and comprehensive recovery. The Storage Data Protector supports a comprehensive list of backup clients (disk agents) and backup device servers (media agents) to provide broad compatibility.

Storage Data Protector 5.0 is supported on HP-UX, Windows, and Sun Solaris platforms. A Disk Agent and Media Agent are required for each server sharing a library across a SAN.

---

**INTERNET**

For the most current information on the operating systems supported by the Storage Data Protector, visit: <http://support.openview.hp.com/>

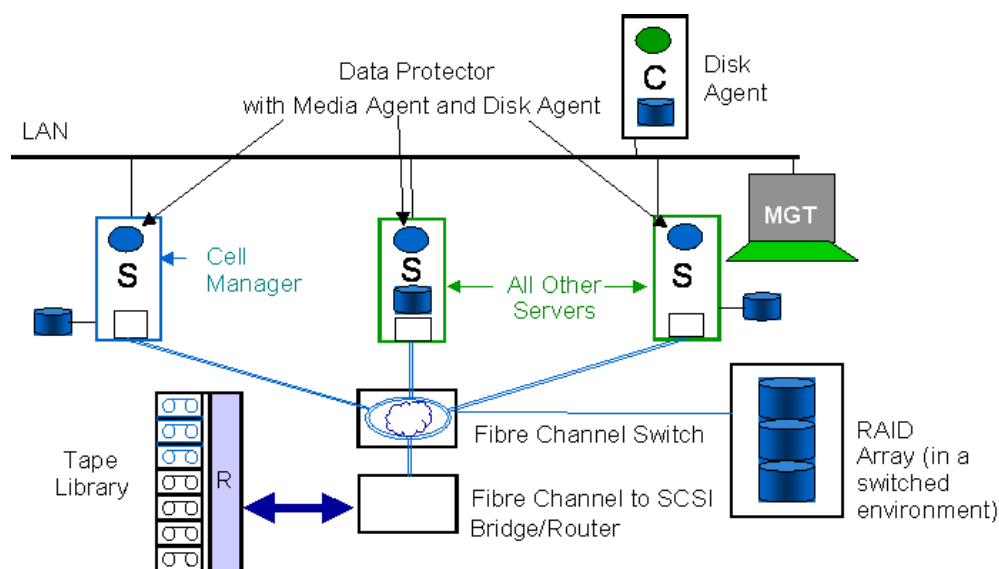
---

### Key features

Whether backup requirements are for small workgroups or a multisite heterogeneous data center environment, Storage Data Protector provides centralized or distributed management, depending on the needs of the organization. The modular design of Storage Data Protector provides scalability as storage environments grow. Other features include:

- **Easy central administration** — Through its easy-to-use graphical user interface (GUI), the Storage Data Protector allows the administrator to manage multiple backup environments from a single system. The GUI can be installed on various systems to allow multiple administrators to access the Storage Data Protector through locally installed consoles. A command line interface (CLI) allows you to manage the Storage Data Protector tool using scripts.
- **High-performance backup** — Storage Data Protector supports parallel data streams. Various types of backups, such as local, network, full, differential, leveled incremental, online, and disk image, provide flexibility in tuning the backups to best fit business requirements.
- **Heterogeneous support** — Storage Data Protector supports heterogeneous environments. The HP-UX, Windows NT, Windows 2000, or Windows Server 2003 Cell Manager can control all client platforms.
- **Monitoring, reporting, and notification** — Reporting and notification capabilities allow the administrator to view the backup status, monitor active backup operations, and customize reports. Reports can be generated using Java-based online web reporting, the Storage Data Protector GUI, or using the omnirpt command on systems running UNIX, Windows NT, Windows 2000, or Windows Server 2003.
- **Integration with online database applications** — Storage Data Protector supports online backup of Microsoft Exchange, Oracle, Informix, SAP, Lotus Domino, and Sybase database objects.

## Storage Data Protector topology



The Storage Data Protector cell is a network environment that has a Cell Manager, client systems, and devices. The Cell Manager is the central control point where Storage Data Protector software is installed. Client systems are then added to the cell to be backed up. Storage Data Protector backs up files to media in the backup devices.

## Cell Manager

## The Cell Manager:

- Manages the cell from a central point
- Contains the Storage Data Protector database, which contains information about backup details, such as backup durations, media IDs, and session IDs
- Runs core Storage Data Protector software
- Runs Session Managers that start and stop backup and restore sessions and write session information to the Storage Data Protector database

**Client systems**

Client systems that need to be backed up must have the Storage Data Protector Disk Agent (also called the *Backup Agent*) installed. The Disk Agent reads or writes data from a disk on the system and sends or receives data from the Media Agent. The Disk Agent is also installed on the Cell Manager, thus enabling you to back up the Storage Data Protector configuration, the Storage Data Protector database, and data on the Cell Manager.

**Devices**

Client systems with access to backup devices attached directly or through a SAN must have the Media Agent installed. Such client systems are also called *drive servers*. The Media Agent reads or writes data from media in the device and sends or receives data from the Disk Agent.

**Installation server**

The installation server holds a repository of the Storage Data Protector software packages for a specific architecture. By default, the Cell Manager is also an installation server. At least two installation servers are needed for mixed environments—one for UNIX systems and one for the Windows systems.

## Other HP backup tools

HP offers a complete range of automated storage management software products that enable you to make the most of the storage devices and servers that HP offers.

Many of these products support backup and restore. The result is increased system performance and improved productivity of system administrators and users, as well as data protection.

### Business Copy

Business Copy, which replaced the StorageWorks Enterprise Volume Manager, is a host-based application that manages controller-based clone and snapshot operations. Clones and snapshots are point-in-time copy functions that can be used to minimize the downtime required for system backups and data migration activities.

StorageWorks Business Copy Server is a GUI from which customers manage their local replication within a SAN. Business Copy Server provides a simple way for you to build jobs that automatically create and mount snapshots and clones.

StorageWorks Business Copy XP helps you leverage critical business data by creating multiple real-time online copies for zero-downtime split-mirror backup, disaster recovery mirroring, decision support, and application development and testing.

### StorageWorks Virtual Replicator

StorageWorks Virtual Replicator capabilities simplify storage management for Windows NT environments. A key component of the Enterprise Network Storage Architecture (ENSA), StorageWorks Virtual Replicator delivers two core capabilities across the ENSA family—storage pooling/ virtualization and snapshots.

Snapshots can be used to facilitate backups. The Snapshot for Backup wizard schedules a backup operation of a snapshot to tape. The snapshot and backup operations are scheduled using the HP Batch Scheduler.

The Restore from Snapshot wizard schedules the re-creation of a virtual disk from a snapshot. If specified, after a virtual disk is re-created, the previous virtual disk and snapshot are both deleted. The restore operation is scheduled using the HP Batch Scheduler.

## Microsoft Exchange backup/restore solution

The Restore Accelerator tool delivers the highest levels of recoverability for Microsoft Exchange. Through the integration of StorageWorks Virtual Replicator and customizable scripts, you can perform a quick, complete, and fully automated recovery of the Exchange environment.

By using a snapshot of the Exchange information store, you can resume full operation of the Exchange environment quickly. This solution also increases data protection by providing a recovery alternative to restoring from tape.

Incorporating StorageWorks Virtual Replicator with the HP Batch Scheduler or NT Scripting and a scripted procedure that requires minimal customization enables the following actions:

- Quiesce Exchange by shutting it down briefly.
- Snapshot the Exchange disks that contain the information store.
- Restart Exchange and resume operations.

## HP management support

- **HP Systems Insight Manager** — Systems Insight Manager enables you to monitor the Fibre Channel-to-SCSI bridge and the libraries. You can see the WorldWide Name, firmware revision, and status of each component, including the tape drives located in the library.
- **Media robot utility (MRU)** — MRU software enables you to monitor and control libraries and mini-libraries conveniently from your desk. MRU provides an application programming interface (API), a GUI on OpenVMS and Tru64 UNIX, a GUI on Windows NT, and a CLI. MRU enables you to:
  - Ease storage management — Verifies the installation and configuration of a library or mini-library.
  - Facilitate diagnostics — Verifies that hardware and control communication paths are functioning properly. If an operation such as backup fails, MRU isolates the library and verifies its operations.
  - Customize storage management — Enables you to write storage management solutions for your libraries.
  - Use simple interfaces — Enables you to choose the CLI, GUI, or API to move cartridges.

## Replication

Businesses, particularly those that operate in national and world markets, must have business continuance at all times. That means their data must be available continuously with complete data integrity regardless of circumstances.

One way to ensure that data is always available is to replicate it, which is different from backing it up. Online data replication happens in real time, so at least two copies of data are always available. The two copies are physically separated so that if disaster strikes in one location, the data will still be available at the other.

If you find that tape backups are running continuously, then replication might be the answer. Data replication is more expensive than routine backups but could be necessary in situations such as web transactions, wire transfers, and supply chain applications.

HP offers several products for replicating data, including:

- StorageWorks NAS Data Copy
- StorageWorks Continuous Access EVA and XP
- OpenView Continuous Access Storage Appliance (CASA)

### NAS Data Copy

NAS Data Copy is a real-time data replication and failover software product that augments an existing data protection strategy by reducing downtime and data loss with minimal impact on existing network and communication resources.

NAS Data Copy enables you to specify mission-critical data that must be protected. This data is replicated in real time from a production machine, known as the *source*, to a backup machine, known as the *target*. The target machine, on a local network or at a remote site, stores the copy of the critical data from the source.

NAS Data Copy monitors changes to the critical data and sends them to the target machine. By replicating only the file changes rather than copying an entire file, NAS Data Copy enables you to use resources more efficiently.

NAS Data Copy enables you to implement various data protection solutions, including local high-availability services, off-site disaster recovery services, and enhanced centralized backup using third-party backup systems.

NAS Data Copy performs four basic types of operations:

- **Mirroring** — The initial copy or subsequent resynchronization of selected data
- **Replication** — The ongoing capture of byte-level file changes
- **Failure monitoring and failover** — The ability to monitor and stand in for a machine in the event of a failure
- **Restoration** — A mirror of selected data from the target back to the source

## **Continuous Access EVA and XP**

Properly configured, Continuous Access can be a complete disaster-tolerant storage solution that guarantees data integrity in the event of a storage system or site failure.

Continuous Access EVA is a controller-based application that performs real-time replication between HP StorageWorks enterprise virtual array (EVA) storage systems.

StorageWorks Continuous Access XP and XP Extension are high-availability data and disaster recovery solutions that deliver host-independent real-time remote data mirroring between XP disk arrays. Continuous Access XP and XP Extension can be deployed in solutions ranging from data migration to high-availability server clustering, available over Enterprise Systems Connection (ESCON) or Fibre Channel.

Data replication can be bidirectional, meaning that a storage array can be both a source and a destination. Data replication is performed at the storage subsystem controller level and is transparent to the host, alleviating unnecessary host cycles to perform the data mirroring functions. Unlike a fabric-based or host-based solution, the storage-based solution dedicates its resources to managing the replication process between arrays, with minimal impact to other data or devices on the SAN.

In the event of a disaster, Continuous Access ensures that an online copy or exact mirror of data is available at the alternate site to support rapid resumption of critical processing at the alternate site. Resumption can usually occur within minutes instead of hours or days, as with other disaster recovery plans.

## **StorageWorks Data Replication Manager**

StorageWorks Data Replication Manager (DRM) provides customers the widest choice of performance, distance, and availability to meet their enterprise-level network storage requirements.

Using DRM software, you can perform data replication at the storage system level and in the background to any host activity. The total business continuance solution includes the MA8000 and the ESA/EMA12000 external storage systems. The MA8000 and ESA/EMA12000 systems running DRM can either replicate data up to 25 miles (40km) through an extended SAN over direct Fibre Channel links at full Fibre Channel speeds (100MB/s), or go unlimited distances with Fibre Channel-to-Asynchronous Transfer Mode (ATM) gateways.

For maximum protection from wide-area disasters or where direct Fibre Channel is not available, DRM supports Fibre Channel-to-ATM connectivity into telecommunication networks. Data replication can be performed across hundreds or thousands of miles.

## Comparing DRM and Continuous Access

DRM and Continuous Access each have unique advantages. The choice of one solution over the other depends on the customer environment and requirements.

Feature	DRM	Continuous Access
Failover	One LUN at a time	Multiple LUNs at a time
Number of LUNs/servers supported	96/192	1,024/512
Bidirectional capability	Requires two sets of arrays	Can share the same array

## CASA

Enabled by virtualization technology, CASA pools enterprise-wide storage assets to deliver local and long-distance data replication, point-in-time images, and data migration services throughout multiple-vendor environments. The performance of CASA increases to 1GB/s and connectivity expands from 16 to 24 ports, providing support for larger network storage environments.

CASA features device plug-ins to integrate with OpenView Storage Area Manager (SAM) for centralized management of virtualized storage environments through one console. With OpenView SAM, customers can increase the return on their IT investment by automating manual processes, increasing IT staff productivity, eliminating downtime attributed to human error, and reallocating unused storage capacity.

CASA allows customers to integrate with four of the modules within OpenView SAM:

- OpenView Storage Node Manager for device management
- OpenView Storage Builder for capacity management
- OpenView Storage Allocator for access control and LUN assignment
- OpenView Storage Accountant for metering and billing

---

**INTERNET** More information about HP OpenView CASA is available at:  
<http://www.hp.com/storage/virtualization>

---

## Summary

You have helped Bob and Carla understand the vulnerabilities that their data centers face. RC Engineering has decided to implement a corporate continuity plan and it is analyzing the ways it can set up its data centers to make them more disaster-tolerant. It has also created a backup strategy to help them survive any storm that comes its way.

## Learning check

1. What is an application domain?  
.....  
.....
2. Continuity plans should be:
  - a. Updated every two years
  - b. Updated only when major or minor changes occur
  - c. Exercised, not tested
  - d. Tested, not exercised
3. When compressed data is passed through a compression process a second time, it typically occupies more space rather than less because of the overhead of the compression algorithm.  
☐ True  
☐ False
4. In general, to configure backup software so that multiple servers are sharing one or more tape libraries, you must:
  - a. Install a primary server.
  - b. Create a tape library in one switch zone in the SAN.
  - c. Create a separate switch zone that contains one of the servers in the data zone.
  - d. Configure a data zone with maximum support.
  - e. Configure a data zone with the minimum supported number of servers and tape drives.

5. What is the difference between backing up data and data replication?
- .....
- .....
- .....
6. The \_\_\_\_\_ of the equipment in the data center determines the number and capacity of HVAC units required.
7. Optional external battery modules that increase the available battery backup time in case of power failure are called:
- a. EBM batteries
  - b. A unity power rated battery
  - c. Extended runtime modules
  - d. Wireless paralleling modules
8. Almost all HP equipment cools from front to rear.
- ☐ True
- ☐ False



### Introduction

As with any network-related project, the most important part of designing and implementing an Enterprise Backup Solution (EBS) is solid documentation.

All work done on the project should be documented and kept in a safe place. This documentation can be used by others to recreate the work you have performed. It can also be used to troubleshoot problems during or after installation of the EBS.

### Performing a needs analysis

Before you can begin the EBS design process, you must have a thorough understanding of the customer's business. The best way to familiarize yourself with the network and its data is to perform a needs analysis. From the needs analysis, you will have all the information necessary to design an EBS.

## EBS site survey form

Use this form to help gather information for the needs analysis and design processes. If the information required for this form differs between servers on the EBS, fill out a separate form for each server.

1. Server information.

	Server Type	Capacity (GB)		Server Type	Capacity (GB)
1.	Backup Server		11.		
2.			12.		
3.			13.		
4.			14.		
5.			15.		
6.			16.		
7.			17.		
8.			18.		
9.			19.		
10.			20.		

2. Backup server network topology:

- a. 10Mb Ethernet
- b. 100Mb Ethernet
- c. Gigabit Ethernet

3. Operating system:

- a. Novell NetWare
- b. HP Tru64 UNIX
- c. Microsoft Windows NT
- d. SUN Solaris
- e. Microsoft Windows 2000
- f. IBM AIX

4. Tape management system software:
  - a. Computer Associates ARCserve
  - b. VERITAS Backup Exec
  - c. VERITAS NetBackup
  - d. Legato NetWorker
  - e. Tivoli Storage Manager
  - f. CommVault Galaxy

5. Backup scheme

Day of Week	Backup Window (Hrs)	Start Time	End Time	Backup Type (Full/Incr/Diff)
Sunday				
Monday				
Tuesday				
Wednesday				
Thursday				
Friday				
Saturday				

6. What is the expected compression ratio for the backups?

.....

7. What is the expected annual growth rate for the data?

.....

8. How many years of expected growth should be planned for?

.....

9. Granularity of cartridge retention:

- a. Current week: ! daily
- b. Current month: ! daily ! weekly
- c. Current year: ! daily ! weekly ! monthly

10. How many weeks or days of historical data is required to be kept “nearline” within the tape library?

.....

11. Will you be using RAIT within the tape library?

☐ Yes

☐ No

12. List any other important facts about this environment.

.....

.....

13. Preferred backup window:

from .....

to .....

14. Maximum backup window:

from .....

to .....

15. Can the server be taken offline to perform the backup?

☐ Yes

☐ No

## A

### **Accredited Integration Specialist (AIS)**

This level of certification measures the competencies required for hands-on planning and the integration and support of technical solutions in SMB class environments. Given a set of customer business requirements, this individual is expected to be able to design, support, and integrate platform, operating system, storage, network, and option components to solve business needs.

### **Accredited Systems Engineer (ASE)**

This level of certification measures the competencies required for hands-on integration and support of technical solutions in complex, enterprise class environments. This individual is expected to be able to design, support, and integrate platform, operating system, storage, network, and option components to solve business needs.

### **agent**

A piece of code that can be incorporated into the element driver or written as a separate program that is loaded to interact with the driver.

### **arbitrated loop physical address (ALPA)**

The address of a Fibre Channel node in an arbitrated loop.

### **application**

A piece of software that monitors a specific piece of the network.

### **Application Programming Interface (API)**

A set of protocols, routines, and tools for building software applications.

### **arbitration**

The process of selecting one respondent from a collection of several candidates that request service at the same time.

### **Array Configuration Utility (ACU)**

A web-based application used to configure array controllers, add additional disk drives to an existing configuration, or completely reconfigure disk drive arrays.

**asynchronous**

Delayed over time.

**Automatic Server Recovery (ASR)**

A server feature designed to restart the server automatically after a critical hardware or software error has occurred. If a critical error occurs, the server records the error in the Server Health Logs, restarts the system, and sends a page if a modem is installed at the server.

**B**

**bandwidth**

A measure of the information-carrying capacity of an optical fiber, normalized to a unit of MHz km. This term is used to specify capacity of multimode fibers only. (For single-mode fibers, use dispersion). The difference between the highest and lowest frequencies in a specific range of frequencies.

**baseline**

A snapshot of the current state of a server or network; this information is used as the starting point for comparing future performance levels.

**baud**

A unit of signaling speed equal to the number of signal intervals per second, which might or might not be equal to the data rate in bits per second. The encoded bit rate per second.

**Berkeley Internet Name Domain**

A program invoked during system generation that creates TNS object (file code 100) system code files and system library files.

**BIOS (Basic Input Output System)**

Software coded into the computer chips to accomplish various tasks.

**bits per second (b/s or bps)**

The number of energy pulses passing a given point in a transmission medium in one second; raw transmission speed before encoding.

**boot BIOS**

Software coded into the chips on the host bus adapter (HBA). A special program used to boot and control the computer.

**broadband**

In data transmission, it denotes transmission facilities capable of handling frequencies greater than those required for high-grade voice communications. The higher frequency allows the carrying of several simultaneous channels.

**C****Cache Coherent Non-Uniform Memory Access (ccNUMA)**

In this design, multiple cells are “clicked” together to form larger systems based on a hierarchical memory architecture. The result is a ccNUMA design that scales effectively from tens to hundreds of processors, gigabytes of memory, and large I/O slot counts.

**central management server (CMS)**

The CMS is central to the systems management architecture and includes the HP Systems Insight Manager core, which aggregates fault, asset, performance, and configuration data from all discovered systems attached to the network. It also manages groups of systems through queries and tasks that control operations, such as SNMP status polling, email and paging notification, and system software update.

**cladding**

The dielectric material surrounding the core of an electrical fiber or material surrounding the core of a fiber optic cable. It usually refers to diameter, often 125m measured in microns.

**client**

A computer connected to a server on the network.

**cluster**

1) A certain number of disk sectors that are treated as a unit, or the smallest manageable unit of storage for an operating system. On a PC with a 200MB hard drive, the smallest cluster would be eight sectors, or 8 by 512 bytes, or 4KB. On a 2GB drive, the cluster would be 32KB. 2) A set of loosely coupled servers used as a single, unified computing resource.

**cluster file system (CFS)**

A technology built into the operating system, rather than layered onto it, that enables IT staff to manage all members of the cluster as if they were a single system, dramatically increasing productivity. This technology allows cluster management to be more comprehensive, better integrated and simpler, with fewer management tasks.

**command line interface (CLI)**

The area on screen where commands are typed into a command-driven system.

**Common Information Model**

A DMTF standard, the Common Information Model is a conceptual information model for describing management that is not bound to a particular implementation.

**Common Internet File System (CIFS)**

A protocol that defines a standard for remote file access using millions of computers at one time. With CIFS, users with different platforms and computers can share files without installing new software.

**Common Management Information Protocol (CMIP)**

A management protocol that defines a set of commands. The management application uses these commands to retrieve or change the values of items that the management agent provides.

**community string**

A string of characters that is similar to a password. A community string offers a limited amount of protection for the SNMP data on a server.

**compiler**

1) Software that translates a program written in a high level programming language such as C/C++ into machine language. 2) Software that alters a set of high-level language statements into a lower-level representation.

**connector**

A mechanical device used to align and join two fibers together to provide a means for attaching and decoupling it to a transmitter, receiver, or another fiber.

**core**

The central region of an optical fiber through which light is transmitted.

**corrected memory log**

A list of corrected memory errors.

**countdown time**

The time in minutes after the end of wink time in which the system waits before starting a shutdown sequence.

**Custom Device Module**

A Custom Device Module is the driver component in the NetWare Peripheral Architecture used to drive specific storage devices attached to the host bus adapter (HBA). The NetWare Peripheral Architecture separates NetWare driver support into two components: a Host Adapter Module, and a Custom Device Module.

## D

**daemon**

A UNIX program that runs unattended in the background and when required performs a specified operation. It functions like an extension to the operating system and is initiated at startup. Microsoft Windows refers to daemons as *system agents* and *services*.

**DAS to SAN (DtS)**

An architecture that allows users to migrate servers from a direct attached storage (DAS) to a storage area network (SAN) environment

**data rate**

The maximum number of bits of information that can be transmitted per second in a data transmission link. It is typically expressed as Megabits per second (Mbps or Mb/s).

**desktop management interface (DMI)**

A software interface between desktop-resident management programs and the manageable hardware desktop components.

**Desktop Management Task Force (DMTF)**

An industry body whose stated objective is to provide a working definition for desktop workstation management.

**direct attached storage (DAS)**

Disk drives located within the computer cabinet and connected through PCI or another peripheral bus to the processor.

**DOS**

An operating system that is stored on a hard drive disk.

**Domain Name Service (or Server) (DNS)**

The name resolution system that allows users to locate computers on a UNIX network or the Internet by domain name by maintaining a database of domain names (host names) and their corresponding IP addresses.

**Domain Name System (DNS)**

A system that defines a hierarchical, yet distributed database of information about hosts on a network. The network administrator configures the DNS with a list of hostnames and IP addresses, allowing users of workstations that are configured to query the DNS to specify remote systems by hostnames rather than by IP addresses. DNS domains should not be confused with Windows NT networking domains.

**duplex cable**

A two-fiber cable suitable for duplex transmission or transmissions in two directions.

**duplex transmission**

Transmission in both directions, either one direction at a time (half duplex) or both directions simultaneously (full duplex).

**dynamic link library (DLL)**

A Windows component that collects information in the Windows registry and allows the management SNMP agent to directly retrieve the information.

## E

**emergency repair diskette (ERD)**

A bootable recovery disk that stores an uncorrupted system configuration and can be used to restore the system to its initial setup state. This disk is used if the system files, boot variable, or boot sector are corrupted, and if a user is unable to recover the previous startup configuration.

**enhanced battery management (EBM)**

A proprietary, microprocessor-based system for monitoring and managing uninterruptible power source (UPS) batteries, with such features as quick recharging, doubled battery life, and up to 60 days of advanced warning before the batteries need replacement.

**enterprise**

The entirety of an organization, including its subsidiaries.

**Enterprise Backup Solution (EBS)**

A data protection solution from HP that combines a SCSI-based automated tape library with Fibre Channel technology and storage management software.

**Enterprise Network Storage Architecture (ENSA)**

A portfolio of modular, scalable, and highly available products from HP.

**event**

An action instigated by either the user or the computer. For example, a user event would include any mouse movement or keystroke. An internally generated event would include a notification sent at a certain time of the day.

**eXtensible Markup Language**

An open standard for describing data from the W3C used for defining data elements on a Web page and business-to-business documents. XML has a similar tag structure to HTML: where HTML defines how elements are displayed and uses predefined tags, XML defines what those elements contain and allows tags to be defined by the developer of the page.

## F

**fabric**

A Fibre Channel interconnection method that allows multiple simultaneous and concurrent data transfers between multiple hosts and/or storage devices connected with a multi-port hub.

**fault tolerance**

The ability to continue without pause in the instance of a hardware failure. Multiples of all critical components, such as processors, memory, disks, and power supplies are included in a fault-tolerant system to ensure reliability. If one component fails, another takes its place.

**fiber**

Thin filament of glass. An optical waveguide consisting of a core and cladding, which is capable of carrying information in the form of light. Fiber is also a general term used to cover all physical media types supported by Fibre Channel, such as optical fiber, twisted pair, and coaxial cable.

**Fiber Distributed Data Interface (FDDI)**

A network based on the use of optical-fiber cable to transmit data at 100Mbit/sec.

**Fibre Channel**

A high-speed, serial, bidirectional, topology independent, multi-protocol, highly scaleable interconnection between computers, peripherals, and networks.

**Fibre Channel Arbitrated Loop (FC-AL)**

An interconnection scheme that supports from 1 to 126 ports on a loop in a shared medium topology.

**Fibre Channel Protocol**

A protocol that supports the SCSI upper-level transport protocol.

**Fibre Channel Switched Fabric (FC-SW)**

Interconnected, non-blocking switches that allow full connectivity between all ports.

**G**

**gigabit interface converter (GBIC)**

A hot-swappable hardware module that converts serial electrical signals to serial optical signals and vice versa. The GBIC attaches network devices to transmission systems such as Fibre Channel and Gigabit Ethernet.

**gigabit link module**

A transmitter and receiver that provides high-speed serial links, enabling continuous throughput in each direction simultaneously.

**globally unique identifier (GUID)**

A 128-bit number used for computing object identifiers (OIDs) from Microsoft. A GUID is created out of the serial number in the local Ethernet card with the date and time.

**gnome**

GNU (Gnu's Not UNIX) Network Object Modeling Environment. A GUI-based interface for Linux and other UNIX environments born of the GNU project that serves as an alternative to the KDE interface.

**graphical user interface (GUI)**

A program interface that takes advantage of the graphics capabilities of a computer to make the program easier to use.

## H

**heartbeat**

1) A recurring signal generated by hardware for the purposes of activation and/or synchronization. 2) A recurring signal generated by hardware or software to demonstrate that it is still active.

**high availability**

Refers to a fault resilient multiprocessing system that quickly recovers from failure.

**host**

1) A computer that acts as a source of information or signals, usually accessed by a remote user. 2) A computer with a unique IP address connected to a TCP/IP network. 3) To provide the infrastructure for a computer service.

**Host Adapter Module (HAM)**

A HAM is the driver component used to drive specific HBA hardware in the NetWare Peripheral Architecture.

**host bus adapter (HBA)**

A device, typically an expansion card that plugs into the bus, connecting one or more peripheral units to a computer. Also called a *controller*.

**HP-UX**

The HP version of UNIX that runs on the HP 9000 server family. Based on SVID, it uses features from BSD UNIX in addition to several HP innovations.

**HyperText Transport Protocol Secure**

The protocol for accessing a secure web server.

# I

**Insight Management Agents**

Server software that queries a manageable device and provides information that responds to SNMP requests for data.

**instrumented devices**

Elements that have the agents installed and are manageable.

**integrated Lights-Out (iLO)**

An independent microprocessor that runs an embedded operating system and is integrated on the system board of select HP ProLiant servers.

**interconnect**

1) To attach one device to another. 2) A physical or wireless port used to attach one device to another, such as a plug, socket, transmitter, or receiver.

**Internet Control Message Protocol (ICMP)**

An IP extension used for out-of-band messages related to network operation. ICMP messages, delivered in IP packets, announce network errors, timeouts, and congestion. The ping command uses ICMP to test an Internet connection.

**Internet Service Provider (ISP)**

An organization that provides access to the Internet, and server space where websites can be created and maintained.

**interrupt request (IRQ)**

Network boards are supplied with default settings for the IRQ, I/O port, and base memory address. An IRQ is a signal protocol used by hardware devices, such as printers and modems, to let the computer know that they need some attention. When an IRQ is invoked, the processor puts its other work on hold and services the needs of the interrupting device.

**IP address**

A unique number that identifies a device on a network. The IP address is often automatically assigned by DHCP or AutoIP. However, a device can be manually assigned a static IP address.

**IPX**

A Novell NetWare native transport protocol that is used to transfer data between server and/or client programs running on different network nodes. IPX packets are not related to packets used in other systems such as Ethernet or token ring.

## L

**latency**

Delay in transmission; can be caused by arbitration time prior to winning control of the network.

**light amplification by stimulated emission of radiation (Laser)**

A device that produces coherent light with a narrow range of wavelengths.

**light-emitting diode (LED)**

A device used in a transmitter to convert information from electrical to optical form. It typically has a large spectral width.

**Lightweight Directory Access Protocol (LDAP)**

A protocol used to access a directory listing, implemented in Web browsers and e-mail programs.

**link**

A fiber optic cable with connectors attached to a transmitter (source) and receiver (detector).

**logical unit number (LUN)**

A three-bit identifier for a Logical Unit, a device in a daisy chain of drives. The maximum number of LUNs per SCSI target is 8, numbered LUN0 through LUN7.

**longwave**

Refers to length of the wave or frequency in the spectrum of light. 780 nm is the operating range of short wave lasers, while 1300 nm describes the range of long wave lasers.

**longwave gigabit interface converter (GBIC-LW)**

A hot-swappable hardware module that attaches network devices and provides a transmission rate of 100MB/s in each direction and a transmission distance of up to 10km between components using 9-micron, single-mode fiber optic cables.

## M

**managed device**

A device managed by a management console. Devices include servers, clients, routers, switches, and hubs. Servers and clients cannot be managed devices unless they have HP Insight Management Agents installed.

**managed object**

Any device on which an SNMP agent has been installed.

**management agent**

A software component within the SNMP that performs in-depth monitoring of a device.

**management console**

The PC workstation or server that runs Insight Manager 7.

**management information base (MIB)**

An SNMP structure that describes the particular device being monitored.

**management information format (MIF)**

A DMI file format that describes a hardware or software component used in a PC. It can contain data, code or both. See *desktop management interface* (DMI).

**management protocol**

A set of rules designed for management transport. Examples of management protocols are SNMP and CMIP.

**Mb/s**

Megabits per second or 1,000 bits per second. A reference to processing speed.

**MB/s**

Megabytes per second or 1,000 bytes per second. A reference to processing speed.

**media interface adapter**

A device that allows a conversion from fiber optic cable to copper media.

**Media Manager**

A database built into NetWare that keeps track of all peripheral storage devices and media attached to NetWare servers, and allows applications to gain access and get information. The Media Manager receives application I/O requests and converts them to messages compatible with the NPA architecture.

**memory access control (MAC) address**

A hardware address that uniquely identifies each node of a [network](#).

**micron ( $\mu\text{m}$ )**

Another term for micrometer, equal to one millionth of a meter.

**Microsoft Database Engine**

A database that stores all the collected information about the systems on a network as well as all configuration options.

**Microsoft Management Console (MMC)**

A network management software from Microsoft for Windows NT and Windows 2000 that provides a hierarchical view of resources similar to an Explorer view.

**mode**

A term used to describe a light path through a fiber, as in multi-mode or single-mode.

**Mozilla**

Originally built as the first Netscape Navigator browser; now an open-source Internet client program supported on Linux, Mac, and Windows PCs.

**multi-initiators**

Two different Fibre Channel host bus adapters (HBAs) in one arbitrated loop sharing the same storage devices, but not communicating with each other.

**multi-mode fiber**

An optical waveguide in which light travels in multiple modes. Typical core/cladding sized (measured in microns) are 50/125, 62.5/125, and 100/140.

**multiple-system aware**

A run type that supports multi-node operations. Tools with this run type operate on the target nodes using their own internal mechanisms instead of using the distributed task facility. The MSA run type uses the distributed task facility to launch the tool on a single node prior to the tool interacting with the other managed nodes.

**multiplexing**

The process by which two or more signals are transmitted over a single communications channel.

## N

**NetWare Loadable Module (NLM)**

Executable files that run on a NetWare server. NLMs link disk drivers, LAN drivers, name space, management applications, and file server enhancements with the operating system.

**network attached storage (NAS)**

A specialized file server that allows additional storage to be added by plugging it into a network hub or switch.

**network interface card (NIC)**

An adapter card installed in a PC, workstation, or server that enables the PC, workstation, or server to communicate with other devices connected to the same network. This term usually implies a local area network (LAN) adapter card.

**Non-Uniform Memory Access (NUMA)**

A multiprocessing architecture in which memory is separated into close and distant banks, and is accessed faster than the memory on other processing boards.

**nonvolatile RAM (NVRAM)**

Memory chips that hold their content without power.

**Novell Storage Services (NSS)**

A 64-bit file storage system that significantly expands file capacity and reduces mounting time allowing users to store a single file as large as 8TB.

## O

**Open Database Connectivity (ODBC)**

A database programming interface from Microsoft that provides a common language for Windows applications to access databases on a network. ODBC is made up of the function calls programmers write into their applications and the ODBC drivers themselves.

**Open Fiber Control**

A power monitoring/control system that continuously monitors the optical fiber link between two ports and prevents any laser emission from exceeding Class 1 levels in the event of a break anywhere in the path.

**optical waveguide**

A dielectric waveguide with a core consisting of optically transparent material or low attenuation (usually silica glass) and with cladding consisting of optically transparent material of lower refractive index than that of the core. It is used for the transmission of signals with light waves and is frequently referred to as fiber.

**Option ROM Configuration for Arrays (ORCA)**

An HP utility that enables users to configure array controllers. ORCA is accessible by pressing *F8* after the system has booted. It enables the user to create and delete logical drives, set interrupts, and set the boot controller order.

**out-of-band**

Synonymous with *off-the-network*; the capacity to deliver information over a modem.

## P

**PCI Bus #**

An internal Peripheral Connect Interface bus number in a computer. The PCI Bus # is set by the system BIOS, and in most cases, the number is zero.

**PCI Device #**

The PCI slot number in a computer.

**Performance Management Pack (PMP)**

The HP ProLiant Essentials Performance Management Pack (PMP) is an integrated management software solution that detects and analyzes hardware bottlenecks on ProLiant servers.

**ping**

1) **P**acket **I**nternet **G**roper; a process of determining if a particular IP address is online by sending out a packet and waiting for a response. It is also used to test and debug a network or verify if a user is online.

**Point-To-Point**

A connection established between two specific locations, as between two buildings.

**Point-to-Point Protocol**

A standard defined by the Internet Engineering Task Force. PPP provides a standard method for transporting multiple protocols over a point-to-point link.

**polling**

1) A communications technique that establishes when a terminal is ready to send data. 2) A technique that repeatedly interrogates a peripheral device to find out if it has data that must be transferred.

**power-on self-test (POST)**

A series of built-in diagnostics performed by the BIOS in a PC when the computer is first started.

**ProLiant Support Pack (PSP)**

A PSP is a bundle of software that includes ProLiant optimized drivers, utilities, and management agents.

**protocol**

A data transmission convention encompassing timing, control, formatting, and data representation.

## Q

### **query**

To interrogate a collection of data, such as records in a database. In addition to obtaining lists of records that match the search criteria, queries to a database allow for counting items and summing amounts. A web query yields only a list of matching pages and is more frequently deemed a *search*.

## R

### **remote access server**

A computer in a network that provides access to remote users using analog modem or ISDN connections. Including dial-up protocols and access control (authentication), it may be a regular file server with remote access software or a proprietary system such as the Shiva LANRover.

### **Remote Access Service (RAS)**

A Microsoft Windows NT Server feature that allows remote users to access the network from their Windows laptops or desktops by means of modem.

### **Remote Insight Lights-Out Edition (RILOE)**

An option card that provides remote server management in corporate data centers and remote sites by allowing browser access to ProLiant servers through a hardware-based operating system.

### **Remote Insight Lights-Out Edition (RILOE) II**

The next-generation RILOE board, which provides a way to remotely access and manage ProLiant servers. It has its own integrated hardware components, that allow full access and control of the server, independent of the state of the operating system or server hardware.

## S

### **Samba**

A free, open source implementation of the CIFS file sharing protocol evolving from SMB (Server Message Block). Samba permits a UNIX server to act as a file server to Windows clients.

### **Samba Web Administration Tool**

Allows a Samba administrator to configure the complex smb.conf(5) file using a web browser. A SWAT configuration page has help links to the configurable options in the smb.conf file, permitting an administrator to verify the effects of any change.

### **schema**

Refers to the definition of an entire database, defining both the structure and the type of content that each data element within the structure can contain.

### **script**

1) A program written in a general-purpose programming language. 2) A program written in a special-purpose language, as is used in a communications program or word processor. 3) A typeface that resembles continuously flowing handwriting or calligraphy.

### **Secure Shell**

Software from SSH Communications Security, Inc., Palo Alto, CA ([www.ssh.com](http://www.ssh.com)) that provides secure logon for Windows and UNIX clients and servers. SSH replaces Telnet, FTP, and other remote logon utilities with an encrypted alternative.

### **Secure Sockets Layer (SSL)**

The leading security protocol on the Internet developed by Netscape. At the onset of an SSL session, the server sends its public key to the browser, which the browser then uses to send a randomly generated secret key back to the server, creating a secret key exchange for that session.

### **SelectID**

Used to configure Seagate drives. A matrix for relating ALPA (hex) numbers to SelectID (hex) numbers.

**Server Message Block (SMB)**

The file sharing protocol in DOS, OS/2 and earlier versions of Windows. SMB originated with the NetBIOS protocol used in early DOS networks. In the 1996/97 time frame, SMB evolved into CIFS.

**Service Pack (SP)**

An executable file that may contain patches, fixes, or newly supported drivers or options for a software application

**shortwave**

Refers to length of the wave or frequency in the spectrum of light. 780nm is the operating range of short-wave lasers, while 1300nm describes the range of long wave lasers.

**shortwave gigabit interface converter (GBIC-SW)**

A hot swappable hardware module that attaches network devices and provides a transmission rate of 100MB/s in each direction, a transmission distance of up to 200 meters between components using 62.5 micron, multi-mode fiber optic cable and up to 500 meters between components using 50-micron, multi-mode fiber optic cables.

**Simple Mail Transfer Protocol (SMTP)**

A protocol for sending email messages between servers, especially over the Internet.

**Simple Network Management Protocol (SNMP)**

SNMP defines a set of commands that a management application uses to retrieve or change the values of items that a management agent makes available.

**simplex cable**

A term sometimes used for a single-fiber cable.

**simplex transmission**

Transmission in one direction only.

**single-mode fiber**

An optical waveguide (or fiber) with a small core diameter in which only a single mode is capable of propagation. This type of fiber is particularly suitable for wideband transmission over large distances, since its bandwidth is limited only by chromatic dispersion.

**single-system aware (SSA)**

A run type that does not support multi-node operations. Tools with this run type are only aware of the node they are running on.

**Small Computer Systems Interface (SCSI)**

A computer industry interface standard used for connecting peripherals to computers.

**Small and medium business (SMB)**

A category of user that is generally larger than a small office, home office user, but smaller than an enterprise company.

**SNMP Multiplexing (SMUX) Manager**

Extends the SNMP Enterprise MIB to include HP ProLiant MIB data and supports get, set, and trap operations on data items defined in the HP ProLiant MIB.

**storage area network (SAN)**

A network of storage disks which connects multiple servers to a centralized pool of disk storage.

**Support Pack (SP)**

An executable file that may contain patches, fixes, or newly supported drivers or options for applications or operating systems.

**symmetric multiprocessing (SMP)**

A multiprocessing architecture in which multiple processors, residing in one cabinet, share the same memory.

**systems management**

Monitoring, evaluating, or improving the performance of the network and its essential elements. Processes include configuration management, performance management, security management, fault management, and desktop management.

## T

**topology**

The logical and/or physical arrangement of stations on a network.

**transducer**

A device for converting energy from one form to another, such as optical energy to electrical energy.

**trap**

Synonymous with *alarm*. Traps are indicators of changes or error conditions.

## U

**uninterruptible power supply (UPS)**

A battery that supplies continuous power to a computer system in the event of a power failure.

**user ID (UID)**

The name you use to identify yourself when logging onto a computer system or online service, consisting of both a username and a password.

## V

**Version Control Agent (VCA)**

An Insight Management Agent installed on a server to allow you to see what HP software and firmware is currently installed. The VCA can be configured to allow for easy version comparison and software updates from a repository to the server on which the Version Control Agent is installed.

**Version Control Repository Manager (VCRM)**

An HP Insight Management Agent that enables the management of stored HP software. The VCRM catalogs system software and firmware that is stored where the Version Control Repository Manager is installed.

**virtual interface architecture (VIA)**

A memory to memory transport protocol used for high-speed transfer of data between machines that enables long blocks of data to be sent from one application in one machine directly to another application in a remote machine without being broken up into packets by a transport protocol.

**virtual LAN (VLAN)**

A logical subgroup within a local area network that is created by software rather than manually moving cables in the wiring closet.

**virtual private network (VPN)**

A private network configured within a public network to take advantage of the economies of scale and management facilities of large networks. Often used by enterprises to create WANs spanning large geographic areas.

## W

**waveguide**

Structure that guides electromagnetic waves along its length. An optical fiber is an optical waveguide.

**wavelength**

The distance between two crests of an electromagnetic waveform.

**Web-Based Enterprise Management (WBEM)**

An umbrella term for the use of Internet technologies to manage systems and networks throughout the enterprise. Built into Windows 98 and 2000, WBEM uses the Common Information Model (CIM) as the database for information about computer systems and network devices.

**web-enabled agent**

An agent that manages information that can be displayed in a web browser.

**web-launch aware (WLA)**

A run type for tools that are launched in a web browser using a web server. WLA tools can be designed to work with multiple systems.

**Windows Management Instrumentation**

Windows 2000 and Windows 98 extensions that provide an interface between the operating system and instrumented components, allowing the instrumented components to provide information and notifications.

### Module 1 — Heterogeneous server deployment and integration

1. List the configuration modifications you must implement to make your deployment server function optimally before you begin to use RDP.  
Configure PXE to process new computers automatically.  
Synchronize the console name with the operating system name.  
Change the primary lookup key to the serial number.  
Configure PSPs.  
Create PXE images, boot diskettes, and multi-NIC boot images.  
Configure ProLiant BL enclosures (if applicable).  
Remotely install Deployment Agent for Windows (formerly *AClient*).
2. Name the three main configuration and installation methods for RDP.  
Simple  
Custom  
Enterprise or Distributed
3. Samba allows connectivity between Linux operating systems and Windows computers by using the Windows native file sharing protocol Server Message Block (SMB).
4. Which ProLiant servers are well suited as firewalls? (Select **two**.)
  - a. ML330
  - b. ML370
  - c. DL380
  - d. DL360
  - e. BL server blades
  - f. DL580
5. Without routing, there is no DNS.

## Module 2 — Enterprise management

1. Which layer of the Systems Insight Manager architecture serves as the primary means of accessing management functionality?  
A web browser
2. Which Systems Insight Manager usage scenario described in this module requires VCRM to be installed on regional servers?  
Enterprise
3. The server designated for Systems Insight Manager is running HP-UX. Which database can be installed on this server to support Systems Insight Manager?
  - a. Oracle
  - b. SQL Server 2000
  - c. SQL Server 7
  - d. PostgreSQL
4. List the three options for customizing the Systems Insight Manager home page.
  - a. Select the desired destination for the Home link
  - b. Disable the display of the Did You Know? image on the introductory page
  - c. Disable the DO THIS NOW to finish the install display on the introductory page.
  - d. View system and event lists
5. An organization has five system administrators who are responsible for 100 different systems in six different buildings. Which feature of Systems Insight Manager will enable each administrator to view only those systems for which they are responsible?  
Monitoring
6. Why add a rule to the System Type Manager?  
To customize the type and product name of systems  
To create rules that map the system object identifiers to product categories and names of their choice
7. You need to compile MIBs. However, you do not want to affect the performance of Systems Insight Manager. What must you do?  
The MIB compile is separate from Systems Insight Manager so it can be run at any time without adverse performance.

8. When would it be useful to copy a report?  
The Copy Report options can be used to make a copy of an existing report before making changes. Modifying a predefined layout is easier than creating an entirely new report.
9. List the steps you should take to integrate management processors with a directory service.  
Upgrade the management processor firmware.  
Extend the directory services schema.  
Install the management snap-in in the management application of the operating system.  
Create and manage objects and roles in the directory service.
10. List the tasks that you must complete when you add a new server to the network.  
After deploying the operating system and configuring network settings, install OpenSSH.  
Ensure that trust relationship exists between the new server and Systems Insight Manager.  
Manually add the server to Systems Insight Manager.  
If a server does not have the VCA installed, use the Initial ProLiant Support Pack Install feature of Systems Insight Manager to install the PSP supported.  
If a server does have the VCA installed, run Software Version Status Polling on that server.  
Add new Lights-Out devices to the directory service.  
Assign new servers to an IT administrator in Systems Insight Manager.

## Module 3 – Security

1. Security is a balance between maintaining ease of use and controlling access.
2. Two-factor authentication involves something the user knows, such as a password, and either: (Select **two**.)
  - a. Something that would be difficult for an intruder to guess (a user's birthday, maiden name, or favorite movie)
  - b. Something difficult to remember (a long string of numbers or a complex combination of symbols)
  - c. Something the user has (mobile phone, smart card, or pager)
  - d. Something that would use biometrics (fingerprints or retina scans)
  - e. Something that changes regularly (random codes, text against a patterned field)
3. What are the three levels of trust configuration?
  - a. Trust All
  - b. Trust By Name
  - c. Trust By Certificate
4. Which protocol is used to encrypt communication among the browser, the CMS, and the managed device?  
SSL
5. When you browse to Systems Insight Manager, a message displays indicating a trust problem with the certificate. What are your two options to prevent this message from displaying again?
  - a. Add the Systems Insight Manager server to its list of trusted certification authorities
  - b. Obtain a certificate from a third-party certification authority and import it into Systems Insight Manager.
6. What should a complete antivirus solution begin with?  
A complete solution should begin with an assessment of all systems for known issues regarding system vulnerabilities and poorly configured or misconfigured systems.

7. To prevent downtime and data loss, what steps should you take when you learn about a new virus threat? (Select **four**.)
- a. Scan the subject line of all incoming email messages for the word *virus*
  - b. Gather data about the possible attack.
  - c. Assess the vulnerability and deploy any temporary workarounds.
  - d. Configure the supported agents on managed devices to trust the Systems Insight Manager server certificate.
  - e. Notify management and users.
  - f. Test and deploy the update from your antivirus software vendor.
  - g. Email employees about the nature of the threat, with the update procedure documentation in an attachment.
  - h. Move important files off the mail server.

## Module 4 — Performance management

1. To use PMP, you must license the servers of interest and enable monitoring.  
True  
False
2. When you use PMP to configure performance-based alerts, when are alerts triggered?
- a. An alert is triggered every time a server or subsystem performance changes between any of the following status indicators:
  - b. Normal performance — Green status indicator
  - c. Approaching bottleneck — Amber status indicator
  - d. Confirmed bottleneck — Red status indicator
3. How do you design an application to execute within a cluster environment? (Select **two**.)
- a. The application containing the business logic must communicate with a database that supplies and stores data.
  - b. Partition the data and allowing each node to access only the assigned partition.
  - c. Reconfigure the application cache to increase performance.
  - d. Put a data access manager in place to solve the challenges of concurrent access to the same data.

4. Write cache is beneficial in random I/O environments; read cache improves performance in sequential I/O environments.
5. Asynchronous I/O is more efficient and faster than synchronous I/O.  
True  
False
6. What should you do if multiple solutions must be implemented or multiple variables changed at a time?  
If multiple solutions must be implemented or multiple variables changed at the time, you should do so in the smallest increments possible
7. Name three steps you should take if you have exhausted all possibilities in your action plan without achieving the desired result.  
Try an alternate solution.  
Re-evaluate the order of the tasks in your action plan.  
Diagnose the mode of failure again.
8. What does an amber status indicator alert you to in PMP?
  - a. Normal performance
  - b. Unknown device
  - c. Approaching bottleneck
  - d. Critical condition
  - e. Confirmed bottleneck
9. When using RPM to restart a partition that houses a troubled application, you must reboot the server.  
True  
False
10. Match the types of NIC teaming with their characteristics.

a. Network fault tolerance	b. Is switch-independent and can be split across Layer 2 switches
b. Transmit load balancing	c. Is supported on IEEE 802.3ad-capable switches
c. Switch-assisted load balancing	a. Provides simple redundancy with two to eight NICs

## Module 5 — High availability and clusters

1. What is the primary objective of the HP Enterprise Network Storage Architecture (ENSA)?
  - a. To integrate open standards and industry-standard approaches toward managing storage
  - b. To create a portfolio of modular, scalable, and highly available products
  - c. To establish a direct connection between storage resources
  - d. To automate the storage management processes that manage data placement and protection through the information life cycle
2. What is a SAN?  
A high-speed system of servers accessing a common or shared pool of heterogeneous storage devices
3. Which component is not a SAN layer?
  - a. Client
  - b. Server
  - c. Storage
  - d. Interconnect
  - e. Fabric
4. List three benefits of implementing a SAN.
  - Centralized storage — Consolidating storage (primary and secondary) in a SAN and sharing the resource across multiple servers reduces the cost of storage management.
  - Elimination of server downtime while adding storage — Using Fibre Channel-based storage, storage resources can be added or deleted without interrupting the production environment.
  - Improved availability — Implementing advanced SAN designs enables fault-tolerant, and disaster-tolerant configurations that are ideal platforms for clustered, mission-critical systems.

- Modular scalability — Providing support for an unpredictable environment that allows changing the infrastructure as business needs evolve. Bandwidth, availability, redundancy, and capacity, can be dynamically scaled on demand, providing maximum flexibility to accommodate business growth
  - Serverless backup — Allowing for direct backup from disk to tape, without going through the host that offloads data from the network. The host will initiate the process but another intelligent device, such as the Modular Data Router (MDR) can transfer the data.
  - Online storage migration — Storage can be dynamically allocated and re-allocated among hosts without interruption, resulting in improved storage utilization.
5. The MSA1000 is the only storage system in the industry that enables a cost-effective migration to SANs by offering DtS capability.
- True
6. Which Systems Insight Manager utility focuses on a computing environment from the perspective of clusters?
- Cluster Monitor
7. What is a cluster?
- A cluster is a set of loosely coupled servers used as a single, unified computing resource
8. Which component provides initialization to Fibre Channel devices?
- a. HBA
  - b. Disk array
  - c. SAN management
  - d. SAN switch
9. Which applications benefit most from migrating to storage hosted on a SAN?
- Transaction processing
  - Email
  - Groupware
  - Enterprise resource planning
  - Multimedia file serving
  - Database acceleration

10. According to the latest research, what are the leading causes of downtime, from most common to least common?
  - a. Infrastructure problems (building, power, and network)
  - b. Software failures (operating system, application, tools, and drivers)
  - c. Operational and administrative activities (procedures, personnel, and maintenance activities)
  - d. Hardware failures (disk, power supply, and memory)
11. List three utilities used to troubleshoot HP SANs and clusters.
  - Array Configuration Utility
  - Array Diagnostic Utility
  - The Fibre Channel Fault Isolation Utility
12. What happens when more than one cluster or node resource is in an abnormal state?
  - a. A pager alert is automatically generated.
  - b. The parent cluster or node icon reflects the most severe state because the resource statuses are propagated upward.
  - c. The central management server administrator receives an email notification.
  - d. Systems Insight Manager restarts because the server in which it is installed is unreachable.

## Module 6 — Business continuity and disaster recovery

1. What is an application domain?  
Application domains use load balancing routers, redundant communication lines, or other methods, to split transactions between multiple servers running in multiple sites.
2. Continuity plans should be:
  - a. Updated every two years
  - b. Updated only when major or minor changes occur
  - c. Exercised, not tested
  - d. Tested, not exercised

3. When compressed data is passed through a compression process a second time, it typically occupies more space rather than less because of the overhead of the compression algorithm.
- True
- False
4. In general, to configure backup software so that multiple servers are sharing one or more tape libraries, you must:
- Install a primary server
  - Create a tape library in one switch zone in the SAN
  - Create a separate switch zone that contains one of the servers in the data zone
  - Configure a data zone with maximum support
  - Configure a data zone with the minimum supported number of servers and tape drives.
5. What is the difference between backing up data and data replication?
- Online data replication happens in real time, so at least two copies of data are always available. The two copies are physically separated so that if disaster strikes in one location, the data will still be available at the other.
6. The heat load of the equipment in the data center determines the number and capacity of HVAC units required.
7. Optional external battery modules that increase the available battery backup time in case of power failure are called:
- EBM batteries
  - A unity power rated battery
  - Extended runtime modules
  - Wireless paralleling modules.
8. Almost all HP equipment cools from front to rear.
- True
- False