HP Medical Archive solution

Software version: 8.0.5

grid primer

Document release date: October 2009
Software release date: October 2009

Legal notices

Warranty

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.

The information contained herein is subject to change without notice.

Restricted rights legend

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Licensing

The use of HP products is governed by the terms and conditions of the applicable End User License Agreement (EULA).

Copyright notices

Trademark notices

Adobe® and Acrobat® are trademarks of Adobe Systems Incorporated.

Microsoft, Windows, Windows NT, and Windows XP are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

Contents

About this document

The *grid primer* provides an introduction to the HP MAS product. This manual contains:

- an overview of the features and components of the HP MAS product
- an introduction to the NMS web interface
- a description of the data flow when objects are ingested, replicated, retrieved, modified, and deleted
- monitoring tips

Intended audience

This *grid primer* is for users and system administrators new to the HP MAS product, and for grid operators whose main function is to monitor the grid on a day-to-day basis, in particular operators who log in to the NMS web interface using user-level permissions.

Prerequisites

You are assumed to have a sound understanding of the nature of audited activities within the HP MAS system. To use the text log file, you are assumed to have access to the configured audit share on the Admin Node server hosting the AMS service.

This document assumes familiarity with many terms related to computer operations and programming, network communications, and operating system file operations. There is wide use of acronyms.

Related documentation

In addition to this guide, please refer to other documents for this product:

- *HP Medical Archive solution audit message reference*
- *HP Medical Archive solution user guide*
- *HP Medical Archive solution DICOM conformance statement*

These and other HP documents can be found on the HP documents web site:

http://www.hp.com/support/

Document conventions and symbols

Convention	Element
`[ ]`	Indicates that the enclosed element is optional and may be left out.
`{ }`	Indicates that you must specify one of the listed options.
|	Separates alternatives.
<parameter_name>	You must supply a value for a variable parameter.
. . .	• Indicates a repetition of the preceding parameter. • Example continues after omitted lines.
Medium blue text: Figure 1 http://www.hp.com	• Cross-reference links • E-mail addresses • Web site addresses
Bold	• Key names or key sequence • GUI elements that are clicked or selected, such as menu and list items, buttons, and check boxes • Text typed into a GUI element, such as into a box
Italics	• Document titles, for example *HP Medical Archive solution user guide* • Text emphasis • You must supply a value for a variable in a GUI element.
`Monospace`	• File and directory names • Text displayed on the screen, such as system output and application messages • Command or reserved keyword in a CLI, API, program language, or operating system • Script or code example
`Italic monospace`	You must supply a value on the command line.

Convention	Element
`Bold monospace`	• Text typed at the command line • Emphasis of file and directory names, system output, and code

⚠ WARNING! Indicates that failure to follow directions could result in bodily harm or death.

△ CAUTION Indicates that failure to follow directions could result in damage to equipment or loss of data.

NOTE Provides additional information.

TIP Provides helpful hints and shortcuts.

RECOMMENDATION Provides guidance from HP for a best practice or for optimum performance.

Documentation updates

The title page of this document contains the following identifying information:

- Software version number

  Indicates the software version.

- Document release date

  Changes each time the document is updated.

- Software release date

  Indicates the release date of this version of the software.

To check for recent updates, or to verify that you are using the most recent edition of a document, go to:

http://ovweb.external.hp.com/lpe/doc_serv/

http://h20230.www2.hp.com/selfsolve/manuals

You can also receive updated or new editions if you subscribe to the appropriate product support service. For details, contact your HP sales representative.

Subscription service

HP strongly recommends that customers sign up online using the Subscriber's choice web site:

http://www.hp.com/go/e-updates

- Subscribing to this service provides you with e-mail updates on the latest product enhancements, newest versions of drivers, and firmware documentation updates as well as instant access to numerous other product resources.
- After signing up, you can quickly locate your products under Product Category.

Support

You can visit the HP Software Support web site at:

http://www.hp.com/go/hpsoftwaresupport

HP Software Support Online provides an efficient way to access interactive technical support tools. As a valued support customer, you can benefit by using the support site to:

- Search for knowledge documents of interest
- Submit and track support cases and enhancement requests
- Download software patches
- Manage support contracts
- Look up HP support contacts
- Review information about available services
- Enter into discussions with other software customers
- Research and register for software training

Most of the support areas require that you register as an HP Passport user and sign in. Many also require a support contract.

To find more information about access levels, go to:

http://h20230.www2.hp.com/new_access_levels.jsp

For more information about HP Passport, go to:

http://h20229.www2.hp.com/passport-registration.html

1 Introduction to the HP MAS software

This chapter contains an overview of the HP MAS system. It covers product features, architecture, deployment scenarios, and basic configuration.

About the HP MAS System

The HP MAS system is a storage management system that stores, protects, and preserves fixed-content data over its lifetime.

The HP MAS system creates a unified storage interface across multiple facilities and heterogeneous storage hardware. Through automation and sophisticated data management, the HP MAS system provides multi-site accessibility, high uptime and reliability, data protection and preservation, and simplified operation.

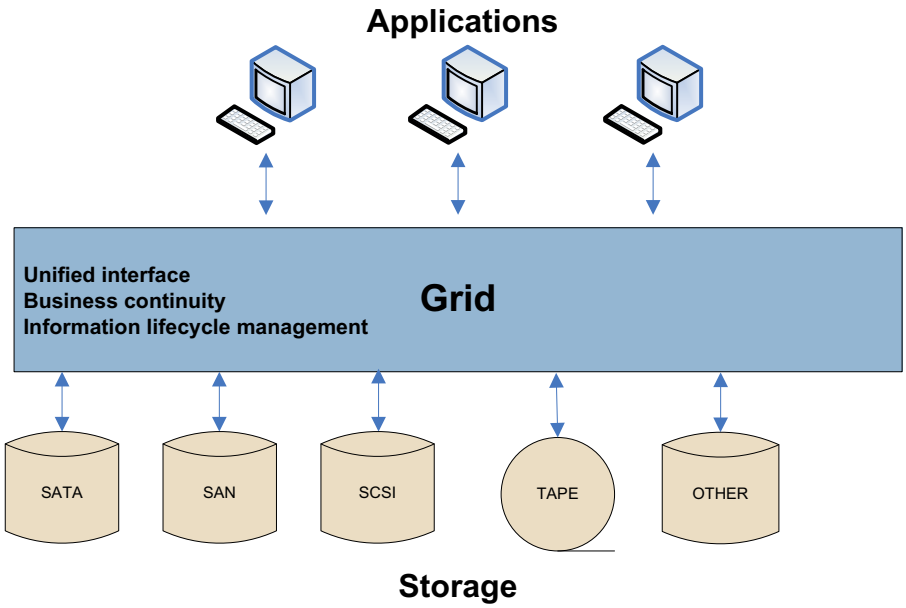


Figure 1 Storage Management System for Massive Volumes of Fixed-Content Data

Building Blocks

The basic building blocks of an HP MAS deployment are called grid nodes. A grid node consists of one or more grid services running on a server. A grid service is a software component that performs a specific function.

The basic types of grid nodes are:

- Admin Node
- Tape Node
- Control Node
- Gateway Node
- Storage Node

Grid nodes can be combined to reduce the number of servers deployed in a grid.

Figure 2 (page 11) shows how grid nodes and services are arranged within the grid.

- Applications, or clients, communicate with the grid via the Gateway Nodes which provide the interface to the grid.
- The Storage Nodes manage data storage on spinning disks.
- The Control Nodes store and manage content metadata.
- The Admin Nodes provide grid management services such as grid monitoring, logging, and grid configuration.
- For environments where content is also stored on archive media, the Tape Node provides an interface to the middleware that manages the archive media storage device such as a tape library.

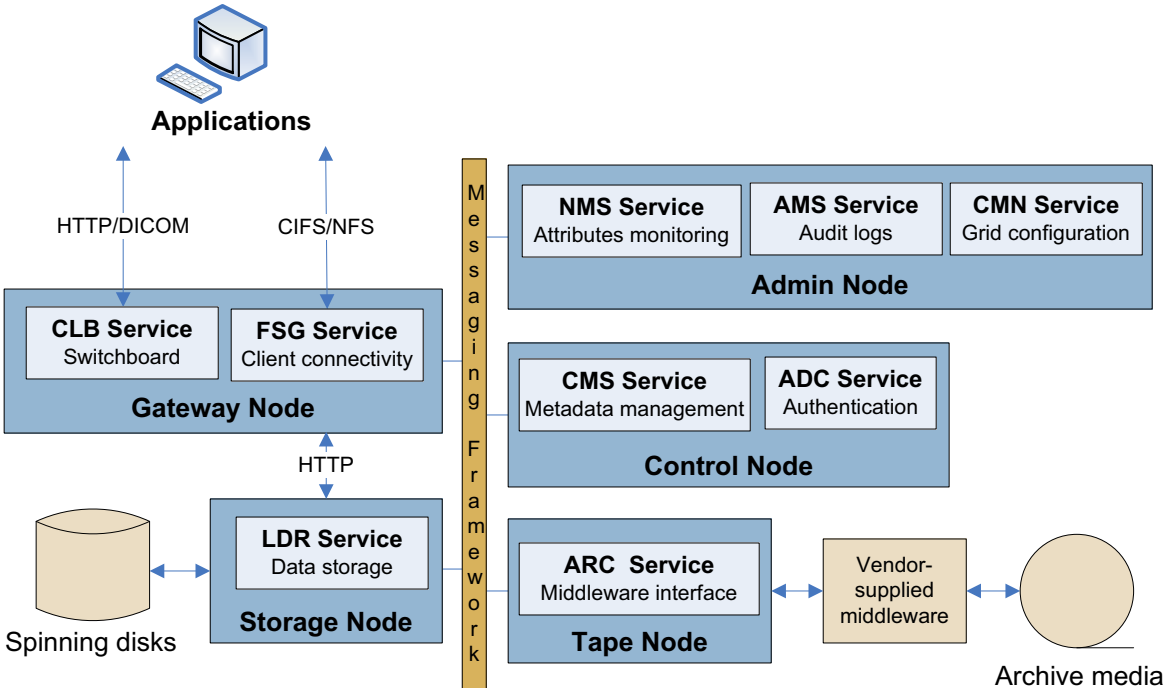


Figure 2 Grid Building Blocks

Grid Nodes

Table 1 (page 11) lists the grid nodes and what grid services they host.

Table 1 Grid Nodes

Grid Node	Function	Services	Notes
Admin Node	Grid management	NMS CMN AMS SSM	The Admin Node provides tools for grid administration. Although a grid can have more than one Admin Node, there is only one CMN service per grid.
Tape Node	Data storage on archive media	ARC SSM	The Tape Node manages storage of data to nearline data storage devices such as such as tape libraries (via IBM Tivoli® Storage Manager).
Control Node	Content management	CMS ADC SSM	The Control Node manages content. Each grid has at least two Control Nodes for redundancy of metadata storage.
Gateway Node	Client connectivity	CLB FSG SSM	The Gateway Nodes provide an interface between the client applications and the grid. Each grid must have at least two Gateway Nodes for redundancy. Depending on what client services are required, a Gateway Node may include an FSG service (for CIFS/NFS), a CLB service (for HTTP/DICOM) or both.
Storage Node	Data storage on spinning disks	LDR SSM	The Storage Nodes manage data storage on spinning disks. Each grid must have at least two Storage Nodes for redundancy.

Table 1 Grid Nodes *(continued)*

Grid Node	Function	Services	Notes
Admin/Gateway Node	Grid management and client connectivity	NMS CMN AMS CLB FSG SSM	Combined Admin Node and Gateway Node functionality on a single server.
Control/Storage Node	Content management and Data storage	CMS ADC LDR SSM	Combined Control Node and Storage Node functionality on a single server.
Custom nodes	Your system may also include custom node servers, for example, an Admin/Gateway/Control/Storage Node or a Gateway/Control/Storage Node that combines the functionality of multiple grid nodes on a single server.		

Services

Each grid service performs a special function. See Table 3 (page 17) for a summary of the grid services. To get an understanding of how these services work together during object ingest, retrieval and delete, see Data Flow (page 51). For more details, see the *user guide*.

Table 2 Grid Services

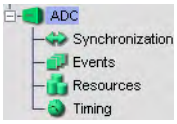
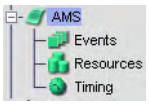
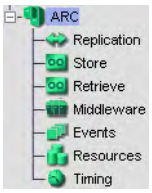
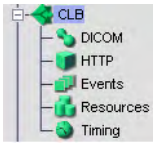
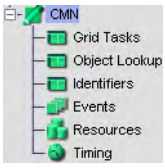
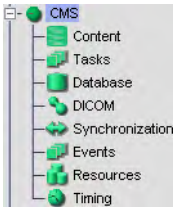
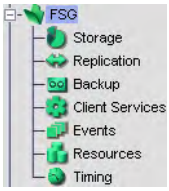
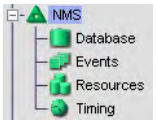
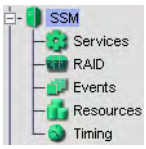
Acronym	Name	Description	Grid Node
ADC 	Administrative Domain Controller	Maintains topology information and provides authentication services.	Control Node
AMS 	Audit Management System	Keeps logs of grid activity and events.	Admin Node
ARC 	Archive	Communicates with archiving middleware to store and retrieve data to and from archive media.	Tape Node
CLB 	Connection Load Balancer	Acts as switchboard for connecting remote entities to the most efficient LDR. Primary connection point for remote entities using the DICOM or HTTP protocols.	Gateway Node
CMN 	Configuration Management Node	Manages grid-wide configurations: connection profiles, FSG replication groups, grid tasks, grid options.	Admin Node
CMS 	Content Management System	Keeps track of what data is stored on the grid. Stores content metadata and manages content replication based on ILM rules.	Control Node

Table 2 Grid Services *(continued)*

Acronym	Name	Description	Grid Node
FSG 	File System Gateway	Allows connections to the grid via a standard file system (NFS or CIFS).	Gateway Node
LDR 	Local Distribution Router	Stores, moves, verifies, and retrieves object data stored on disks.	Storage Node
NMS 	Network Management System	Provides a window into the grid. Used to monitor grid status and to configure the grid.	Admin Node
SSM 	Server Status Monitor	Monitors hardware performance such as key operating system metrics and network metrics.	Present on all nodes

Deployment Topologies

The HP MAS system can be deployed in a number of ways. A large grid could consist of hundreds of servers. Servers are connected using LAN and WAN network links to form a unified storage system. This section describes these basic deployment topologies:

- Single site
- Data Center (DC) + Disaster Recovery (DR) Site

The deployments shown in Figure 3 (page 15) and Figure 4 (page 16) are simplified examples and do not represent a complete grid deployment.

Single Site

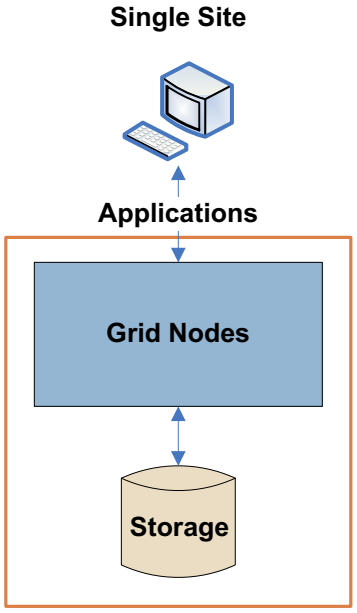


Figure 3 Single Site Deployment

In a single site deployment, the infrastructure and operations are centralized in a single data center. There is no off-site disaster recovery facility.

DC+DR

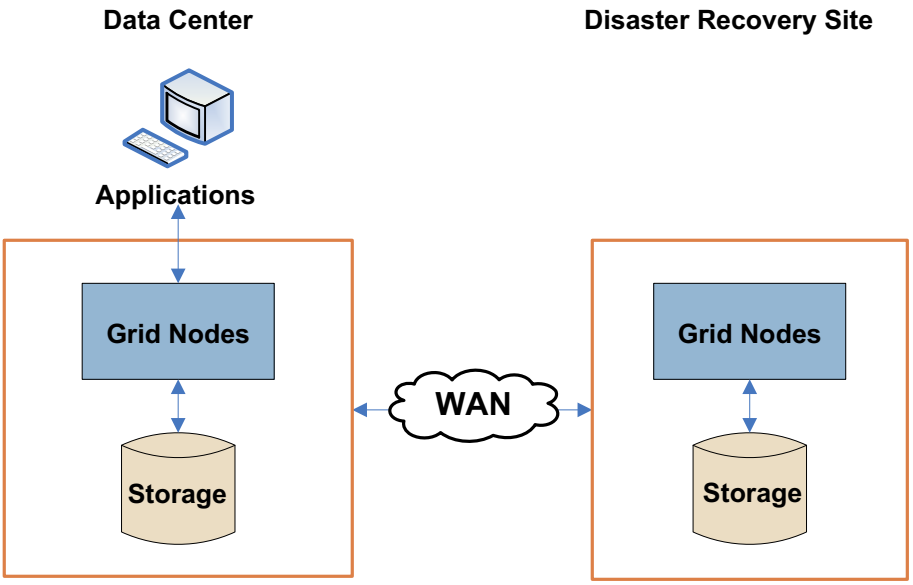


Figure 4 DC+DR Deployment

In a DC+DR deployment, the infrastructure is consolidated in a single data center and replicated to another location for disaster recovery.

Grid Options

The HP MAS system has a number of configurable options. The key ones are listed in Table 3 (page 17).

Table 3 HP MAS Options

Option	Description
DICOM	Enables the HP MAS system to communicate with medical imaging devices and applications using the DICOM (Digital Imaging and Communications in Medicine) protocol.
Audit	Provides user access to the audit logs which contain a complete record of grid activity.
Compression	Compresses objects saved to the grid, reducing file size by roughly 50% for content that is not already in a compressed format.
Deduplication	Deletes unnecessary identical copies of an object from the grid.
Deletion protection	Prevents clients from deleting any content that has been stored to the grid.
Distributed CMS	Enables replicating content metadata to a subset of CMSs instead of synchronizing all CMS databases. Distributed CMS operation improves scalability and performance.
Dual commit	Forces two copies of an object to be saved to two Storage Nodes on initial ingest to provide an additional level of data protection.
Encryption	Enables encrypted storage of all data. Content is encrypted during ingest and objects are stored in an encrypted form so that if a server is compromised no data can be retrieved in any readable form.
Parallel loading	Enables Gateway Nodes to preload in their cache all the files in a directory upon an initial file request. This enhances performance for systems that store related files in a single directory
Secondary preloading	Enables caching on all FSGs in the FSG replication group as files are ingested or retrieved to speed access via the Secondary Gateway Node.
Security partitions	Provides the ability to isolate content ingested from different FSG replication groups into different partitions, such that each client only has access to data stored in its own partition.

Grid Configuration Information

License Agreement

The text of the HP license agreement is located in a file stored in `/var/local/install/` for HP MAS 7.5 or earlier and `/var/local` for HP MAS 8.0.

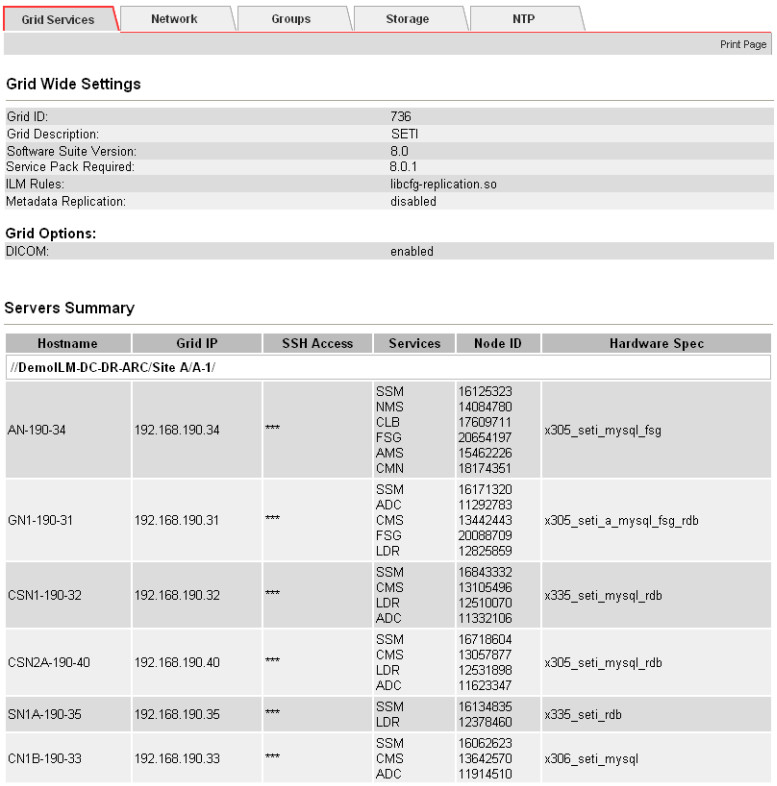
SAID Package

You may be asked to look up grid configuration information to help troubleshoot problems.

Specific configuration and integration information about the grid such as IP addresses and group assignment is contained in the Software Activation and Integration Data (SAID) package that is used to install the grid.

NOTE The SAID package contains highly confidential passwords and encryption keys needed during system maintenance, updates, and expansion. Store the SAID package in a secure location.

The grid configuration information is in the `index.html` file in the `\Doc` directory. See Figure 5 (page 18) for an example.



Grid Services | Network | Groups | Storage | NTP

Print Page

Grid Wide Settings

Grid ID:	736
Grid Description:	SETI
Software Suite Version:	8.0
Service Pack Required:	8.0.1
ILM Rules:	libcfg-replication.so
Metadata Replication:	disabled

Grid Options:

DICOM:	enabled

Servers Summary

Hostname	Grid IP	SSH Access	Services	Node ID	Hardware Spec
//DemoILM-DC-DR-ARC/Site A/A-1/					
AN-190-34	192.168.190.34	***	SSM NMS CLB FSG AMS CMN	16125323 14084780 17609711 20654197 15462226 18174351	x305_seti_mysql_fsg
GN1-190-31	192.168.190.31	***	SSM ADC CMS FSG LDR	16171320 11292783 13442443 20088709 12825859	x305_seti_a_mysql_fsg_rdb
CSN1-190-32	192.168.190.32	***	SSM CMS LDR ADC	16843332 13105496 12510070 11332106	x335_seti_mysql_rdb
CSN2A-190-40	192.168.190.40	***	SSM CMS LDR ADC	16718604 13057877 12531898 11623347	x305_seti_mysql_rdb
SN1A-190-35	192.168.190.35	***	SSM LDR	16134835 12378460	x335_seti_rdb
CN1B-190-33	192.168.190.33	***	SSM CMS ADC	16062623 13642570 11914510	x306_seti_mysql

Figure 5 index.html File

Passwords used to access the grid are in the `Passwords.txt` file.

NOTE Only trained and authorized service personnel should have access to the Passwords.txt file.

Server Consoles

During regular day-to-day operations, you do not need to access the server consoles. However, occasionally, you may be required to run commands directly from the server console in order to troubleshoot problems or execute maintenance procedures.

Server Manager

Each server in a HP MAS system runs the Server Manager application. Server Manager is used to supervise the starting and stopping of services on the server, ensuring services gracefully join and leave the grid. Server Manager also monitors services on the server and attempts to restart any that report errors.

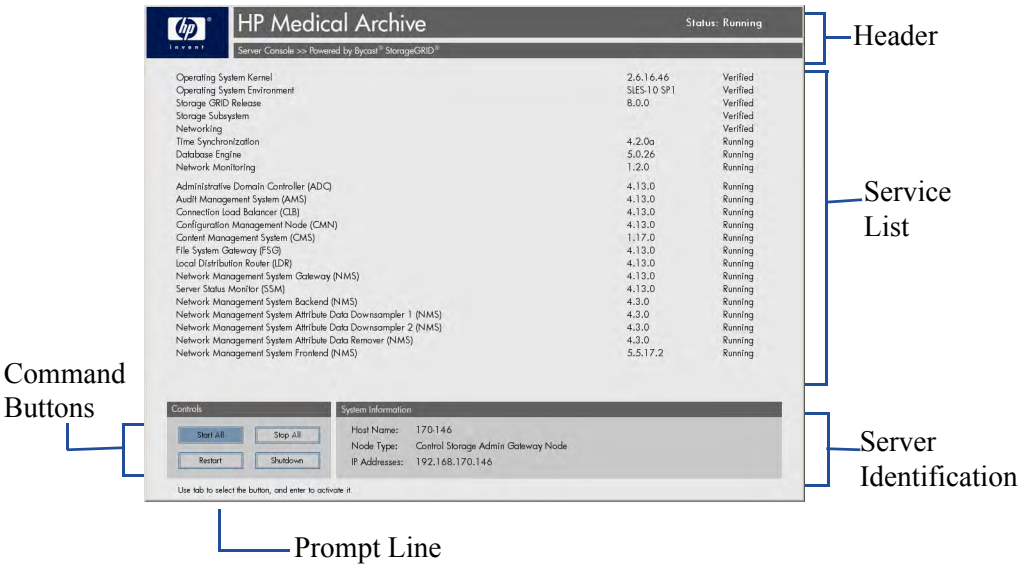


Figure 6 Server Manager Interface

The body of the display is the list of services being monitored by the Server Manager on this server. Some services may appear that are not identified through the NMS. These are services that can run independently and provide support capabilities to the grid services.

For more information on Server Manager, see the *user guide*.

Command Shell Access

Occasionally, you may be asked to run commands directly from a command shell on the server console for troubleshooting.

Log In

To log in at the console of a grid server:

1. Press **<Alt>+<F1>** to access a command shell on the server.
2. Enter the account name **`root`**.
3. Enter the password for the server specified in the `Passwords.txt` file.

Log Out

To log out of a command shell session:

1. Enter **`exit`** to close the command shell session.
2. Press **<Alt>+<F7>** to return to the Server Manager GUI.

2 NMS Web Interface

This chapter contains an overview of the NMS web interface that you use to monitor the grid. The chapter explains how to log in and out, describes the interface elements, and contains procedures to configure your account, monitor alarms, and create reports.

Browser Requirements

The only supported browser is Microsoft Internet Explorer v6.0 SP2 or v7.0. JavaScript and cookies must be enabled.

Configure Internet Options Settings

Internet Explorer settings for temporary internet files, security, and privacy must be set correctly.

To verify the Internet Explorer settings:

1. Go to **Tools ▶ Internet Options ▶ General**.
2. In the Temporary Internet files box, click **Settings**.
3. In the Check for newer versions of stored pages section, verify that **Automatically** is selected.

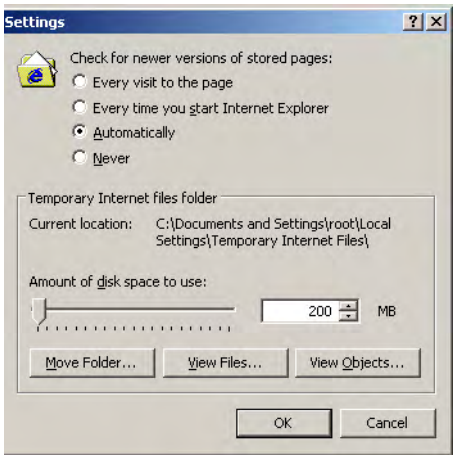


Figure 7 Temporary Files Setting

4. Go to **Tools ▶ Internet Options ▶ Security ▶ Custom Level** and ensure that the Active scripting setting is **Enable**.

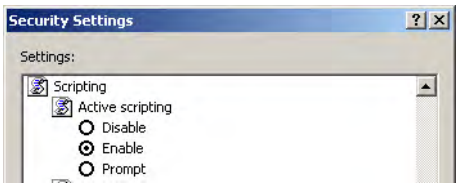


Figure 8 Active Scripting Setting

5 Go to **Tools ▶ Internet Options ▶ Privacy** and ensure that the privacy setting is **Medium** or lower (cookies must be enabled).

Enable Pop-ups

To make changes to passwords, you must ensure that Internet Explorer has the Pop-up Blocker turned **off**.

NOTE The menu option is a toggle. If the blocker is already disabled, the menu option is Turn On Pop-up Blocker.

To enable pop-ups:

- Select **Tools ▶ Pop-up Blocker ▶ Turn off Pop-up Blocker** from the Internet Explorer main menu.

Log In and Out

Accessing the NMS requires a web browser with grid access to a known address.

You will need a user name and password to access the system. Each system user is assigned a user name and password when first introduced to the NMS.

Log In

To log in to the NMS:

1. Launch the web browser.
2. Enter the address: **https://<IP_Address>**
3. If you are prompted with a Security Alert dialog, do one of the following:
   - Click **Yes** to proceed with this session. The alert will appear again the next time you access this URL.
   - Click **View Certificate** and install the certificate using the installation wizard so that you no longer receive the alert.

Figure 9 Security Alert Window

4. Enter your user name and password in the Login window and click **log in**. Both user name and password are case sensitive. Keystrokes appear as bullet characters (•) to protect your password.

Figure 10 NMS Login Window

If the user name or password you entered cannot be validated, the following message appears in red below the password field: `Invalid credentials`. Repeat the login process to correct your entry. If you have forgotten your password, contact Support to have your password reset.

Log Out

When you have finished your NMS session, log out to keep the system secure.

To log out:

1 Click the **Logout** button Logout located at the top right corner of the screen. The logging out message appears.

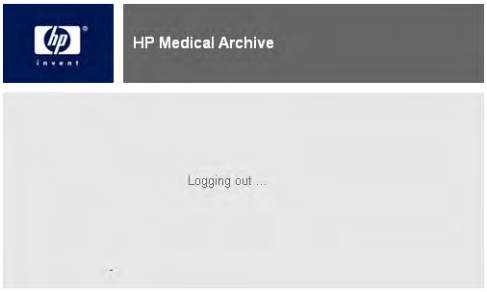


Figure 11 Successful Logout Window

2 You may safely close the browser or continue using other applications.

NOTE Failure to log out may give unauthorized users access to your NMS session. Simply closing your browser is *not* sufficient to log out of the session.

NMS Interface

The NMS interface provides basic operational data, alarm status, reporting functionality, and configuration options for each grid node, service, and component.

The main elements of the NMS interface are:

- Header
- Grid Management Menu
- Grid Topology Tree
- Content Tabs

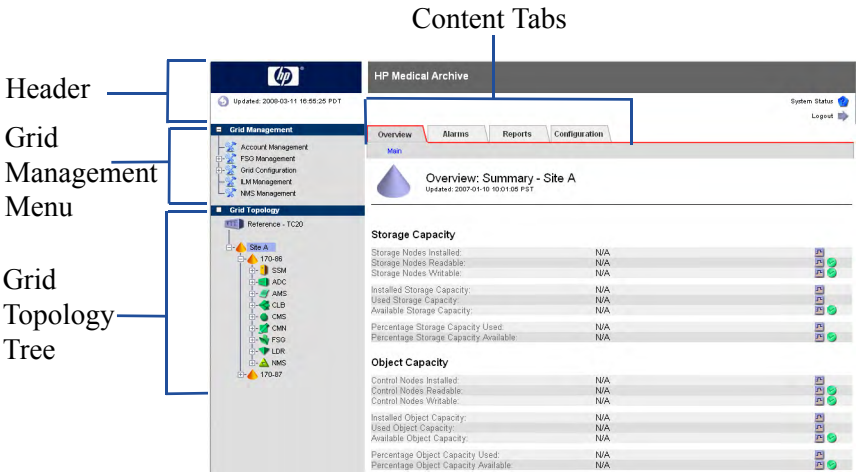


Figure 12 NMS Interface Elements

Header

The Header contains high-level grid status information. The latest browser refresh time is shown on the left. The Hewlett-Packard Product Name logo on the left serves as a button to access NMS version information. The right side contains the System Status indicator and the Logout button.

Figure 13 NMS Interface — Header

Refreshing the Display

Information presented in the NMS interface is time-sensitive. Since the interface is delivered as HTML “pages”, the content shown is static. The Updated timestamp indicates when the data shown was collected, that is, the time at which the last grid status “snapshot” was taken. Local time is shown as determined from the preferences set in the user account. The information is refreshed automatically at set intervals (the default is 15 seconds).

To refresh the display manually, do one of the following:

- Click **Refresh Page** .
- Click the **timestamp**.
- Click the **Refresh** button on the browser (for Microsoft Internet Explorer, press **<F5>**).

  After the page has finished reloading, the content and the timestamp are updated.

If you leave the NMS management window open with no activity, the session expires after the configurable timeout period. If you leave an NMS management interface window open and minimized on the desktop for an extended period of time (greater than one week for example), the display of the current time and attribute values may not be correct when you return to it. Refresh the display using the procedure given above.

Displaying Version Information

To display version information:

- Click the HP logo in the Header.

  A box displays the interface version number, the software build number, and copyright information.

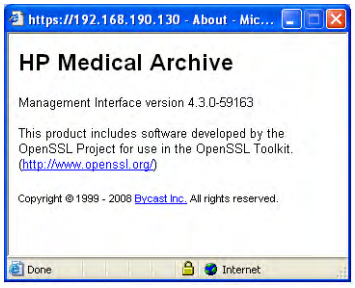


Figure 14 NMS Version Information

Grid Management Menu

The Grid Management menu provides access to a number of configuration pages.

Figure 15 Grid Management Menu

Table 4 Grid Management Menu

Item	Used to Configure
Account Management	User Accounts. For more information on accounts and user groups, see User Accounts (page 32).
FSG Management	FSG settings such as content protection options, cache space, and backups. Configuration is restricted to user accounts that have Maintenance permissions such as the Admin and Vendor accounts.
Grid Configuration	Grid options such as Audit levels, HTTP and DICOM profiles, link cost groups, and storage grades. Configuration is restricted to user accounts that have Grid Management permissions such as the Vendor account.
ILM Management	ILM (Information Lifecycle Management) policies. Configuration is restricted to user accounts that have Grid Management permissions such as the Vendor account.
NMS Management	Custom alarms, alarm notifications, and GUI timeout period. Configuration is restricted to user accounts that have Maintenance permissions such as the Admin and Vendor accounts.

Grid Topology Tree

The Grid Topology Tree provides quick access to any element of the grid.

- To expand and collapse the tree, click the ⊞ and ⊟ buttons.
- To view detailed information about an element of the tree, click its name.

NOTE Occasionally the NMS fails to display the entire navigation tree. The tree will be restored when the navigation frame is automatically refreshed. To prevent this, ensure the browser settings are correct. See Configure Internet Options Settings (page 21).

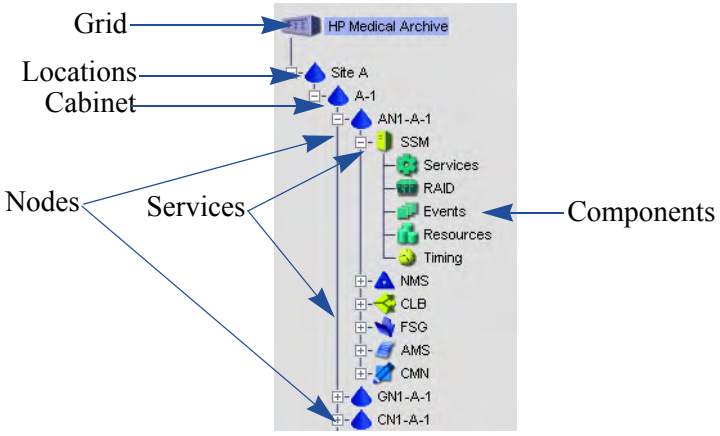


Figure 16 Grid Topology Tree

The highest level of detail is the grid as a whole, shown in the NMS as the root of the Grid Topology tree. Together, all of the other elements shown below constitute the grid. From the highest (big picture) to the lowest (most granular detail) the elements are:

- Grid
- Locations
- Grid nodes
- Grid services
- Service components

The HP MAS can be deployed as a Single Site (Site A) or Single Site + DR (Site A and Site B) in a standard appliance configuration. In an Enterprise deployment, locations may be cities throughout the country, buildings within a city, or any other grouping. Each cabinet can be expanded to reveal one or more grid nodes, a grid node being a server hosting a collection of one or more grid services. A grid service consists of software components that deliver a particular capability.

Naming Conventions

The names in the topology tree follow the following convention:

- Locations:
  - Site A—The primary site for the HP MAS.
  - Site B —An optional Disaster Recovery (DR) site.
- Cabinets:
  - A-1 through A-8—Cabinets at the primary site.
  - B-1 through B-4—Cabinets at an optional DR site.
- Nodes—Named using the following elements:
  - Two or three-letter code for the type of node (such as: `AN` for Admin Node, `GN` for Gateway Node, or `CSN` for Control/Storage Node)
  - A sequence number (`1` through `4`) within the cabinet

— The cabinet identifier (such as `B-1`, `A-3`, and so on)

Example: `GN1-A-1` is the primary Gateway Node in cabinet `A-1`.

- Services—Named by the software with a three letter acronym, such as `ADC` for an Administrative Domain Controller service.
- Components—Named by the software

Content Tabs

Content is organized in tabs. Every element of the grid topology tree has four tabs: `Overview`, `Alarms`, `Reports`, `Configuration`.

Figure 17 NMS Content Tabs

- Overview: The Overview tab is used to monitor the grid attributes. Each attribute represents a property, for example number of managed objects, free storage space, backup size, or service state. These attributes are used to monitor normal grid operation and to detect and troubleshoot abnormal conditions. While there are hundreds of attributes, most of them are used for troubleshooting and only a small number must be monitored on a regular basis to ensure smooth operation. For examples on how to work with attributes, see Operations (page 93).
- Alarms: The Alarms tab is used to view and acknowledge alarms. For more information, see Alarms and State Indicators (page 36).
- Reports: The Reports tab is used to create charts and text reports. For more information, see Reports (page 45).
- Configuration: The Configuration tab is used to change configuration settings at the location, node, service or component level. Configuration is restricted to user accounts that have Maintenance permissions such as the Admin and Vendor accounts.

Some tabs contain multiple pages. Click the page name to access the content. The page currently selected is shown in blue and the other pages in black.

Figure 18 NMS Page Selection

Attribute Description

The NMS interface contains a description of each attribute.

To find out more about each attribute:

- Click the attribute name to display its description. Click  to close the description.

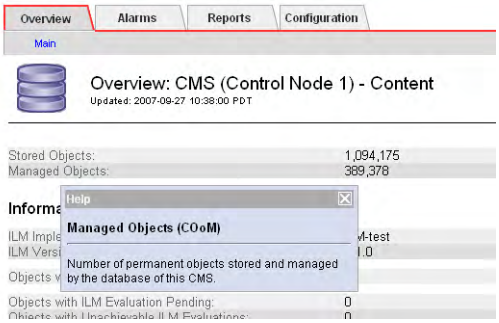


Figure 19 NMS Attribute Description

Timing

The reporting of NMS attributes is subject to propagation delays within the grid. Update values for most attributes, except for state attributes, are sent to the NMS at fixed intervals. Therefore, it may take a few seconds before an update is visible in the NMS, and two attributes that change more or less simultaneously may be reported at slightly different times.

Valid Characters

The NMS accepts only valid UTF-8 characters as user input in text fields.

Units of Measure

For units of “Seconds” or “Bytes”, the values displayed in the NMS are scaled to a suitable unit. For example, durations scale to microseconds, milliseconds, seconds, minutes, hours, or days; bytes scale to kilobytes, megabytes, or gigabytes.

NOTE The scale of bytes displayed by the NMS uses the “natural” measure of powers of 10. For example 3 MB = 3 x 10^6 = 3,000,000 bytes. This is not the same as powers of 2 normally used for computing, where 3 MiB = 3 x 2^{20} = 3,145,728 bytes.

Apply Changes Button

To commit changes, for example to acknowledge alarms or change configuration settings, you must click the **Apply Changes** button at the bottom of the page. After you click the button, the button becomes dimmed. Changes may take time to

process. Do not click **Apply Changes** more than once. Wait for the page to refresh. You can tell the changes have been committed when the **Apply Changes** button disappears completely.

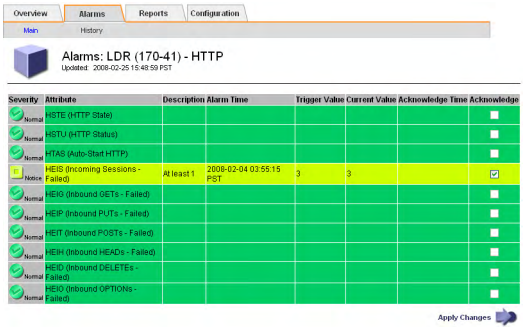


Figure 20 Apply Changes Button

To abort changes prior to clicking **Apply Changes**, simply refresh the page using the **Refresh Page** button at the top left in the header or the browser's refresh button.

User Accounts

The HP MAS system has two built-in user accounts:

Admin	Responsible for grid maintenance. The admin account can configure services and components but cannot make grid-wide changes.
Vendor	Responsible for grid configuration. The vendor account has full permissions.

The built-in accounts cannot be deleted. Additional accounts may exist on your grid depending on how it is configured. For example, the grid could have a read-only access intended for people who simply monitor the grid.

Permissions

Three built-in user groups (group accounts) have been configured for the HP MAS system: Vendor, Admin, and User.

Built-in groups are granted a collection of permissions. There are four types of permissions:

- Grid management
- Maintenance
- Alarm acknowledgement
- Accounts

Table 5 (page 32) describes the allowable tasks for each set of permissions.

Table 5 User Groups Permissions

Permission Set	Allowable Tasks
Grid Management	• Configure grid-wide options • Configure ILM
Maintenance	• Configure FSGs • Configure the NMS (customize alarms, configure e-mail notifications, and configure GUI time-out) • Configure services and components
Alarm acknowledgement	• Acknowledge alarms
Accounts	• Create new accounts, configure existing accounts, and delete accounts • Create new user groups, configure existing user groups, and delete user groups

Figure 21 (page 33) shows the built-in user accounts and group accounts. Depending on the Accounts permission on your account, you may not see all the accounts. See Figure 22 (page 33) for a comparison.

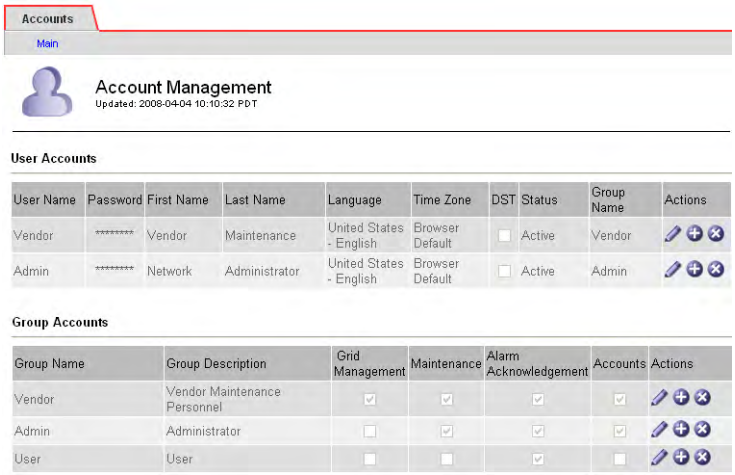


Figure 21 Account Management for Vendor Account

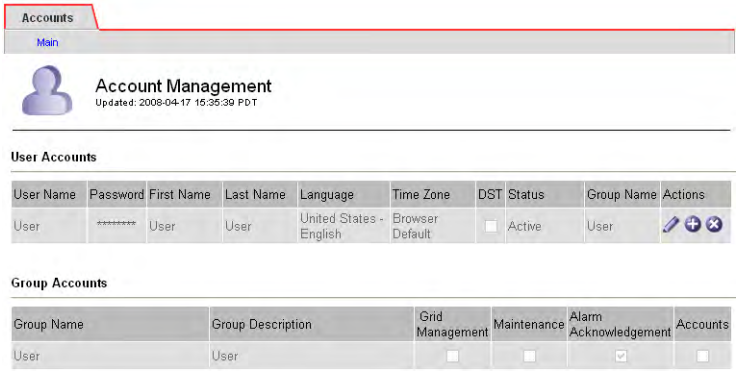


Figure 22 Account Management for Account Without Accounts Permissions

Configuring Your Account

You can configure your account to change your password, first name, last name, and time zone.

To configure your account:

1. If you intend to change the password, make sure pop-ups are enabled. See Enable Pop-ups (page 22).
2. Expand the **Grid Management** menu and click **Account Management**. See Table 6 (page 34) for a description of the fields.

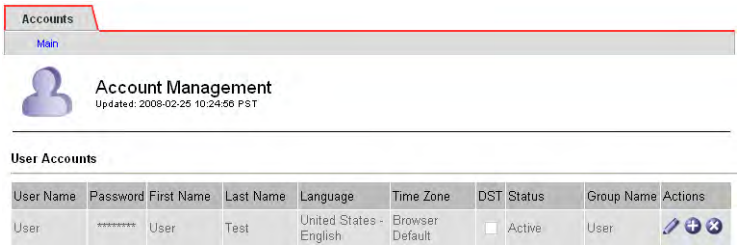


Figure 23 Account Management

3 Click **Edit** .

4 To change your password:

a Double-click the **Password** entry to select the complete field.

b Type a new password. Your password must contain between 8 and 32 characters and is case-sensitive.

c Press **<Tab>**. A confirmation pop-up window appears.

Figure 24 Password Confirmation Pop-up Window

d Re-enter the password and click **Confirm Password**. If the password fails to match, re-enter the password as prompted.

5 To change your name, edit the First Name and Last Name boxes.

6 To change the time zone, select a new time zone from the list.

7 Click **Apply Changes**.

Table 6 Account Management Fields

Field	Notes
User Name	The user name entered at login. Read-only.
Password	Masked password for the account; shown as a string of asterisks.
First Name	User's first name.
Last Name	User's last name.
Language	The default language to be used for this user. At this time, only English is supported.

Table 6 Account Management Fields *(continued)*

Field	Notes
Time Zone	User's time zone. The default is Browser Default. If you change the time zone settings of your computer, the change is effective the next time you log in.
DST	Click to set to Daylight Saving Time. You can modify the DST field only if Time Zone is not set to Browser Default. When Time Zone is set to Browser Default, daylight saving time is determined by the browser settings.
Status	Active—the user can log in and use the NMS. Disabled—the user is prevented from logging in. Depending on your permissions, this field may be read-only.
Group Name	Profile that governs the permitted activities for this user. Depending on your permissions, this field may be read-only.

Creating accounts and modifying accounts of other users is restricted to accounts that have "Accounts" permission such as the Admin and Vendor accounts.

Alarms and State Indicators

The color of the icon next to each location, cabinet, grid node, grid service, and service component in the grid topology tree reflects the overall status of that part of the grid.

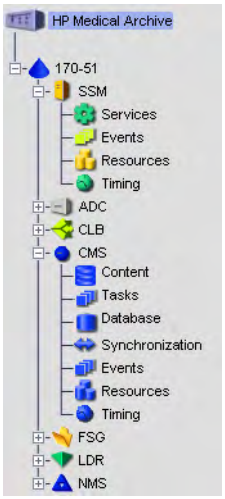


Figure 25 Grid Topology Tree with State and Alarm Colors

If there are no alarms and all services are connected, the icon appears in the normal (green) state. If there is an alarm, the color of the icon reflects the most severe alarm currently active on that branch of the tree. Locations display the highest alarm level of the grid nodes on that branch. Grid nodes display the color of the most severe state or alarm among their hosted services. Each individual service reflects the highest alarm severity of its components.

Table 7 Alarm Severity and Service State

Priority Level	Icon Color	State	Alarm Severity
Highest	Blue	Unknown	Not applicable
	Red	Connected	Critical
	Dark Orange		Major Alert
	Light Orange		Minor Alert
	Yellow		Notice
	Green		Normal
Lowest	Gray	Administratively Down	Not applicable

Service State Indicators

A service can have one of three states: Unknown, Connected, or Administratively Down. A service that is Connected is operating normally and displays the color of its highest alarm severity—either itself or its components. A service that is Administratively Down has been deliberately shut down for maintenance by a grid administrator. A service that is Unknown is problematic and must be investigated.

Table 8 Service States

Icon	Color	State	Meaning
	Green	Connected	All services are working normally.
	Blue	Unknown	An unknown condition exists that has stopped normal operation. Requires immediate attention. The “Unknown” state is considered the most severe. It is typically used to indicate loss of connection between the NMS and a service.
	Gray	Administratively Down	A service has been purposefully stopped. All alarms on the stopped service including acknowledged alarms are removed.

Alarm Indicators

NOTE A change in the value of an attribute can trigger an alarm. A change in the state of a service does not trigger an alarm.

An alarm is triggered when the value of an attribute reaches the alarm threshold value. When an alarm is triggered, the alarm information is displayed on the NMS interface and an e-mail notification is automatically sent to designated personnel.

Alarms are generated at the attribute level. There are five alarm severity levels in the NMS. Each alarm level has an associated color and icon (see Table 9 (page 37)).

Table 9 Alarm Severity and Indicators

Icon	Color	Severity	Meaning
	Green	Normal	All functions are working normally.
	Yellow	Notice	An unusual condition exists that does not affect normal operation.

Table 9 Alarm Severity and Indicators *(continued)*

Icon	Color	Severity	Meaning
	Light Orange	Minor Alert	An abnormal condition exists that could affect operation in the future; should be investigated to prevent escalation.
	Dark Orange	Major Alert	An abnormal condition exists that currently affects operation; requires prompt attention to prevent escalation.
	Red	Critical	A critical alert of an abnormal condition that has stopped normal operation; should be addressed immediately.

Propagation

Alarm Indicators

Alarms are generated at the attribute level. When an issue is detected, the alarm is propagated up through the grid hierarchy. The associated attribute, component, service, node, and location information on the NMS all change to reflect the alarm's severity. The color displayed reflects the most severe alarm currently active on that branch of the tree. As a result, you can view the general alarm severity level at the grid level, then drill down through the service components to locate the specific details.

For example, in Figure 26 (page 38), the SSM service has at least two alarms: the Events component has at least one alarm with a severity of Notice and the Resources component has at least one alarm with a severity of Minor Alert. Minor Alert is the more severe of the two alarms and therefore it propagates up the tree so that the SSM service takes on the Minor Alert alarm color, light orange.

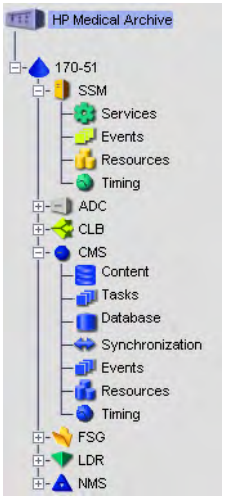


Figure 26 Propagation of Alarm and State Indicators

State Indicators

State indicators are displayed at the services level and above and take priority over alarm indicators. When a service enters either an Administratively Down or Unknown state, the state color is propagated down through the service to its components. This overrides any component alarm indicators displayed in the grid. For example, a service state of "Unknown" supersedes an alarm severity of Critical and results in the service displaying the Unknown state color and not the Critical alarm severity color. The state of the service also propagates up to the node level.

For example, in Figure 26 (page 38), the CMS service has a state of "Unknown." This state overrides any alarms that may have been raised on any of it components. "Unknown" is a more severe state than any alarm and therefore propagates down the navigation tree so that all CMS components take on the "Unknown" state color, blue. The node also displays the "Unknown" state color of blue as this is the most critical state of any of its services. The state condition of a node's services propagates up the navigation tree in the same manner that alarms do.

E-mail Notifications

E-mail notifications to designated personnel automatically alert recipients that an alarm has been triggered or a service state has changed. Managing e-mail notifications is restricted to user accounts that have Maintenance permissions such as the Admin and Vendor accounts.

Alarm Customization

The NMS is configured with a set of default alarms. In addition, it is possible to create custom alarms at the service or component level, or at the grid level.

The Configuration ▶ Alarms page of each service or component is used to view configured Default alarms and Global Custom alarms and to create Custom alarms for a service. Access to this page is restricted to user accounts that have Maintenance permissions such as the Admin and Vendor accounts.

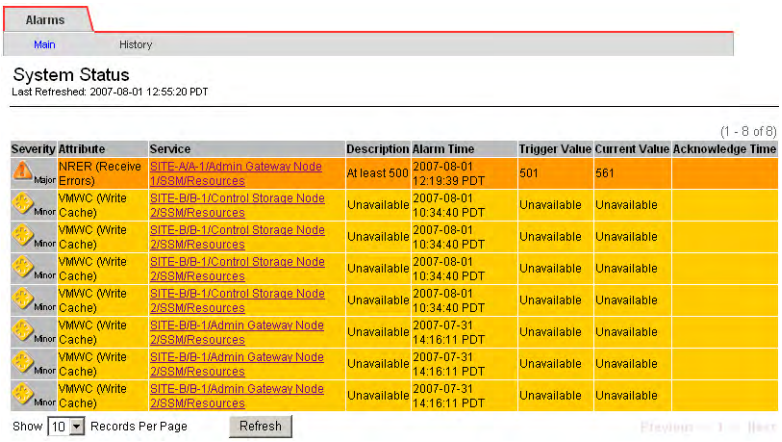
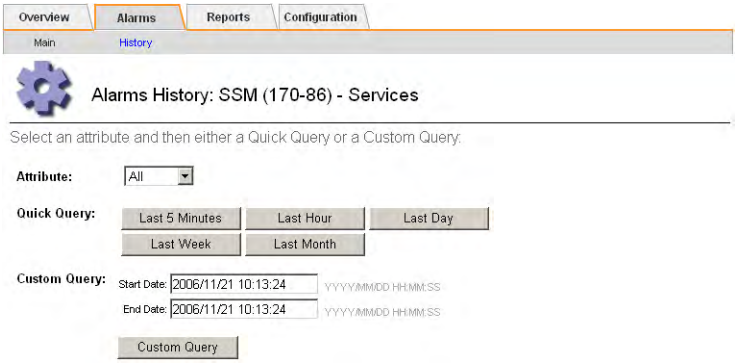
The Grid Management ▶ NMS Management area is used to create Global Custom alarms and to enable or disable Default alarms globally. Access to this page is restricted to user accounts that have Grid Management permissions such as the Vendor account.

Reviewing Alarms

Table 10 (page 40) summarizes how to review the information on the current status of alarms that is found in the following places:

- System Status page
- Alarms tab for each component and service
- Overview tab for each component and service

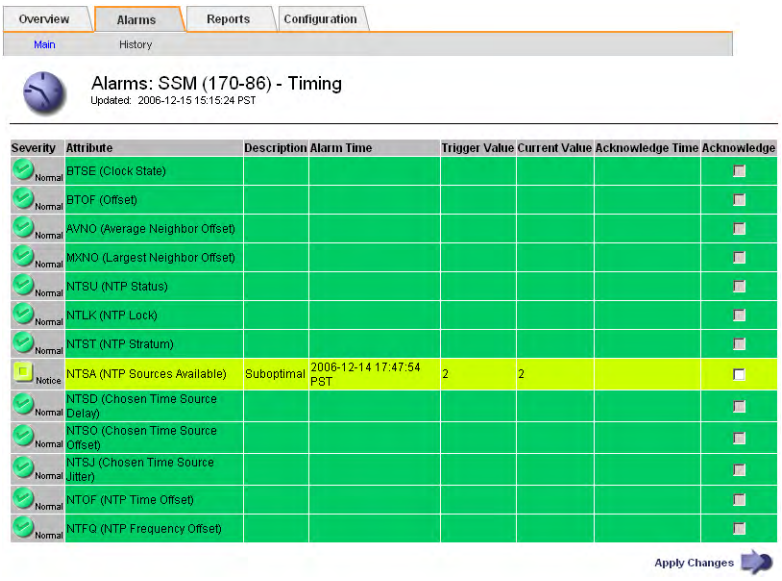
Table 10 Reviewing Alarms

To	Do this
Get a list of all current alarms in the grid	Click **System Status**  in the header to display the grid alarms sorted by severity. 
Get a list of all alarms triggered over a period of time	1 Click **System Status** in the header. 2 Click the **History** page.  3 Do one of the following: — Click one of the time periods. — Enter a custom range and click **Custom Query**.

NOTE The first thing to look at in the NMS is the System Status indicator. It immediately tells you the most serious status (state or alarm) of the grid.

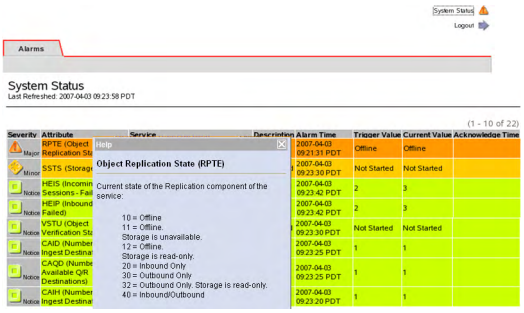
Table 10 Reviewing Alarms *(continued)*

To	Do this
Find out more about an alarm	1 Click the service path in the System Status table to go to the Alarms tab for the selected alarm (or use the Grid Topology Tree). The Alarms tab displays the current status of the attributes for the selected service or component. The colors and icons reflect the severity levels of the alarms. Once an alarm is resolved, the alarm returns to the green “Normal” severity level.



See Table 11 (page 42) for a description of the fields.

2 Click the alarm to display a description of the attribute.



3 Look up the four-character attribute in the reference table in the Troubleshooting chapter of the *user guide*.

4 Click the **Overview** tab.

5 Locate the alarm, and if it can be charted, click the chart button to view a trend of the attribute over the last hour. Adjust the time period as required. See Displaying Charts (page 47) for more information.

Table 10 Reviewing Alarms *(continued)*

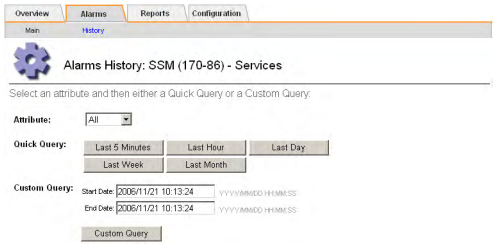
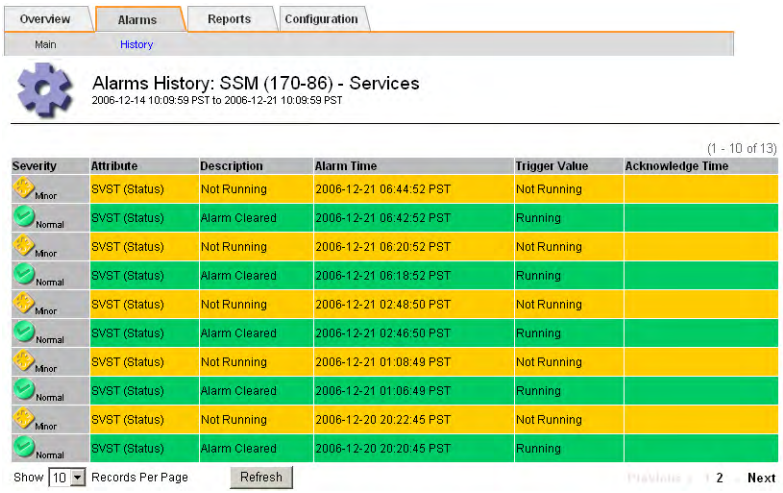
To	Do this
Find out how often alarms have been triggered for a particular attribute	1 Go to the service or component that has the attribute. 2 Click the **Alarms** tab and then the **History** page.  3 Select the attribute from the list. 4 Do one of the following: — Click one of the time periods. — Enter a custom range and click **Custom Query**.  The alarms are listed in reverse chronological order. See Table 11 (page 42) for a description of the fields. 5 To return to the alarms history request form, click **History**.

Table 11 Alarms Table Fields

Field	Description
Severity	Color icon indicating the alarm severity level.
Attribute	Code that identifies the attribute and issue being monitored. See the Troubleshooting chapter in the *user guide* for an alphabetical listing of the attributes and reference information on each alarm.
Description	Brief details about the cause of the alarm.

Table 11 Alarms Table Fields *(continued)*

Field	Description
Alarm Time	The date and time in your local time zone at which the trigger value was reported. This field is blank if the alarm has been acknowledged.
Trigger Value	Value that triggered the alarm.
Current Value	Value of the attribute as last reported by the service or component.
Acknowledge Time	Date and time the alarm was acknowledged.
Acknowledge	Selecting the **Acknowledge** check box acknowledges the alarm. See the procedure in Acknowledging Alarms (page 43).

Acknowledging Alarms

Depending on the situation, you may choose to acknowledge alarms while you are trying to resolve the underlying issue.

Acknowledging alarms is restricted to user accounts that have Alarm Acknowledgement permissions such as the Admin and Vendor accounts.

An acknowledged alarm continues to display as an alarm at the component level on the `System Status` page. However, once an alarm has been acknowledged, it no longer propagates up the tree. The tree is displayed as “Normal” (green) or the color of the next most severe unacknowledged alarm or more severe service state. See Figure 27 (page 43).

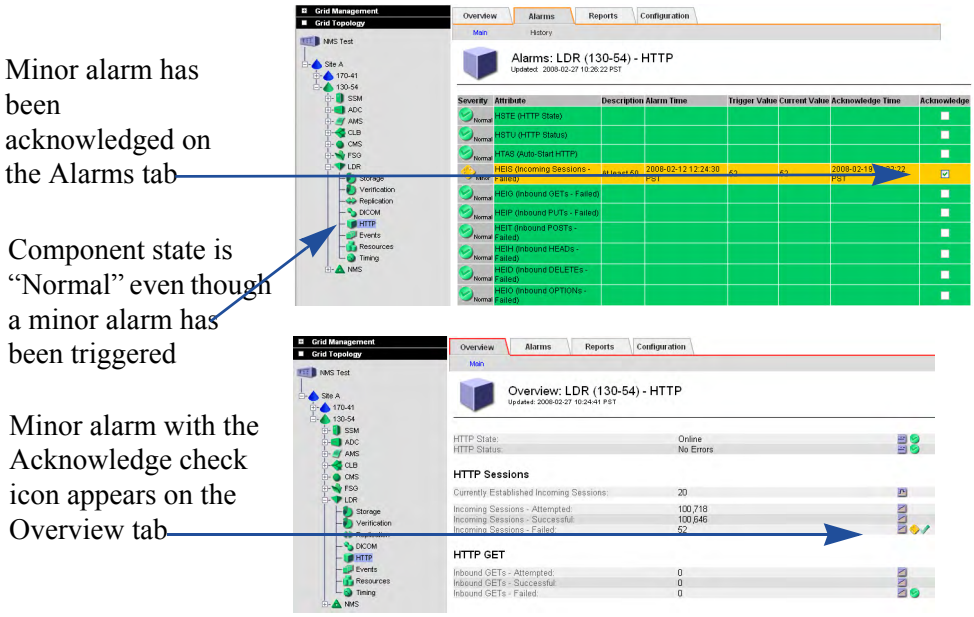


Figure 27 Acknowledging Alarms

There are many reasons why you may want to acknowledge an alarm. For instance, while testing or troubleshooting the grid, you may want to hide (by acknowledging) alarms that you are aware of in order to better track unknown issues. Or, you may, because of time constraints, want to acknowledge an alarm that you can more effectively attend to later.

When a grid is restarted, all unacknowledged alarms are reset to “Normal.” Acknowledged alarms remain unchanged unless the condition that has triggered the alarm has changed and the alarm has reached a new severity level. If the previously acknowledged alarm is triggered, it is considered a new alarm and can be re-acknowledged.

To acknowledge an alarm:

1 Go to the service or component that you are interested in.

2 Click the **Alarms** tab.

3 Select the **Acknowledge** box next to the alarm.

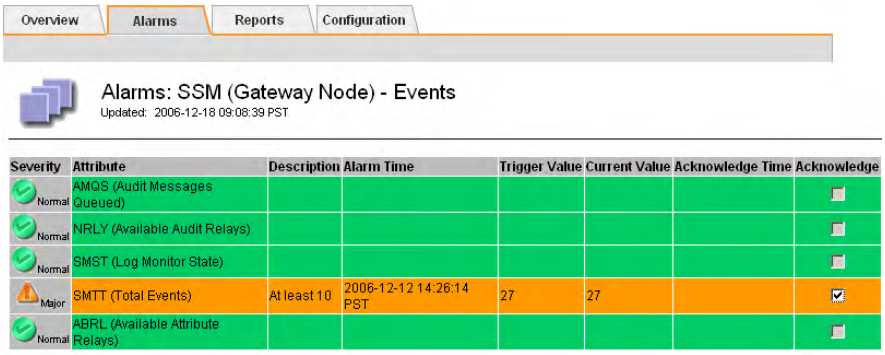


Figure 28 Sample Acknowledged Alarm

4 Click **Apply Changes**. The alarm is acknowledged and a notification is sent to designated personnel.

To unacknowledge an alarm:

1 Go to the service or component that you are interested in.

2 Click the **Alarms** tab.

3 Clear the **Acknowledge** box next to the alarm.

4 Click **Apply Changes**. The alarm is unacknowledged and a notification is sent to designated personnel.

Reports

Reports are an invaluable tool to monitor the state of the grid and to troubleshoot problems. There are two types of reports: chart reports and text reports.

Charts

Chart reports present the data with the attribute value (vertical axis) over a specified time span (horizontal axis).

Chart Types

There are three types of charts:

- line graph
- area graph
- state graph

Line Graph

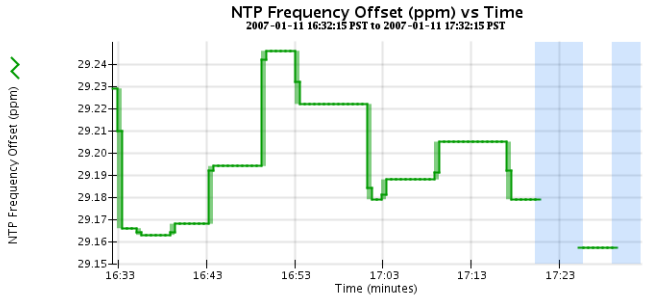


Figure 29 Line Graph

Line graphs are used to plot the values of an attribute that has a “unit” value (such as `NTP Frequency Offset`, in ppm). The changes in the value are plotted in bins at regular intervals over time.

Area Graph

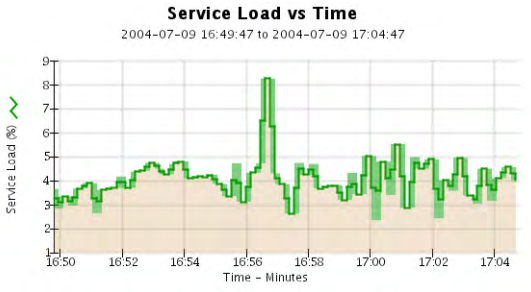


Figure 30 Area Graph

Area graphs are used to plot volumetric quantities, file count or service load values for instance. Area graphs are similar to line graphs but include a light brown shading below the line. The changes in the value are plotted in bins at regular intervals over time.

State Graph

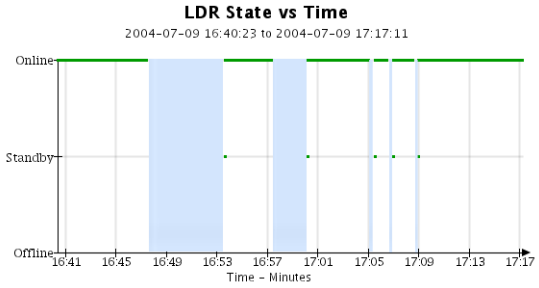


Figure 31 State Graph

State graphs are used to plot values that represent distinct states such as a service state that can be online, standby, or offline. State graphs are similar to line graphs but the transition is discontinuous, that is, the value jumps from one state value to another.

Interpreting Chart Colors

The chart colors have a specific meaning. Table 12 (page 46) describes how to interpret the various colors and line types.

Table 12 Chart Colors and Shading

Sample	Meaning
	Reported attribute values are plotted using dark green lines.
	Light green shading around dark green lines indicates that the actual values in that time range vary and have been “binned” for faster plotting. The dark line represents the weighted average. The range in light green indicates the maximum and minimum values within the bin. Light brown shading is used for area graphs to indicate volumetric data.
	Blank areas (no data plotted) indicate that the attribute values were unavailable. The background may be blue, gray, or a mixture of gray and blue depending on the state of the service reporting the attribute

Table 12 Chart Colors and Shading *(continued)*

Sample	Meaning
	Light blue shading indicates that some or all of the attribute values at that time were indeterminate; the attribute was not reporting because the service was in an unknown state.
	Gray shading indicates that some or all of the attribute values at that time were not known because the service reporting the attributes was administratively down.
	A mixture of gray and blue shading indicates that some of the attribute values at the time were indeterminate (because the service was in an unknown state), while others were not known because the service reporting the attributes was administratively down.

Displaying Charts

The fastest way to create a chart is to click the chart button on the Overview tab. This is called an immediate report. You can also create reports from the Report tab.

NOTE There are some attributes for which it is not possible to create charts, for example, text attributes such as Node ID, version number, and build number.

To create a chart from the Overview tab:

1 Go to the component or service that has the attribute you are interested in.

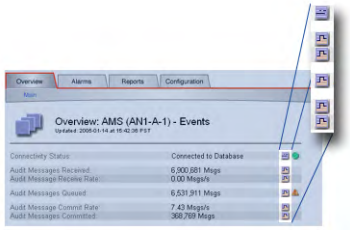


Figure 32 Sample Immediate Report Buttons

2 Click the chart button next to the attribute to display a chart. The display automatically changes to the `Reports` tab. The chart displays the attribute's data over the past hour. To view other time ranges, follow the procedure described below.

To create a chart from the `Reports` tab:

1 Click the **Reports** tab of the service or component you are interested in.

2 Choose an attribute from the **Attribute** list.

3 To force the Y-axis to start at zero, clear **Vertical Scaling**.

4 To show values at full precision, select **Raw Values**. To round values to a maximum of three decimal places (for example, for attributes reported as percentages), clear **Raw Data**.

5 Choose the time period for the chart from the Quick Query list. The chart appears after a few moments. Allow several minutes for long time ranges.

6 To display a chart for a custom time period:

   a Choose **Custom Query** from the **Quick Query** list.

   b Enter the **Start Date** and **End Date**.

   Use the format YYYY/MM/DD HH:MM:SS in local time. Leading zeros are required to match the format. For example, 2007/4/8 7:30:00 fails validation; the correct format is 2007/04/08 07:30:00.

   c Click **Update**.

Displaying Multiple Charts

When you generate a chart report, it is often useful to compare it to another chart. The NMS provides the ability to view the data in a new window in order to hold it while you use the main interface to select other reports or views.

To open a new chart window:

- Click  to display the current view in new window.

  You may open as many of these windows as you wish.

  When you are done, click  to close the chart windows.

Text Reports

Text reports present the data in a simple table with the local date and time in the first column and the attribute values in the second column. For most attributes, data is sent to the NMS only when the value changes. Text reports do not list any data during periods when the attribute was either unavailable or was unreported (indeterminate).

Gray text indicates the service was administratively down during that time. Blue text indicates the service was in an unknown state.

Text Results for Active Session ID
2008-03-23 12:44:16 PDT To 2008-04-23 12:44:16 PDT

Time	Value
2008-04-22 20:02:40	1208919755488938
2008-04-21 20:02:37	1208833348536926
2008-04-20 20:02:22	1208746940730243
2008-04-19 20:02:19	1208660532166872
2008-04-18 20:02:05	1208574123839836
2008-04-17 20:01:56	1208487716300372
2008-04-16 20:01:57	1208401308989401
2008-04-15 20:01:47	1208314901386964
2008-04-16 01:01:22	1208314901386964
2008-04-15 20:01:47	1208314901386964
2008-04-14 20:01:39	1208228493805419

Figure 33 Simple Text Report

When data is reported over longer time periods, the information is downsampled and the report displays the average, minimum, and maximum values of the attributes in the bin for each time period.

Text Results for Inbound Replications - Completed
2008-02-11 12:22:59 PST To 2008-03-11 12:22:59 PDT

Time	Average Value	Minimum Value	Maximum Value
2008-03-11 12:18:09	1,236,776	1,236,561	1,237,343
2008-03-11 12:12:09	1,235,585	1,235,302	1,236,346
2008-03-11 12:04:09	1,234,073	1,233,753	1,235,151
2008-03-11 11:56:09	1,231,467	1,231,462	1,232,971
2008-03-11 11:48:09	1,229,868	1,229,856	1,230,488
2008-03-11 11:40:08	1,228,081	1,227,902	1,228,752
2008-03-11 11:34:08	1,226,534	1,226,357	1,227,497
2008-03-11 11:26:08	1,225,393	1,225,304	1,226,064
2008-03-11 11:18:08	1,223,584	1,223,361	1,224,433
2008-03-11 11:10:08	1,221,501	1,221,200	1,222,340
2008-03-11 11:04:08	1,219,800	1,219,800	1,220,403
2008-03-11 10:56:07	1,218,780	1,218,674	1,219,177
2008-03-11 10:48:07	1,217,532	1,217,377	1,218,109
2008-03-11 10:42:07	1,216,344	1,216,164	1,217,027
2008-03-11 10:34:07	1,214,916	1,214,913	1,215,712
2008-03-11 10:26:06	1,213,698	1,213,551	1,214,176
2008-03-11 10:20:06	1,212,262	1,211,956	1,213,252
2008-03-11 10:12:06	1,210,664	1,210,664	1,211,401
2008-03-11 10:04:06	1,209,499	1,209,483	1,209,888
2008-03-11 09:56:06	1,208,856	1,208,856	1,208,990

Figure 34 Downsampled Text Report

Displaying Text Reports

To display a text report:

1. Click the **Reports** tab of the service or component you are interested in.
2. Click the **Text** page.
3. Select an attribute from the **Attribute** list.
4. Select the number of values reported per page from the **Results per Page** list.
5. To round values to a maximum of three decimal places (for example, for attributes reported as percentages), clear **Raw Data**.
6. Select the time period for the report from the **Quick Query** list.

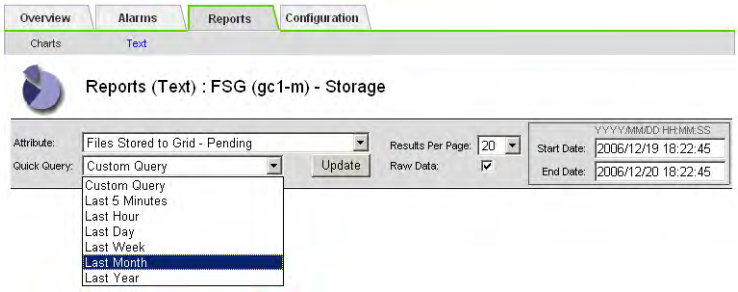


Figure 35 Report Request Form—Quick Text Reports

The report appears after a few moments. Allow several minutes for tabulation of long time ranges.

7 To display a report for a custom time period:

a Select **Custom Query** from the **Quick Query** list.

b Enter the **Start Date** and **End Date**.

Use the format YYYY/MM/DD HH:MM:SS in local time. Leading zeros are required to match the format. For example, 2007/4/8 7:30:00 fails validation; the correct format is 2007/04/08 07:30:00.

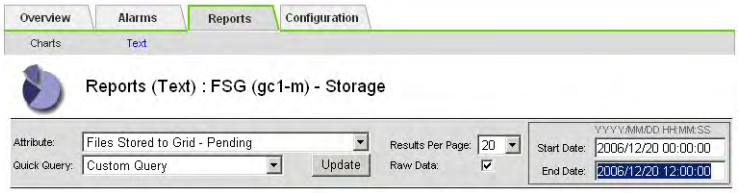


Figure 36 Report Request Form—Custom Text Report

c Click **Update.**

Printing

To print a chart or text report:

1 Create the desired chart or text report.

2 Optionally, open charts in a new window. See Displaying Multiple Charts (page 48).

3 Right-click the chart or report to display the context menu.

4 Click either **Print** or **Print Picture** to open the Print dialog.

5 Use the dialog to select what you need. Depending on the context menu, the Options tab in the dialog may include the option to print **Only the selected frame**. This is the recommended option.

6 Click **Print**.

3 Data Flow

This chapter describes the grid activities that take place as objects are ingested, replicated, retrieved, modified, and purged.

Key Concepts

To follow objects through the grid as they are processed, you need to understand these concepts:

- Client shares and FSG managed file system
- File System Gateway (FSG) replication groups
- Topology queries
- HTTP protocol commands
- ILM policy
- Owner CMS and metadata replication
- Object content handle

Client Shares and FSG Managed File System

In order to communicate with the FSG, client applications map a network drive to the FSG file share, for example, `/fsg/myDirectory`. The FSG supports CIFS and NFS file share protocols.

Applications interface with the grid via the FSG's "managed file system". There is one managed file system per FSG and it is mounted at `/fsg`. The managed file system contains "file stubs" that point to the file's location in the grid. The FSG managed file system shown in Figure 37 (page 52) contains a file stub for the object `/fsg/myDirectory/image.jpg`. The FSG uses the object's unique identifier to communicate with the CMS (via the LDR) which tracks where the object is stored on the Storage Nodes (LDR service) and on the Tape Nodes (ARC service).

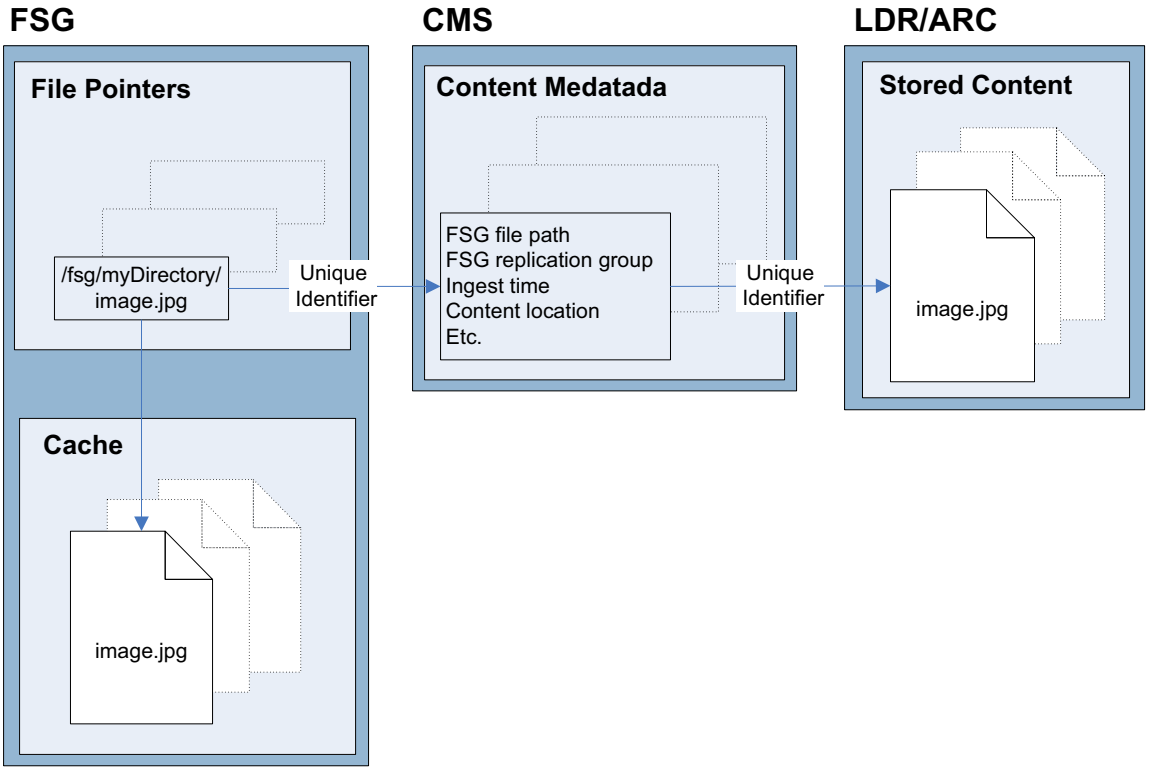


Figure 37 FSG File Pointers

The FSG cache contains recently stored and retrieved files. The purpose of the cache is to speed up data access for clients. Unaccessed files in the FSG cache are swapped out to make room for new files, leaving a stub that allows the file data to be re-cached if the client later retrieves the file. Files are swapped out of the FSG cache in approximately least-recently accessed order.

FSG Replication Groups

All FSGs belong to replication groups. A replication group contains a Primary FSG and one or more Secondary FSGs. The Primary FSG provides read and write access to the clients. The Secondary FSG provides a mirror of the Primary FSG's managed file system for redundancy in case the Primary FSG fails or must be taken out of service.

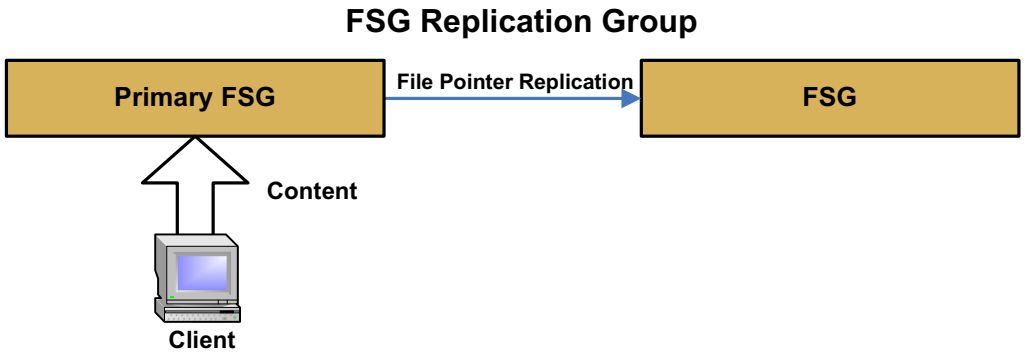


Figure 38 FSG Replication Group

The Primary FSG replicates its file pointers (for instance Figure 39 (page 53) shows the pointer for `/fsg/myDirectory/image.jpg`) to the Secondary FSG.

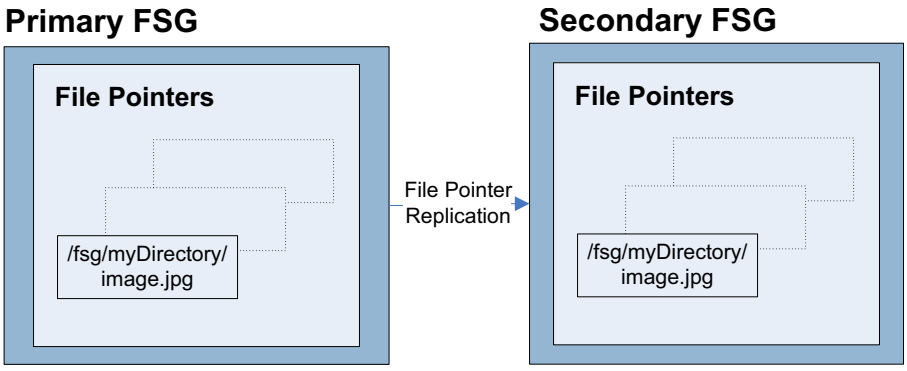


Figure 39 FSG File Pointer Replication

Depending on the configuration, the Secondary FSG may also provide read-only client access and may perform backups of the file system into the grid for additional redundancy.

There are two types of FSG replication groups:

- Unclustered
- High Availability Gateway Cluster

For a detailed review of each type of replication group, see the *user guide*.

Topology Queries

When a grid service needs information from another service or needs an action to be performed by another service, it contacts the ADC to find the best service to process the request. This is called a topology query. The ADC responds to each query with the latest information received from the grid. The information maintained by the ADC includes CPU load, amount of available disk space, supported services, location, and so on.

HTTP Protocol Commands

Internally, FSGs use standard HTTP protocol commands to communicate with LDRs to store, retrieve, and purge objects. For example, the FSG issues a PUT command to the LDR to store an object into the grid, a GET command to retrieve the object, and a DELETE command to purge the object.

ILM Policy

The ILM policy defines how and where objects are stored in the grid. At a high level, ILM policies dictate:

- geography—the location of the files
- storage grade—what type of storage to use
- replication—the number of copies to make

Location, storage grade, and number of copies can vary over time. In the ILM example shown in Figure 40 (page 54), a file is ingested in the grid via an FSG. At ingest, two copies of the object are stored in the Data Center on Fibre Channel disks and one copy is stored in the Disaster Recovery site on SATA disks. One year after ingest, one copy at the DC is deleted and one copy is created on archive media at the DR site.

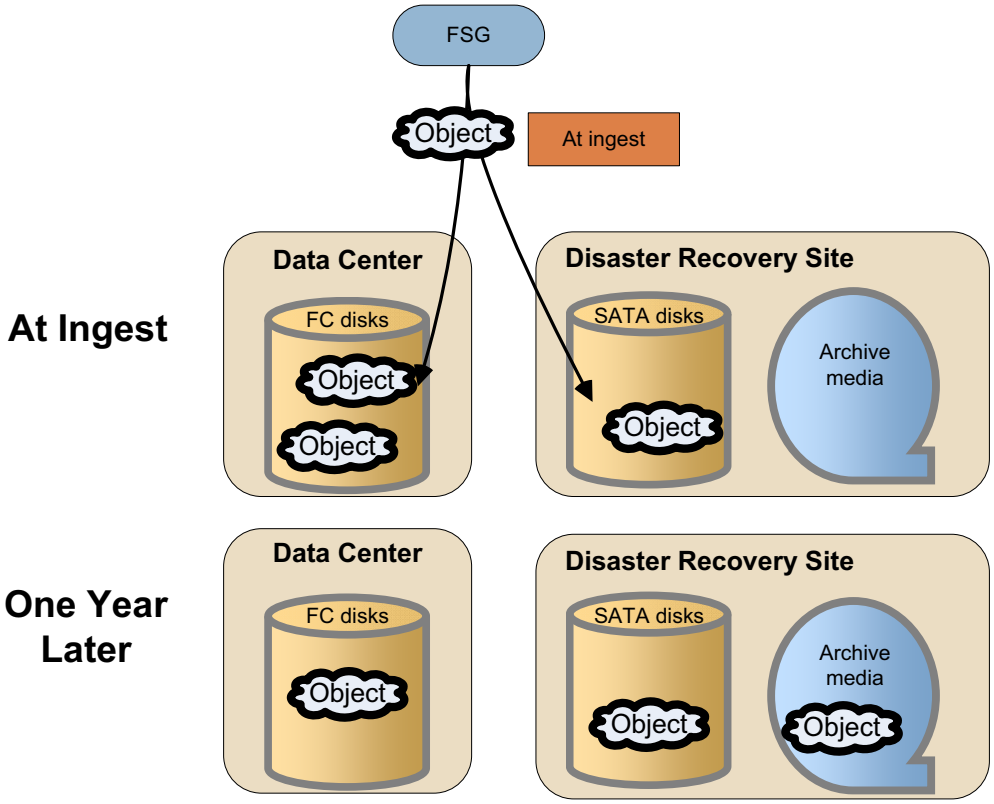


Figure 40 Information Lifecycle Management (ILM) Example

Owner CMS and Metadata Replication

Each object ingested into the grid has a set of associated metadata. Metadata is information related to the object, for instance the ingest file path or the file ingest time. Content metadata is managed by the CMS. The first CMS to get the object metadata is called the owner CMS. The owner CMS then replicates the metadata to other CMSs. If grid has distributed CMSs, the metadata is replicated to the CMSs in the same replication group as the owner CMS and to the CMS in

whatever groups the data is stored. Otherwise, the CMS databases are synchronized: all CMSs store all content metadata. (In a synchronized grid that has been expanded to add metadata storage capacity, the metadata is stored on all CMSs of the same generation).

Content Handle

The grid assigns a unique identifier to each object ingested into the grid. This identifier is called a “content handle”. The grid uses the content handle to refer to the object. As long as the object is referenced by the client application, it is said to have a content handle. When the client application deletes the object, the object’s content handle is said to be released.

Object Lifecycle

Figure 41 (page 57) follows an object as it is ingested, retrieved, becomes inactive, and is finally deleted:

- The client application creates the file over CIFS/NFS. The file is ingested into the grid in the background from the FSG to an LDR over HTTP and is replicated according to the ILM policy. The file is also stored in the FSG cache.
- When the client reads the file, the file is retrieved either from the FSG cache over CIFS/NFS or from the LDR to the FSG and out to the client over CIFS/NFS.
- An inactive file may be swapped out of the FSG cache to make room for more recent or more active files
- When the client deletes the file, removal from the FSG triggers a removal notification to the LDR.

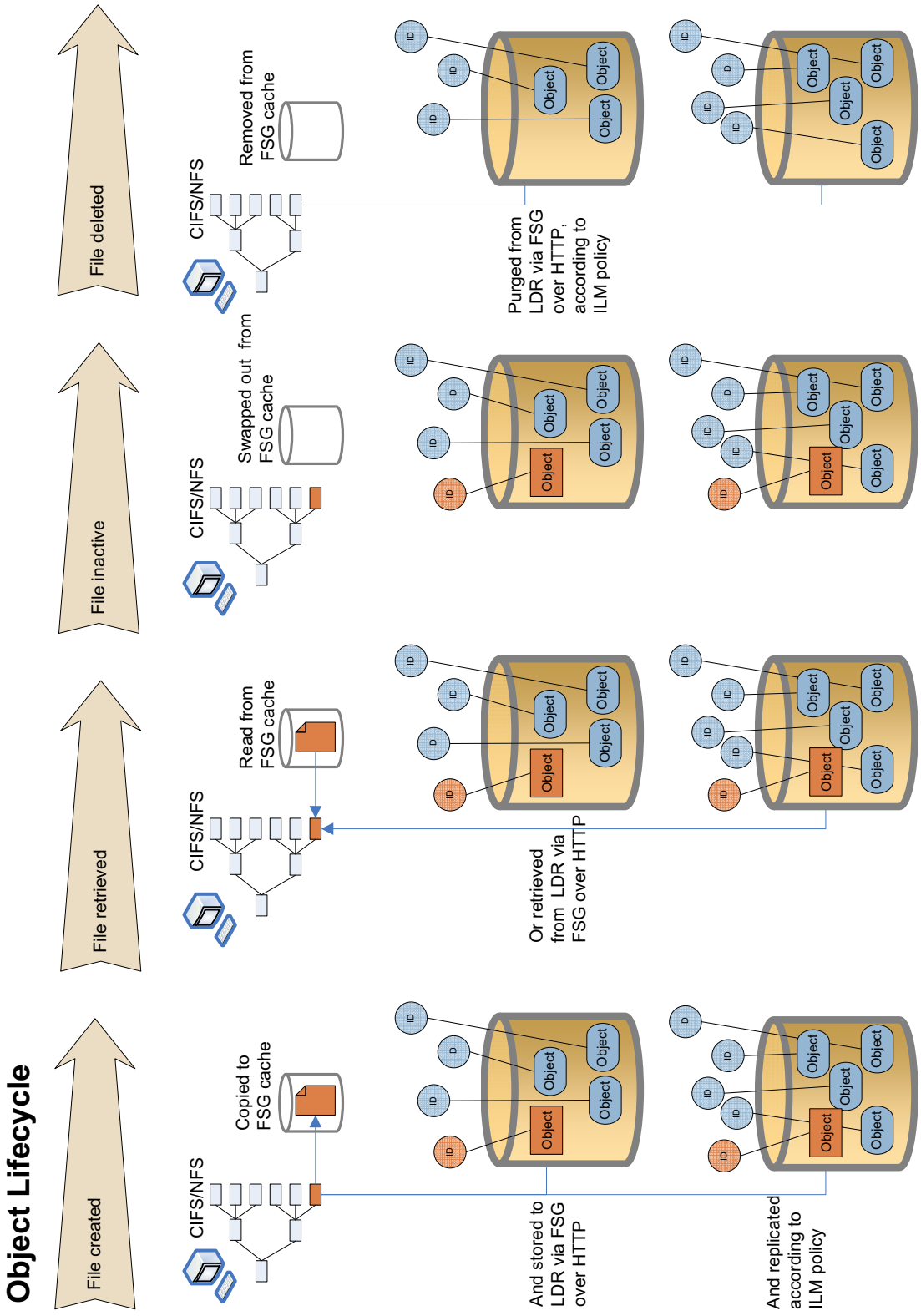


Figure 41 Object Lifecycle

Ingest

Object ingest refers to what happens when client applications store content into the grid.

Data Flow

See Figure 42 (page 58) for a simplified step-by-step description of what happens when objects are ingested into the grid.

When a client saves a file to the FSG file system (for example to the mapped network drive /fsg/myDirectory), the FSG stores a local copy in its cache and streams the file to permanent storage via an LDR. The LDR assigns a unique identifier called “content handle” to the file and transmits this information to the FSG. The LDR also transmits the object metadata to the CMS. The Primary FSG replicates the content handle and file pointers to the other FSGs in its replication group, and the CMS replicates the metadata to other CMSs in the grid.

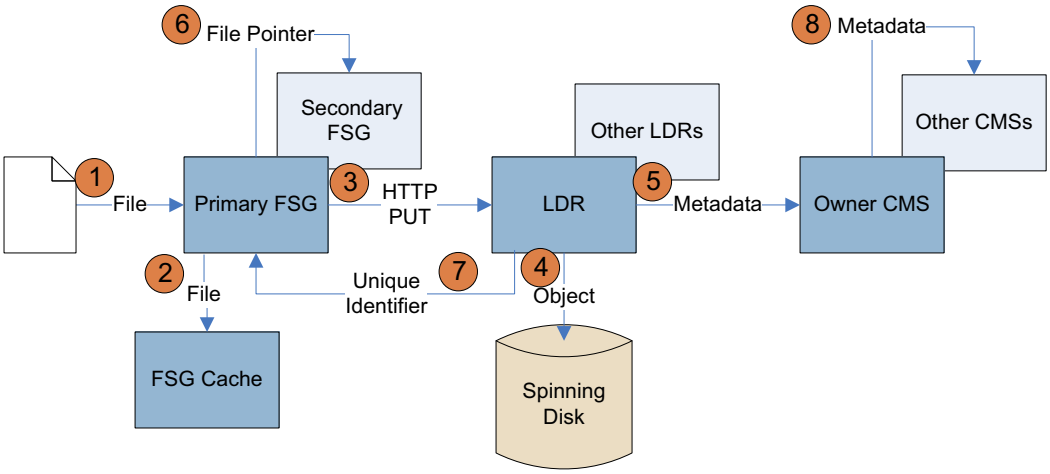


Figure 42 Ingest Data Flow

1	The client application saves a file to the Primary FSG managed file system via its mapped network share, for example /fsg/myDirectory. This triggers an FSG “create operation” and as a result the FSG creates a file pointer for this object.
2	The Primary FSG saves a copy of the object in its cache.
3	The Primary FSG stores the file to an LDR via a HTTP PUT command. What LDR is chosen depends on the result of an ADC topology query.
4	The LDR saves the object to spinning disk and allocates a “content handle”, that is, a unique identifier, to the object.

5	The LDR notifies a CMS that a new piece of content has been ingested and sends the object metadata, which includes the unique identifier, to the CMS. This CMS becomes the owner CMS. What CMS is chosen depends on the result of a topology query.
6	The Primary FSG replicates the file pointer to the other FSGs in its replication group.
7	The LDR sends the unique identifier to the Primary FSG.
8	The owner CMS replicates the metadata and content location to the other CMSs. If the grid has distributed CMSs, metadata is replicated to a subset of the CMSs; otherwise, content metadata is synchronized across all CMSs.

Related Attributes

NOTE The actual shape of the trends varies with each grid and depends on ingest, replication, and purging rates.

Table 13 (page 60) lists some of the NMS attributes used to track what happens when a single object is ingested into the grid.

Table 13 Object Ingest Attributes

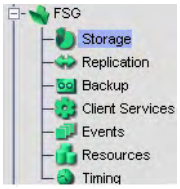
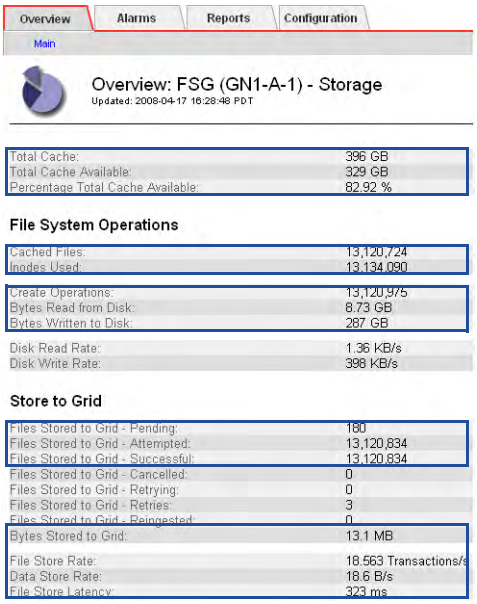
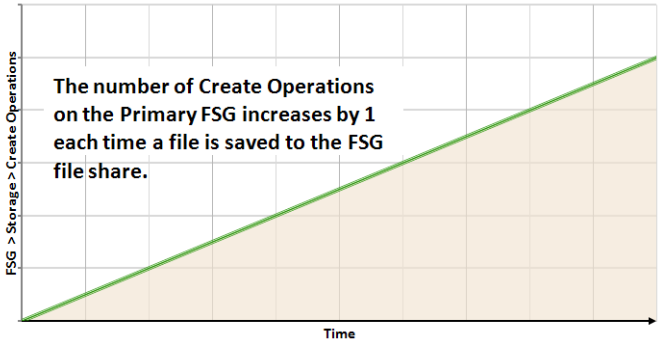
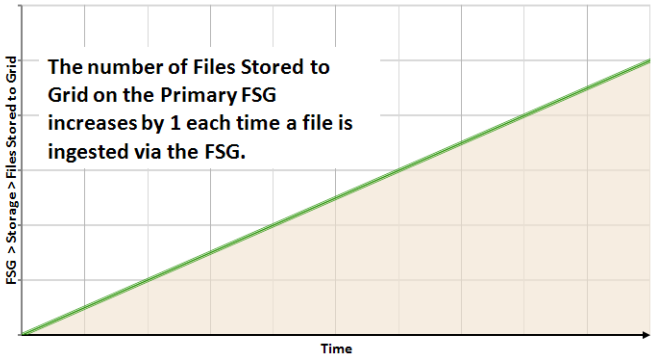
Component	Attribute Changes
Primary FSG ▶ Storage   	Create Operations (FRCO): The number of new files or folders created in the file system increases by 1.  Inodes Used (FSIU): The number of inodes (files and directories) used on the FSG file system increases by 1. Inodes Used and Create Operations can differ if objects are deleted and their inodes are later re-used for new objects. Files Stored to Grid - Successful (FSGC): The number of files stored persistently on the grid increases by 1. The value increases after the FSG receives an acknowledgment from the LDR that the file has been successfully stored into the grid (which is at the same time as it receives the unique identifier from the LDR).  Files Stored to Grid - Attempted (FSGA): The number of initiated file transfers to the grid increases by 1.

Table 13 Object Ingest Attributes *(continued)*

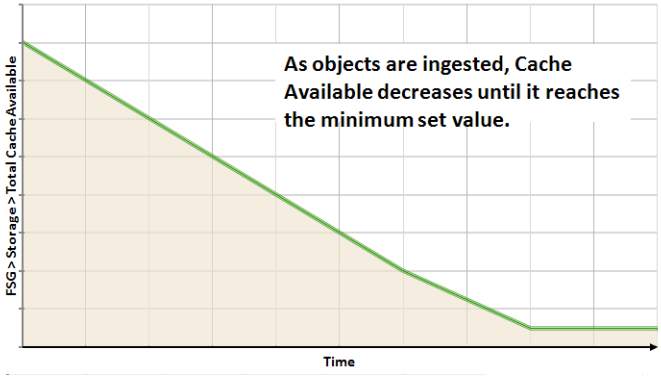
Component	Attribute Changes
	Files Stored to Grid - Pending (FSGP): The number of new files cached locally and waiting for transfer to the grid for persistent storage increases by 1 temporarily, and decreases again once the file has been stored. Bytes Stored to Grid (FSGB): The number of bytes ingested successfully into the grid increases by an amount equivalent to the file size. Bytes Read from Disk (FSRB): The number of bytes read from disk increases. File ingest generates multiple read operations as the FSG reads the file and stores it to the grid. Bytes Written to Disk (FSWB): The number of bytes written to disk increases. File ingest generates multiple write operations to disk for the creation of the file and the saving of the file content and metadata. Total Cache Available (FSTA): The total local cache space on the Primary FSG still available for use decreases unless the minimum value has been reached. After that, the Primary FSG swaps files out of the cache to make room for the file that has been ingested. The cache may dip below the minimum value temporarily if files are being created faster than existing files can be swapped out. 

Table 13 Object Ingest Attributes *(continued)*

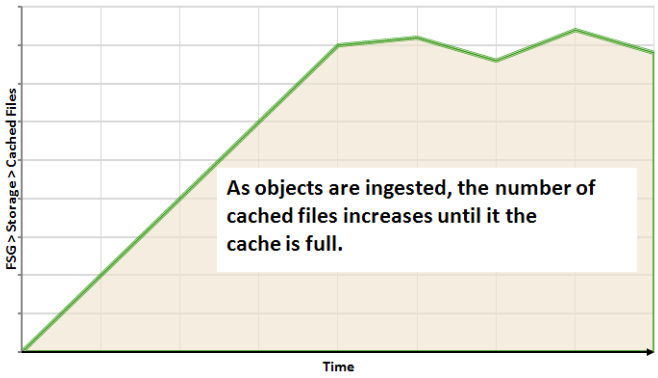
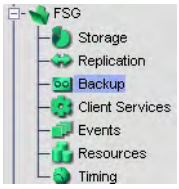
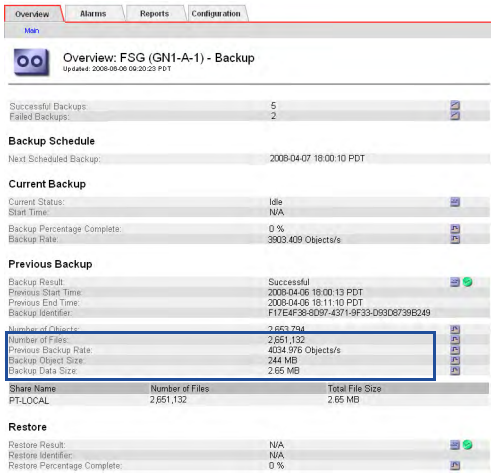
Component	Attribute Changes
	Cached Files (FSCF): The number of cached files on the Primary FSG increases by 1 unless the cache is full. The number of cached files will generally stabilize once the Total Cache Available reaches its minimum, but may fluctuate as files are added and removed from the cache depending on file size. 
Primary FSG ▶ Backup  	Number of Files (PBSF): The number of files included in the backup for this share will increase by 1 the next time a backup takes place. Total File Size (PBSB): The total size of all files referenced by the backup for this share will increase by an amount equivalent to the file size the next time a backup takes place.

Table 13 Object Ingest Attributes *(continued)*

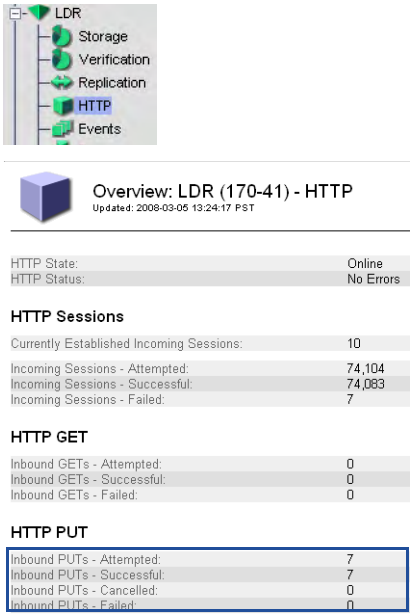
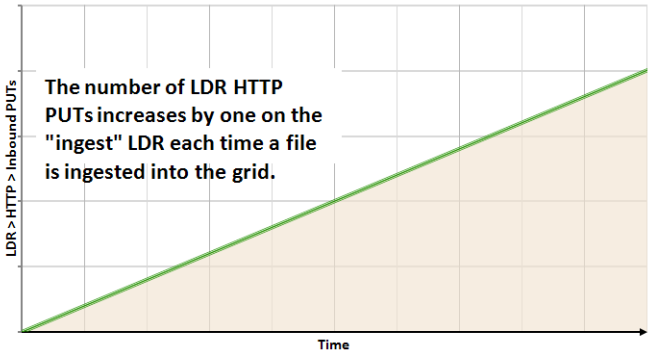
Component	Attribute Changes
LDR ▶ HTTP 	Inbound PUTs - Successful (HIPC): The total number of HTTP PUT (“content store”) requests that have been completed successfully by the LDR increases by 1.  Inbound PUTs - Attempted (HAIP): The total number of HTTP PUT (content store) requests that have been received by the LDR also increases by 1.

Table 13 Object Ingest Attributes *(continued)*

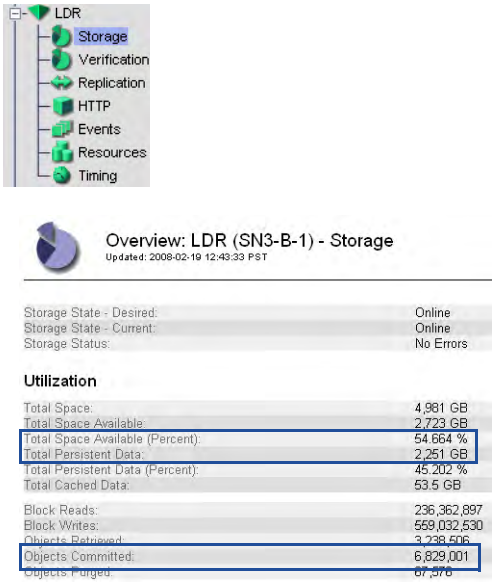
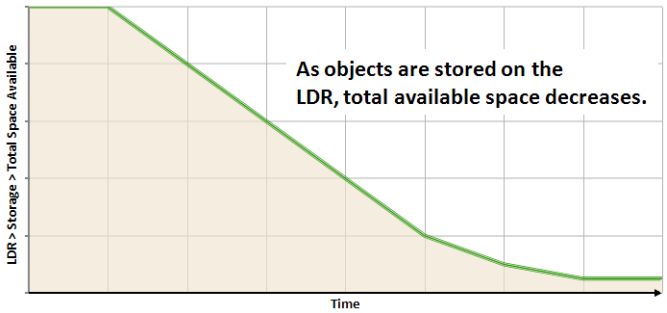
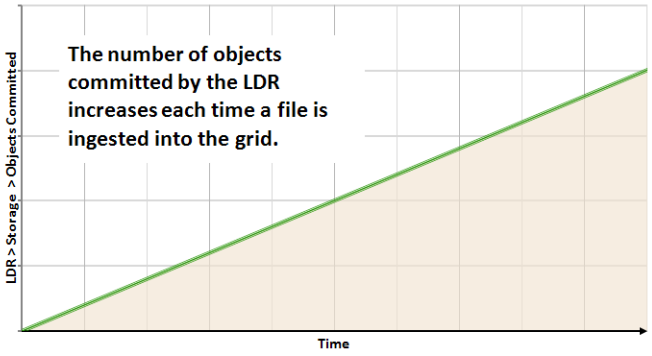
Component	Attribute Changes
LDR ▶ Storage 	Total Space Available (Percent) (SAVP): The total object storage space currently available (unused) decreases by an amount roughly equivalent to the file size.  Objects Committed (OCOM): The number of object store operations that have been processed by the LDR increases by 1.  Total Persistent Data (SPSD): The estimate of the size of the persistently stored data increases by an amount roughly equivalent to the file size.

Table 13 Object Ingest Attributes *(continued)*

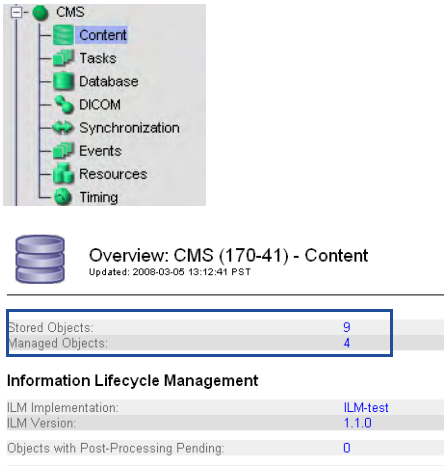
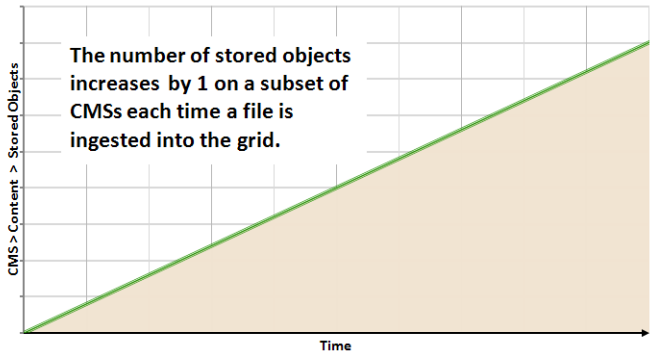
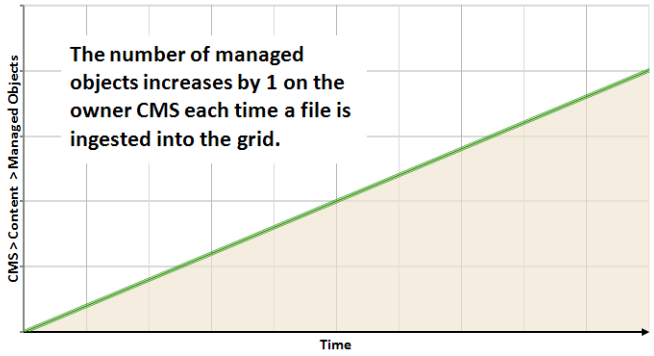
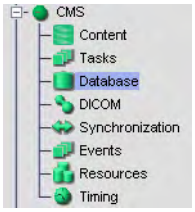
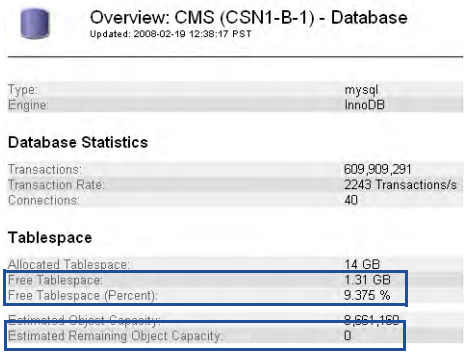
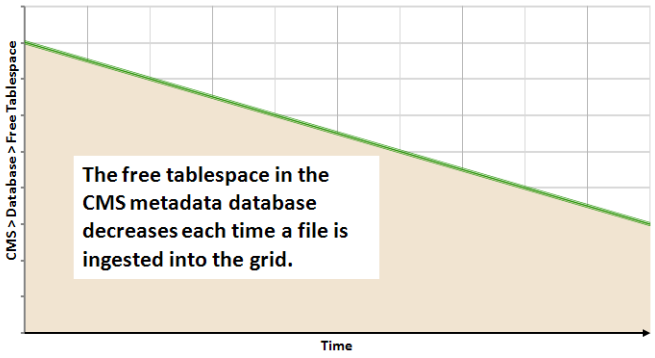
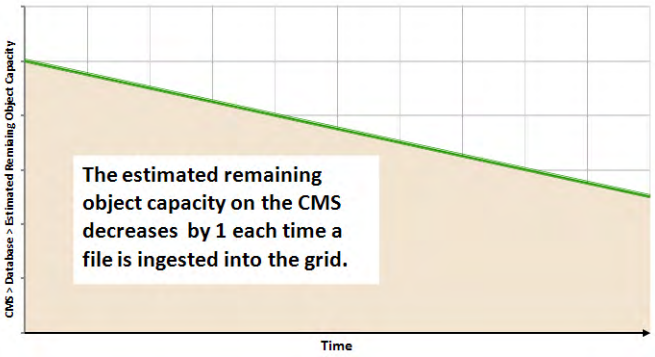
Component	Attribute Changes
CMS ▶ Content  NOTE For grids with synchronized CMSs, the attributes Stored Objects and Managed Objects are shown under the Content component. For grids with distributed CMSs, these two attributes appear under the Metadata component.	Stored Objects (COoT): The number of objects in the CMS metadata database increases by 1 on all CMSs of the same generation in a grid that uses distributed CMSs and all CMSs in the same CMS replication group in a grid that uses distributed CMSs.  Managed Objects (COoM): The number of objects owned by the owner CMS increases by 1. 

Table 13 Object Ingest Attributes *(continued)*

Component	Attribute Changes
CMS ▶ Database   	Free Tablespace Percent (DBSP): The amount of space remaining in the metadata database decreases. MySQL manages free tablespace in chunks. Ingesting (or deleting) a single object may not change the reported free table space. Ingesting (or deleting) many objects will eventually change the free table space.  Estimated Remaining Object Capacity (CORS): The estimate of how many more objects can be tracked in the CMS metadata database decreases by 1. 

Content Replication

Following ingest, the object is replicated according to the grid's ILM policy. Content replication refers to the process of making copies of the object in order to satisfy the ILM policy.

Data Flow

The owner CMS, which is the first CMS to receive the object metadata from the LDR, controls the replication, that is, ensures that the correct number of copies are stored in the correct locations for the duration specified by the ILM policy. See Figure 43 (page 67) for a simplified step-by-step description of what happens as objects are replicated.

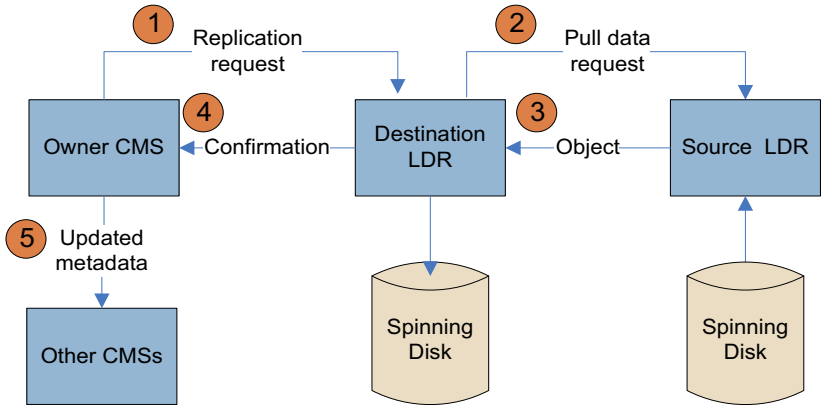


Figure 43 Replication Data Flow

1	The owner CMS queries the ADC to determine the best destination LDR within the storage pool defined by the ILM policy, and sends that LDR a command to initiate replication.
2	The destination LDR queries the ADC for the best source location and sends a replication request to the source LDR.
3	The source LDR sends a copy of the object to the destination LDR.
4	The destination LDR notifies the CMS that the object has been stored.
5	The owner CMS updates the location information and distributes that information to the other CMSs that store metadata for this object.

Related Attributes

Table 14 (page 68) lists some of the NMS attributes used to track what happens when a single object is replicated.

Table 14 Object Replication Attributes

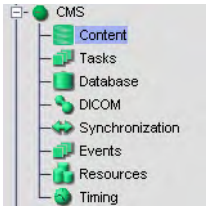
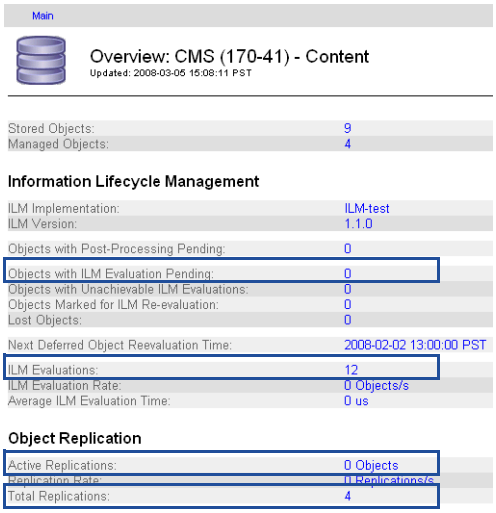
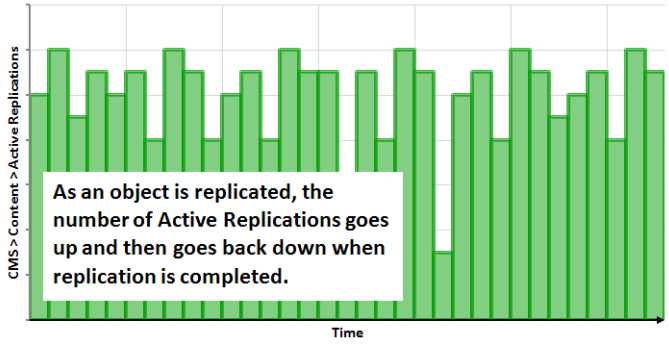
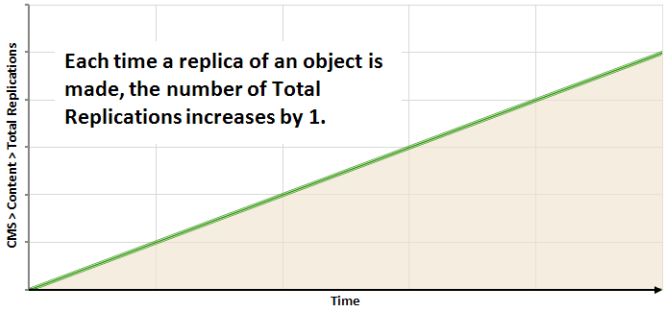
Component	Attribute Changes
CMS ▶ Content   	Objects with ILM Evaluation Pending (ORpe): The number of objects waiting to be processed through the business rules for replication increases by 1 when the object is ingested into the grid and decreases by 1 once replication is complete. `Active Replications (DCdA)`: The number of objects in the process of being replicated increases by 1 when replication starts and decreases by 1 when replication is complete.  ILM Evaluations (ILev): The total number of ILM evaluations that have been performed to date increases when the object is evaluated after ingest and again when the object has been replicated. `Total Replications (DCdT)`: The total number of object replications performed by the owner CMS since grid startup increases by 1 each time a copy is made. 

Table 14 Object Replication Attributes *(continued)*

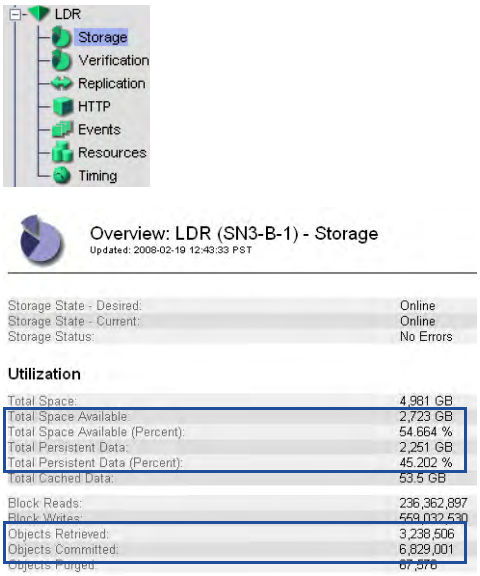
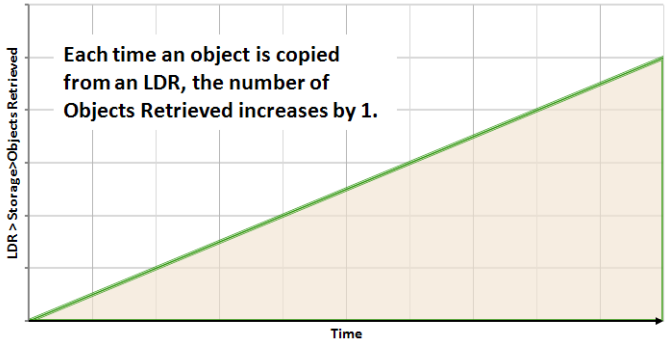
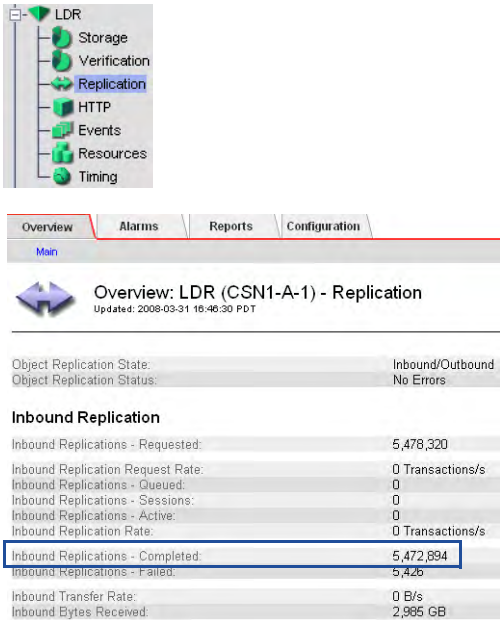
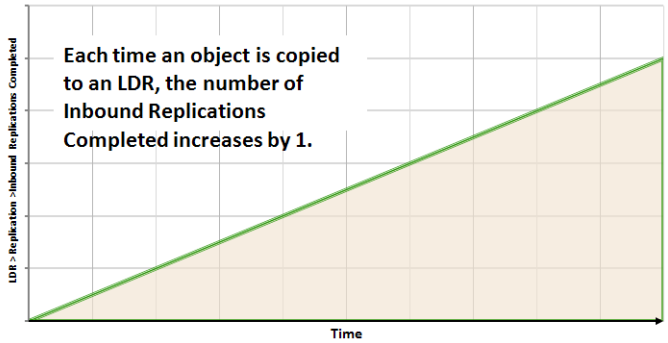
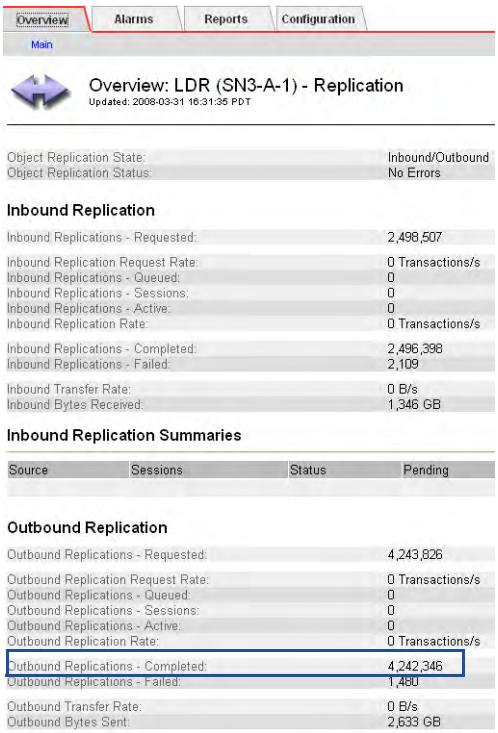
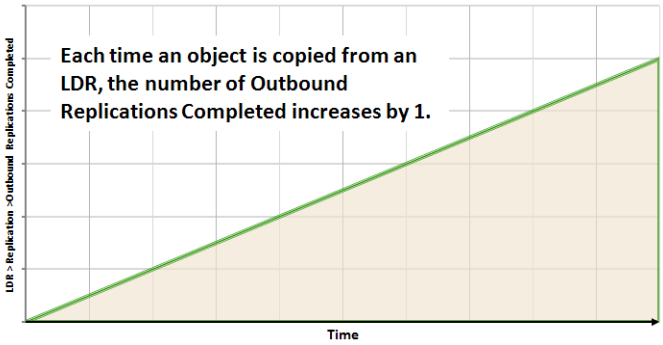
Component	Attribute Changes
LDR ▶ Storage 	`Objects Retrieved (ORET)`: The number of persistent objects retrieved from the source LDR increases by 1.  `Objects Committed (OCOM)`: The number of persistent objects stored on each destination LDR increases by 1. `Total Space Available (SAVP)`: The percentage of object storage space available for use on each destination LDR decreases by an amount roughly equivalent to the size of the replicated object. `Total Persistent Data Percent (SPDP)`: The percentage of the total storage space used by persistent data on each destination LDR increases by an amount roughly equivalent to the size of the replicated object.
LDR ▶ Replication 	`Inbound Replications Completed (RIRC)`: The total number of objects replicated to the destination LDR increases by 1. 

Table 14 Object Replication Attributes *(continued)*

Component	Attribute Changes
	`Outbound Replications Completed (RORC)`: The total number of objects replicated from the source LDR increases by 1. 

Content Replication to Archive Media

The Tape Node provides an interface between the grid and an archival media device which is external to the grid. The Tape Node communicates with a middleware layer that manages access to the physical storage device. Currently supported archive devices include any storage device managed by Tivoli® Storage Manager (such as a tape library).

Data Flow

If the ILM policy requires an object to be stored on archive media, the CMS sends a request to the Tape Node which in turn sends the object to the middleware, see Figure 44 (page 71).

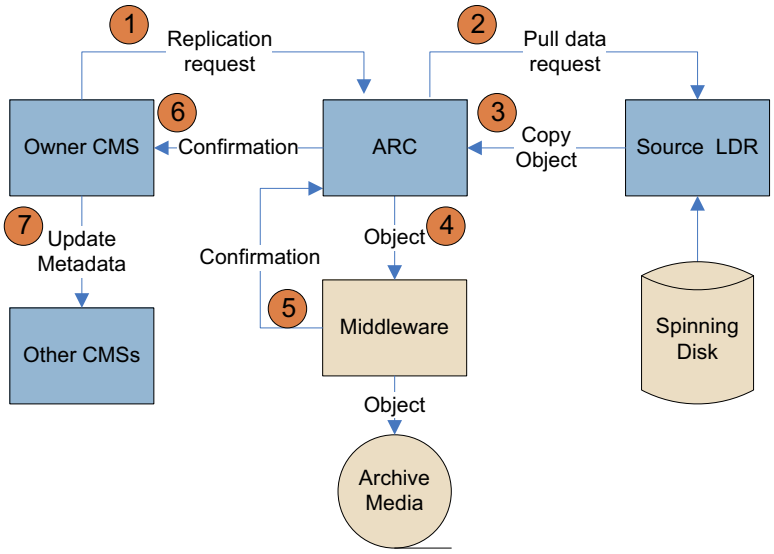


Figure 44 Archiving Data Flow

Related Attributes

1	The owner CMS sends a request to the ARC to store a copy of the object on archive media.
2	The ARC queries the ADC for the best source location and sends a request to the source LDR.
3	The ARC retrieves the object from the LDR.
4	The ARC sends the object to the archiving middleware which in turn copies it to the archive media.

5	The middleware notifies the ARC that the object has been stored.
6	The ARC notifies the CMS that the object has been stored.
7	The owner CMS updates the location information and distributes that information to the other CMSs that store metadata for this object.

Table 15 (page 72) lists some of the NMS attributes used to track object replication to archive media.

Table 15 Object Replication to Tape Node Attributes

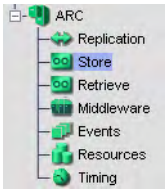
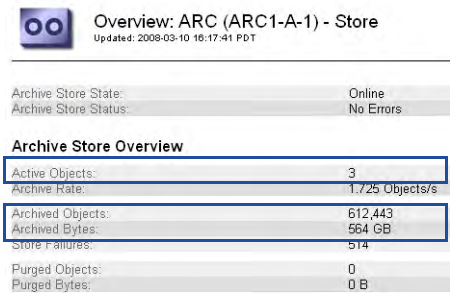
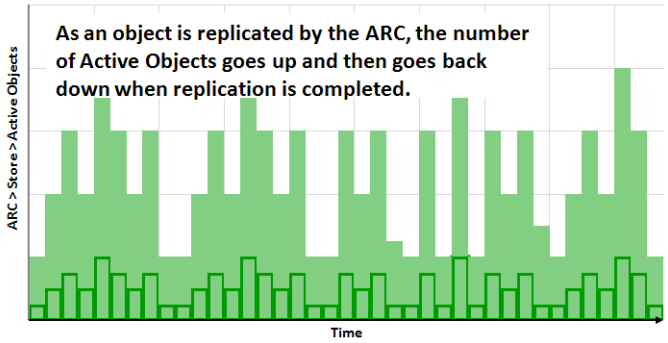
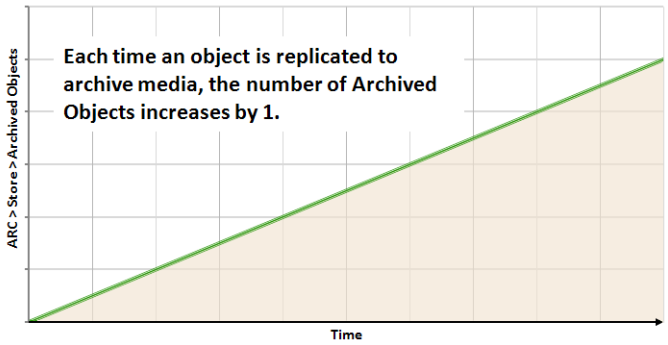
Component	Attribute Changes
ARC ▶ Store   	`Active Objects (AROP)`: The number of objects in the process of being written to archiving media increases by 1 as the object is being archived and decreases by 1 once the object is archived.  `Archived Objects (AROA)`: The total number of objects written to archive media by this ARC increases by 1.  `Archived Bytes (ARBA)`: The total amount of content written to archive media increases by an amount equivalent to the file size. The grid does not know how much installed and available storage is on the archival media device attached to the Tape Node.

Table 15 Object Replication to Tape Node Attributes *(continued)*

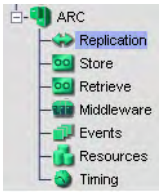
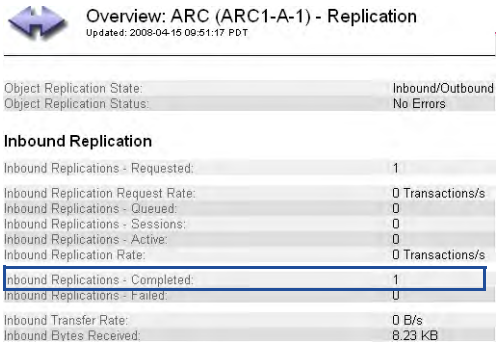
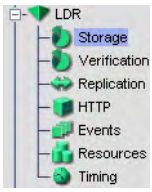
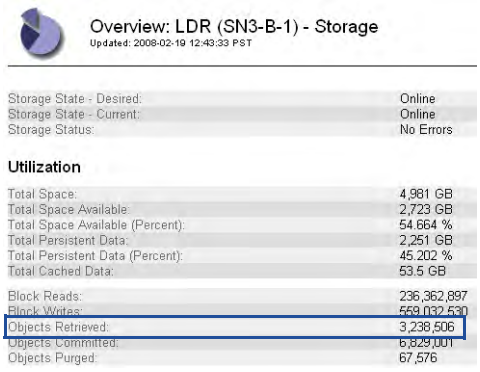
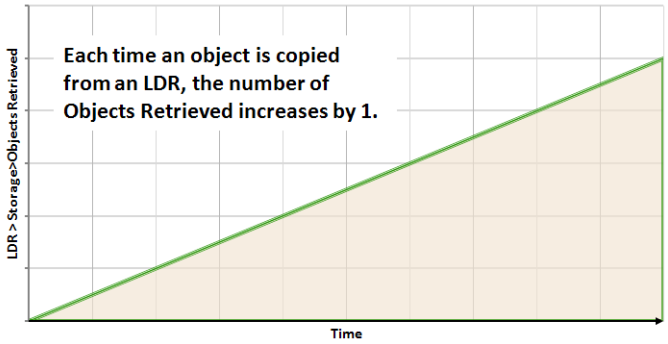
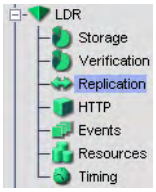
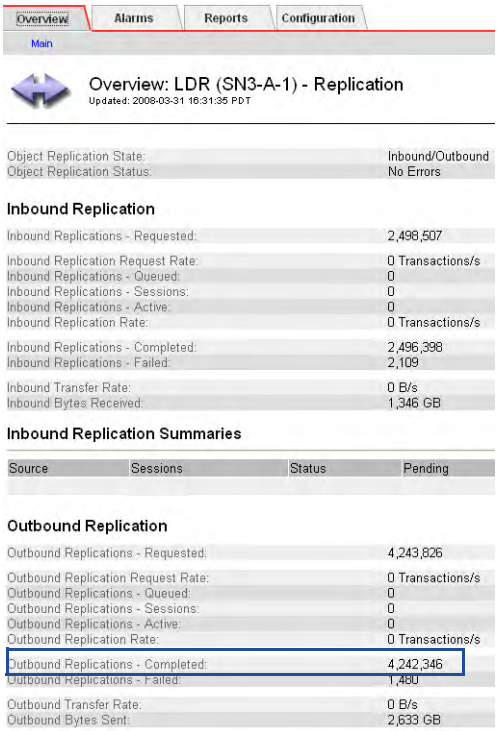
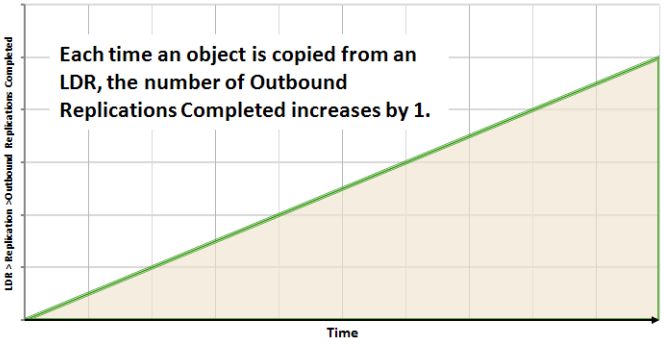
Component	Attribute Changes
ARC ▶ Replication  	`Inbound Replications Completed (RIRC)`: The total number of objects replicated to the destination ARC increases by 1.
LDR ▶ Storage  	`Objects Retrieved (ORET)`: The number of persistent objects retrieved from the storage system of the source LDR increases by 1 each time the object is replicated from this source LDR to the Tape Node. 

Table 15 Object Replication to Tape Node Attributes *(continued)*

Component	Attribute Changes
LDR ▶ Replication   	`Outbound Replications - Completed (RORC)`: The total number of objects replicated from the source LDR increases by 1 each time the object is replicated from this source LDR to the Tape Node. 

Retrieval

Retrieval refers to what happens when a client application accesses a file stored in the grid. There are two scenarios:

- The file is in the FSG cache.
- The file is not in the FSG cache.

Data Flow

File in FSG Cache

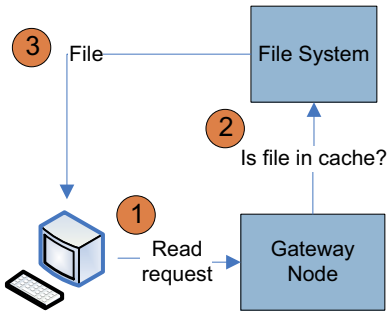


Figure 45 Retrieval Data Flow - File in Cache

1	The client requests the file.
2	The file system (CIFS/NFS) finds the file in the FSG cache.
3	The client reads the file.

File Not in FSG Cache

If the file is not in the FSG cache, the FSG sends a request to an LDR. The LDR returns the file if it has it. Otherwise, the LDR retrieves it from another LDR or an ARC after getting the file location from the CMS, see Figure 46 (page 76). Retrieval preferentially goes to spinning media under normal performance load balancing. When there is no higher grade copy accessible, the retrieval request is directed to the Tape Node.

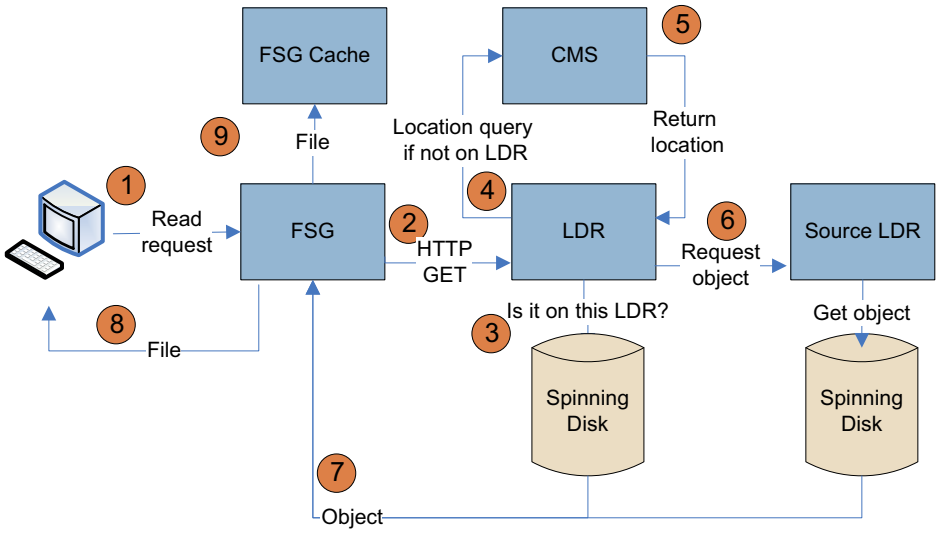


Figure 46 Retrieval Data Flow - File Not in Cache

1	The FSG receives a read request from a client.
2	Since the FSG cannot find the file in its cache, the FSG submits an HTTP GET command to an LDR. The LDR is chosen based on the result of a topology query.
3	The LDR checks if it has the object. If yes, it sends it to the FSG.
4	If the LDR does not find the object, it requests the location from a CMS. The CMS is chosen based on the result of a topology query
5	The CMS returns the object location to the LDR.
6	The LDR retrieves the content from the LDR or ARC that has it and sends it to the FSG.
7	The LDR starts streaming the object to the FSG.
8	The FSG sends the file to the client as soon as it has retrieved enough data from the LDR (it does not wait until it has retrieved the entire file).
9	The FSG stores the file in its cache for future use by the client.

Related Attributes

Table 16 (page 77) lists some of the NMS attributes used to track what happens when a client retrieves an object stored in the grid.

Table 16 Object Retrieval Attributes

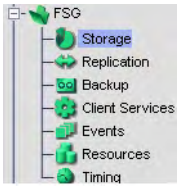
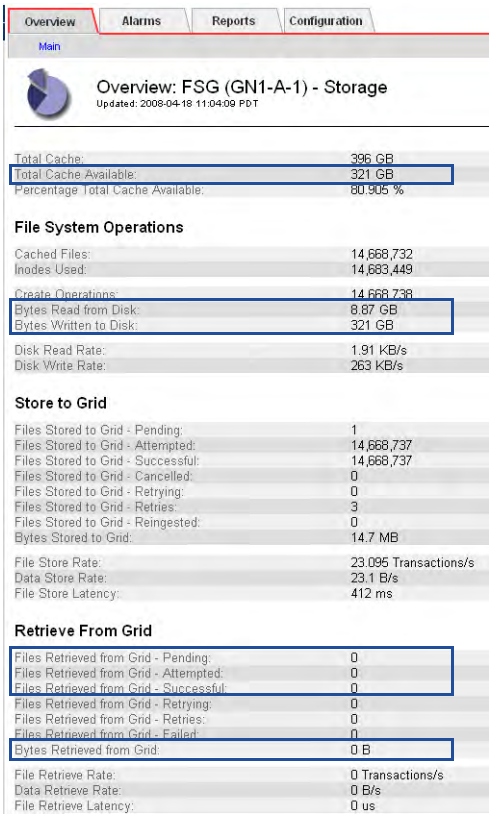
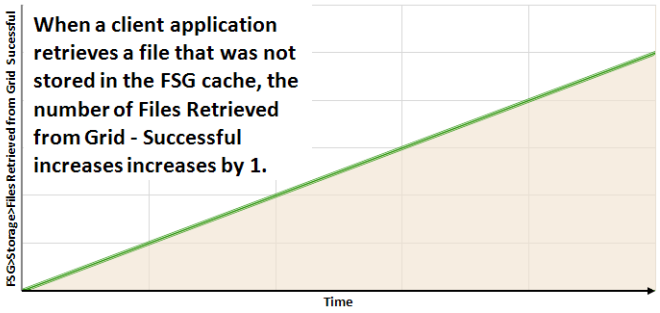
Component	Attribute Changes
FSG ▶ Storage   	`Files Retrieved from Grid - Successful (FRGC)`: If the file is not in the FSG cache, the number of file transfer requests completed successfully increases by 1.  `Files Retrieved from Grid - Attempted (FRGA)`: If the file is not in the FSG cache, the number of file retrieval requests waiting for a response from the grid increases by 1. `Files Retrieved from Grid - Pending (FRGP)` also increases by 1 when the transfer is requested and then decreases once it is completed. `Bytes Retrieved from Grid (FRGB)`: If the file is not in the FSG cache, the number of bytes retrieved successfully from the grid increases by an amount equivalent to the file size. `Bytes Read from Disk (FSRB)`: If the file is not in the FSG cache, the number of bytes read from disk increases. File retrieval generates multiple read operations as the FSG retrieves the file from the grid and the client application accesses the file. `Bytes Written to Disk (FSWB)`: If the file is not in the FSG cache, the number of bytes written to disk increases. File retrieval generates multiple write operations as the FSG retrieves the file from the grid. `Total Cache Available (FSTA)`: The total local cache space that is still available for use does not change if the file was already in cache. Otherwise, total cache available decreases unless the minimum value has been reached.

Table 16 Object Retrieval Attributes *(continued)*

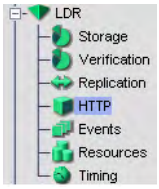
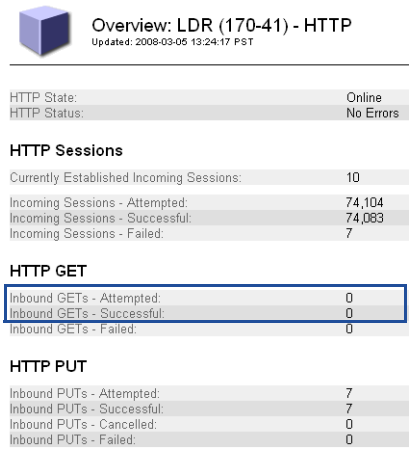
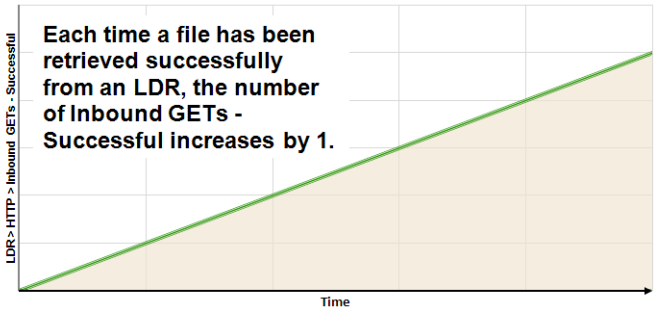
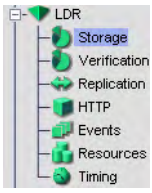
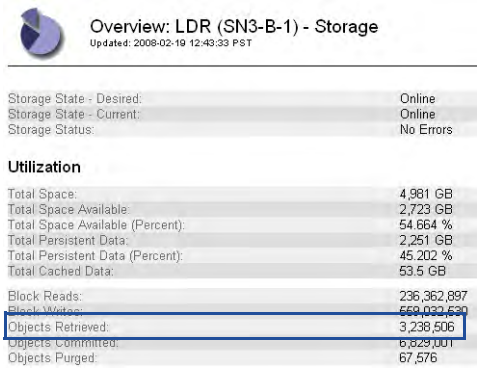
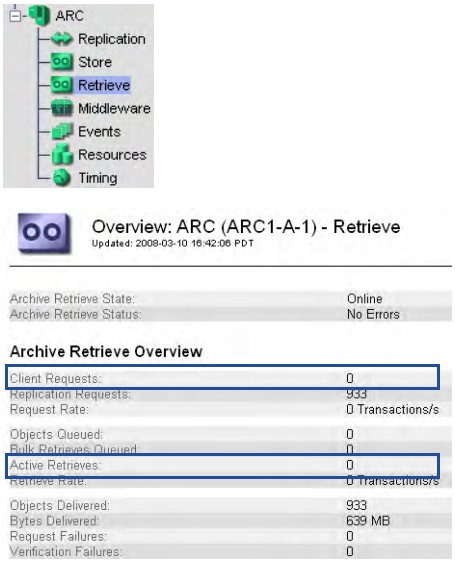
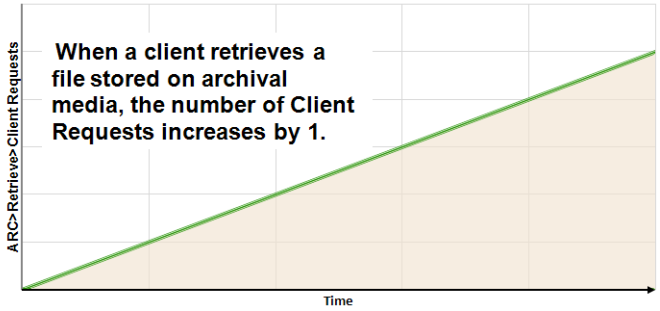
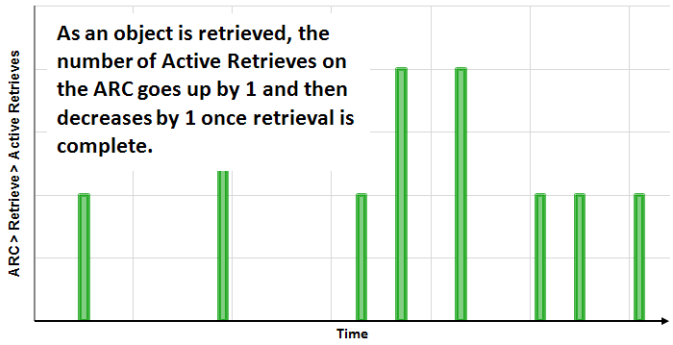
Component	Attribute Changes
LDR ▶HTTP   	Inbound GETs - Successful (HIGC): If the file is not in the FSG cache, the total number of HTTP GET (“content retrieve”) requests that have completed successfully increases by 1.  Inbound GETs - Attempted (HAIG): The total number of HTTP GET (content retrieve) requests that have been received by the LDR also increases by 1.
LDR ▶ Storage   	Objects Retrieved (ORET): If the file is not in the FSG cache, the number of persistent objects retrieved from the source LDR increases by 1.

Table 16 Object Retrieval Attributes *(continued)*

Component	Attribute Changes
ARC ▶ Retrieve 	`Client Requests (ARCR)`: The total number of requests received from clients for objects stored on the ARC increases by 1 each time the grid attempts to retrieve an object from this Tape Node. This happens only if the file is not in the FSG cache or on an LDR.  `Active Retrieves (ARAR)`: The number of object retrievals in progress increases by 1 while retrieval is taking place on this Tape Node and then decreases by 1 once retrieval is complete. 

Purging

Removing an object from the grid's Storage Nodes and Tape Nodes is called "purging". Purging an object changes it from persistent content to transient content. Transient content refers to cached content that can be removed to free up disk space.

Data Flow

See Figure 47 (page 80) for a simplified step-by-step description of what happens when the FSG receives a client request to delete a file.

NOTE In some grids, in order to protect against accidental or malicious object deletions, the ILM policy prevents content from being purged from the grid even if the client deletes the file on the FSG.

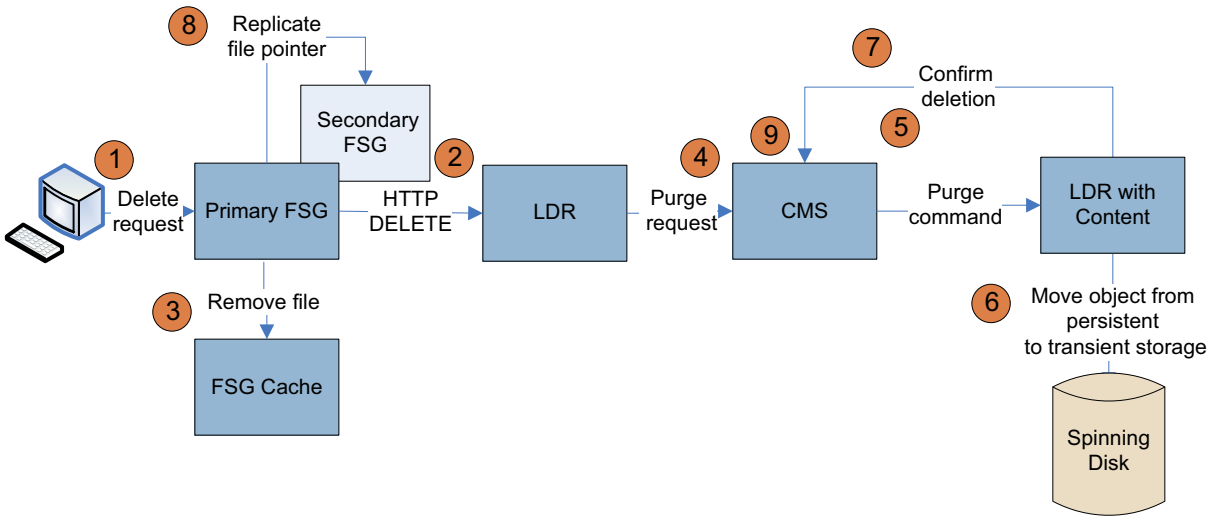


Figure 47 Purging Data Flow

1	The client sends a file delete request to the FSG which removes the file stub from the FSG file system.
2	The FSG sends an HTTP DELETE command to an LDR. (This releases the content handle.) What LDR is chosen depends on the result of a topology query.
3	The file is removed from the FSG cache (that is, the file stub and data are deleted).
4	The LDR notifies a CMS that the content handle has been released.
5	A purge command is sent to an LDR that has a copy of the object.

6	The LDR receives the purge command and moves the content from persistent to transient storage.
7	The LDR notifies the CMS that the content has been purged.
8	The Primary FSG replicates the file pointer information to the other FSGs in its replication group.
9	In a grid that uses distributed CMSs, the metadata is purged. Metadata is never purged when CMSs are synchronized.

Related Attributes

Table 17 (page 82) lists some of the NMS attributes used to track what happens when a client deletes an object stored in the grid.

Table 17 Object Purging Attributes

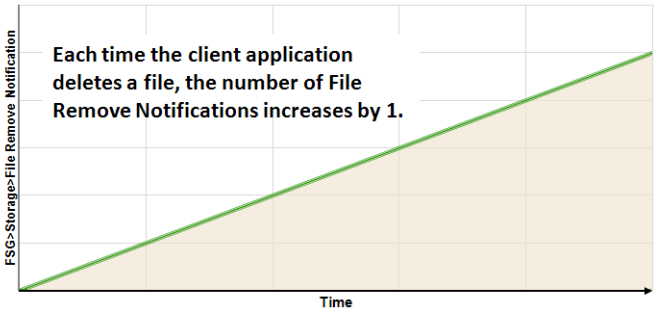
Component	Attribute Changes
FSG ▶ Storage 	File Remove Notifications (FRGN): The number of content handle release notifications sent by the FSG to LDRs increases by 1 when the FSG receives a request to delete an object. 

Table 17 Object Purging Attributes *(continued)*

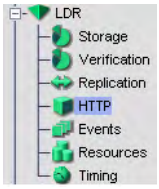
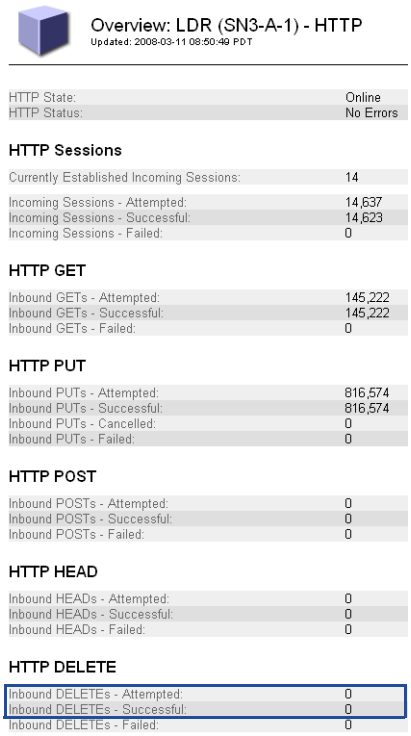
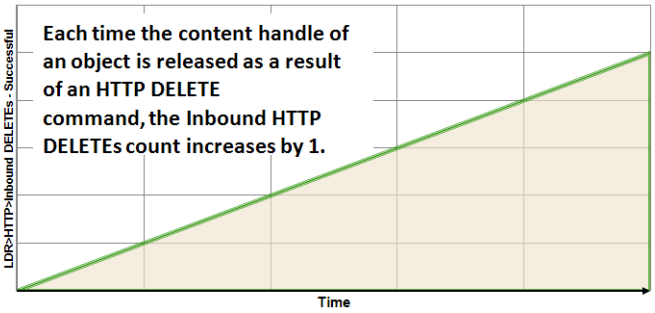
Component	Attribute Changes
LDR ▶ HTTP   	Inbound DELETEs - Successful (HIDC): The total number of objects for which the content handle has been released increases by 1 after the HTTP DELETE command has been completed successfully.  Inbound DELETEs - Attempted (HAID): The total number of HTTP DELETE (content handle release) requests that have been received by the LDR also increases by 1.

Table 17 Object Purging Attributes *(continued)*

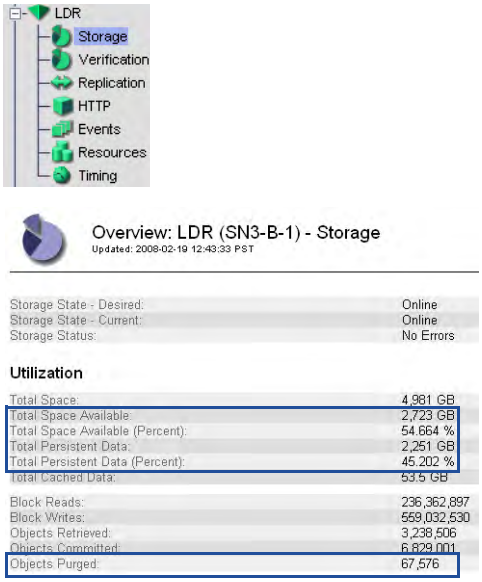
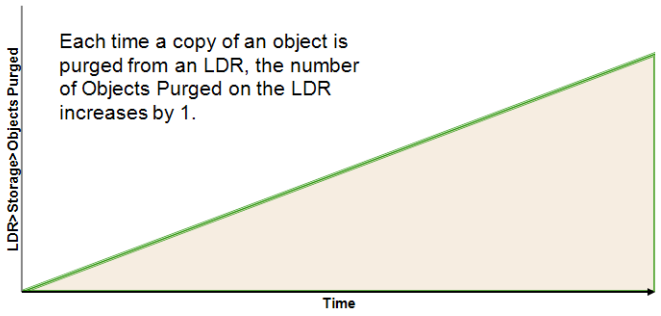
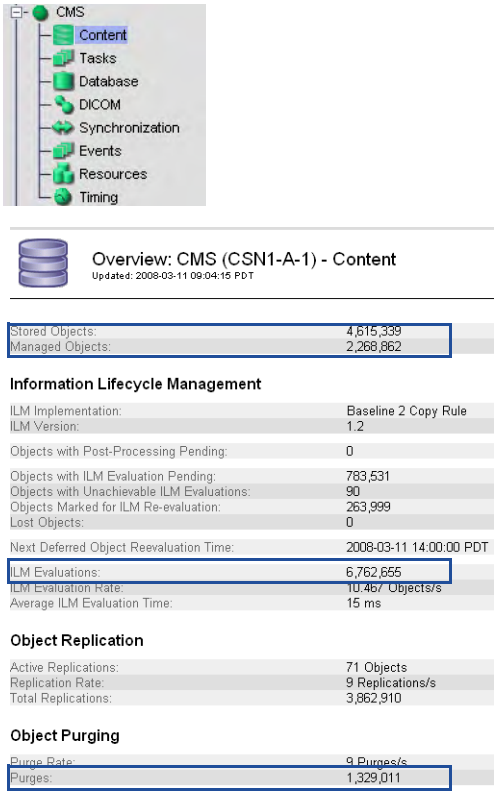
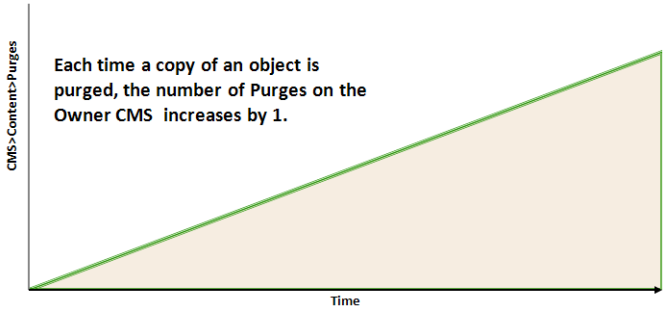
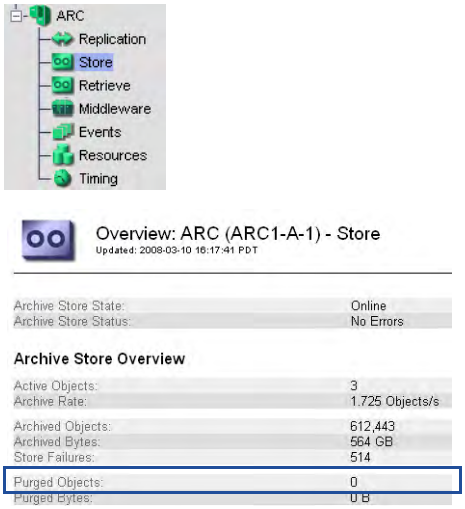
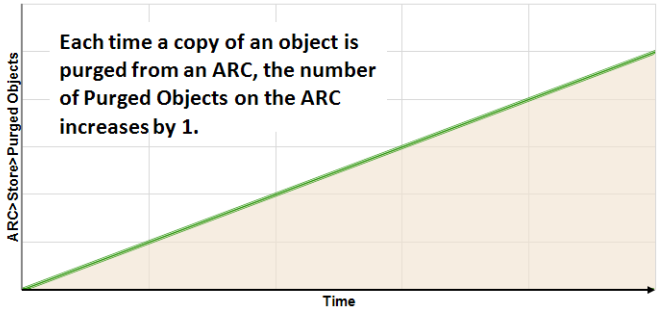
Component	Attribute Changes
LDR ▶ Storage 	Objects Purged (OPUR): The number of persistent objects purged from this LDR increases by 1 for each copy of the object purged.  Total Space Available (Percent) (SAVP): The percentage of object storage space available for use increases by an amount roughly equivalent to the size of the purged object. Total Persistent Data (Percent) (SPSD): The percentage of the total storage space used by persistent data decreases by an amount roughly equivalent to the size of the purged object.
CMS ▶ Content 	Purges (DCpT): The number of object copies deleted increases by 1 on the owner CMS for each copy of the object purged from the LDRs and ARCs.  Free Tablespace Percent (DBSP) and the number of Managed Objects (COoM) and Stored Objects (COoT) do not change in a grid with synchronized CMSs. When distributed CMSs are used, free tablespace and the number of stored objects decrease. (With distributed CMSs, COoM is listed under the Metadata component, not the Content component.) ILM Evaluations (ILev): The total number of ILM evaluations that have been performed to date increases when an object is purged because an ILM evaluation is triggered each time a content handle is released.

Table 17 Object Purging Attributes *(continued)*

Component	Attribute Changes
ARC ▶ Store 	Purged Objects (ADOP): The total number of objects purged from the middleware server by the ARC increases by 1. Whether objects purged from the ARC are actually deleted from archive media depends on the retention settings defined in the middleware server. 

Deletion Protection

Different levels of deletion protection can be enabled to protect files from being altered or removed after they have been ingested into the grid.

Purging Initiated by ILM Policy

Purging can also happen without a client request. For instance, an ILM policy could mandate that all content be automatically deleted two years after ingest. However, if objects are deleted from the grid without being deleted from the FSG first, links from the application to the objects will be broken and attempts to retrieve the objects from the FSG will fail.

File Modification

If the grid configuration allows it, client applications may modify content that has already been ingested into the grid.

When a file is modified, the content handle of the original file is released and the object is purged from the grid according to the ILM policy. The modified file is assigned a different unique identifier and the object is treated like a new file ingest. The grid does not track the multiple versions of the object.

Data Flow

See Figure 48 (page 86) for a simplified step-by-step description of what happens when the client modifies a file stored in the grid.

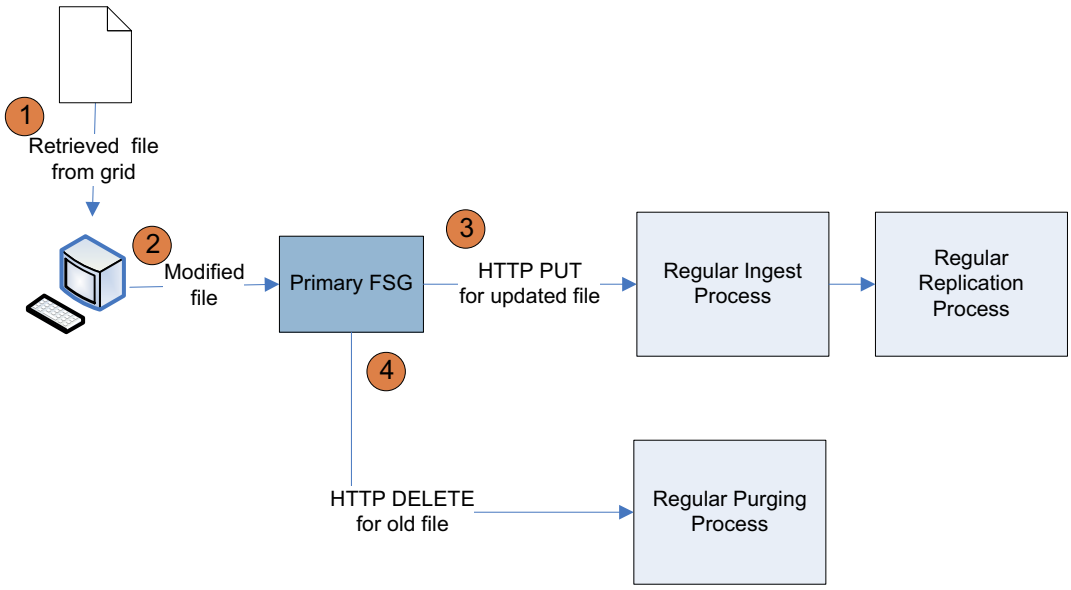


Figure 48 Object Modification Data Flow

1	The client application retrieves the file from the grid via the FSG file share. See Retrieval (page 75).

2	The client application saves the modified file to the grid via the FSG file share.
3	The Primary FSG saves a copy of the object in its cache and sends an LDR a HTTP PUT command. The data flow is identical to that for ingest. The LDR saves the object to spinning disk and allocates a new unique identifier to the object. The LDR notifies a CMS that a new piece of content has been ingested and sends the object metadata to the CMS. This CMS becomes the owner CMS. The object is then replicated according to the ILM policy. See Ingest (page 58) and Content Replication (page 67).
4	The Primary FSG also sends a HTTP DELETE command to an LDR to delete the old version of the object. What happens next is identical to what happens when an object is purged. See Purging (page 80).

Related Attributes

The key attributes that change when a file is modified are basically the same as when a file is ingested and replicated, and then purged. One notable exception is `Create Operations`. You can tell when a file has been modified if `File Remove Notifications` and `Files Stored To Grid` have both increased by 1 but `Create Operations` has not changed.

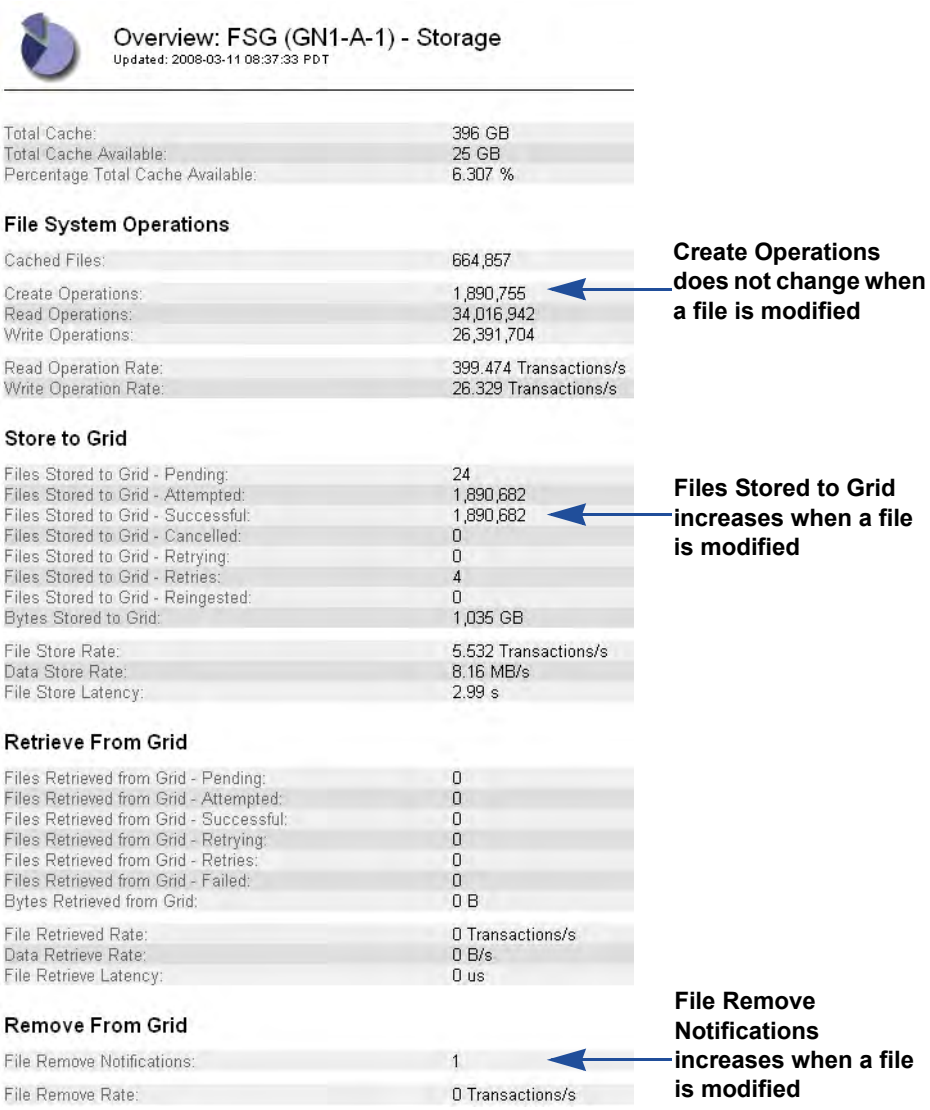


Figure 49 FSG Storage Attributes For Modified Object

FSG Replication

The Primary FSG maintains a system of file pointers to the objects stored in the grid. As seen above, the file pointer system is modified each time a file is ingested, changed, or deleted.

The Primary FSG must replicate its file pointer system to the Secondary FSG to ensure redundancy in case the Primary FSG becomes unavailable.

During normal operation, the file pointers are replicated in real time, as files are ingested, modified or deleted. A backlog of replication operations may form during periods of increased grid activity or during the daily backup of the FSG system.

The backup together with the active session file can be used to restore the managed file system should it become corrupted. The active session file is a log of the FSG activity.

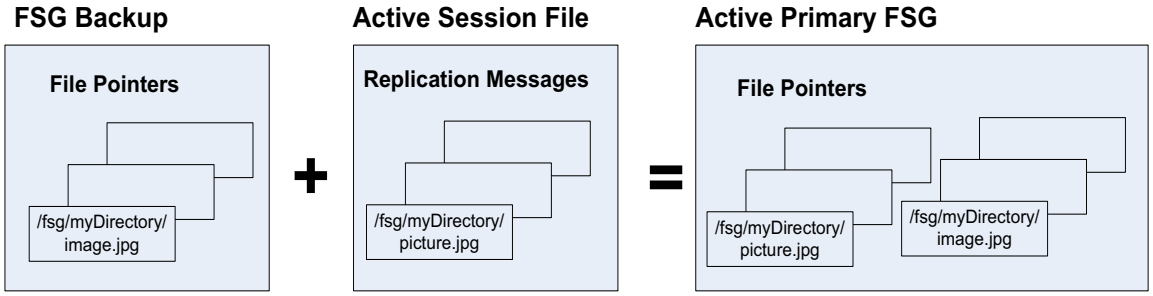


Figure 50 FSG Backups and Active Session File

See Figure 51 (page 90) for a simplified description of the FSG replication message flow when a file is ingested, deleted, or modified.

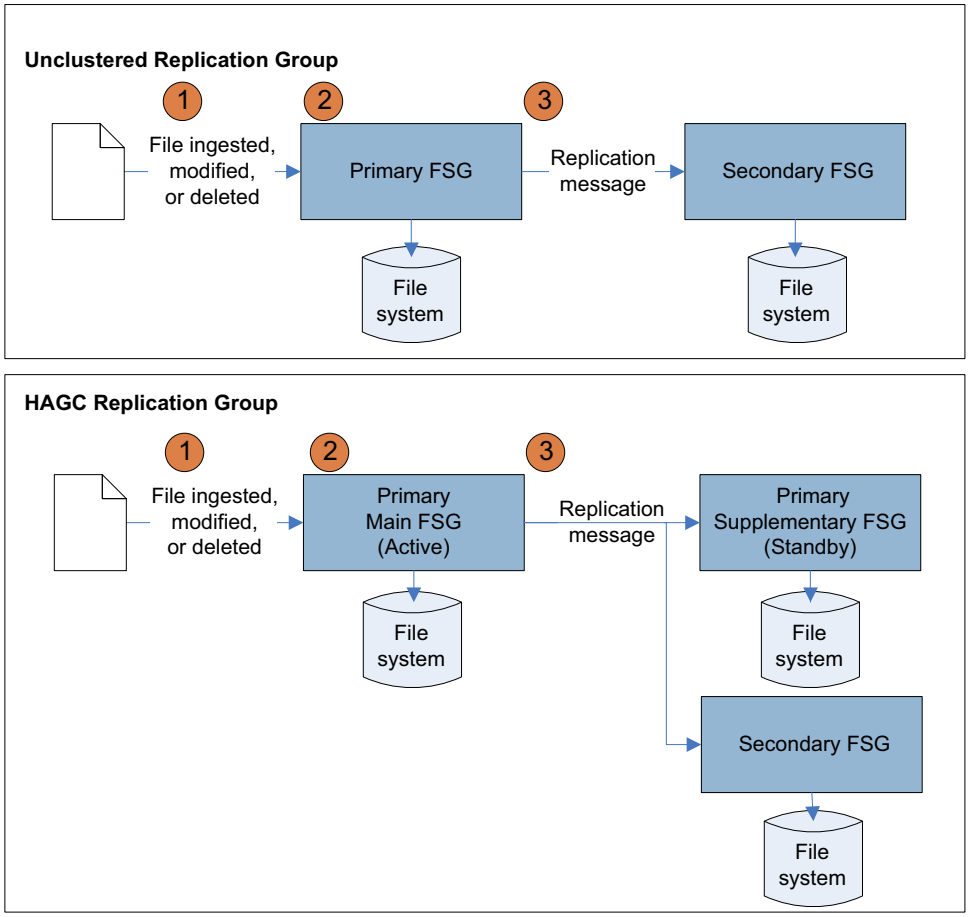


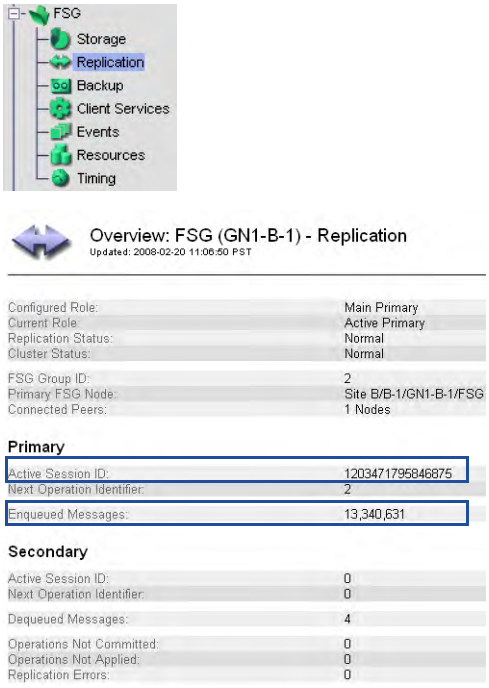
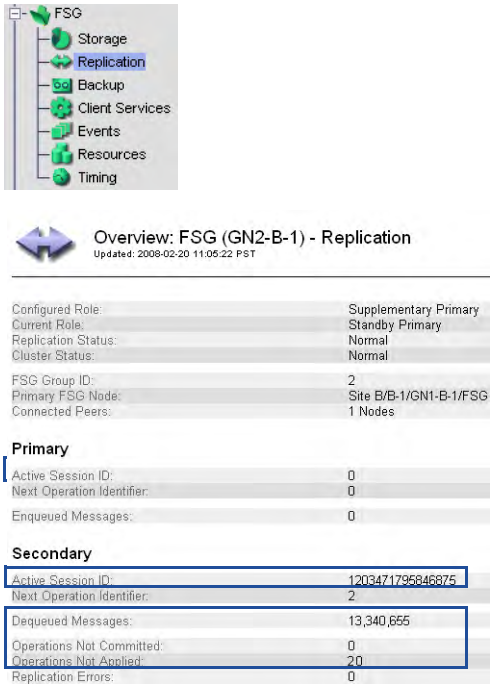
Figure 51 FSG Replication Data Flow

1	The client creates, modifies, or deletes a file via the FSG file share.
2	The Primary FSG either creates a file pointer, modifies the file pointer, or deletes the file pointer from its file system.
3	The Primary FSG sends a replication messages to the other FSGs in its replication group.
4	The other FSGs in the replication group process the replication messages in real time and update their file system. A backlog may form during periods of increased grid activity or during the daily FSG backup

Related Attributes

Table 18 (page 91) lists some of the NMS attributes used to track FSG replication.

Table 18 FSG Replication Attributes

Component	Attribute Changes
Primary FSG ▶ Replication 	`Primary Active Session ID (PAID)`: A new unique identifier for the current replication session is assigned each time a new session is started by the active primary FSG. A new session is started when a FSG failover occurs or when the size or age of the current session file exceeds an internal threshold. This number is the same as the Secondary Active Session ID (SAID) on the Secondary FSG (see below). `Enqueued Messages (PEOP)`: The total count of replication messages generated increases each time a file is ingested, modified or purged. The count increases by more than one for some operations. For example, ingest generates two replication messages (one for the initial file creation and a second message to associate the UUID after the file has been ingested. Modify also generates two replication messages (one to release the old UUID and one to assign the new UUID after ingest). Attribute events on the file (change permissions, etc.) may also generate additional replication messages. The number of enqueued messages matches the number of dequeued messages at the Secondary FSG (see below). Because this FSG is the Primary FSG, the fields in the “Secondary” section do not apply.
Secondary FSG ▶ Replication 	`Secondary Active Session ID (SAID)`: The unique identifier for the current replication session from which messages are being processed is the same as the Primary Active Session ID (PAID) on the Primary FSG (see above). `Dequeued Messages (SDOP)`: The total count of dequeued replication messages increases each time a message has been processed by the Secondary FSG. The number of dequeued messages matches the number of enqueued messages at the Primary FSG (see above). `Operations Not Committed (SUOP)`: The number of replication messages from the Primary FSG that have not yet been written to the Secondary FSG increases temporarily during periods of high grid activity. `Operations Not Applied (SPOP)`: The number of messages to be processed on the Secondary FSG in order to catch up to the Primary FSG increases during backups. This applies to older systems where “offline backups” are used. With online backups, SPOP does not generally increase except temporarily during periods of high grid activity. Because this FSG is the Standby Primary FSG in an HAGC, the fields in the “Primary” section do not apply. The same would be true for a Secondary FSG.

4 Operations

This chapter describes common routine tasks that you perform as a grid operator:

- Monitor trends
- Monitor FSG backups
- Monitor LDR verification
- Monitor the Tape Node
- Monitor grid tasks

Top Attributes

The NMS interface displays hundreds of attributes. However, most of these attributes are required only for troubleshooting. The list of attributes to monitor routinely, shown in Table 19 (page 93), is much shorter. Tips on how to analyze these attributes are described in the remainder of this chapter.

Table 19 Key Attributes to Monitor

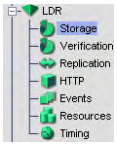
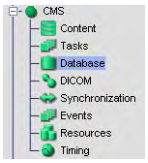
Category	Component	Code	Description	See
LDR content storage capacity	Grid Overview 	PSCU	Percentage Storage Capacity Used: The percentage of installed storage capacity that has been used up for the entire grid.	Content Storage Capacity (page 97)
	LDR Storage 	SAVP	Total Space Available (Percent): Object storage capacity that is still available for use on the LDR.	
CMS metadata storage capacity	CMS Database 	CORS	Estimated Remaining Object Capacity: The number of additional objects that can be managed by this CMS.	Metadata Storage Capacity (page 99)
		DBSP	Free Tablespace (Percent): The metadata storage capacity that is still available for use on this CMS.	
	Grid Overview 	PMCA	Percentage Metadata Capacity Available: An estimate of how much metadata capacity remains in the CMS databases in the grid.	

Table 19 Key Attributes to Monitor *(continued)*

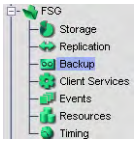
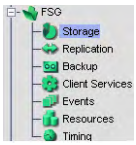
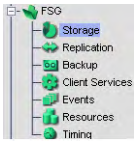
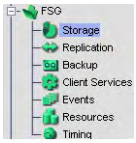
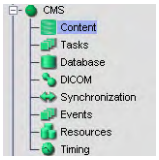
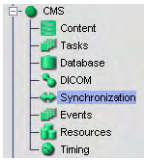
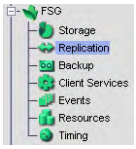
Category	Component	Code	Description	See
FSG capacity	FSG Backup 	PBNF	Number of Files: The number of files included in the last FSG backup.	FSG Capacity (page 101)
		PBDS	Backup Data Size: The total size of all files referenced by the last FSG backup.	
		PBSF	Number of Files: The number of files included in the last backup for each share.	
		PBSB	Total File Size: The total size of all files referenced by the last backup for each share.	
	FSG Storage 	FSIU	Inodes Used: Number of inodes (files and directories) used on the FSG filesystem.	
Ingest load	FSG Storage 	FSGP	Files Stored to Grid - Pending: The number of new ingested files cached locally that are waiting for transfer to the grid for persistent storage.	Increasing Ingest Load (page 104)
Retrieve load	FSG Storage 	FRTM	File Retrieve Latency (FRTM): The average amount of time required to retrieve a file from the grid. Look at this attribute along with the related attributes Bytes Retrieved from Grid, File Retrieve Rate, and Data Retrieve Rate.	Retrieve Load (page 106)
Object replication	CMS Content 	ORun	Objects with Unachievable ILM Evaluations: The number of objects whose ILM business rules cannot be met because the topology or operational state of the grid prevents the rules from being satisfied.	ILM Replication (page 108)
		ORpe	Objects with ILM Evaluation Pending: The number of objects waiting to be processed through the business rules for replication.	
Metadata synchronization (only for synchronized CMSs)	CMS Synchronization 	CsQT	Queue Size: The number of outgoing metadata synchronization messages queued to be sent to other CMSs.	Metadata Synchronization and Replication (page 109)

Table 19 Key Attributes to Monitor *(continued)*

Category	Component	Code	Description	See
FSG replication	Secondary FSG Replication 	SUOP	Operations Not Committed: The number of replication messages from the Primary FSG that have not been written to the Secondary FSG yet.	FSG Replication (page 111)
		SPOP	Operations Not Applied: The number of messages to be processed by the Secondary FSG in order to catch up to the Primary FSG.	
Amount of content written to archive media	ARC Store 	ARBA	Archived Bytes: The total amount of content written to archive media by this ARC.	Tape Node Capacity (page 115)

Regular Tasks

Table 20 (page 96) lists the tasks to be performed on a regular basis.

Table 20 Daily Activities

Task	Frequency	See
Monitor System Status. Note what has changed from previous day.	Daily	Reviewing Alarms (page 40)
Monitor system status lights on hardware.	Daily	
Monitor the rate at which LDR storage capacity is being used up.	Weekly	Content Storage Capacity (page 97)
Monitor the rate at which content metadata storage capacity on the CMS is being used up.	Weekly	Metadata Storage Capacity (page 99)
Monitor FSG capacity.	Weekly	FSG Capacity (page 101)
Check available space on the archive media.	Weekly	Tape Node Capacity (page 115)

Monitor the key attributes regularly to become familiar with the grid operation and spot trends before they turn into problems. The important attributes to monitor relate to:

- Content storage capacity on LDRs
- Metadata storage capacity on CMSs
- FSG capacity
- Attribute storage capacity on NMSs
- Ingest load on FSGs
- Retrieve load on FSGs
- Metadata synchronization/replication on CMSs
- ILM replication on CMSs

In the case of all the capacity attributes, for example LDR content storage space, you must not only look at the absolute value but also at the rate at which capacity is being consumed.

Content Storage Capacity

The LDRs on the Storage Nodes are responsible for storing objects in the grid. You need to monitor the total space available on Storage Nodes to make sure the grid does not run out of space to store content. This information is available at the grid level (see Figure 52 (page 97) for example), at the site level, and at the node level (see Figure 53 (page 97) for example).

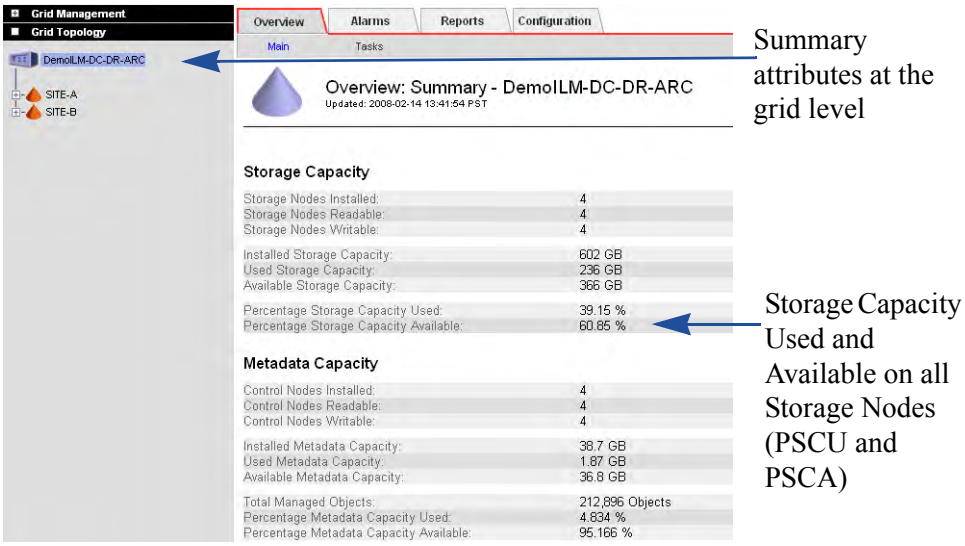


Figure 52 Overall Storage Node Capacity

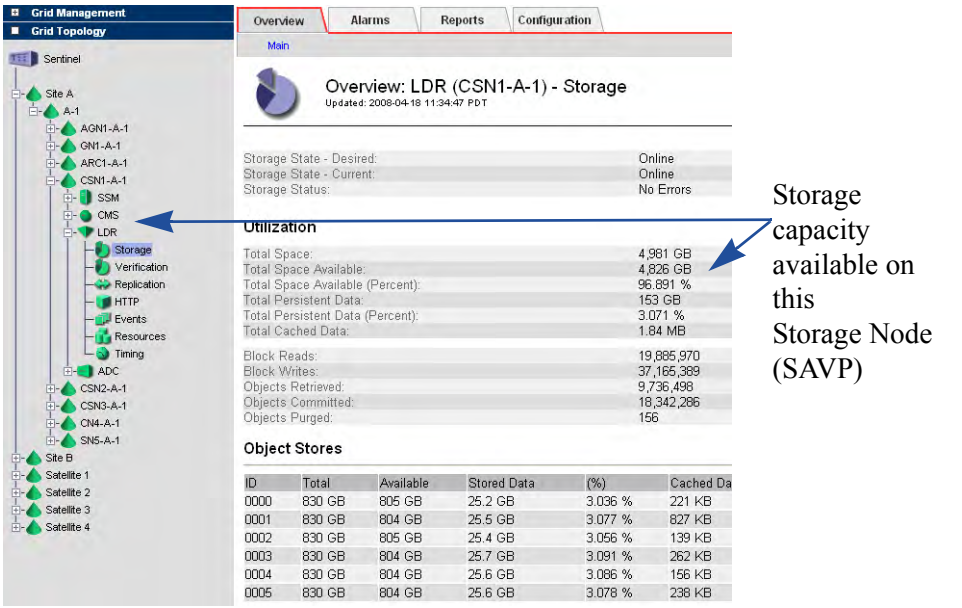


Figure 53 Individual LDR Storage Capacity

Track `LDR Storage Total Space Available` over a period of time to estimate the rate at which available object storage space is being consumed. To maintain normal grid operations, you have to add Storage Nodes, or add storage volumes, or migrate content to archive media before the storage disks fill up.

For the example shown in Figure 54 (page 98), content storage space is being consumed at a rate of approximately 4% per month, which means that there are 8 months left before this LDR runs out of storage space.

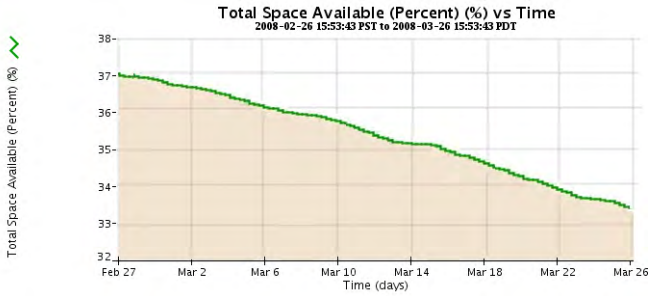


Figure 54 Content Storage Capacity Available

Metadata Storage Capacity

The CMS databases on the Control Nodes are responsible for storing metadata, e.g., the information about the objects ingested into the grid. You need to monitor the total space available on Control Nodes to make sure the grid does not run out of space to store metadata. If all CMS databases in the grid fill up, the grid can no longer ingest files until the grid's CMS capacity is increased.

Metadata storage capacity information is available at the grid level (see Figure 55 (page 99) for an example), site level, and node level (see Figure 56 (page 99) for an example).

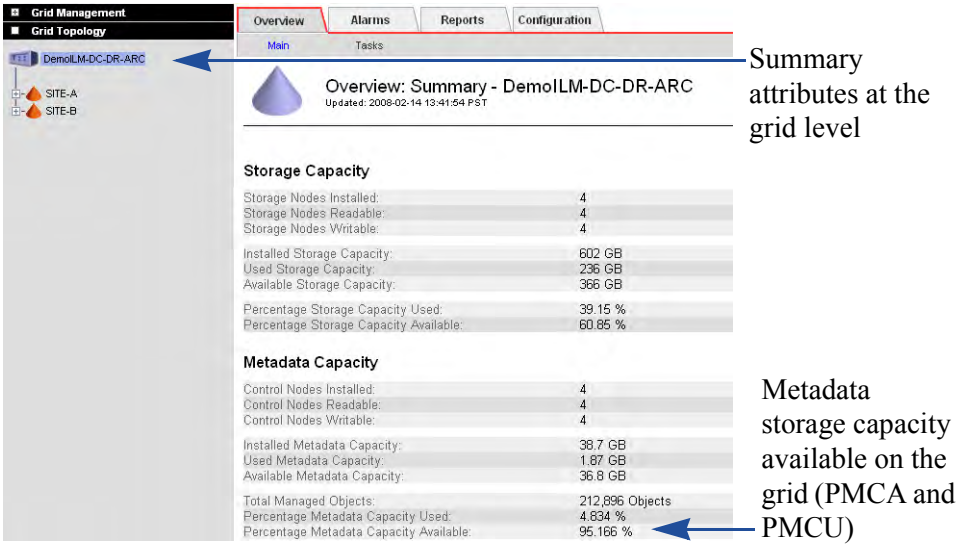


Figure 55 Summary CMS Metadata Storage Capacity

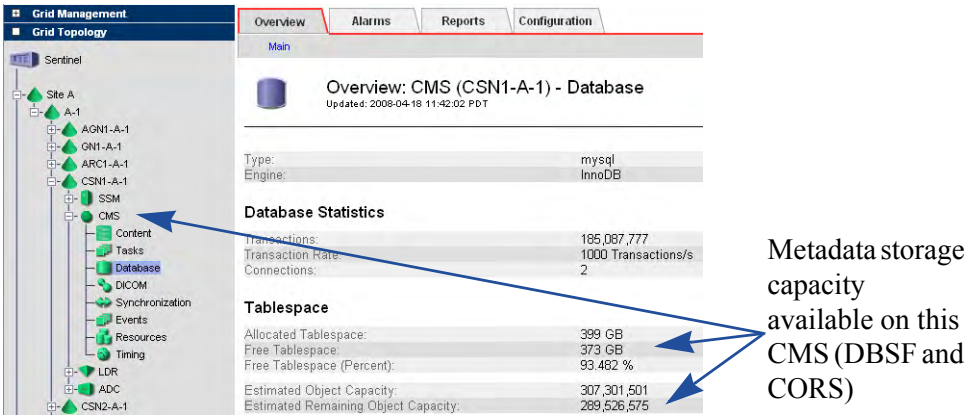


Figure 56 Individual CMS Metadata Storage Capacity

Track `CMS Database Free Tablespace` and `CMS Database Estimated Remaining Object Capacity` over a period of time to estimate the rate at which the available database space is being consumed. The databases of synchronized CMSs fill at approximately the same time unless the grid uses multiple generations of CMSs. The databases of distributed CMSs that are in the same CMS replication groups also fill at approximately the same time.

To maintain normal grid operations, you have to add Control Nodes before the metadata database fills up. The CMS databases go into “read-only” mode when Free Tablespace drops below 10%.

NOTE Figure 57 is for the same grid as the one used in Figure 54, Content Storage Capacity Available. The migration effect is not noticeable in Figure 54 because the object size is small: the metadata of a large quantity of small objects use more space proportionally than the actual content.

For the example shown in Figure 57 (page 100), metadata storage capacity is being consumed at a rate of approximately 20 GB per month. However, note how the utilization rate increases towards the end of the period. This could be due, for example, to a data migration that is happening in parallel with regular grid ingest.

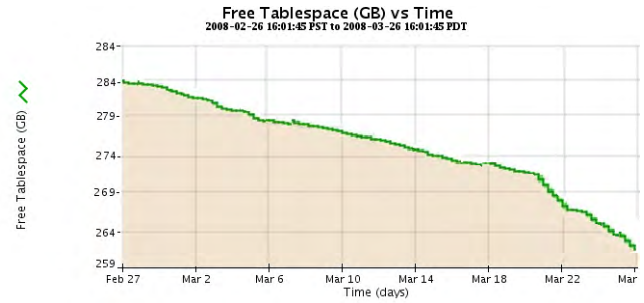


Figure 57 Available CMS Content Storage Metadata Capacity

Gateway Node Capacity and Load

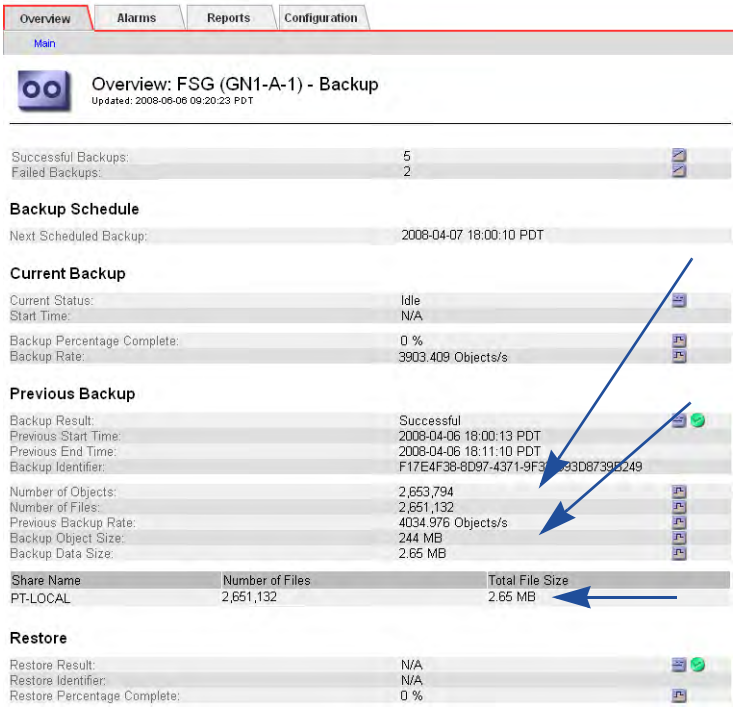
FSG Capacity

You can get an estimate of how much content is managed by each FSG replication group by looking at the FSG backup information: `FSG Backup Number of Objects` and `FSG Backup Data Size`, Figure 58 (page 101). The `Backup Number of Objects` should increase steadily.

The number of files and total file size is also broken up by FSG file share (`Number of Files` and `Total File Size`).

The backup values reflect the FSG managed file system as of the most recent backup. This includes files that are pending for ingest and files for which ingest into the grid is disabled through FSG profiles.

Since FSG capacity limits are not actively enforced, proactive monitoring is necessary to identify when an FSG has reached its capacity. At that point, client ingests should be directed to a new FSG replication group to avoid problems that may occur by exceeding the supported capacity.



Number of files backed up from this FSG (PBNF)

Total size of files referenced by the last backup (PBDS)

Number of files and size per FSG file share (PBSF and PBSB)

Figure 58 FSG Backup Size

The capacity of an FSG replication group is two hundred million objects. When an FSG Replication group reaches this limit, it is full and the grid must be expanded. To estimate the number of objects per FSG replication group (and thus estimate when you must expand the grid), monitor the number of inodes used.

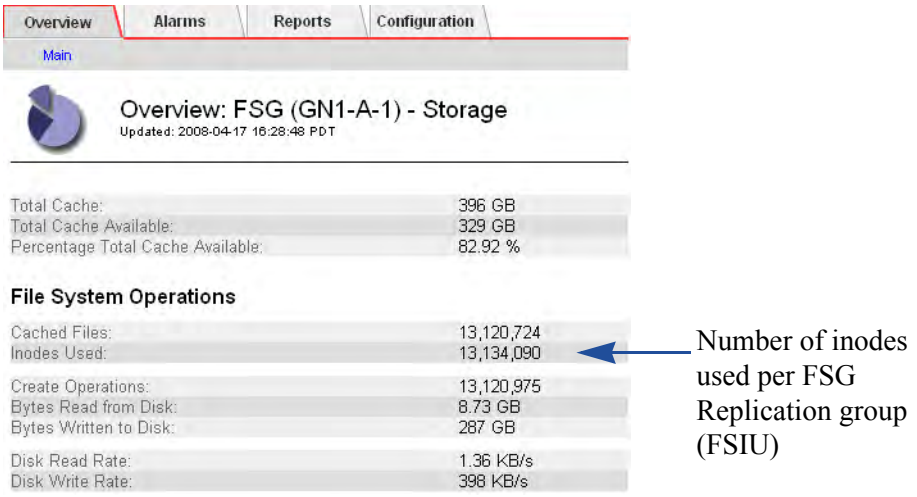


Figure 59 Inodes Used

Summary Attributes

FSG information is also available at the grid and site level on the Overview tab, see Figure 60 (page 102).

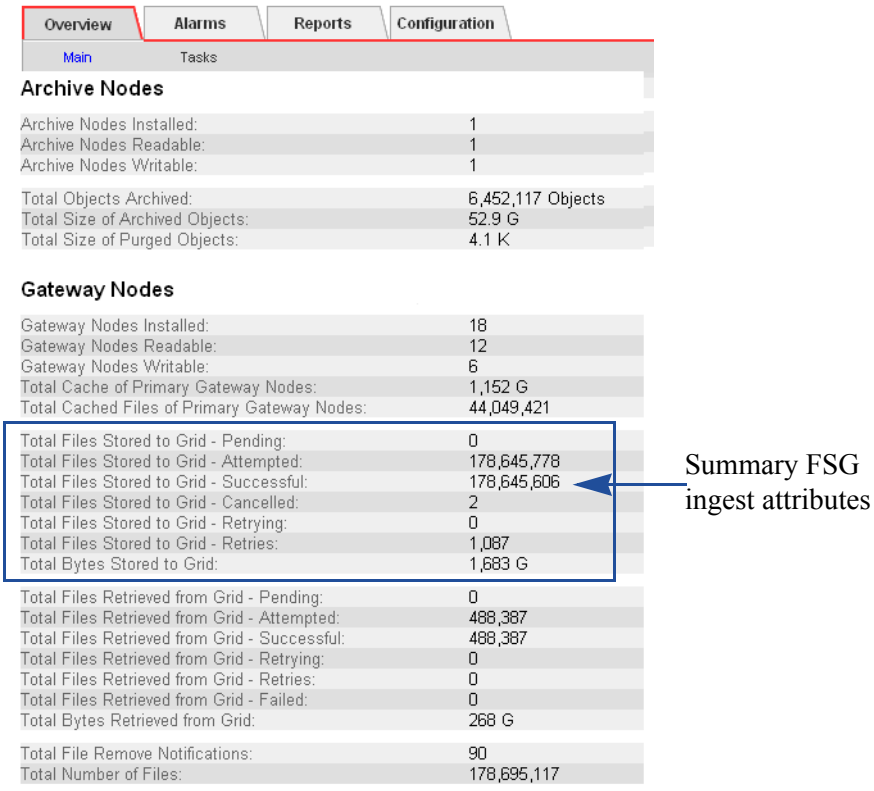


Figure 60 Gateway Node Summary Attributes

Ingest Load

To monitor the ingest load on the grid, analyze the trends of these four attributes over time, see Figure 61 (page 103):

- Bytes Stored to Grid (FSGB): The number of bytes ingested successfully into the grid.
- File Store Rate (FSRA): The rate at which files are successfully stored to the grid (number of transactions per second).
- Data Store Rate (FSBA): The rate at which data is successfully stored to the grid (in bytes per second)
- File Store Latency (FSTM): The average amount of time required to store the entire file into the grid. The average is calculated over the last sampling period.

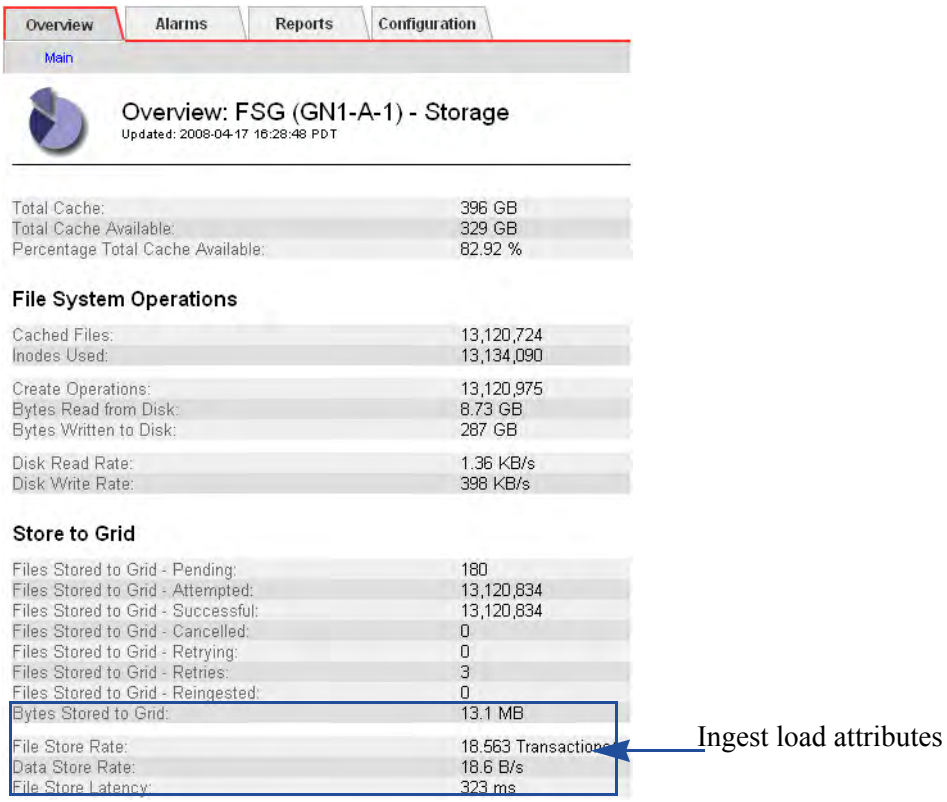
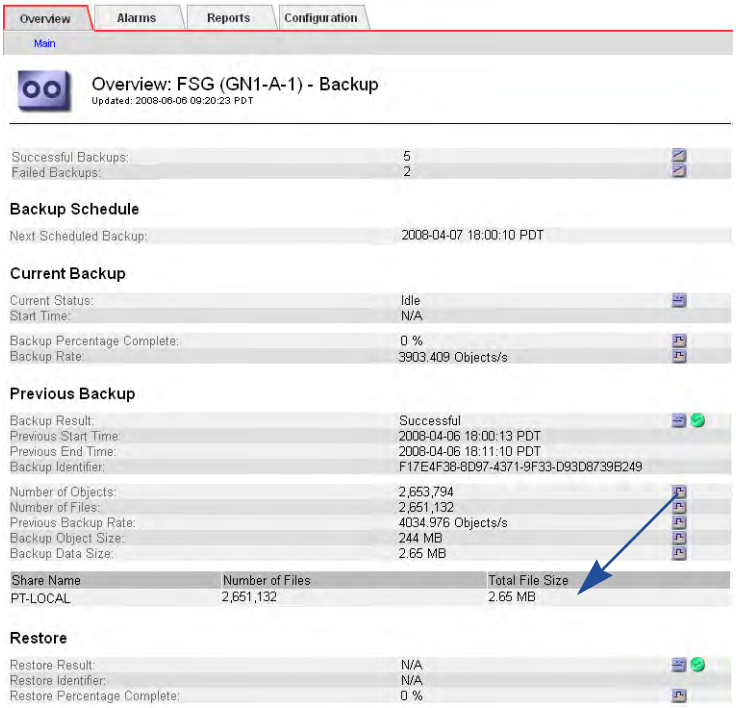


Figure 61 Ingest Load Attributes

Ingest Load per Application

Usually, each client application is assigned an FSG share. You can monitor how much each application is storing to the grid by looking at the FSG backup values for `Number of Files` and `Total File Size` per share.



Files ingested per share/application (number PBSF and size PBSB)

Figure 62 Ingest per Share

Increasing Ingest Load

During normal operations, it is possible for the ingest load to exceed the rate at which the services on the grid process the objects. When this happens, services may queue operations that can no longer be fulfilled in real time. For instance, the value of `File Stored to Grid - Pending` (that is, the number of ingested files cached locally that are waiting for transfer to the grid for persistent storage) may increase temporarily.

Figure 63 Files Stored to Grid Pending

NOTE `Files Stored to Grid - Pending` includes files actively being written but for which the ingest delay period has not expired. Therefore, `Files Stored to Grid - Pending` will always be non-zero when ingests are active.

In the example shown in Figure 64 (page 105), the number of files waiting to be stored goes up and down but remains fairly low. Such a trend could indicate that there was a short term overload due to network throughput, disk I/O performance, grid services availability, and so on.

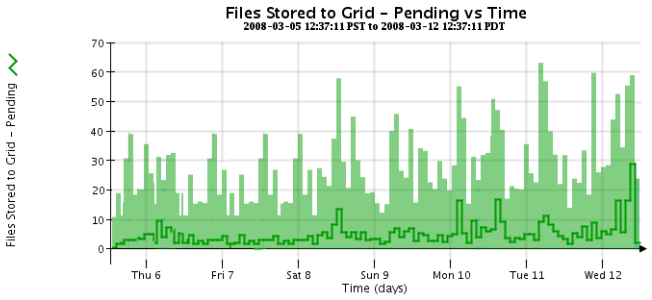


Figure 64 Files Stored to Grid Pending

In contrast, the trend shown in Figure 65 (page 106) is not sustainable. If the number of files waiting to be stored to the grid starts to increase, make sure that all CMS and LDR services are operating normally. It is also possible that the ingest rate is exceeding the throughput of the grid and that a grid expansion is required.

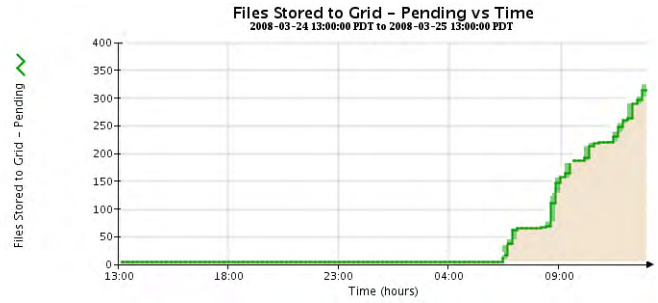


Figure 65 Non Sustainable Ingest Load

`Total Cache Available` may also be monitored: if new files are ingested faster than existing cached files can be swapped out, the amount of cache available may dip below the "`Swapout No Create Watermark`" (defined in `FSG Management`). When this happens, the creation of new files is temporarily disallowed until enough space is freed. This is uncommon but may occur for example when the grid has ingested many small files followed by many large files. In this case, the FSG may not be able to swap out the small files fast enough to make room for the large files. If this happens, the client must throttle its ingest rate.

Retrieve Load

To monitor the retrieve load on the grid, analyze the trends of these four attributes over time, see Figure 66 (page 107).

- Bytes Retrieved from Grid (FRGB): The number of bytes retrieved successfully from the grid.
- File Retrieve Rate (FRRA): The rate at which files are successfully retrieved from the grid (number of transactions per second).
- Data Retrieve Rate (FRBA): The rate at which data is successfully retrieved from the grid (in bytes per second).
- File Retrieve Latency (FRTM): The average amount of time required to retrieve the entire file from the grid. The average is calculated over the last sampling period.

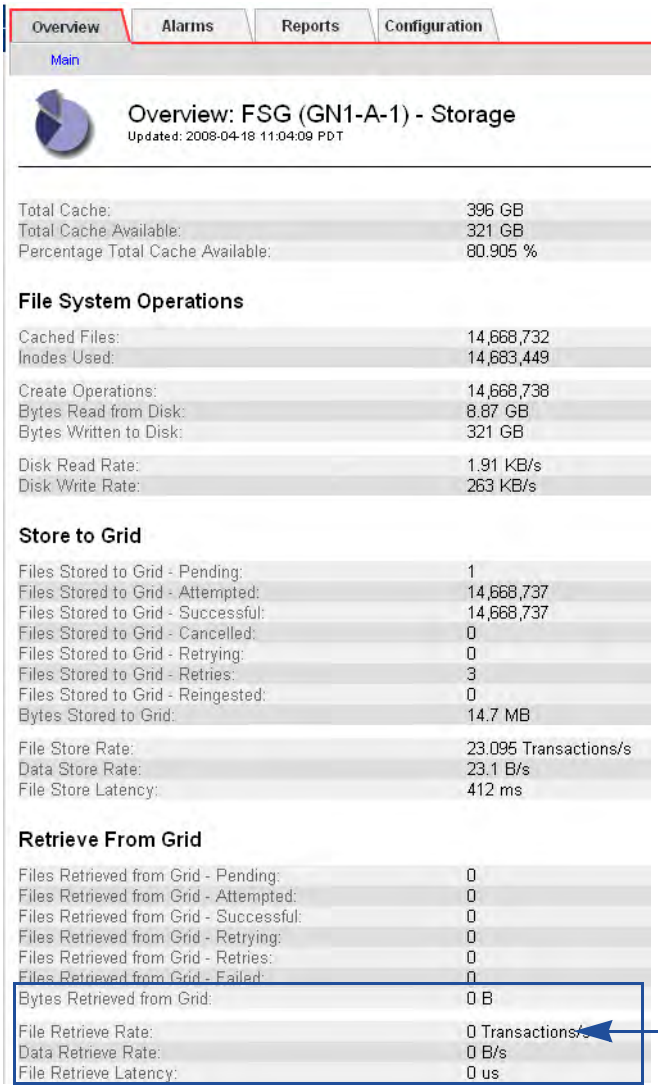


Figure 66 Retrieve Load Attributes

ILM Replication

The CMSs manage ILM replication. ILM replication refers to the process of making copies of the object and keeping the copies in the appropriate storage locations for a pre-determined length of time. You can track what is happening with ILM replication by looking at the number of objects in each of these categories:

Pending	Objects with ILM Evaluation Pending (ORpe) The number of objects waiting to be processed.
Unachievable	Objects with Unachievable ILM Evaluations (ORun) The number of objects that require additional copies to be made but the storage resources specified by the ILM policy are unavailable to store the additional copies.
Future	Objects Marked for ILM Re-evaluation (ORde) The number of objects that currently satisfy the ILM policy but are due to be re-evaluated at a scheduled point in the future, for instance because of a rule that says “store a copy to archive media two years after ingest”.

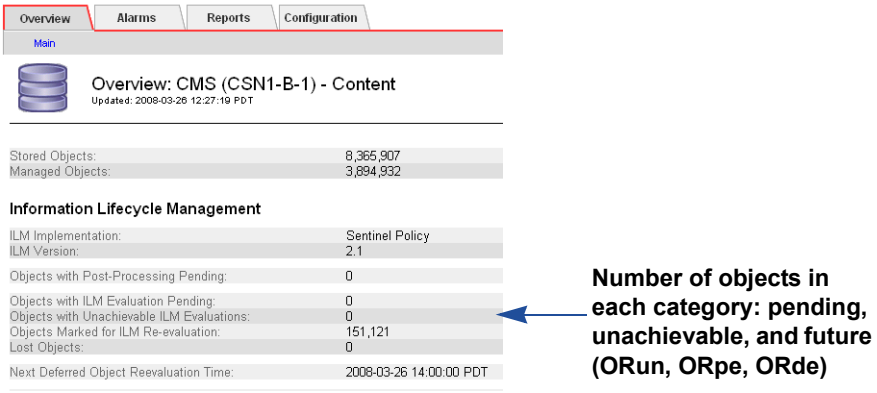


Figure 67 ILM Evaluation

Metadata Synchronization and Replication

Metadata management depends on whether the CMS operation is distributed or synchronized.

Synchronized CMSs

When an owner CMS receives new metadata, it stores the metadata in its local database and sends synchronization messages to the other CMSs it must replicate metadata to. The attribute `Queue Size` tracks the number of messages to be sent to another CMS. The corresponding attribute `Incoming Messages` tracks the number of synchronization messages coming from another CMS and waiting to be processed, see Figure 68 (page 109).

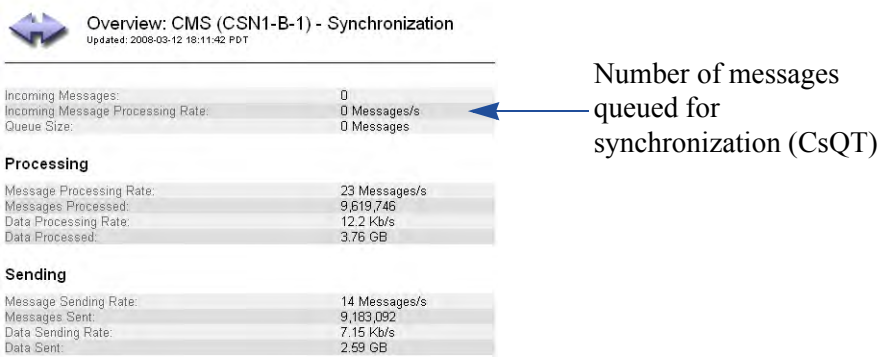


Figure 68 CMS Metadata Synchronization

During normal operations, it is possible for the ingest load to exceed the rate at which the CMSs can synchronize metadata. This temporary overload solution will resolve itself over time. However, if synchronization messages start to accumulate, ensure that the other CMS services are running normally and if the trend continues, escalate the issue as it could be that the ingest rate is exceeding the throughput of the grid.

Metadata CMSs

NOTE The Metadata component is only displayed for distributed CMSs.

With distributed CMSs, the number of replication messages is drastically reduced. The information about metadata replication is shown in the Metadata component, see Figure 69 (page 110).

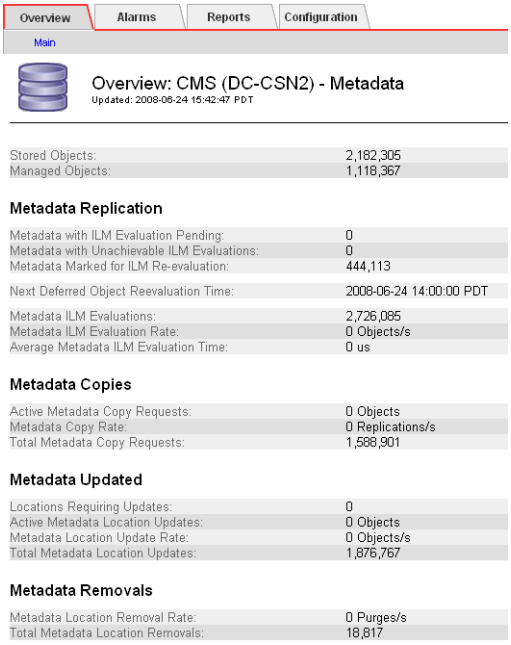


Figure 69 CMS Metadata Component for Distributed CMSs

FSG Replication

When an FSG ingests a file, it creates a file pointer to reference the object and it replicates the file pointer to the other FSGs in its replication group. To verify that FSG replication is proceeding normally, look at the attributes `Operations Not Committed` and `Operations Not Applied` in the FSG Replication component, see Figure 70 (page 111).

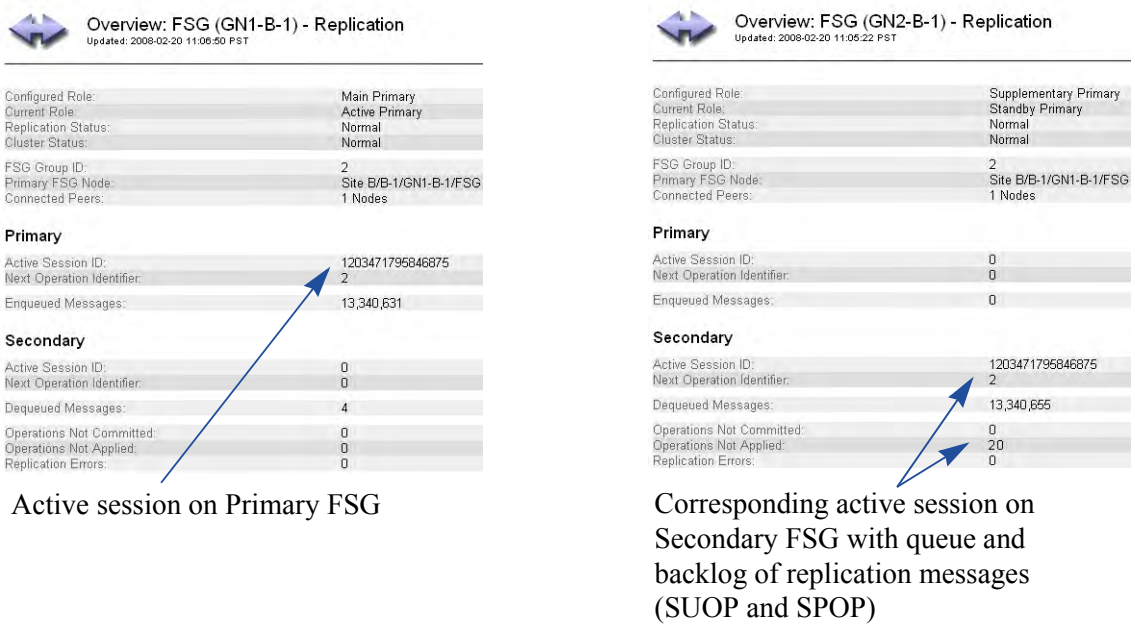


Figure 70 FSG Replication Messages

- Operations Not Committed (SUOP): The number of replication messages from the Primary FSG session that have not been written to the Secondary FSG yet.

  An upwards trend indicates that the grid is ingesting files faster than the Secondary FSG can process transactions. The number will go back down during periods of reduced grid activity. If it does not go down, escalate the issue as it could be that the ingest rate is exceeding the throughput of the grid.

- Operations Not Applied (SPOP): The number of messages to be processed on the Secondary FSG in order to catch up to the Primary FSG.

  An upwards trend indicates that the replication backlog is growing continuously. This may occur because FSG backups take too long to complete. If this is the case, the backup frequency may need to be changed. SPOP backlogs is an issue mostly with FSGs deployed prior to release 7.5. With online backups, SPOP does not generally not increase except temporarily during periods of high grid activity.

Gateway Node Failovers

If the Primary FSG fails, the `RSTU FSG Replication Status` alarm is triggered, displaying a status of `No Primary` or `No Session`. What happens next depends on the type of replication group.

Unclustered

If the Primary FSG in an unclustered replication group fails, immediately notify a grid administrator who has access to the Admin or Vendor account so that a manual failover procedure can be performed if the grid supports business continuity.

In grids that support business continuity, if the Primary FSG fails, a Secondary FSG can be manually configured to act as a Primary. After clients are manually redirected to the acting Primary, they can continue to read and write to the grid. This is a temporary measure to maintain service while the Primary FSG is repaired: grid access is interrupted until manual failover is completed, and the redundancy of file system information in the grid is reduced while the Secondary FSG is acting as a temporary Primary FSG. If the Secondary is also the backup FSG for the replication group, backups are not performed while it acts as the Primary.

In grids that do not support business continuity failover, clients can continue to access files via the read-only file system on the Secondary FSG while the Primary is repaired, but cannot write data to the grid.

High Availability Gateway Cluster (HAGC)

If the Primary Active FSG in an HAGC replication group fails, the supplementary FSG starts acting as the active Primary FSG without any manual intervention. The cluster status changes to Vulnerable and the alarm FCST Cluster Status is triggered.

When a failover occurs, any client operations that are in progress fail, as do client operations initiated while the Standby FSG makes the transition to Active. Once the CIFS service starts on the newly Active FSG, Windows CIFS clients should be able to process new operations without remapping their connections to the Primary FSG cluster. NFS clients must remount shares before they can continue to store and retrieve data to the Gateway cluster. Full grid functionality and full grid access is maintained while the second FSG is Active.

You need to investigate the cause of the failure as soon as possible as another FSG failure in the replication group will render FSG services unavailable.

To restore the grid to full redundancy, a recovery procedure must be performed even if the failed FSG in the cluster recovers automatically. After the failed FSG is restored to service, manually failing back to the restored FSG makes it the Active FSG and also restores the second FSG to Standby status.

For more details on failovers, see the *user guide*. Contact Support for assistance with file pointer replication and failover procedures.

FSG Backups

Within an FSG replication group, one FSG is designated as the backup FSG. This FSG backs up the replication group managed file system (the file pointer references) daily.

The backup files are ingested into the grid and by default automatically deleted after 14 days.

The greater the number of objects, the longer it takes to back up the FSG. The backup must complete each day with enough time remaining to process queued replication messages.

You can view information about the FSG backups (for example, backup schedule, backup duration, backup size, backup status) on the FSG Backup component of the Gateway Node, see Figure 71 (page 113).

The value for Number of Files is the total number of files on the FSG managed file system as of the most recent backup. This includes files that are pending for ingest and files for which ingest into the grid is disabled through FSG profiles.

In addition, the backup page gives you an idea of how much content is managed by each FSG.

- `Backup Data Size` and `Number of Objects` refer to all files managed by the FSG.
- `Number of Files` and `Total File Size` break down the information by file share.

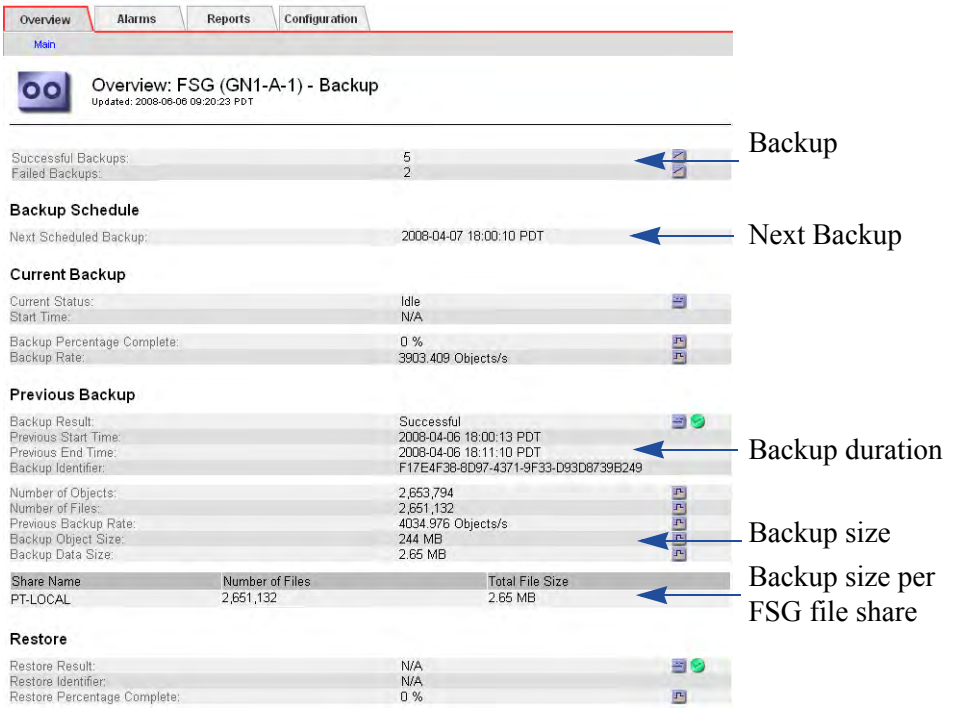


Figure 71 FSG Backup Attributes

For more information on backups, see the *user guide*.

LDR Verification

The grid is said to be self-healing. This means that the grid checks the integrity of the ingested objects via a process called background verification. If a corrupt object is found on an LDR, the object is quarantined and replaced with a copy of an uncorrupted object stored somewhere else on the grid. The existence of corrupt objects can indicate disk corruption or data tampering. Background verification operates at an adaptive priority to avoid interfering with normal grid operations.

You can view information about LDR verification, for example progress and number of corrupt objects found, on the Verification component of the LDR, see Figure 72 (page 114). Any corrupt object should be investigated.

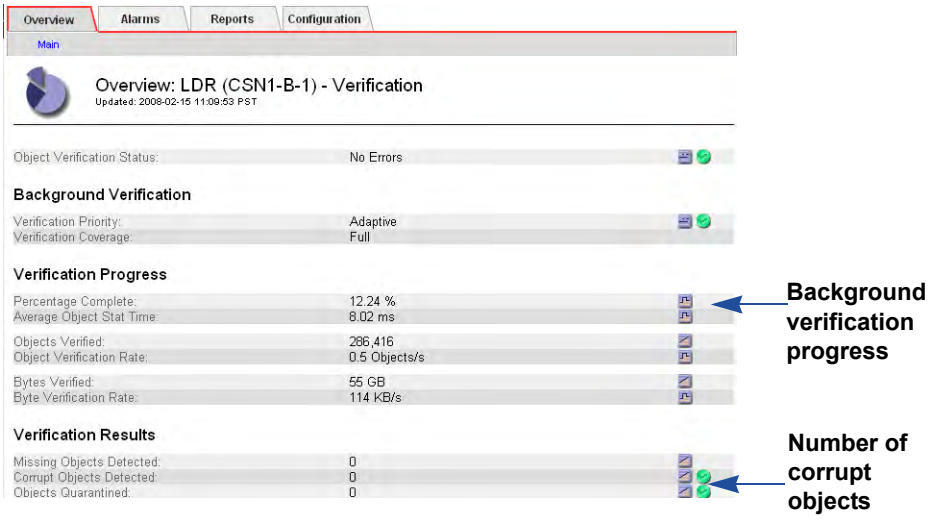


Figure 72 LDR Verification Attributes

There is another type of LDR verification called foreground verification. Foreground verification detects whether objects are missing. The LDR foreground verification procedure is initiated by a grid task and is used mostly during maintenance. See Table 21 (page 114) for a comparison of the two types of LDR verification.

Table 21 LDR Verification

Background	Foreground
Runs continuously at a low level	Used in maintenance procedures
Automatic	Initiated by a grid task
Identifies corrupt objects	Identifies missing objects
Performed by the Storage Nodes	Performed by Control Nodes
Slower process	Faster process
Adaptive, lower priority	Higher priority

All attributes on Figure 72 (page 114) above refer to LDR background verification except for Missing Objects which is updated by foreground verification.

For more information on LDR verification, see the *user guide*.

Tape Node Capacity

Each Tape Node can interface with the supported type of archival storage: archival media managed by Tivoli® Storage Manager (TSM).

In a grid that includes a TSM Tape Node, the TSM middleware has no way to inform the Tape Node when the TSM database or the archive media managed by the TSM is near capacity. The Tape Node will continue to accept objects for archiving after the TSM stops accepting new content and an alarm will be triggered.

Grid Tasks

A grid task is a program that performs grid procedures that involve several grid services automatically. For instance, LDR foreground verification is performed via a grid task. Most maintenance and expansion procedures involve running grid tasks.

You can follow the progress of a grid task from the `CMN Grid Tasks Overview` tab, see Figure 73 (page 115) and Table 22 (page 116). Running grid tasks is restricted to accounts with `Maintenance` permissions such as the `Admin` and `Vendor` accounts.

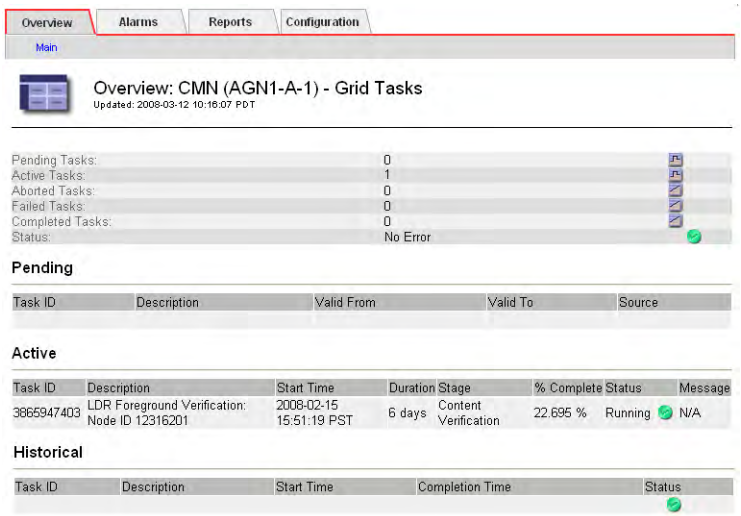


Figure 73 Grid Tasks Overview

Grid tasks go through three distinct phases:

Pending — The grid task has been submitted but not started yet.

Active	The grid task has been started. It can be either actively running or temporarily paused.
Historical	A historical grid task is a task that has been submitted but is no longer active. This includes grid tasks that completed successfully, grid tasks that were rejected (for example because the valid time period had expired), grid tasks that were cancelled or aborted, and grid tasks that terminated in error.

Table 22 Grid Tasks Overview Fields

Field	Description
Task ID	Unique identifier assigned when the task is created.
Description	Brief description of the purpose of the task.
Valid From	Date from which the task is valid. The grid task will be rejected if it is submitted before this date.
Valid To	Date until which the task is valid. The grid task will be rejected if it is submitted after this date.
Source	The author of the grid task.
Start Time	Date and time on which the grid task was started.
Stage	Description of the current stage of the active task.
% Complete	Progress indicator for active tasks.
Duration	Estimated amount of time since the grid task was started.

Table 22 Grid Tasks Overview Fields *(continued)*

Field	Description
Status	Current status of the active or historical task. For active tasks, one of: • Starting • Running • Pausing • Paused • Error: An error has been encountered. User action is required. • Aborting • Abort Paused: Task failed to be aborted and is paused in error. For historical tasks, one of: • Successful • Expired • Aborted • Cancelled • Duplicate • Invalid
Message	Information about the last stage of the active task.
Completion time	The date and time on which the grid task completed (or failed or expired or was aborted).

For more information on grid tasks, consult the *user guide*.

Common Alarms

Table 23 (page 118) lists common alarms that are usually no cause for concerns as long as trends do not develop.

Table 23 Common Alarms

Category	Code	Service	Notes
HTTP	HEIS HEIP HEIG HEID	LDR	HTTP protocol alarms such as Incoming Sessions Failed (HEIS), Inbound PUTs Failed (HEIP), Inbound GETs Failed (HEIG), and Inbound DELETEs - Failed (HEID) are not usually a cause for concern unless the error count escalates. In general, these alarms occur during periods of high load or due to temporary network disruptions. The operations that have failed are retried automatically and usually succeed. A failed GET will usually be visible to a client as a delayed, failed, or timed out retrieve. A failed GET may indicate that an object has been lost from the grid although this typically triggers other alarms on the FSG or elsewhere.

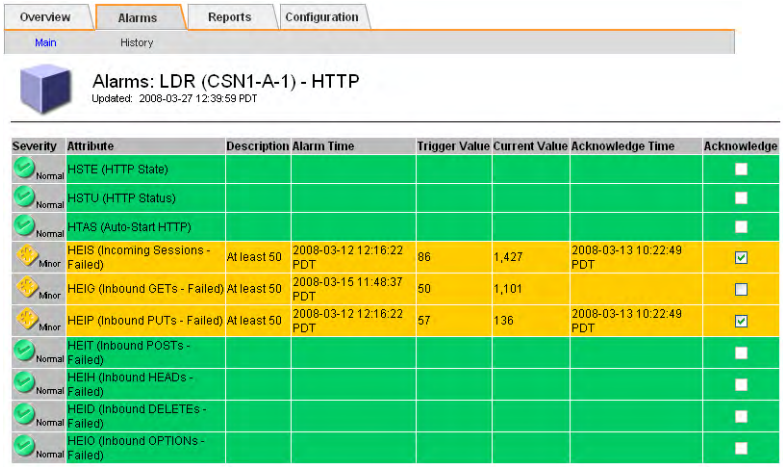


Figure 74Common HTTP Alarms: HEIS, HEIP, HEIG, HEID

Content replication	RIRF RORF	LDR ARC	Replication alarms (Inbound Replications - Failed RIRF and Outbound Replications - Failed RORF) occur in general during periods of high load or due to temporary network disruptions. After grid activity goes back down, these alarms should clear. If the count of failed replications continues to increase, look for network problems and verify that the source and destination LDRs and the ARCs are online and available.

Table 23 Common Alarms *(continued)*

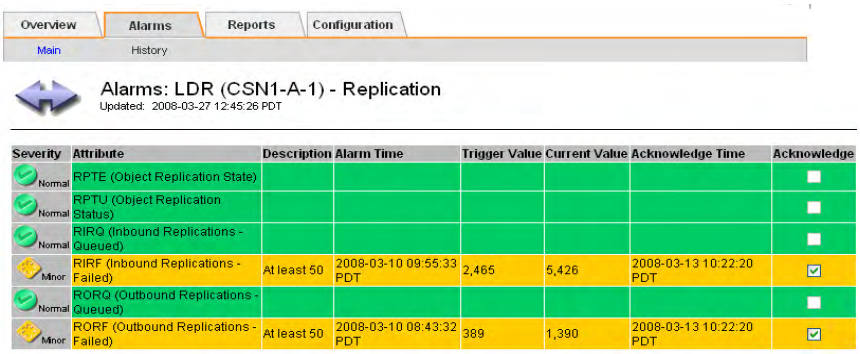
Category	Code	Service	Notes
			

Figure 75Common LDR Replication Alarms: RIRF, RORF

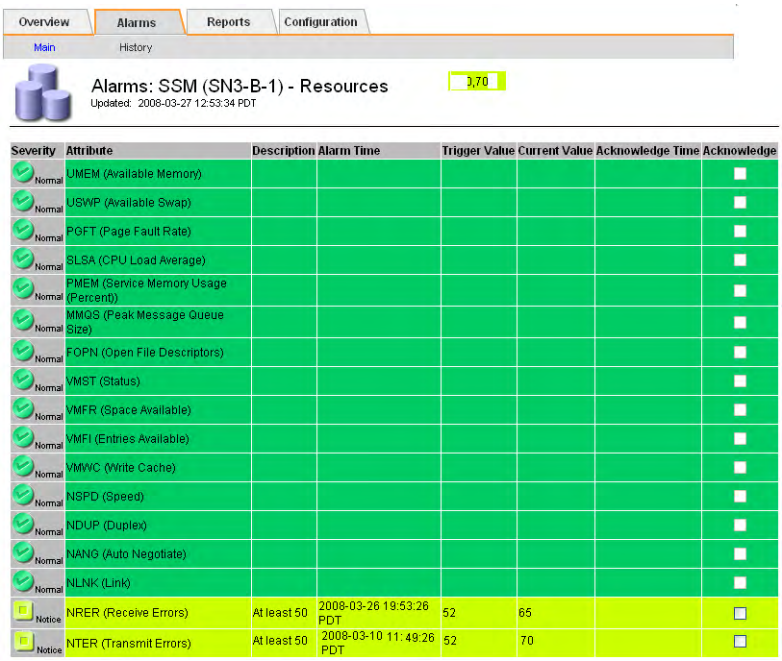
Network	NRER NTER	SSM	Network interface errors (`Receive Errors NRER` and `Transmit Errors NTER`) are fairly common with some network interface adapters. These errors may clear without being manually reset. If they don't clear, check the network hardware.
			

Figure 76Common Network Alarms: NRER, NTER

Resource utilization	UMEM	SSM	Minor alarms for Available Memory (the amount of system RAM available for system operations) set by default at 100 MB are not a cause for concern unless available memory continues to decrease. This could indicate a serious problem. If available memory falls below 50 MB, contact Support.

Table 23 Common Alarms *(continued)*

Category	Code	Service	Notes
Total events	SMTT	SSM	The total number of logged error or fault events (`Total Events SMTT`) includes errors such as network errors and FSG replication errors. Unless these errors have been cleared (that is, the count has been reset to 0), total events alarms may be triggered. NOTE This alarm is safe to ignore only if the events that triggered the alarm have been investigated.

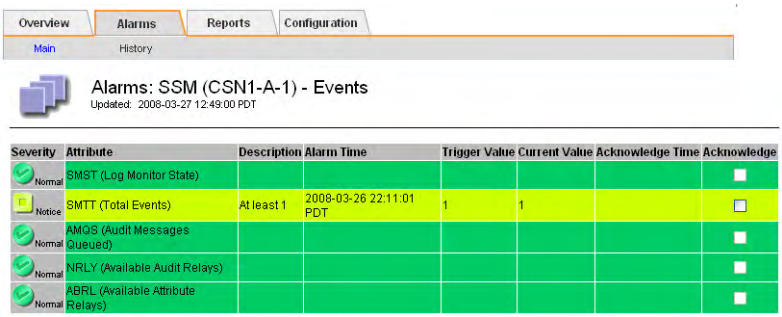


Figure 77Common Events Alarm: Total Events

For more information on these and other alarms, consult the Troubleshooting chapter of the *user guide*.

Glossary

ACL — Access control list—Specifies what users or groups of users are allowed to access an object and what operations are permitted, for example read, write, and execute.

ADC — Administrative Domain Controller—A software component of the HP MAS product. The ADC service maintains topology information, provides authentication services, and responds to queries from the LDR, CMS, and CLB. The ADC service is found on the Control Node.

ADE — Asynchronous Distributed Environment—Proprietary development environment used as a framework for grid services within the HP MAS product.

Admin Node — A building block of the HP MAS product. The Admin Node provides services for the web interface, grid configuration, and audit logs.

AE title — Application Entity Title—The identifier of a DICOM node communicating with other DICOM AEs.

AMS — Audit Management System—A software component of the HP MAS product. The AMS service monitors and logs all audited system events and transactions to a text log file. The AMS service is found on the Admin Node.

API — Application Programming Interface—A set of commands and functions, and their related syntax, that enable software to use the functions provided by another piece of software.

ARC — Archive—A software component of the HP MAS product. The ARC service manages interactions with archiving middleware that controls nearline archival media devices such as tape libraries. The ARC service is found on the Tape Node.

Association — A connection protocol between two DICOM Application Entities (AEs), typically a local and remote AE. The AEs use the Association Establishment to negotiate the type of data to exchange and the format of data encoding.

audit message — Information about an event occurring in the HP MAS system that is captured and logged to a file.

atom — Atoms are the lowest-level component of the container data structure, and generally encode a single piece of information. (Containers are sometimes used when interacting with the grid via the HTTP API).

AutoYaST — An automated version of the Linux installation and configuration tool YaST (“Yet another Setup Tool”), which is included as part of the SUSE Linux distribution.

BASE64

A standardized data encoding algorithm that enables 8-bit data to be converted into a format that uses a smaller character set, enabling it to safely pass through legacy systems that can only process basic (low order) ASCII text excluding control characters. See RFC 2045 for more details.

bundle

A structured collection of configuration information used internally by various components of the grid. Bundles are structured in container format.

Bycast Enablement Layer

The Bycast Enablement Layer CD is used during installation to customize the Linux operating system installed on each grid server. Only the packages needed to support the services hosted on the server are retained, which minimizes the overall footprint occupied by the operating system and maximize the security of each grid node.

cabinet

Used to house hardware components, a cabinet includes the physical rack and all power and network wiring required for an installation.

cabinet connectivity kit

Used to link cabinets. One Cabinet Connectivity Kit is required at Single Site installations that have more than one cabinet, and one is required at each site in a Single Site + DR installation. The Cabinet Connectivity Kit is installed in the Base Cabinet. See also: WAN Connectivity Kit.

CBID

Content Block Identifier—A 64-bit number that uniquely identifies a piece of content within the HP MAS system. CBIDs are represented as a zero-padded, 16-character, hexadecimal number when used to refer to a unique piece of content using the HTTP interface.

CIDR

Classless Inter-Domain Routing—A notation used to compactly describe a subnet mask used to define a range of IP addresses. In CIDR notation, the subnet mask is expressed as an IP address in dotted decimal notation, followed by a slash and the number of bits in the subnet. For example, 192.168.110.0/24.

CIFS

Common Internet File System—A file system protocol based on SMB (Server Message Block, developed by Microsoft) which coexists with protocols such as HTTP, FTP, and NFS.

CLB

Connection Load Balancer—A software component of the HP MAS product. The CLB service provides a gateway into the grid for clients connecting via DICOM and HTTP protocols. The CLB service is part of the Gateway Node.

CMN

Configuration Management Node— A software component of the HP MAS product. The CMN service manages system-wide configuration and grid tasks. The CMN service is found on the Admin Node.

CMS

Content Management System—A software component of the HP MAS product. The CMS service manages content metadata and content replication according to the rules specified by the ILM policy. The CMS service is found on the Control Node.

command

In HTTP, an instruction in the request header such as GET, HEAD, DELETE, OPTIONS, POST, or PUT. Also known as an HTTP method.

connectivity kit See cabinet connectivity kit and WAN Connectivity Kit.

container A container is a data structure used by the internals of grid software. In the HTTP API, an XML representation of a container is used to define queries or audit messages submitted using the POST command. Containers are used for information that has hierarchical relationships between components. The lowest-level component of a container is an atom. Containers may contain 0 to N atoms, and 0 to N other containers.

content block ID See CBID.

Control Node A building block of the HP MAS product. The Control Node provides services for managing content metadata and content replication.

C-STORE A DICOM operation to send data between devices.

CSTR Null-terminated, variable length string.

DC Data Center.

DICOM Digital Imaging and COmmunications in Medicine—A standard developed by ACR-NEMA (an alliance of the American College of Radiology and the National Electrical Manufacturer's Association) for communications between medical imaging devices.

DR Disaster Recovery.

EVA Enterprise Virtual Array—an HP product that uses virtual arrays to allocate SAN or Fibre Channel storage resources to different uses.

Fibre Channel A networking technology primarily used for storage. The standard connection type for SAN.

FCS Fixed Content Storage—a class of stored data where the data, once captured, is rarely changed and must be retained for long periods of time in its original form. Typically this includes images, documents, and other data where alterations would reduce the value of the stored information.

FSG File System Gateway—A software component of the HP MAS product. The FSG service enables standard network file systems to interface with the grid. The FSG service is found on the Gateway Node.

FSG replication group A replication group is a group of FSGs that provide grid access to a specified set of clients. Within each replication group, one FSG is a primary and all others are secondaries. The primary FSG allows clients read and write access to the grid, while storing file system information (stubs) for all files saved to the grid. The secondary FSG "replicates" file system information, and backs up this information to the grid on a regular schedule.

Gateway Node — A building block of the HP MAS product. The Gateway Node provides connectivity services for NFS/CIFS file systems and the HTTP and DICOM protocols.

grid node — The name of the HP MAS product building blocks, for example Admin Node or Control Node. Each type of grid node consists of a set of services running on a server.

Grid Specification File — An XML file that provides a complete technical description of a specific grid deployment. It describes the grid topology, and specifies the hardware, grid options, server names, network settings, time synchronization, and gateway clusters included in the grid deployment. The Deployment Grid Specification file is used to generate the files needed to install the grid.

Grid Task — A managed sequence of actions that are coordinated across a grid to perform a specific function (such as adding new node certificates). Grid Tasks are typically long-term operations that span many entities within the grid. See also Task Signed Text Block.

HP MAS — HP Medical Archive solution — Fixed-content grid storage system from Hewlett-Packard. The solution is sold under the HP brand and is serviced and supported by the HP services/support organization worldwide. The HP MAS Solution is powered by Bycast® StorageGRID® software.

HTTP — Hyper-Text Transfer Protocol—A simple, text based client/server protocol for requesting hypertext documents from a server. This protocol has evolved into the primary protocol for delivery of information on the World Wide Web.

HTTPS — Hyper-Text Transfer Protocol, Secure—URIs that include HTTPS indicate that the transaction must use HTTP with an additional encryption/authentication layer and often, a different default port number. The encryption layer is usually provided by SSL or TLS. HTTPS is widely used on the internet for secure communications.

ILM — Information Lifecycle Management—A process of managing content storage location and duration based on content value, cost of storage, performance access, regulatory compliance and other such factors.

inode — On UNIX/Linux systems, data structure that contains information about each file, for example, permissions, owner, file size, access time, change time, and modification time. Each inode has a unique inode number.

instance — A DICOM term for an image. One or more instances for a single patient are collected in a “study”. For example, each “slice” of an MRI is an instance; together, the full set of slices is a study.

KVM — Keyboard, Video, Mouse—A hardware device consisting of a keyboard, LCD screen (video monitor), and mouse that permits a user to control all servers in a cabinet.

LAN Local Area Network—A network of interconnected computers that is restricted to a small area, such as a building or campus. A LAN may be considered a node to the Internet or other wide area network. Contrast with WAN.

latency Time duration for processing a transaction or transmitting a unit of data from end to end. When evaluating system performance, both throughput and latency need to be considered. See also throughput.

LDR Local Distribution Router—A software component of the HP MAS product. The LDR service manages the storage and transfer of content within the grid. The LDR service is found on the Storage Node.

metadata Information related to or describing an object stored in the grid, for example file ingest path or ingest time.

namespace A set whose elements are unique names. There is no guarantee that a name in one namespace is not repeated in a different namespace.

nearline A term describing data storage that is neither "online" (implying that it is instantly available like spinning disk) nor "offline" (which could include offsite storage media). An example of a nearline data storage location is a tape that is loaded in a tape library, but is not necessarily mounted.

NFS Network File System—A protocol (developed by SUN Microsystems) that enables access to network files as if they were on local disks.

NMS Network Management System—A software component of the HP MAS product. The NMS service provides a web-based interface for managing and monitoring the HP MAS system. The NMS service is found on the Admin Node.

node ID An identification number assigned to a grid service within the HP MAS system. Each service (such as an CMS or ADC) in a single grid must have a unique node ID. The number is set during system configuration and tied to authentication certificates.

NTP Network Time Protocol—A protocol used to synchronize distributed clocks over a variable latency network such as the internet.

object store A configured file system on a disk volume. The configuration includes a specific directory structure and resources initialized at system installation.

PACS Picture Archiving and Communication System—A computerized system of patient records management responsible for short and long term (archival) storage of images.

presentation context A combination of a DICOM SOP Class and a transfer syntax; the type and format of a DICOM transaction.

provisioning The process of editing the Grid Specification File (if required) and generating or updating the SAID package. This is done on the Admin Node using the provision command. The new or updated SAID package is saved to the Provisioning USB flash drive. See Grid Specification File and SAID.

SAID Software Activation and Integration Data—Generated during provisioning, the SAID package contains site-specific files and software needed to install a grid.

Samba A free suite of programs which implement the Server Message Block (SMB) protocol. Allows files and printers on the host operating system to be shared with other clients. For example, instead of using telnet to log into a Unix machine to edit a file there, a Windows user might connect a drive in Windows Explorer to a Samba server on the Unix machine and edit the file in a Windows editor. A Unix client called “smbclient”, built from the same source code, allows FTP-like access to SMB resources.

SAN Storage Area Network—A high-speed network connecting heterogeneous storage devices to servers.

SATA Serial Advanced Technology Attachment—A connection technology used to connect servers and storage devices.

SCP Storage Class Provider—A device that provides images (a storage class) to a DICOM compliant system. Contrast with “SCU”.

SCSI Small Computer System Interface——A connection technology used to connect servers and peripheral devices such as storage systems.

SCU Storage Class User—A device that receives (uses) images (a storage class) from a DICOM compliant system. Contrast with “SCP”.

server Used when referring specifically to hardware.

service A unit of the HP MAS software such as the ADC, CMS or SSM.

SLES SUSE Linux Enterprise Server—A commercial distribution of the SUSE Linux operating system, used with the HP MAS software.

SOP class Service-Object Pair Class—The combination of an information object description (IOD) and the set of Services that are useful for a given purpose.

SOP instance Service-Object Pair (SOP) Instance—A specific occurrence of an Information Object.

SQL Structured Query Language— An industry standard interface language for managing relational databases. An SQL database is one that supports the SQL interface.

SSAM IBM® System Storage™ Archive Manager.

ssh Secure Shell— A Unix shell program and supporting protocols used to log in to a remote computer and execute commands over an authenticated and encrypted channel.

SSM Server Status Monitor—A unit of the HP MAS software that monitors hardware conditions and reports to the NMS. Every server in the grid runs an instance of the SSM. The SSMS service is present on all grid nodes.

SSL Secure Socket Layer—The original cryptographic protocol used to enable secure communications over the internet. See TLS.

Storage Node A building block of the HP MAS product. The Storage Node provides storage capacity and services to store, move, verify, and retrieve objects stored on disks.

StorageGRID A registered trademark of Bycast Inc. for their fixed-content storage grid architecture and software system.

study A DICOM term for a collection of images (instances) related to an individual patient or subject.

SUSE See SLES—SUSE Linux Enterprise Server.

Tape Node A building block of the HP MAS product. The Tape Node manages storage of data to nearline data storage devices such as such as tape libraries (via IBM Tivoli® Storage Manager).

Task Signed Text Block A BASE64 encoded block of cryptographically signed data that provides the set of instructions that define a grid task.

TCP/IP Transmission Control Protocol / Internet Protocol—A process of encapsulating and transmitting packet data over a network. It includes positive acknowledgement of transmissions.

throughput The amount of data that can be transmitted or the number of transactions that can be processed by a system or subsystem in a given period of time. See also latency.

TLS Transport Layer Security—A cryptographic protocol used to enable secure communications over the internet. See RFC 2246 for more details.

transfer syntax The parameters, such as the byte order and compression method, needed to exchange data between systems.

TSM Tivoli® Storage Manager—IBM storage middleware product that manages storage and retrieval of data from removable storage resources.

URI Universal Resource Identifier—A generic set of all names or addresses used to refer to resources that can be served from a computer system. These addresses are represented as short text strings.

UTC

A language-independent international abbreviation, UTC is neither English nor French. It means both “Coordinated Universal Time” and “Temps Universel Coordonné”. UTC refers to the standard time common to every place in the world.

UUID

Universally Unique Identifier—Unique identifier for each piece of content in the HP MAS. UUIDs provide client applications with a content handle that permits them to access grid content in a way that does not interfere with the grid’s management of that same content. A 128-bit number which is guaranteed to be unique. See RFC 4122 for more details.

XFS

A scalable, high performance journaled file system originally developed by Silicon Graphics.

WAN

Wide Area Network—A network of interconnected computers that covers a large geographic area such as a country. Contrast with LAN.

WAN Connectivity Kit

Required for linking the primary and DR sites of an HP MAS deployment. The WAN Connectivity Kit is installed in the Base Cabinet at both locations. See also cabinet connectivity kit.

XML

eXtensible Markup Language—A text format for the extensible representation of structured information; classified by type and managed like a database. XML has the advantages of being verifiable, human readable, and easily interchangeable between different systems.

Index

B

C

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y