Administration Guide

HGST Active Archive System SA-7000

September 2015

1ET0032

Revision 1.1

Long Live Data ™ | www.hgst.com

Copyright

The following paragraph does not apply to the United Kingdom or any country where such provisions are inconsistent with local law: HGST a Western Digital company PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer or express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This publication could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. HGST may make improvements or changes in any products or programs described in this publication at any time.

It is possible that this publication may contain reference to, or information about, HGST products (machines and programs), programming, or services that are not announced in your country. Such references or information must not be construed to mean that HGST intends to announce such HGST products, programming, or services in your country.

Technical information about this product is available by contacting your local HGST representative or on the Internet at: www.hgst.com/support

HGST may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents.

HGST, a Western Digital company
3403 Yerba Buena Road
San Jose, CA 95135
Produced in the United States

Long Live Data™ is a trademark of HGST, Inc. and its affiliates in the United States and/or other countries.

HGST trademarks are authorized for use in countries and jurisdictions in which HGST has the right to use, market and advertise the brands.

Other product names are trademarks or registered trademarks of their respective owners.

One MB is equal to one million bytes, one GB is equal to one billion bytes, one TB equals 1,000GB (one trillion bytes) and one PB equals 1,000TB when referring to storage capacity. Usable capacity will vary from the raw capacity due to object storage methodologies and other factors.

References in this publication to HGST products, programs or services do not imply that HGST intends to make these available in all countries in which HGST operates.

Product information is provided for information purposes only and does not constitute a warranty.

Information is true as of the date of publication and is subject to change. Actual results may vary. This publication is for general guidance only. Photographs may show design models.

Contents

List of Figures.. 8

Chapter 1 About this Guide.. 9
1.1 Conventions.. 9
1.2 Storage Notations.. 9
1.3 Admonitions..10
1.4 Related Documents..10
1.5 Document Map.. 11

Chapter 2 Using the Administrator Interfaces.. 12
2.1 Using a Console.. 12
2.2 Changing the Root Password.. 12
2.3 Using the CMC.. 12
2.3.1 Managing CMC Users and Groups..13
2.4 Using the Q-Shell..15
2.5 Using the OSMI Menu..16
2.5.1 The OSMI Menu.. 16
2.6 Troubleshooting the Administrator Interfaces.. 18
2.6.1 General..18

Chapter 3 Starting and Stopping the Active Archive System.. 20
3.1 Powering on the Active Archive System..20
3.1.1 Powering on a Single-Rack Active Archive System..20
3.1.2 Powering on a Multi-Rack Active Archive System..20
3.2 Shutting Down the Active Archive System..21
3.2.1 Shutting Down a Rack Using the CMC..21
3.2.2 Stopping Services Using the OSMI.. 21
3.2.3 Shutting Down the Entire Active Archive System.. 21
3.3 Troubleshooting Startup and Shutdown Issues..22

Chapter 4 Managing Storage..26
4.1 Creating S3 Users..26
4.1.1 Creating S3 Users Through the Q-Shell.. 26
4.1.2 Creating S3 Users Through OSMI..26
4.1.3 Creating S3 Users Through S3 API..27
4.2 Adding an S3 Client..27
4.3 Creating S3 Buckets (Name Spaces).. 28
4.4 Deleting Buckets (Name Spaces)..29
4.4.1 Deleting Buckets Through S3 API.. 29
4.4.2 Deleting Buckets Through the Q-Shell..29
4.5 Deleting Objects.. 30
4.5.1 Deleting Objects Through S3 API.. 30
4.5.2 Deleting Objects Through the Q-Shell..30
4.6 Disabling Bucket Operations.. 31
4.7 Configuring S3 Multipart Support.. 31

4.8 Metering.....32
4.8.1 Enabling Metering.....32
4.8.2 Disabling Metering.....32
4.8.3 Metering Log Details.....32
4.8.4 Accessing Metering Logs.....33
4.9 Using Encryption.....33
4.9.1 Supported RSA Ciphers.....33
4.10 Managing MetaStores.....34
4.10.1 Marking a MetaStore as Full.....34
4.10.1.1 Automatically Marking a MetaStore as Full.....34
4.10.1.2 Automatically Reactivating a MetaStore.....34
4.10.1.3 Manually Marking a MetaStore as Read Only.....35
4.10.1.4 Reactivating a MetaStore.....35
4.10.2 MetaStore Recovery.....35
4.10.2.1 Fully Automated tlog Collapse.....35
4.10.2.2 Fully Automated Recovery from Unclean Shutdown.....35
4.10.2.3 Backups for Disaster Recovery and RCA Investigation.....35
4.10.3 Automated Master Change.....36
4.10.4 Batch Processing of MetaStore Transactions.....36
4.10.5 Rolling Updates.....36
4.11 Troubleshooting Storage Issues.....36
4.11.1 s3cmd Errors.....36
4.11.2 Cyberduck Errors.....36
4.11.3 Spread Width, Safety Policies, and Storage Policies.....37

Chapter 5 Managing Networks.....41
5.1 The Active Archive System LANs.....41
5.2 Updating the Public Network Settings.....41
5.3 Troubleshooting Network Issues.....41
5.3.1 General.....42

Chapter 6 Managing Hardware.....46
6.1 Using the Intelligent Platform Management Interface (IPMI).....46
6.1.1 Remote Power Cycle.....47
6.1.2 Remote Capture of Screen and Keyboard.....47
6.1.3 Viewing Sensor Readings.....47
6.1.4 Toggling a Location LED.....47
6.2 Handling Blacklists.....48
6.3 Handling a Degraded Disk Notification.....48
6.3.1 Troubleshooting a Degraded Disk on a Storage Enclosure Basic.....49
6.3.2 Troubleshooting a Degraded Disk on a Storage Node.....49
6.3.3 Troubleshooting Multiple Degraded Disks on a Storage Node.....49
6.3.4 Troubleshooting Degraded Disks on a Controller Node.....51
6.3.5 Performing a Detailed SMART Analysis.....52
6.3.6 Degraded Disk Troubleshooting Flowchart.....53
6.3.7 Reviewing the Number of Blacklists of a Degraded Disk.....54
6.3.8 Further Checks.....54
6.3.9 Decommissioning a Degraded Disk.....55
6.3.10 Enhanced Secure Erase.....56
6.4 Handling a Read-Only File System.....56

PostgreSQL Disk....56
Non-PostgreSQL Disk....57
6.5 Handling Node Failures....57
6.5.1 Decommissioning a Node in the CMC....57
6.5.2 Executing a Failover....58
6.6 Managing Unmanaged Disks....59
6.7 Handling Unverified Objects....60
6.8 Troubleshooting Hardware Issues....61
6.8.1 Field Replaceable Units....61
6.8.2 System Expansion....62
6.8.3 General....62

Chapter 7 Configuring Maintenance Policies....66
7.1 Back Up Model Database Policy....66
7.2 Back Up Software Repository Policy....66
7.3 Blockstore Automatic Disk Decommissioning Policy....67
7.3.1 About Degraded and Decommissioned Disks....67
7.3.2 Viewing Auto Decommissioned Drives....68
7.3.3 Enabling the Auto Decommission Policy....68
7.3.4 Disabling the Auto Decommission Policy....68
7.3.5 Editing the Auto Decommission Policy....68
7.4 Collapse MetaStore Transaction Logs Policy....69
7.5 Defragment MetaStores Policy....69
7.6 Phone Home Policy....70
7.7 System Maintenance Policies....70
7.7.1 Aggregate Storagepool Info Policy....70
7.7.2 Clean Up Old Versions Policy....71
7.7.3 Monitor Blockstores Policy....71

Chapter 8 Logging....72
8.1 Collecting Log Files....72
8.2 Default Log Behavior....73
8.3 Normal Log Collection....73
8.4 IPMI Log Collection....75
8.5 Emergency Log Collection....75
8.6 Extracting Log Files....76
8.7 Other Log Options....77
8.8 Log File Management....77
8.9 Client Daemon Log File Statistics....79
8.10 Arakoon Log File Statistics....81
8.11 Troubleshooting Logging Issues....83
8.11.1 General....83

Chapter 9 Monitoring the Active Archive System....85
9.1 About Monitoring....86
9.2 About Events....86
9.3 Monitoring Repair Tasks....87
9.4 Monitoring of Long Running Jobs....88
9.5 Monitoring the System Using the CMC Dashboard....88

9.6 Monitoring the System Using Q-Shell....89
9.7 Event Viewer....93
9.7.1 Event Severity....93
9.7.2 Event Details....94
9.7.3 Live Events vs All Events....94
9.7.4 Removing Events from the List of Events....94
9.7.5 Exporting Events.... 94
9.8 Configuring Event Escalation.... 95
9.8.1 Configuring SNMP Traps.... 95
9.9 Checking System Health.... 96
9.10 Node Monitoring.... 97
9.11 Storage Pool Monitoring.... 98
9.11.1 Monitoring the Storage Pool.... 98
9.11.2 Monitoring the Storage Policy.... 98
9.11.3 Disk Safety Details....99
9.12 SNMP Polling....99
9.12.1 Which Data Can Be Polled?.... 99
9.12.2 Downloading the MIB Variables.... 100
9.12.3 Configuring SNMP Polling.... 100
9.13 Dead Man Timer.... 100
9.14 Collecting Telemetry.... 101
9.14.1 About Telemetry Collection.... 101
9.14.2 Displaying Telemetry Collection Categories.... 101

Chapter 10 Tuning the Active Archive System....104
10.1 Tuning for Optimal Connection Management.... 104
10.2 Tuning for Maximum Throughput....104
10.3 Tuning for Optimal Repair Performance.... 105
10.4 Optimal Performance Vs Maximum Number of Concurrent Users....106
10.4.1 Tuning for Maximum Number of Concurrent Users.... 106
10.4.2 Tuning for Optimal Performance.... 106
10.4.3 In-depth Information....107
10.5 Updating the Time Settings....108
10.5.1 Date and Time in the Active Archive System.... 108
10.5.2 Selecting an NTP Server.... 108

Appendix A Glossary.... 109
A.1 Application Server....109
A.2 Blacklists....109
A.3 Cache Daemon....110
A.4 Client Daemon....110
A.5 Cloud Management Center (CMC)....110
A.6 Controller Node.... 111
A.7 Disk Safety.... 111
A.8 HTTPS Proxy Server....111
A.9 Intelligent Platform Management Interface (IPMI).... 111
A.10 Maintenance Agent....112
A.11 Management Node....112
A.12 MetaStore....112
A.13 Object Store Management Interface (OSMI)....113

A.14 Repair Spread........ 113
A.15 Storage Daemon........ 113
A.16 Storage Node........ 113
A.17 Superblock........ 113

Appendix B Ports, Protocols, and Services........ 115
B.1 General Services........ 115
B.2 PXE Services........ 116
B.3 NFS (RPC)........ 116
B.4 Application Server........ 116
B.5 Arakoon........ 116
B.6 BitSpread/DSS........ 117

Appendix C Events........ 118
C.1 Events Explained........ 118
C.2 Complete List of Events........ 118
C.3 Events in Detail........ 131
C.3.1 Application Events........ 131
C.3.2 Disk Events........ 136
C.3.3 DSS Events........ 143
C.3.4 Physical Machine Events........ 149
C.3.5 MetaStore Events........ 163
C.3.6 Network Events........ 170
C.3.7 Generic Events........ 173
C.3.8 Other Events........ 177

Appendix D Object Verification........ 180
D.1 About Object Verification........ 180
D.2 Object Verification in Detail........ 181

List of Figures

Figure 1: Document Map.. 11

Figure 2: A Storage Node Pane in the CMC... 50

Figure 3: The Shutdown Button in the Commands Pane.. 50

Figure 4: Degraded Disk Flowchart...53

Figure 5: Decommissioning Status... 57

Figure 6: Verification Progress: Unverified Objects.. 60

Figure 7: Network Connectivity...64

Figure 8: The Active Archive System Monitoring System... 86

Figure 9: Maintenance Agent Repair Tasks... 88

Figure 10: The CMC Dashboard.. 89

Figure 11: Exporting Events... 95

Figure 12: Downloadable Content: Exported Events... 95

Figure 13: Disk Safety Details..99

Figure 14: Verification Progress: First Iteration..181

Figure 15: Verification Progress: Sixth Iteration...181

1 About this Guide

Topics:

- Conventions
- Storage Notations
- Admonitions
- Related Documents
- Document Map

The HGST Active Archive System is a fully integrated, tested, and assembled storage appliance in an industry-standard 42RU rack.

The Active Archive System can be deployed with minimal effort, integrated with your existing S3-aware applications, and expanded in one-rack increments. It provides a web-based GUI, a command-line interface, and a menu-driven interface. This guide explains how to use these interfaces for executing system management, monitoring, and analytics tasks.

1.1 Conventions

Element	Sample Notation
OS shell or Q-Shell commands (user input)	`rm -rf /tmp`
OS shell or Q-Shell system output	`Installation successful!`

Commands longer than one line are split with "\\"	`q.dss.manage.setPermissions('/manage', \`
	`[....])`

User-supplied values	*`ManagementNodeVirtualIPAddress`* or `<ManagementNodeVirtualIPAddress>`
File and directory names	The file `aFile.txt` is stored in `/home/user`.
Any graphical user interface label	Click **OK**.
Keyboard keys and sequences	To cancel the operation, press `Ctrl+c`.
Menu navigation in a GUI	Navigate to **Dashboard** > **Administration** > **Hardware** > **Servers**.

1.2 Storage Notations

Convention	Prefix	Size (bytes)
KB	kilobyte	1,000
KiB	kibibyte	1,024
MB	megabyte	1,000,000
MiB	mebibyte	1,048,567
GB	gigabyte	1,00,000,000
GiB	gibibyte	1,073,741,824
TB	terabyte	1,000,000,000,000
TiB	tibibyte	1,099,511,627,776

- Sizes of disks are expressed with *SI prefixes* (kilo, mega, tera, peta, exa)
- Space, size of partitions and file systems are expressed with the *binary prefixes* (kibi, mebi, tebi, pebi, exbi)
- A comma (",") is used for digit grouping, for example 1,000 is 1 thousand.
- A period (".") is used as decimal mark, for example 12.5 %.

1.3 Admonitions

Type	Usage
Note:	Indicates extra information that has no specific hazardous or damaging consequences.
Tip:	Indicates a faster or more efficient way to do something.
Caution:	Indicates an action that, if taken or avoided, may result in hazardous or damaging consequences.
Warning:	Indicates an action that, if taken or avoided, may result in data loss or unavailability.

1.4 Related Documents

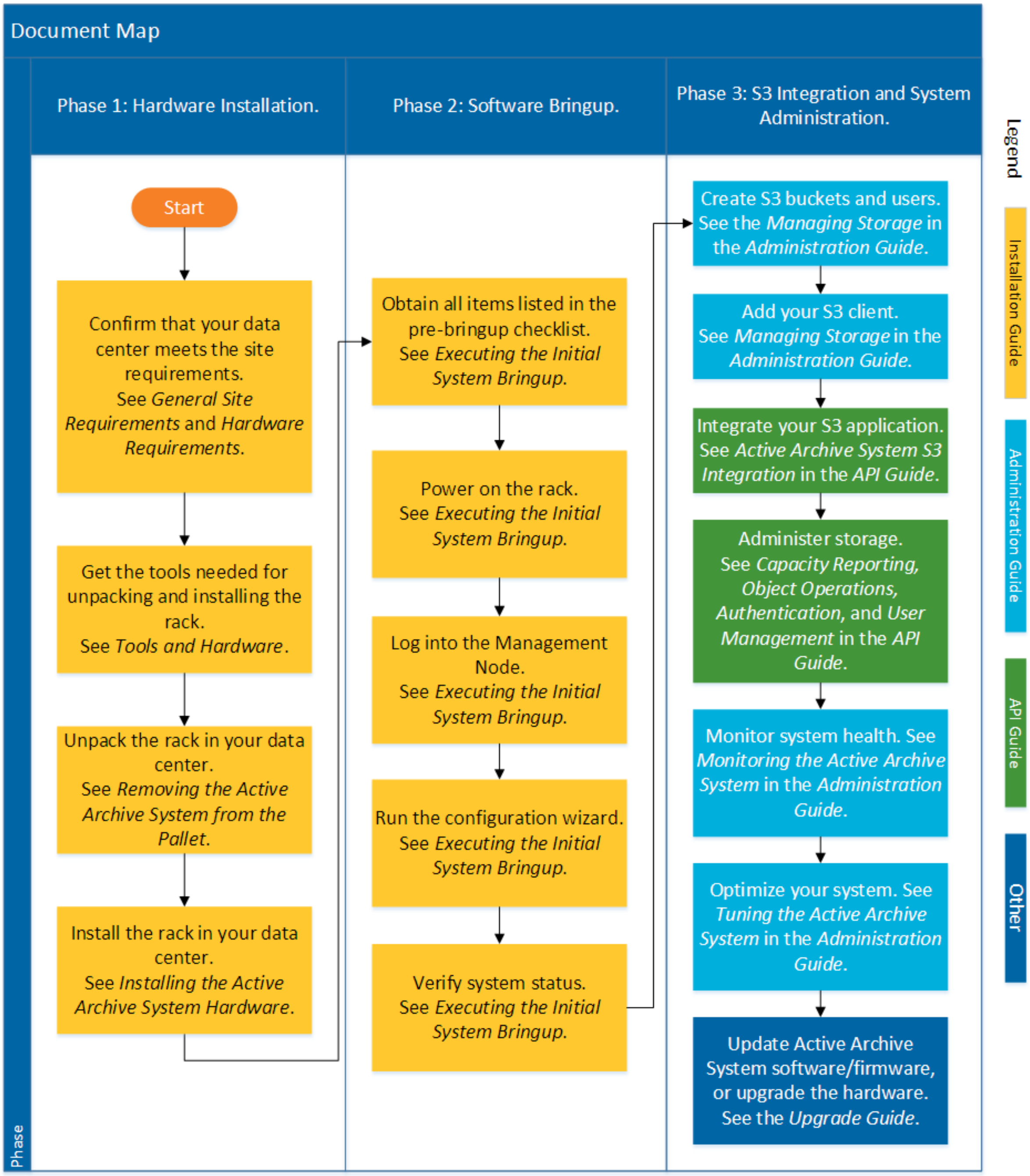
For more information about the Active Archive System, please consult the following documents:

- The *HGST Active Archive System Administration Guide* explains how to use the Active Archive System interfaces for executing system management, monitoring, and analytics tasks.
- The *HGST Active Archive System API Guide* provides a reference for the Active Archive System S3 API.
- The *HGST Active Archive System FRU Replacement Guide* provides procedures for replacing hardware components of the Active Archive System.
- The *HGST Active Archive System Installation Guide* provides instructions for the installation of the Active Archive System in the data center, and its initial bringup.
- The *HGST Active Archive System Release Notes* provide important information about changes, new features, and known limitations.
- The *HGST Active Archive System Site Requirements Document* contains data center requirements for the Active Archive System.
- The *HGST Active Archive System Troubleshooting Guide* provides help for issues you might encounter.
- The *HGST Active Archive System Upgrade Guide* provides instructions for software and firmware updates, and system expansion.

For the latest or online version of any of these documents, visit http://www.hgst.com/support.

1.5 Document Map

Figure 1: Document Map


Document Map
Phase
Phase 1: Hardware Installation.
Phase 2: Software Bringup.
Phase 3: S3 Integration and System Administration.
Start
Confirm that your data center meets the site requirements. See *General Site Requirements* and *Hardware Requirements*.
Get the tools needed for unpacking and installing the rack. See *Tools and Hardware*.
Unpack the rack in your data center. See *Removing the Active Archive System from the Pallet*.
Install the rack in your data center. See *Installing the Active Archive System Hardware*.
Obtain all items listed in the pre-bringup checklist. See *Executing the Initial System Bringup*.
Power on the rack. See *Executing the Initial System Bringup*.
Log into the Management Node. See *Executing the Initial System Bringup*.
Run the configuration wizard. See *Executing the Initial System Bringup*.
Verify system status. See *Executing the Initial System Bringup*.
Create S3 buckets and users. See the *Managing Storage* in the *Administration Guide*.
Add your S3 client. See *Managing Storage* in the *Administration Guide*.
Integrate your S3 application. See *Active Archive System S3 Integration* in the *API Guide*.
Administer storage. See *Capacity Reporting*, *Object Operations*, *Authentication*, and *User Management* in the *API Guide*.
Monitor system health. See *Monitoring the Active Archive System* in the *Administration Guide*.
Optimize your system. See *Tuning the Active Archive System* in the *Administration Guide*.
Update Active Archive System software/firmware, or upgrade the hardware. See the *Upgrade Guide*.
Legend
Installation Guide
Administration Guide
API Guide
Other

2 Using the Administrator Interfaces

Topics:

- Using a Console
- Changing the Root Password
- Using the CMC
- Using the Q-Shell
- Using the OSMI Menu
- Troubleshooting the Administrator Interfaces

The Active Archive System includes a web-based GUI, a command-line interface, and a menu-driven interface for system management, monitoring, and analytics. This chapter provides information on using these interfaces.

2.1 Using a Console

You can log into the console of any Controller Node by connecting a monitor to its VGA port and a keyboard to its USB port.

2.2 Changing the Root Password

To change the password of `root` in the Active Archive System database, do the following:

Run the following code in the Q-Shell on any node:
This code changes the root password on all nodes to *`new_password`*.

Note: You only need to run this code one once, on any one node.

```
api = i.config.cloudApiConnection.find('main')
machine_list = api.machine.find()['result']
for machine in machine_list:
    api.machine.changePassword(machine, 'root', 'new_password')
```

2.3 Using the CMC

To connect to the CMC, you need the following:

- A supported browser:
  - Internet Explorer
  - Safari
  - Mozilla Firefox
  - Google Chrome
- Adobe Flash Player 13.0.0.214 or lower
- Valid login credentials

To connect to the CMC, proceed as follows:

1. Open one of the supported browsers.
2. Navigate to `http://`*`ManagementNodeVirtualIPAddress`*`/flash/CMC/cmc.swf`, where *`ManagementNodeVirtualIPAddress`* is the virtual IP address of the Management Node.
3. Enter your username and password.

The default username and password are `admin` and `HGST`.

The CMC dashboard appears.

> **Note:** You are automatically logged out after ten minutes of inactivity.

2.3.1 Managing CMC Users and Groups

This section describes the management of CMC users and user groups.

> **Tip:** These users and groups are associated with the CMC only, not the S3 interface. For information on creating S3 users, see the *Managing Storage* chapter.

2.3.1.1 Adding a User

To add a CMC user:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Users**.
2. In the right column, click **Add**.
   The **Add Clouduser** wizard appears.
3. On the **General** tab, fill out the form:
   a) In the **Login** field, set a login name for the new user.
   b) In the **Password** and **Confirm Password** fields, set a password for the new user.
   c) In the **Account Name** field, set a meaningful account name for the new user.
4. (Optional) Fill out the form on the **Information** tab.
5. Click **Next** to complete the wizard.

2.3.1.2 Editing a User

To edit a CMC user:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Users**.
2. Select the desired account.
3. In the right column, click **Edit**.
   The **Edit User** dialog appears.
4. To update the account name, click the **General** tab.
5. To update the account's personal information, such as **e-mail**, **address**, **phone**, and so on, click the **Information** tab.
6. Click **Next**, and confirm the changes.

2.3.1.3 Updating a User Password

To update the password of a CMC user:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Users**.
2. Select the desired account.
3. In the right column, click **Change Password**.
   The **Change Password** wizard appears.
4. In the **Old password** field, fill in the current password of the user.
5. Set a new password in the **New password** and **Confirm Password** fields.
6. Click **Next** to complete the wizard.

2.3.1.4 Removing a User

To remove a CMC user, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Users**.
2. Select the desired account.
3. In the right column, click **Remove**.
4. Click **Yes** to confirm the removal of the account.

2.3.1.5 Adding a Cloud Group

To add a cloud group, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Groups**.
2. In the right column, click **Add**.
   The **Add Cloudgroup** wizard appears.
3. Fill out the form:
   a) In the **Name** field, specify a name for the new group.
   b) (Optional) In the **Description** field, specify a description for the new group.
   c) In the **Role** pull-down menu, select a role for the new group.
4. Click **Next**, and confirm the changes.

2.3.1.6 Adding Users to a Cloud Group

To add a user to a cloud group, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Groups**.
2. Select the desired cloud group.
3. In the right column, click **Add Clouduser**.
   The **Add Clouduser** dialog appears.
4. In the list of users, select the users that you want to add to the group.
5. Click **Next**, and confirm the changes.

2.3.1.7 Removing Users from a Cloud Group

To remove a user from a cloud group, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Groups**.
2. Select the desired cloud group.
3. In the right column, click **Remove Clouduser**.
4. In the list of users, select the users that you want to remove from the group.
5. Click **Next**, and confirm the changes.

2.3.1.8 Adding Cloud User Groups to Another Cloud Group

To add a group to an existing cloud group, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Groups**.
2. Select the desired cloud group.
3. In the right column, click **Add Cloudusergroup**.
4. In the list of groups, select the groups that you want to add to this cloud group.

5. Click **Next**, and confirm the changes.

2.3.1.9 Removing Cloud User Groups from Another Cloud Group

To remove a group from an existing cloud group, proceed as follows:

1. In the list of groups, select the desired groups you want to remove from this group.
2. Click **Next** and confirm.

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Groups**.
2. Select the desired cloud group.
3. In the right column, click **Remove Cloudusergroup**.
4. In the list of groups, select the groups that you want to remove from this cloud group.
5. Click **Next**, and confirm the changes.

2.3.1.10 Editing a Cloud Group

To edit a cloud group, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Groups**.
2. Select the desired cloud group.
3. In the right column, click **Edit**.
   The **Edit Cloudgroup** wizard appears.
4. Update the desired fields in the form:
   a) In the **Name** field, specify a new name for the group.
   b) (Optional) In the **Description** field, specify a new description for the group.
   c) In the **Role** pull-down menu, select a new role for the group.
5. Click **Next**, and confirm the changes.

2.3.1.11 Removing a Cloud Group

To remove a cloud group, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Groups and Users** > **Groups**.
2. Select the desired cloud group.
3. In the right column, click **Remove**.
4. Click **Yes** to confirm the changes.

2.4 Using the Q-Shell

The Q-Shell is an interactive Python shell that is available on every Controller Node.

To start a Q-Shell session on any Controller Node, do the following:

1. Open an SSH session to the Controller Node.
   The OSMI menu appears.
2. Exit the OSMI menu.
   The Linux prompt appears.
3. Run the following code at the Linux prompt:

```
/opt/qbase3/qshell
```

2.5 Using the OSMI Menu

The Object Store Management Interface (OSMI) is a menu-based management interface that does not require knowledge of Q-Shell or interactive Python.

When you log into the Management Node over SSH, OSMI is automatically started. To start OSMI manually, use the following command at the node's Linux prompt:

```
/opt/qbase3/apps/osmi/osmi
```

> **Note:** Any deprecation warning you may encounter when starting OSMI can be safely ignored.

You can start OSMI with the following command line options:

Option	**Description**
`-h, –help`	Shows the help messages and exits.
`-p, –preload`	Preloads daemon configuration files at startup. If you notice that OSMI reacts very slowly to certain selections in large environments, use this option. Using this option results in a performance increase (especially for tasks in the **Machines and services** submenu), but the initial menu takes longer to load. (With this option, the OSMI contacts all running nodes and builds an in-memory cache of configuration files for all daemons. This speeds up the listing of the daemons in OSMI at run-time, but slows down the startup process.) Preloading is done typically in large environments when you want to get better response times when doing large queries. If the preload is not specified, the configuration file is loaded on each listing.
`-d, –debug`	Enables debug logging. **Note:** The debug mode is for debugging purposes only. Use this mode with caution.
`-v, –version`	Displays version information.
`-l, –list`	Lists the menu and exits.

2.5.1 The OSMI Menu

Select the number of the item you would like to navigate to or use `0` to return to the previous menu. To select a menu item, do any of the following:

- Type a single numeric value (for example, `2`)
- Type a comma-separated list of values (for example, `1,4,5`)
- Type `a`, to indicate all items

> **Note:** If the output should become too long, it will not be shown on screen, but written in a file on the following location: `/opt/qbase3/var/tmp/`.

The OSMI menu is shown in the output below.

```
1) Environment
   1) List installed packages
   2) List management policy
   3) Execute management policy
   4) Show storage tasks
   5) Turn off all location LEDs
   6) Update public LAN settings
   0) Return to Main Menu
2) Policies and Namespaces
   1) List namespaces
   2) Delete namespace
   3) Repair namespace
   4) Find files with disk safety
   5) Show statistics
   6) Show usage
   7) Show permissions
   8) Set permissions
   0) Return to Main Menu
3) Machines and services
   1) Machines
      1) List machines
      2) Start all services
      3) Stop all services
      4) Restart all services
      5) Locate machine
      0) Return to Main Menu
   2) Metastores
      1) List metastores
      2) List metanodes
      3) Start metanode
      4) Stop metanode
      0) Return to Main Menu
   3) Client Daemons
      1) List client daemons
      2) Start client daemon
      3) Stop client daemon
      0) Return to Main Menu
   4) Storage Daemons
      1) List storage daemons
      2) Start storage daemon
      3) Stop storage daemon
      4) Start repair on storage daemon
      5) List blacklisted storage daemons
      0) Return to Main Menu
   5) Cache daemons
      1) List cache daemons
      2) Start cache daemon
      3) Stop cache daemon
      0) Return to Main Menu
   6) Maintenance Agent
      1) List maintenance agents
      2) Start maintenance agent
      3) Stop maintenance agent
      0) Return to Main Menu
   7) Block Stores
      1) List block stores
      2) List block store usage
      3) Decommission block store
      4) Verify block store
      0) Return to Main Menu
   0) Return to Main Menu
4) Users and permissions
   1) List users
   2) Add user
   3) Delete user
   4) Show permissions
   5) Set permissions
   0) Return to Main Menu
5) Events and logging
   1) Test SMTP configuration
   2) Test SNMP configuration
   3) Test Phone Home configuration
   0) Return to Main Menu
0) Exit
```

2.6 Troubleshooting the Administrator Interfaces

2.6.1 General

<table>
<tr><th>Problem</th><th>Recommended Action</th></tr>
<tr><td>Cannot determine the virtual IP address of the Management Node.</td><td>1. Open an SSH session to any Controller Node.
2. Use the following command to determine the virtual IP address of the Management Node.<pre>grep dmachine.amplistor.com /etc/hosts | grep -v 127.0.0.1 | awk
 '{print $1}'</pre>The output of this command is the virtual IP address of the Management Node. For example,<pre>172.16.63.154</pre></td></tr>
<tr><td>Cannot determine the hostname and physical IP addresses of the Management Node.</td><td>1. Open an SSH session to any Controller Node.
2. Use the following command to determine the virtual IP address of the Management Node.<pre>grep dmachine.amplistor.com /etc/hosts | grep -v 127.0.0.1 | awk
 '{print $1}'</pre>The output of this command is the virtual IP address of the Management Node. For example,<pre>172.16.63.154</pre>3. Open an SSH session to the virtual IP address of the Management Node, obtained in the previous step.
4. Exit the OSMI menu.
5. Note the hostname in the Linux command prompt.
6. Use <code>ifconfig</code> to gather all unique IP addresses for the Management Node.</td></tr>
<tr><td>Cannot access the CMC.</td><td>The CMC runs on the Management Node. If the Management Node has failed over to another Controller Node, you must access the CMC using:
• the IP address of the new Management Node, or
• the virtual IP address of the Management Node, which remains the same despite a failover.</td></tr>
<tr><td>Cannot access the OSMI menu.</td><td>To start OSMI manually, use the following command at the Linux prompt:<pre>/opt/qbase3/apps/osmi/osmi</pre></td></tr>
</table>

<table>
<tr><th>Problem</th><th>Recommended Action</th></tr>
<tr><td>Cannot log into CMC with correct credentials.</td><td>

Check to see if there is a defective SSD on your Management Node. See *Managing Hardware* in the *HGST Active Archive System Administration Guide*.

</td></tr>
<tr><td>Cannot print failed drive map from CMC.</td><td>

If **Export Details as PDF** does not respond, you may be using an incompatible version of Adobe Flash Player.

Use Adobe Flash Player 13.0.0.214 or lower.

</td></tr>
<tr><td>Cannot identify the Management Node.</td><td>

There are two ways to determine which node is the Management Node:

1. Through OSMI:

A. Log into any Controller Node.
B. In the OSMI menu select option 3, then option 1, then option 1:

```
3) Machine and Services -> 1) Machines -> 1) List Machines
```

A list of machines is displayed. The Management Node is the one that has the Management Framework running. For example,

```
------------
- Machines -
------------
1) Machine HGST-S3-DC01-R01-CN01 (type: CPUNODE, status:
 RUNNING)
        Components: 1 management framework(s), 4 client
 daemon(s)
...
```

2. Through the base OS of the Management Node.

A. Open an SSH session to the Management Node using the virtual IP address.
B. Exit the OSMI menu by pressing `0` twice.
C. The Linux prompt is the hostname of the Management Node. Part of the hostname is Controller Node (`CN01`, `CN02`, or `CN03`)

</td></tr>
<tr><td>Upon rebooting or shutting down a Controller Node through the CMC, the connection to the CMC is lost.</td><td>

If you reboot or shut down a Controller Node, without realizing that the CMC is running on the same physical node, the CMC session is lost.

Workaround: Ensure that the node you wish to reboot or shut down is not the Management Node: determine the physical IP addresses of the Management Node by following the steps in this guide to identify the Management Node.

</td></tr>
</table>

3 Starting and Stopping the Active Archive System

Topics:

- Powering on the Active Archive System
- Shutting Down the Active Archive System
- Troubleshooting Startup and Shutdown Issues

3.1 Powering on the Active Archive System

3.1.1 Powering on a Single-Rack Active Archive System

Power on the entire rack.

a) Connect the external power cords of the rack to two different power distribution networks.
The rack begins to power up as soon as the power cords are connected. The intelligent programmable PDUs control the bring-up sequence.

b) Confirm that all nodes power on in the right order. There is a short gap between each segment:

   a. Network switches
   b. Controller Nodes
   c. Storage Enclosure Basic
   d. Storage Nodes

c) Log into the CMC.

d) Wait until the CMC displays the status of the Management Node as **RUNNING**; in other words, its startup is complete.

e) Verify that the CMC dashboard indicates that the system status is good:

   Disk Safety is **5**.
   Controller Nodes indicate the correct number are **UP**.
   Storage Nodes indicate the correct number are **UP**.
   MetaStores indicate the correct number are **OK**.
   Disks displays the correct number for your system, and none are degraded or decommissioned.
   No status indicator is red.

f) Verify that the CMC displays the status of at least 5 Storage Nodes as **RUNNING**:
Navigate to **Dashboard** > **Administration** > **Hardware** > **Servers** > **Storage Nodes**. Check the **Status** field.

3.1.2 Powering on a Multi-Rack Active Archive System

To power on a multi-rack Active Archive System, proceed as follows:

1. Connect the two external power cables of Rack 1 to two different power distribution networks.
The rack begins to power up as soon as the power cables are connected. The intelligent programmable PDUs control the bring-up sequence.
2. Confirm that all nodes power on in the right order. There is a short gap between each segment. In a multi-rack setup, the startup order of Rack 1 is:

   a) Controller Nodes
   b) Network switches

c) Storage Enclosure Basic

d) Storage Nodes

3. Connect the two external power cables of Racks 2-5 to two different power distribution networks.
   The racks begins to power up as soon as the power cables are connected. The intelligent programmable PDUs control the bring-up sequence.
4. Verify that the CMC dashboard indicates that the system status is good:

a) Log into the CMC.

b) Look at the following indicators:

Disk Safety is **5**.

Controller Nodes indicate the correct number are **UP**.

Storage Nodes indicate the correct number are **UP**.

MetaStores indicate the correct number are **OK**.

Disks displays the correct number for your system, and none are degraded or decommissioned.

No status indicator is red.

3.2 Shutting Down the Active Archive System

3.2.1 Shutting Down a Rack Using the CMC

To shut down a rack in the Active Archive System using the CMC, proceed as follows:

Shut down the entire rack.

a) In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Servers**.

b) Select each of the six Storage Nodes, and in the right pane, click **Shutdown**.

c) Shutdown Controller Node 3, then Controller Node 2, then Controller Node 1 (in other words, shut down the Management Node last).

3.2.2 Stopping Services Using the OSMI

For debugging purposes, you might need to leave a node running but stop the services on certain or all nodes. This is necessary in order to access log files, for example. The OSMI allows you to stop all services or a specific service on a specific node.

1. Open an SSH session to the Management Node.
   The OSMI menu appears.
2. In the OSMI menu, select **Machines and services** > **Machines** > **Stop all services**.
3. When prompted to select a machine, do one of the following:
   - type a comma-separated list of IDs of multiple machines.
   - type "`a`" to select all machines.

3.2.3 Shutting Down the Entire Active Archive System

To shut down the entire Active Archive System, proceed as follows.

1. Log into the Management Node.
   The OSMI menu appears.
2. Exit the OSMI menu.
   The Linux prompt appears.
3. At the Linux prompt, do a test run of the `shutdown_environment.py` script to verify the order of nodes to be shut down:

Note: This script does not run on a node that is not the Management Node.

```
/opt/qbase3/bin/python  /opt/qbase3/utils/HGST/shutdown_environment.py
```

4. Run the `shutdown_environment.py` script with the `--shutdown` option:

```
/opt/qbase3/bin/python  /opt/qbase3/utils/HGST/shutdown_environment.py
```

3.3 Troubleshooting Startup and Shutdown Issues

This section provides troubleshooting tips for issues you might encounter during when starting or stopping the Active Archive System. For more troubleshooting tips, see the *HGST Active Archive System Troubleshooting Guide*.

3.3.1 General

Problem	**Recommended Action**
There is an unknown problem at startup.	The following are a list of logs and commands that will assist in troubleshooting procedures. • Log: `/var/log/boot.log` - The contents of this log are identical to what is printed on the system console during the boot sequence. This provides a useful alternative to attaching a monitor and keyboard to the system to see the console output. • Log: `/var/log/kern.log` - The contents of this log may indicate if there are any hardware faults on start up. • Q-Shell command: `q.manage.servers.all.start()` - This Q-Shell command starts all services on a system in the correct order. This command is useful in quickly identifying what services are failing to start on bootup, if any. This command can be run multiple times without impacting already running services, so it is also useful to resume the remaining services on a node if any one service failed to start. • Q-Shell command: `q.manage.servers.all.stop()` - This command, combined with the previous command, allows services to be brought up in a clean fashion. This may be required if services need to be restarted after a network change upon booting. • Q-Shell command: `q.amplistor.healthCheck()` - This command runs a health check on the Active Archive System to see if services are running and MetaStores have masters elected. • Q-Shell command: `print q.dss.manage.showLocationHierarchy()` - This command prints a list of the state of all blockstores.
There is an unknown problem at startup.	The following are a list of logs and commands that will assist in troubleshooting procedures. • Log: `/var/log/boot.log` - The contents of this log are identical to what is printed on the system console during the boot sequence. This provides a useful alternative to attaching a monitor and keyboard to the system to see the console output. • Log: `/var/log/kern.log` - The contents of this log may indicate if there are any hardware faults on start up. • Q-Shell command: `q.manage.servers.all.start()` - This Q-Shell command starts all services on a system in the correct order. This command is useful in quickly identifying what services are failing to start on bootup, if any. This command can be run multiple times without impacting already running services, so it is also useful to resume the remaining services on a node if any one service failed to start. • Q-Shell command: `q.manage.servers.all.stop()` - This command, combined with the previous command, allows services to be brought up in a clean fashion. This may be required if services need to be restarted after a network change upon booting.

Problem	Recommended Action
	• Q-Shell command: `q.amplistor.healthCheck()` - This command runs a health check on the Active Archive System to see if services are running and MetaStores have masters elected. • Q-Shell command: `print q.dss.manage.showLocationHierarchy()` - This command prints a list of the state of all blockstores.
A service failed to start.	Active Archive System services are started in a specific order. The startup sequence stops if one service fails to start. This means that if one service is not running, it may not be because there is a problem with that service. Instead, it may indicate that some other service in the start order failed to start. Active Archive System services may be dependent upon other services to be functioning to start correctly, such as the `framework` (management and monitoring) and `env_metastore` (DSS) *system MetaStores*: • The application server requires the `framework` MetaStore to have a master. • The monitoring agent requires the `framework` MetaStore to have a master. • DSS processes (client daemons, Storage Nodes, and maintenance agents) require the env_metastore MetaStore to have a master. The best place to see if services are starting correctly is on the console, in `/var/log/boot.log` of the machine being started, or in `/opt/qbase3/var/log/pylabslogs/autostart.log`.
There are many events immediately after startup.	When you start the Active Archive System, you will encounter the following side effects: • Many events are raised, indicating that: ◆ There are failed jobs, caused by MetaStores which are not fully operational ◆ The disk safety is lowered, because not enough MetaStore nodes are available • There are failed data operations, due to MetaStores that are not yet available Once the MetaStores are available again, the number of events lowers, but for large environments, it may take a couple of hours before the *data MetaStores* are fully operational. The recovery of the *system MetaStores* (`env_metastore` and `framework`) takes less time.
The CMC indicates that a node is down.	Run the Aggregate Storagepool Info policy. This system policy runs only if all services are running correctly on the Management Node. In addition to this policy's role in aggregating monitoring data from all nodes in an environment, it is also responsible for checking the UP/DOWN status and restarting the agent service and monitoring agent all nodes. A node that is powered on may not show as UP in the CMC until this policy runs once. The policy runs by default every 30 minutes, but can be triggered to run immediately through the OSMI.

The Arakoon cluster is corrupted.	Depending upon whether the Controller Nodes that host the Arakoon cluster were shut down gracefully or not, the Arakoon cluster may be impacted by some sort of corruption. The KB article ARA002 describes in detail how to recover from Arakoon corruption.
	When powering up any node, you may see messages in the system console or `/var/log/boot.log` similar to the following.
	`WARNING:root:Unable to connect to 192.168.108.2:9002 (error: '[Errno 113] No route to host')`
	`WARNING:root:Unable to connect to 192.168.109.2:9002 (error: '[Errno 113] No route to host')`
	`WARNING:root:Attempt 0 to exchange message with node node_1_9001 failed with error`
	`        (ArakoonNotConnected: 'No connection available to node at`

Problem	Recommended Action

	`['192.168.108.2', '192.168.109.2'] on port 9002').`
	`WARNING:root:Could not query node 'node_1_9001' to see who is master`
	`WARNING:root:Unable to connect to 192.168.108.3:9002 (error: '[Errno 113] No route to host')`
	`WARNING:root:Unable to connect to 192.168.109.3:9002 (error: '[Errno 113] No route to host')`
	`WARNING:root:Attempt 0 to exchange message with node node_3_9001 failed with error`
	`(ArakoonNotConnected: 'No connection available to node at`
	`['192.168.108.3', '192.168.109.3'] on port 9002').`
	`WARNING:root:Could not query node 'node_3_9001' to see who is master`
	`WARNING:root:Node 'node_0_9001' does not know who the master is`
	`WARNING:root:Node 'node_0_9001' does not know who the master is`
	`ERROR:root:Could not determine master.`
	As part of the power up, a connection is made to the `framework` Arakoon instance. This is to test the health of the Arakoon services. The timing in which Arakoon services come online and elect a master may cause some of these messages to appear in the console or boot log temporarily. The connection will retry for as many as 30 minutes before it times out.
	The following indicates that the powering system is not able to contact an arakoon service to see who the master is. No route to host means that system is not pingable. This may indicate that the system that is being contacted does not yet have network services started or that the network settings on that node are not correct.
	`WARNING:root:Unable to connect to 192.168.108.2:9002 (error: '[Errno 113] No route to host')`
	`WARNING:root:Unable to connect to 192.168.109.2:9002 (error: '[Errno 113] No route to host')`
	`WARNING:root:Attempt 0 to exchange message with node node_1_9001 failed with error`
	`(ArakoonNotConnected: 'No connection available to node at`
	`['192.168.108.2', '192.168.109.2'] on port 9002').`
	`WARNING:root:Could not query node 'node_1_9001' to see who is master`
	The following indicates that the powering system can contact an arakoon service but that service does not know who the master is. This generally means that two out of the three framework arakoons have not elected a master yet.
	`WARNING:root:Node 'node_0_9001' does not know who the master is`
	`WARNING:root:Node 'node_0_9001' does not know who the master is`
	`ERROR:root:Could not determine master.`

There is a problem with DSS.	If the client daemons on the Controller Nodes fail to start due to the `env_metastore` system MetaStore not having a master, you will see the following two error signatures.
	`boot.log` output:
	`***ERROR*** <type 'exceptions.Exception'>`
	`<type 'exceptions.Exception'> Client Daemon \\`
	`/opt/qbase3/cfg/dss/clientdaemons/e359600e-e773-44c1-bd53-21bf4a957f87.cfg`
	`could not be started: utils.execute: execution failed:`
	`command: ['/opt/qbase3/bin/dss', '-d', '--clientdaemon', '/opt/qbase3/cfg/dss/clientdaemons/e359600e-e773-44c1-bd53-21bf4a957f87.cfg']`
	`exit code: 1`

<table>
<tr><th>Problem</th><th>Recommended Action</th></tr>
<tr><td></td><td>

DSS client daemon log:

```
Jun  4 16:32:03.1648 warning [None] could not determine arakoon
 master via node
       node_3_9003 at 192.168.108.3:9004: Unix.Unix_error(Connection
 refused,connect,)


                          ...
Jun  4 16:32:48.1710 error [None] node server: failed: syncstore
 arakoon::env_metastore::
       node_0_9003:192.168.108.1:9004;192.168.109.1:9004,
       node_1_9003:192.168.109.2:9004;192.168.108.2:9004,
       node_3_9003:192.168.109.3:9004;192.168.108.3:9004:
       could not get deployment id: Failure: arakoon command timed
 out
       Fatal error: syncstore arakoon::env_metastore::
       node_0_9003:192.168.108.1:9004;192.168.109.1:9004,
       node_1_9003:192.168.109.2:9004;192.168.108.2:9004,
       node_3_9003:192.168.109.3:9004;192.168.108.3:9004:
       could not get deployment id: Failure: arakoon command timed
 out
```

If you see these types of signatures, first troubleshoot the `env_metastore` system MetaStore. Once `env_metastore` has been corrected, restart processes on that node using the Q-Shell commands to stop and start all services.

</td></tr>
<tr><td>The entire Active Archive System needs to be shut down gracefully.</td><td>

To shut down the entire Active Archive System, proceed as follows.

1. Log into the Management Node.
2. Exit the OSMI menu.
3. At the Linux prompt, do a test run of the `shutdown_environment.py` script to verify the order of nodes to be shut down:

   Note: This script does not run on a node that is not the Management Node.

   ```
   /opt/qbase3/bin/python  /opt/qbase3/utils/HGST/
   shutdown_environment.py
   ```
4. Run the `shutdown_environment.py` script with the `--shutdown` option:

   ```
   /opt/qbase3/bin/python  /opt/qbase3/utils/HGST/
   shutdown_environment.py
   ```

</td></tr>
<tr><td>The application server failed to start.</td><td>

After a power cycle, the Active Archive System does not automatically resume operations because pid files are lingering around. This is observed when there has been an improper shutdown (such as power failures).

When a Controller Node (and more specifically the Management Node) is power cycled (in other words, rebooted in an uncontrolled fashion), upon restart, some of the pid files (used to prevent starting multiple instances of the same process) are not cleaned up, preventing the restart.

Workaround: Identify the process that failed to start and to remove its pid file. In the case of the application server, restart it manually:

```
q.manage.applicationserver.restart()
```

</td></tr>
</table>

4 Managing Storage

Topics:

- Creating S3 Users
- Adding an S3 Client
- Creating S3 Buckets (Name Spaces)
- Deleting Buckets (Name Spaces)
- Deleting Objects
- Disabling Bucket Operations
- Configuring S3 Multipart Support
- Metering
- Using Encryption
- Managing MetaStores
- Troubleshooting Storage Issues

S3 is enabled by default, but there are still some administrative actions you must take in order to enable your users to communicate with the S3 interface.

4.1 Creating S3 Users

To create an S3 user, use the Q-Shell or the OSMI menu.

4.1.1 Creating S3 Users Through the Q-Shell

To create an S3 user through the Q-Shell, do the following:

1. Open an SSH session to the Management Node.
   The OSMI menu appears.
2. Exit the OSMI menu.
   The Linux prompt appears.
3. Start the Q-Shell by running the following command at the Linux prompt.

```
/opt/qbase3/qshell
```

4. In the Q-Shell, invoke `q.dss.manage.addUser()`:

```
q.dss.manage.addUser('login_name','password')
```

5. In the Q-Shell, invoke `q.dss.manage.setPermissions()` to define this user's S3 permissions:

```
q.dss.manage.setPermissions('/manage','login_name', \\
["READ","CREATE","DELETE","LIST","UPDATE"])
```

4.1.2 Creating S3 Users Through OSMI

To create an S3 user through OSMI, do the following:

1. Open an SSH session to the Management Node.
   The OSMI menu appears.
2. At the OSMI prompt, type `4` (for **Users and Permissions**).
3. At the next OSMI prompt, type `2` (for **Add User**).

4.1.3 Creating S3 Users Through S3 API

For instructions on creating users through the S3 API, see the *HGST Active Archive System API Guide*.

4.2 Adding an S3 Client

You need an S3 client in order to communicate with the Active Archive System using the S3 API.

Prerequisites

Some S3 clients are tied into a specific username. If that is the case for your S3 client, you must first create the S3 user with username *`login_name`* and password *`password`* by following the instructions in Creating S3 Users on page 26.

The following instructions explain how to add the `s3cmd` S3 client to a Linux client machine. For more information on `s3cmd`, see http://s3tools.org/s3cmd.

1. Install `s3cmd` on your client machine.
2. Configure `s3cmd` on your client machine as follows:

   You are configuring this instance of `s3cmd` for the specific S3 user you created with username *`login_name`* and password *`password`*.

   a) At the Linux prompt, run `s3cmd` with the `--configure` option to start the configuration wizard:

   ```
   s3cmd --configure
   ```

   For detailed information about the `s3cmd` configuration wizard, see http://s3tools.org/kb/item14.htm and http://knackforge.com/blog/sivaji/my-experience-s3cmd-utility.

   b) When the wizard prompts you to enter an `Access Key`, type the username of the S3 user, *`login_name`*

   c) When the wizard prompts you to enter an `Secret Key`, type the password of the S3 user, *`password`*

   d) When the wizard prompts you to enter an `Encryption password`, leave it blank.

   e) When the wizard prompts you to enter a `Path to GPG program`, leave it blank.

   f) When the wizard prompts you to enter an `HTTP Proxy server name`:

   - If the Active Archive System is configured for TLS/SLL, leave this blank.
   - Otherwise, type the virtual IP address of the Management Node (obtainable from the CMC).

   g) When the wizard prompts you to enter an `HTTP Proxy server port`:

   - If the Active Archive System is configured for TLS/SLL, leave this blank.
   - Otherwise, type the value of the **first** port number that the Active Archive System uses for S3 (obtainable from the CMC).

   h) When the wizard prompts you to `Test access with supplied credentials? [Y/n]`, type `n`.

   i) When the wizard prompts you to `Save settings? [y/N]`, type `y`.

   The configuration for this instance of `s3cmd` is saved in `~/.s3cfg`.
3. Manually edit the `host_base` and `host_bucket` settings in `~/.s3cfg` with the correct S3 domain name for this Active Archive System.

   Tip: Get the Active Archive System S3 domain name from the CMC.

   For example,

   ```
   host_base = s3.hgst.com
   host_bucket = %(bucket)s.s3.hgst.com
   ```

   A sample `~/.s3cfg` is shown below.

   ```
   [default]
   ```

```
access_key = login_name
secret_key = password

access_token =
add_content_encoding = True
add_encoding_exts =
add_headers =
bucket_location = US
cache_file =
cloudfront_host = cloudfront.amazonaws.com
default_mime_type = binary/octet-stream
delay_updates = False
delete_after = False
delete_after_fetch = False
delete_removed = False
dry_run = False
enable_multipart = True
encoding = UTF-8
encrypt = False
follow_symlinks = False
force = False
get_continue = False
gpg_command = /usr/bin/gpg
gpg_decrypt = %(gpg_command)s -d --verbose --no-use-agent --batch --yes --
passphrase-fd %(passphrase_fd)s -o %(output_f
                                            ile)s %(input_file)s
gpg_encrypt = %(gpg_command)s -c --verbose --no-use-agent --batch --yes --
passphrase-fd %(passphrase_fd)s -o %(output_f
                                            ile)s %(input_file)s
gpg_passphrase =
guess_mime_type = True

host_base = s3.hgst.com
host_bucket = %(bucket)s.s3.hgst.com

human_readable_sizes = False
ignore_failed_copy = False
invalidate_default_index_on_cf = False
invalidate_default_index_root_on_cf = True
invalidate_on_cf = False
list_md5 = False
log_target_prefix =
max_delete = -1
mime_type =
multipart_chunk_size_mb = 15
preserve_attrs = True
progress_meter = True
use_https = False

proxy_host = ManagementNodeVirtualIPAddress
proxy_port = 7070
```

4.3 Creating S3 Buckets (Name Spaces)

Note:

Use DNS compliant, globally unique bucket names that comply with following rules:

- A bucket name must be at least 3 and no more than 63 characters long.
- A bucket name must be a series of one or more labels separated by a period (.), where each label:

    - Must start with a lowercase letter or a number.
    - Must end with a lowercase letter or a number.
    - Can contain lowercase letters, numbers, and dashes.
- A bucket name must not be formatted as an IP address, for example `192.168.5.4`.

For more information, see http://docs.amazonwebservices.com/AmazonS3/latest/dev/BucketRestrictions.html.

To create an S3 bucket named *`bucketname`* for an S3 user named *`login_name`*, do the following:

1. On your client machine, use your S3 client to create a bucket named *`bucketname`* for the user *`login_name`*. If you are using `s3cmd`, creating a bucket would be done like this:

```
s3cmd mb s3://bucketname
```

2. On your client machine, add the following lines to the `/etc/hosts` file:

```
virtual_IP_of_management_node S3_domain_name
virtual_IP_of_management_node  bucketname.S3_domain_name
```

For example,

```
192.168.107.1 s3.hgst.com
192.168.107.1 mybucket.s3.hgst.com
```

where *`S3_domain_name`* is the S3 domain name of your Active Archive System (obtainable from the CMC) and is identical to the value you specified in your S3 client configuration file (for `s3cmd`, this is `~/.s3cfg`).

4.4 Deleting Buckets (Name Spaces)

4.4.1 Deleting Buckets Through S3 API

For instructions on deleting buckets through the S3 API, see the *HGST Active Archive System API Guide*.

4.4.2 Deleting Buckets Through the Q-Shell

You can delete buckets with the following Q-Shell command:

```
q.dss.manage.deleteNameSpace(self,nameSpaceName,nodeIP='127.0.0.1',port=23510)
```

This command has the following parameters:

Parameter	Explanation
`namespaceName`	The name of the bucket (name space) you want to delete.
`nodeIP`	The IP address of the node to communicate with. This does not have to be the IP of the storage daemon.
`port`	The port number on nodeIP to communicate with.

Note: This command has no `force` flag like the `deleteObject` Q-Shell command, because this may result in an unavailable Active Archive System.

The process of deleting name spaces goes as follows:

- **The name space itself**: the following actions are done immediately and synchronously in the MetaStore:

  - The `namespace_by_name` record is removed from the metadata. This action is logged in the MetaStore transaction logs.
  - The `namespace_by_id` record is set to "deleted".
- **Data belonging to that namespace**: the data belonging to the deleted name space is removed asynchronously:
  - Every 2 hours, every storage daemon will check if there are name spaces that:
    - are set to be deleted
    - still have data for the blockstore this storage daemon manages
  - If so, the storage daemon will remove the checkblock files on those blockstores.
  - Since a name space corresponds to a directory on the blockstore, a name space deletion corresponds to a directory removal (using the `rm -rf` command).
- **Metadata belonging to that name space**: the metadata belonging to the deleted name space is removed asynchronously:
  - The master storage daemon (this is the storage daemon responsible for that name space) deletes the metadata of the deleted name space. The metadata key and the corresponding value are deleted and a delete entry is written in the MetaStore transaction logs.
  - When completed, it will set the master storage daemon ID to "None" for that name space, so that it will no longer checks for delete tasks.

4.5 Deleting Objects

Note: After deleting a significant amount of data, you may see events in the CMC stating that servers are experiencing a high load average. These events persist until the delete operations have completed.

4.5.1 Deleting Objects Through S3 API

Requesting the deletion of an object through S3 uses the default method of deleting objects. You cannot use a `force` flag. As a result, all deletes are done asynchronously.

For instructions on deleting objects through the S3 API, see the *HGST Active Archive System API Guide*.

4.5.2 Deleting Objects Through the Q-Shell

Default Method

You can delete objects with the following Q-Shell command:

```
q.dss.client.deleteObject(self,nameSpaceName,objectName,force=False,\\
nodeIP='127.0.0.1',port=23510,timeout=None)
```

With the following attributes:

Attribute	**Explanation**
`nameSpaceName`	The name of the bucket (name space) the object is stored on.
`objectName`	The name of the object you want to delete.
`force`	The force flag, forcing an immediate deletion (default is `False`).
`nodeIP`	The IP address of the node to communicate with. This does not have to be the IP of the storage daemon.
`port`	The port number on `nodeIP` to communicate with.

Attribute	Explanation
timeout	The maximum time the command can take (default is no timeout).

The Active Archive System then deletes the object as follows:

1. A delete task is created and put in the delete queue of the name space the object belongs to.
2. The corresponding metadata of that object is immediately deleted.
3. They metadata key and the corresponding value are deleted. This implies that a pointer to the key and value are removed and that the data is garbage collected on the next occurrence of the Metastore Defragment policy. A delete entry is written in the MetaStore transaction logs.
4. Once every 24 hours the delete queue is processed by the storage daemon, responsible for that name space, as part of the repair crawl.
5. The repair crawl creates delete tasks, that are picked up by the maintenance agents.
6. The maintenance agent that picked up a delete task inspects the delete task, figures out on which storage daemon the encoded file segments are residing, and sends delete commands to those storage daemons.
7. These issue a file system unlink for the checkblock files.
8. Once the repair crawl has completed and all delete tasks are processed, the object is completely deleted.

Note: Improving the repair speed is not an easy task. You can follow the guide lines, described in Tuning for Optimal Repair Performance on page 105.

Specifying the Force Flag

If you do not want to wait for the repair crawl, you can delete the object using the force flag set to True:

```
q.dss.client.deleteObject(self,namespaceName,objectName,force=True,\\
nodeIP='127.0.0.1',port=23510,timeout=None)
```

In this case, all the checkblocks are immediately deleted.

Only if some of the blockstores are unavailable at the time of deletion, a delete task is put in the delete queue for the name space that the object belongs to.

4.6 Disabling Bucket Operations

The Active Archive System enables the LIST Bucket, PUT Bucket, and DELETE Bucket requests by default.

You can disable any of these requests by setting the following parameters in the [s3] section of the client daemon configuration file:

```
[s3]
enable_bucket_list = false # disables listing of buckets \\(does not disable listing of
 the content of buckets)
enable_bucket_create = false # disables creation of buckets
enable_bucket_delete = false # disables deletion of buckets
```

If these calls are disabled, the Active Archive System sends the following error message: action is disabled or not allowed on bucket, and clients receive the Exc.MethodNotAllowed exception.

4.7 Configuring S3 Multipart Support

S3 multipart support is enabled by default.

You can configure S3 multipart by editing the following parameters. These parameters are defined in the client daemon configuration file (`/opt/qbase3/cfg/dss/clientdaemons/`*`guid`*`.cfg`) in the section `[s3]`.

Parameter Name	Description	Unit	Default Value
`multipart_part_min_size`	Minimum part size	bytes	5,000,000 (5 MB)
`multipart_object_min_size`	Minimum object size	bytes	5,000,000 (5 MB)
`multipart_max_partnr`	Maximum number of parts		10,000 **Note:** The Active Archive System supports up to a maximum of 10,000 parts per object.

For example,

```
[s3]
multipart_part_min_size=5000000
multipart_object_min_size=5000000
multipart_max_partnr=10000
```

4.8 Metering

When metering is enabled, the client daemons keep track of all S3 requests that are issued to them.

You can use metering logs to do billing based on the number of requests and the bandwidth used by those requests.

4.8.1 Enabling Metering

To enable metering on all client daemons, proceed as follows:

1. Ensure that S3 is enabled.

   Note: While metering is able to work without S3, it does not log anything unless S3 is enabled.

2. Create a name space named `_metering_info`.
   For instructions on creating a name space, see Creating S3 Buckets (Name Spaces) on page 28.
3. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Interfaces** > **S3**.
4. On the **S3 Management** page, in the **Metering** section, click **Enable**.

4.8.2 Disabling Metering

To disable metering on all client daemons, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Interfaces** > **S3**.
2. On the **S3 Management** page, in the **Metering** section, click **Disable**.

4.8.3 Metering Log Details

File Names

When metering is enabled, the client daemon logs all traffic to the file *`epoch`*`.meter`, in CSV format, where *`epoch`* is the current epoch time used. For example, `/metering/1374485550`. Because of file naming convention, the client daemon can create new metering logs without any history.

Log Entries

Entries in the metering have the following format.

- S3 request entry:

```
s3_request;epoch_time;[namespace_or_bucket_id;]
action_type;http_method;[user_id;]result
```

- S3 transfer entry:

```
s3_transfer;epoch_time;[namespace_or_bucket_id;]action_type;
http_method;[user_id;]result;size_of_transfer_in_bytes
```

Log Entry Examples

An entry logging the creation of a bucket:

```
s3_request;1362666411;25b7c6a181154c70b61d7b22b27feda9;bucket;PUT;32;SUCCESS
```

An entry logging a PUT operation of an object of 3 bytes:

```
s3_transfer;1362666412;25b7c6a181154c70b61d7b22b27feda9;object;PUT;32;SUCCESS;3
```

An entry logging a failed GET operation:

```
s3_request;1362666443;481021ffd01843a9b1b8a1e589c093d6;object;GET;32;ERROR
```

An entry logging a failed HEAD operation:

```
s3_request;1362666444;481021ffd01843a9b1b8a1e589c093d6;object;HEAD;32;ERROR
```

4.8.4 Accessing Metering Logs

The metering logs for each client daemon are uploaded into separate directories identified by the GUID of the client daemon.

You can access each client daemon's metering logs through the following path:

```
 http://client_daemon_ip:s3_port/namespace/
_metering_info/client_daemon_GUID/metering_id.gz
```

Note: The Active Archive System rotates the metering logs every hour: a script (`logrotate`) triggers the rotation of the metering logs, compresses the metering logs into a gzip (.gz) file, uploads all gzipped files to the `_metering_info` name space, and removes all gzipped metering files from the file system.

4.9 Using Encryption

4.9.1 Supported RSA Ciphers

The Active Archive System supports all following ciphers for encryption in transit over HTTPS:

- Accepted TLSv1 256 bits DHE-RSA-AES256-SHA
- Accepted TLSv1 256 bits DHE-RSA-CAMELLIA256-SHA
- Accepted TLSv1 256 bits ADH-AES256-SHA
- Accepted TLSv1 256 bits ADH-CAMELLIA256-SHA
- Accepted TLSv1 256 bits AES256-SHA

- Accepted TLSv1 256 bits CAMELLIA256-SHA
- Accepted TLSv1 168 bits EDH-RSA-DES-CBC3-SHA
- Accepted TLSv1 168 bits ADH-DES-CBC3-SHA
- Accepted TLSv1 168 bits DES-CBC3-SHA
- Accepted TLSv1 128 bits DHE-RSA-AES128-SHA
- Accepted TLSv1 128 bits DHE-RSA-CAMELLIA128-SHA
- Accepted TLSv1 128 bits ADH-AES128-SHA
- Accepted TLSv1 128 bits ADH-CAMELLIA128-SHA
- Accepted TLSv1 128 bits AES128-SHA
- Accepted TLSv1 128 bits CAMELLIA128-SHA

The HTTPS functionality is provided by a proxy server, called `pound`. The following browsers support TLS1.1:

- Google Chrome 22 supports and runs TLS 1.1.
- Google Chrome 30 or higher supports TLS 1.1 or higher and is enabled by default
- Microsoft Internet Explorer Version 8 and higher supports TLS 1.1 and higher, but this function is disabled by default.
- Internet Explorer 11 supports TLS 1.1 or higher and is enabled by default.
- Mozilla Firefox 27 or higher supports TLS 1.1 or higher and is enabled by default. In earlier versions TLS 1.1 is disabled by default.
- Safari 7 supports TLS 1.1 or higher.
- Opera version 8-9 supports TLS 1.1. Opera Version 10 and higher support TLS 1.1 and higher.
- As of Opera 17, this function is enabled by default.

4.10 Managing MetaStores

4.10.1 Marking a MetaStore as Full

A MetaStore can reach its limits in two ways:

- The partition which hosts the MetaStore has no more free space.
- The number of keys in the MetaStore has reached its limit.

These situations should be avoided since it may lead to a corrupt MetaStore and eventually data loss. To protect the Active Archive System against corrupt MetaStores, you can mark a MetaStore as `FULL`, automatically upon the occurrence of certain events, or manually, through the CMC.

When a MetaStore is marked as `FULL`, it no longer accepts write operations, but it remains available for read, repair, and delete operations.

4.10.1.1 Automatically Marking a MetaStore as Full

When the following events occur, the Active Archive System automatically marks the associated MetaStore as `FULL`:

- OBS-ARAKOON-0011: MetaStore node database partition is full
- OBS-ARAKOON-0013: MetaStore node tlf partition is low on space
- OBS-ARAKOON-0020: Number of keys in MetaStore exceeds critical threshold

Depending on the event, you can take the necessary actions to free up disk space or delete keys from the database.

4.10.1.2 Automatically Reactivating a MetaStore

When there is again enough free disk space or the number of keys has dropped below the critical threshold, a new event, OBS-ARAKOON-0021, automatically reactivates the MetaStore in order to allow new write operations.

4.10.1.3 Manually Marking a MetaStore as Read Only

Besides the automatic marking of a MetaStore, you can set a MetaStore to `READONLY`, as a reaction on other less critical events, such as:

- OBS-ARAKOON-0010: MetaStore node database partitions is almost full
- OBS-ARAKOON-0012: MetaStore node tlf partition is low on space

The status `READONLY` has the same functionality as the status `FULL`, but `READONLY` is only set by human intervention and `FULL` is set by the Active Archive System monitoring agent.

To manually mark a MetaStore as read only, do the following.

1. In the CMC, navigate to **Dashboard** > **Administration** > **Storage Management** > **MetaStores**.
2. In the **Status** column of the proper MetaStore, select the option **READONLY** from the menu.

   You can also set the status by opening the details of the MetaStore and starting the **Edit** wizard.

4.10.1.4 Reactivating a MetaStore

When a MetaStore is to `READONLY`, you can reactivate it by setting its status back to `READ/WRITE` as follows.

> **Important:** If you have manually set the status of a MetaStore to `READONLY`, you must reactivate the MetaStore manually, even when the values (free disk space or number of keys) are no longer critical.

1. In the CMC, navigate to **Dashboard** > **Administration** > **Storage Management** > **MetaStores**.
2. In the **Status** column of the proper MetaStore, select the option **READ/WRITE** from the menu.

   You can also set the status by opening the details of the MetaStore and starting the **Edit** wizard.

4.10.2 MetaStore Recovery

4.10.2.1 Fully Automated tlog Collapse

The purpose of collapsing transaction logs (tlogs) is to limit the number of tlogs per MetaStore.

A tlog is a log file in which every entry contains the metadata-update along with control information. These tlogs (or `.tfl` files in compressed form) are used to replay in case the database terminates ungracefully.

If a database terminates ungracefully:

1. The database is moved aside (since internal pointers could have been misplaced);
2. A copy of a previously taken consistent database (called `head.db`) is used as a starting point; and
3. The tlog files are replayed onto this database.

At the end of this process, a running database exists that has the state as described in the tlogs.

The duration of this process is highly determined by the number of tlog files. Therefore, a periodic collapse operation is required to merge the tlog files into a new consistent database that can be used as a basis for recovery.

4.10.2.2 Fully Automated Recovery from Unclean Shutdown

In the case of an unclean shutdown, the Active Archive System management framework is capable of automatically restarting the MetaStores, except in the following cases:

- The management framework itself is not capable of restarting.
- There are issues with the file systems that hold the database and tlogs. In this case, manual intervention is needed to resolve these so that automated recovery can proceed.

4.10.2.3 Backups for Disaster Recovery and RCA Investigation

The Active Archive System stores compressed copies of the MetaStore and the associated tlogs in a dedicated name space, called `_metastorebackup`. Every time an Arakoon process terminates ungracefully, the Active Archive System uploads such a copy.

As part of the tlog collapse process, the Active Archive System uploads compressed copies of the `head.db` and the associated `tlf` / tlog files and retains up to five versions of the `head.db` files.

4.10.3 Automated Master Change

Some maintenance operations cannot be executed on a master and therefore need to be able to trigger a change of master so that the software can guarantee that all operations are executed on all nodes of the MetaStore cluster.

4.10.4 Batch Processing of MetaStore Transactions

The Active Archive System supports *batch processing*, meaning that you can push a number of transactions at once to the MetaStore.

The Active Archive System handles a number of updates as a single update, and defines a maximum batch size. As soon as the number of updates equals the maximum batch size, the quorum acknowledges and commits the changes. While the batch is being filled, the Active Archive System already starts exchanging data, so that it can commit the batch sooner if it reaches a quorum on the change set in the batch. As a result, the size of the individual updates grows larger, leading to larger tlogs. When the amount of data is too big (too many set and/or delete operations in one batch), it can slow down again the performance of the MetaStore because then the push of a batch would take too long.

The default number of operations in one batch is set to `196` and should not be changed in your setup.

4.10.5 Rolling Updates

MetaStores are implemented as an Arakoon distributed key-value store *cluster* running on all three Controller Nodes. Since the Arakoon protocol does not support the concept of versions, Arakoon is updated by a rolling update. This process is based upon the Paxos algorithm. For instructions on updating Arakoon, see the *HGST Active Archive System Upgrade Guide*.

4.11 Troubleshooting Storage Issues

This section provides troubleshooting tips for issues you might encounter during when managing the Active Archive System storage. For more troubleshooting tips, see the *HGST Active Archive System Troubleshooting Guide*.

4.11.1 s3cmd Errors

Error	Recommended Action
401	This error indicates a problem with the file, username, or password given to the `s3cmd` tool. Verify that the filename, username, and password given to `s3cmd` is correct, and re-run the command.
405	This error indicates that S3 bucket operations are disabled. On the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Interfaces** > **S3**. Enable the **Enable S3 bucket operations** check box. Click **Save** in the right pane.

4.11.2 Cyberduck Errors

Error	Recommended Action
Cannot make a bucket. Cyberduck returns "Interoperability error."	If you are using version 4.5+ of Cyberduck, the workaround is to set `s3.upload.expect-continue` to `false`. For more information, see https://trac.cyberduck.io/wiki/help/en/howto/preferences.

4.11.3 Spread Width, Safety Policies, and Storage Policies

4.11.3.1 Introducing Spread Width and Safety Keys

As part of the small file support, the key "spread" has been split into the following two keys:

- `wireblock_spread`: the spread width of the object
- `full_copy_spread`: the spread width of the full copy of a file when small file support is enabled

4.11.3.1 Spread Width and Safety Keys Explanation

Key	Explanation	Value (no small file support)	Value (small file support)
combined_spread_width	The combination of the spread_width of the wireblocks and the full copy	spread_width of storage policy	spread_width of storage policy
wireblock_spread_width	The previous spread_width. The number of blockstores the data is spread upon.	spread_width	spread_width - 1
full_copy_spread_width	Only applicable if the small file support is active. The blockstore where a full copy of the small file is made.	0	1
metadata_spread_width	The number of blockstores the metadata is spread upon (backup of Arakoon).	safety + 1	safety + 1
combined_safety	The maximum number of blockstores without which retrieval is still possible.	safety	safety
wireblock_safety	If no full copy is available, the maximum number of blockstore without which retrieval is still possible.	safety	safety - 1
full_copy_safety	If not enough blockstores with wireblocks are available to allow retrieval, the maximum number of blockstores without which retrieval is still possible.	Not applicable	0
metadata_safety	The safety of the metadata.	safety	safety

Note: The mentioned values are those in ideal circumstances (no lost blockstores).

4.11.3.2 Basic Examples

4.11.3.2 Example 1

- Spread width: 16
- Safety: 4

- Small file support active

Key	Value
combined_spread_width	16
wireblock_spread_width	15
full_copy_spread_width	1
metadata_spread_width	5
combined_safety	4
wireblock_safety	3
full_copy_safety	0
metadata_safety	4

In this case a 15/3 is stored and encoded. A full copy is stored on the 16th blockstore.

4.11.3.2 Example 2

- Spread width: 18
- Safety: 7
- No small file support active

Key	Value
combined_spread_width	18
wireblock_spread_width	18
full_copy_spread_width	0
metadata_spread_width	8
combined_safety	7
wireblock_safety	7
full_copy_safety	n/a
metadata_safety	7

In this case, a 18/7 is stored and encoded.

4.11.3.3 Storage Object Examples

4.11.3.3 Example 3: A Storage Object with a Few Lost Blockstores

Policy details:

- Spread width of 20
- Safety of 5
- 3 lost blockstores that were part of the wireblock spread width of its superblock
- Metadata spread has no lost blockstores
- No small file support activated

Key	Value
combined_spread_width	17
wireblock_spread_width	17
full_copy_spread_width	0

Key	Value
metadata_spread_width	6
combined_safety	2
wireblock_safety	2
full_copy_safety	n/a
metadata_safety	5

4.11.3.3 Example 4: A Storage Object with a Few Lost Blockstores and a Lost Full Copy

Policy details:

- Spread width of 16
- Safety of 4
- 2 lost blockstores that were part of the wireblock spread of its superblock
- 4 lost blockstores that were part of the metadata spread
- The blockstore containing the full copy of the first superblock is lost

Key	Value
combined_spread_width	13
wireblock_spread_width	13
full_copy_spread_width	0
metadata_spread_width	1
combined_safety	1
wireblock_safety	1
full_copy_safety	-1
metadata_safety	0

4.11.3.3 Example 5: A Storage Object with a Lost Full Copy

Policy details:

- Spread width of 18
- Safety of 7
- Small file support activated
- Blockstore containing the full copy of the first superblock is lost
- All other blockstores are available

Key	Value
combined_spread_width	17
wireblock_spread_width	17
full_copy_spread_width	0
metadata_spread_width	8
combined_safety	6
wireblock_safety	6
full_copy_safety	-1
metadata_safety	7

4.11.3.3 Example 6: A Storage Object with Negative Safety, but the Full Copy Available

Policy details:

- Spread width of 18
- Safety of 7
- Small file support activated
- 9 blockstores lost which were part of the wireblock spread of its superblock
- 7 of the lost blockstores were part of the metadata spread
- The blockstore containing the full copy still remains

Key	**Value**
combined_spread_width	18
wireblock_spread_width	17 - 9 = 6
full_copy_spread_width	1
metadata_spread_width	1
combined_safety	0
wireblock_safety	-3
full_copy_safety	0
metadata_safety	0

In this case, even with the high amount of blockstores that are lost, you still have a safety of 0, because the full copy still exists.

5 Managing Networks

Topics:

- The Active Archive System LANs
- Updating the Public Network Settings
- Troubleshooting Network Issues

5.1 The Active Archive System LANs

Three LAN segments are configured for each data center:

- One public LAN, used for user and/or application communication.
- One storage management LAN, used for storage and environmental communication
- One secondary storage LAN, used for storage only.
- A fourth LAN type (install LAN) is automatically created on the management private LAN segment.
- A fourth LAN type (IPMI LAN), for IPMI functionality.

5.2 Updating the Public Network Settings

You can use the OSMI menu to change the public LAN subnet or IP addresses assigned to the Controller Nodes.

> **Note:** To cancel the operation in the OSMI menu at any time, press `Ctrl+c`.

1. Log into any Controller Node over SSH.
   The OSMI menu appears.
2. In the OSMI menu, select **Environment** > **Update Public LAN settings**.
3. Enter the name of the public LAN.
   If you entered the correct name, the current settings of this LAN are displayed.
4. Change these settings by entering new values.

   > **Note:** If you do not want to change the LAN settings, but only to the IP addresses that are assigned to your Controller Nodes, simply re-enter the current values, then enter the new values for the IP addresses and virtual IP addresses of the Controller Nodes.

5. Re-enter the IP addresses for those Controller Nodes whose IP addresses you do not want to change.

5.3 Troubleshooting Network Issues

This section provides troubleshooting tips for issues you might encounter during when managing the Active Archive System networks. For more troubleshooting tips, see the *HGST Active Archive System Troubleshooting Guide*.

5.3.1 General

Problem	Recommended Action
The CMC displays a truncated view of the public and private IP addresses associated with any Controller or Storage Node. You cannot view the complete list of IP addresses from the **Controller Nodes** pane or from the any individual **Controller Node** panes.	To work around this problem, do the following: **1.** Open the Controller Nodes pane and click on the desired Controller Node icon. **2.** In the **Controller Node: <Node_Name>** screen, click the **Network Statistics** tab. The IP addresses for all NICs on this Controller Node are displayed.
Shutting down a Controller Node from the CMC fails when the primary private network is down, with a `no route to host` error.	When the primary private network (*private network #1* or *private network left*) is down, management actions, like those taken through the CMC, do not fail over to the secondary private network. Workaround: to shut down the node, log into the node using its private network #2 IP address, and run the following command from the Linux prompt: `shutdown -hy 0`
You cannot communicate with the Active Archive System. Instead, you see a network error (`No route to host`) in your client application.	You may have recently installed an unsupported SFP+ 1G module on a Controller Node, or recently replaced an SFP+ 1G module but connected it to the wrong port on the Controller Node. To fix this problem, obtain a replacement SFP+ 1G module from HGST Support, and follow the replacement procedure in the *HGST Active Archive System FRU Replacement Guide*.
A status 403 response was received on an S3 API call.	A status 403 response on an S3 API call may indicate that there is a time skew between the client system and the Controller Node that is larger than 15 minutes. To fix this problem, ensure that both systems are synchronized to a valid NTP server and try the request again. For more information, see http://docs.amazonwebservices.com/AmazonS3/latest/dev/RESTAuthentication.html.

Opening an SSH session to the new Management Node (after a failover) and then attempting to open an SSH session to a Storage Node, using its virtual IP address, fails.	The error message from the `ssh` command looks like this:
	`root@HGST-Alpha02-DC01-R02-CN01:~# ssh root@10.1.12.154`
	`@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@`
	`@    WARNING: REMOTE HOST IDENTIFICATION HAS CHANGED!     @`
	`@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@`
	`IT IS POSSIBLE THAT SOMEONE IS DOING SOMETHING NASTY!`
	`Someone could be eavesdropping on you right now (man-in-the-middle`
	` attack)!`
	`It is also possible that a host key has just been changed.`
	`The fingerprint for the ECDSA key sent by the remote host is`
	`8d:ad:d9:92:b2:11:18:b5:d9:1b:fc:82:94:6a:1f:35.`
	`Please contact your system administrator.`
	`Add correct host key in /root/.ssh/known_hosts to get rid of this`
	` message.`
	`Offending ECDSA key in /root/.ssh/known_hosts:18`

<table>
<thead>
<tr><th>Problem</th><th>Recommended Action</th></tr>
</thead>
<tbody>
<tr><td></td><td>
<pre>
  remove with: ssh-keygen -f "/root/.ssh/known_hosts" -R 10.1.12.154
ECDSA host key for 10.1.12.154 has changed and you have requested
 strict checking.
Host key verification failed.
</pre>
This error indicates that you need to remove the old ECDSA keys.
To remove old ECDSA keys, copy the exact command shown in the error message, and paste it at the Linux prompt. For example, in the sample error message above, you would paste the following command at the Linux prompt:
<pre>
 ssh-keygen -f "/root/.ssh/known_hosts" -R 10.1.12.154
</pre>
</td></tr>
<tr><td>There are many unnecessary services listening either on public interfaces.</td><td>
To disable unnecessary services from listening on public interfaces, proceed as follows. The following steps are intended only to be executed on Controller Nodes since they are the only public facing nodes.
First, create the following bash script:
<b>Note:</b> This script may cause performance issues when you have a combination of:
• a high number of threads (>= 228)
• existence of small objects (<= 4KB)
• no connection reuse by the client software in its interaction with the Active Archive System
If you run this script under this scenario, your client software may get HTTP 503 errors.
<pre>
#!/bin/bash
s3_axr_ports="7070,7071,7072,7073,7080,7081,7082,7083"
allow_tcp_ports="${s3_axr_ports},80,443,22"
allow_udp_ports="123"
# replace the following public interfaces/ips with the one from the
 actual system
public_interfaces=(eth0 172.31.24.120 eth5 10.0.0.120)
function firewall_interface {
interface=$1
ip=$2
echo iptables -A INPUT -m conntrack --ctstate ESTABLISHED,RELATED -j
 ACCEPT
-i $interface
echo iptables -A INPUT -m multiport -p tcp -d $ip --dport
 $allow_tcp_ports
-j ACCEPT -i $interface
echo iptables -A INPUT -m multiport -p udp -d $ip --dport
 $allow_udp_ports
-j ACCEPT -i $interface
# allow everything for outgoing traffic
echo iptables -A OUTPUT -j ACCEPT
}
for ((c=0; c<${#public_interfaces[@]}; c+=2))
do
interface=${public_interfaces[$c]};
ip=${public_interfaces[$c+1]};
echo "Firewalling ${interface} with ${ip}"
firewall_interface $interface $ip
</pre>
</td></tr>
</tbody>
</table>

Problem	Recommended Action

	`echo iptables -A INPUT -j REJECT -i $interface`
	`done`
	1. Save the above script as `update_fw_rules.sh` on all Controller Nodes.
	2. On each Controller Node, update the variables in the script:
	A. Set `s3_axr_ports` to all S3 and AXR TCP ports that are in use on the Controller Node.
	B. Set `public_interfaces` to the public NIC names and IP addresses.
	C. Make the script executable:
	`chmod +x update_fw_rules.sh`
	3. Increase the maximum number of entries in the `conntrack` table on all Controller Nodes:
	A. Create the file `nf-contrack.conf` in `/etc/modprobe.d`.
	B. Add the following line to this new file:
	`options nf_conntrack hashsize=524288`
	C. Increase the number of entries manually also, by executing the following command at the Linux prompt:
	`echo 524288 > /proc/sys/net/netfilter/nf_contrack_max`
	4. Execute the script on all Controller Nodes.
	5. Save the firewall rules on all Controller Nodes:
	`iptables-save -c > /etc/iptables.rules`
	6. Make sure that the firewall rules are persistent through reboots.
	On all Controller Nodes, do the following:
	A. Create `/etc/network/if-post-down.d/iptablesload` with the following content:
	`#!/bin/sh`
	`if [ -f /etc/iptables.rules ]; then`
	`iptables-restore < /etc/iptables.rules`
	`fi`
	`exit 0`
	B. Create `/etc/network/if-pre-up.d/iptablessave` with the following content:
	`#!/bin/sh`
	`iptables-save -c > /etc/iptables.rules`
	`if [ -f /etc/iptables.downrules ]; then`
	`iptables-restore < /etc/iptables.downrules`
	`fi`
	`exit 0`
	C. Make both scripts executable:
	`chmod +x /etc/network/if-post-down.d/iptablessave`
	`chmod +x /etc/network/if-pre-up.d/iptablesload`
	Important: If the Management Node fails over:
	1. Update the public virtual IP address.

Problem	Recommended Action
	2. Clean the firewall rules by running the following command on all Controller Nodes (including the Management Node): `iptables -F` **3.** Execute the 6 steps above again.

6 Managing Hardware

Topics:

- Using the Intelligent Platform Management Interface (IPMI)
- Handling Blacklists
- Handling a Degraded Disk Notification
- Handling a Read-Only File System
- Handling Node Failures
- Managing Unmanaged Disks
- Handling Unverified Objects
- Troubleshooting Hardware Issues

6.1 Using the Intelligent Platform Management Interface (IPMI)

Active Archive System IPMI Setup

In order to work with IPMI:

- Understand how the Storage and Controller Nodes are connected to the IPMI LAN.
- Understand how the IPMI LAN is configured and how to allocate IP addresses for IPMI purposes.

Finding the IPMI Address of a Node

To see the IPMI IP address of a node, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Servers**.
2. Click **Controller Node** or **Storage Node**.
3. On the **Summary** tab, in the **General** pane, the IPMI IP address of the node is displayed.

Connecting to IPMI

You can connect to IPMI in any of the following ways:

- Connect to the Management Node's virtual IP address using SSH.
- Connect over the public network (only if the public network is secure).

Connecting to the Management Network Switch

1. Connect to the management network.
2. Use an unused IP address from the IPMI LAN range (only first three addresses are allowed).
3. Open a web browser and navigate to the following address: `http://`*`IPMI_IP_address_of_node`*
4. Log into the IPMI console with the credentials (the default user name and password are both `ADMIN`).
5. Execute the desired actions.
6. Once finished, log out of the console.
7. Disconnect the network cable.

Connecting Over the Public Network

Note: This method is only allowed in setups where the public network is secure.

1. Log in to one of the Controller Nodes using secure shell (SSH).
2. Tunnel different IPMI ports to publicly accessible ports by executing the following on your local machine:

   ```
   sudo ssh root@<public_IP_address_of_controller_node \\
   -L 80:IPMI_address_of_node:80 -L 7578:IPMI_address_of_node:7578 -L
    443:IPMI_address_of_node:443
   ```

   The setup with PuTTY (on Windows) is similar .
3. Open a web browser and navigate to: http://localhost:80.
4. Log in to the IPMI console with the credentials (default user name and password are both: admin).
5. Execute the desired actions.
6. Once finished, log out of the console.
7. Stop the SSH session and log out of the controller.

> **Note:** Issues with IPMI need to be investigated using a real keyboard and monitor.

6.1.1 Remote Power Cycle

To toggle the power off and on, proceed as follows:

1. Connect to IPMI.

   For information on connecting to IPMI, see Using the Intelligent Platform Management Interface (IPMI) on page 46.
2. In the IPMI console, navigate to **Remote Control** > **Power Control**.
3. In the **Power Control and Status** window, select **Power Cycle Server**.
4. Click **Perform Action**.

6.1.2 Remote Capture of Screen and Keyboard

To capture the screen and keyboard of the node, proceed as follows:

1. Connect to IPMI.

   For information on connecting to IPMI, see Using the Intelligent Platform Management Interface (IPMI) on page 46.
2. In the IPMI console, navigate to **Remote Control** > **Console Redirection**.
3. In the remote control window, click on **Java Console**.
   A Java viewer launches, emulating the screen and keyboard of the node.

6.1.3 Viewing Sensor Readings

To view the sensor readings of the node, proceed as follows:

1. Connect to IPMI.

   For information on connecting to IPMI, see Using the Intelligent Platform Management Interface (IPMI) on page 46.
2. In the IPMI console, navigate to **Server Health** > **Sensor Readings**.

6.1.4 Toggling a Location LED

The location LED helps you to easily retrieve a node in a data center. The location LED is supported by all Storage Nodes and the following Controller Nodes:

- HGST_CN

1. Connect to IPMI.

For information on connecting to IPMI, see Using the Intelligent Platform Management Interface (IPMI) on page 46.

2. In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Servers** > **Storage Nodes**.
3. Select the desired node.
4. In the **Commands** pane, click **Location LED On** or **Location LED Off**.

6.2 Handling Blacklists

When a node is blacklisted, there is a process that periodically "un-blacklists" that node if the blacklisted node starts to respond again. However, in some cases, the node still does not operate as expected, even if it accepts basic connections again; then an endless loop starts:

- The node first gets blacklisted because is not performing operations as it should.
- Then the node gets un-blacklisted because it is responding to a basic connection.
- Then it gets blacklisted again.
- Then it gets un-blacklisted again.
- ...

The end result is a lot of wasted effort and slow uploads as the system tries to use the bad blockstore time and time again.

To avoid this scenario, there is a self-check mechanism called the *blockstore liveness checker*. This mechanism determines whether a blockstore can be taken off the blacklist as follows:

1. The blockstore liveness checker tries to make a basic connection to the blacklisted blockstore once each thirty seconds.
2. If a connection can be made, the blockstore liveness checker asks the blockstore whether or not it can perform basic read and write operation on its backing store.
3. If either the read or the write operation fails (or both), then the blockstore remains on the blacklist.
4. If the blockstore can perform both read and write operations, it is taken off the blacklist.

When you have blacklists, do the following:

1. First, check if there is a disk related issue.
   a) In the CMC, navigate to **Dashboard** > **Administration** > **Storage Management** > **Storage Services**.
   b) Click **storagepool**.
   c) Click the affected Storage Node.
   d) Select a **dssstoragedaemon**.
   e) Look at the graphics in the right corner to see the blacklists.
   f) To get the exact number of blacklists issued for a specific blockstore, click on a blockstore.
   The number of blacklists on that blockstore is displayed.
   g) Click **Physical Details** to see the device name of the disk that hosts the blockstore.
   h) Compare the blacklists for the degraded disk with other disks in the same node:
      - If that one disk has 100% more blacklists than any other disk in the same node, decommission the disk.
      - If not, run further checks.
2. If there is no disk issue, you most likely are encountering network issues. Check your network.

6.3 Handling a Degraded Disk Notification

A disk becomes *degraded* when:

- One of the file systems on the disk encounters an I/O error.
- The kernel issues an I/O error for the disk.

A degraded disk generally can continue to be used, but its continued use may have a negative impact on performance or disk safety.

As soon as a disk is considered degraded, the disk is added to the *degraded disks list*. You can receive notifications about degraded disks through SNMP and through the Phone Home function.

To handle such notifications, proceed as follows:

1. Check the degraded disks list:

   In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Disks** > **Degraded**.
2. Determine the next step:
   - If only one disk on a Storage Node is degraded, see Troubleshooting a Degraded Disk on a Storage Node on page 49.
   - If multiple disks are degraded on the same Storage Node and around the same time, see Troubleshooting Multiple Degraded Disks on a Storage Node on page 49. When the disks have become degraded with large intervals, then consider the issue as one disk in a Storage Node.
   - If there are degraded disk(s) on a Controller Node, see Troubleshooting Degraded Disks on a Controller Node on page 51.
   - For all other cases, see Degraded Disk Troubleshooting Flowchart on page 53.

6.3.1 Troubleshooting a Degraded Disk on a Storage Enclosure Basic

If you have a degraded disk on a Storage Enclosure Basic, check the disk as follows:

Run diagnostics to check the health of the disk and its I/O Module.

6.3.2 Troubleshooting a Degraded Disk on a Storage Node

If you have one degraded disk on one Storage Node, check the disk as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Disks** > **Degraded**.
2. In the **Degraded Disks** window, click the degraded disk.
3. In the **Commands** pane, click **Diagnose**.
4. Select both the **SMART test** and **Read/Write test** check boxes.
5. Check the test results.

   a) Both test results are successful:

      a. Take note of the number of the disk.

      b. If this is the first time the disk was degraded, click **Reset** in the **Commands** pane.The disk disappears from the degraded disks list.

      c. If this is **not** the first time the disk was degraded, perform a detailed SMART analysis (see Performing a Detailed SMART Analysis on page 52).

   b) If the SMART test failed or timed out, perform a detailed SMART analysis (see Performing a Detailed SMART Analysis on page 52).

   c) If the read/write test failed, check the results of the diagnostics at the bottom of the page or check the event list by navigating to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Logging** > **Events**.

   d) If both tests failed because the operation timed out (it took more than the expected 30 seconds), perform a detailed SMART analysis on that disk (see Performing a Detailed SMART Analysis on page 52).

   e) If there is an event pointing to a read-only partition, see Handling a Read-Only File System on page 56

   f) If there are no such events, perform a detailed SMART analysis.

6.3.3 Troubleshooting Multiple Degraded Disks on a Storage Node

If you have multiple degraded disks on the same Storage Node, proceed as follows:

Note: This procedure only need to be followed when you have multiple disks that are degraded around the *same* time, because this may indicate to a faulty disk controller.

If the interval between the degrading of disks is too large, then follow the procedure for one disk in a storage node per degraded disk.

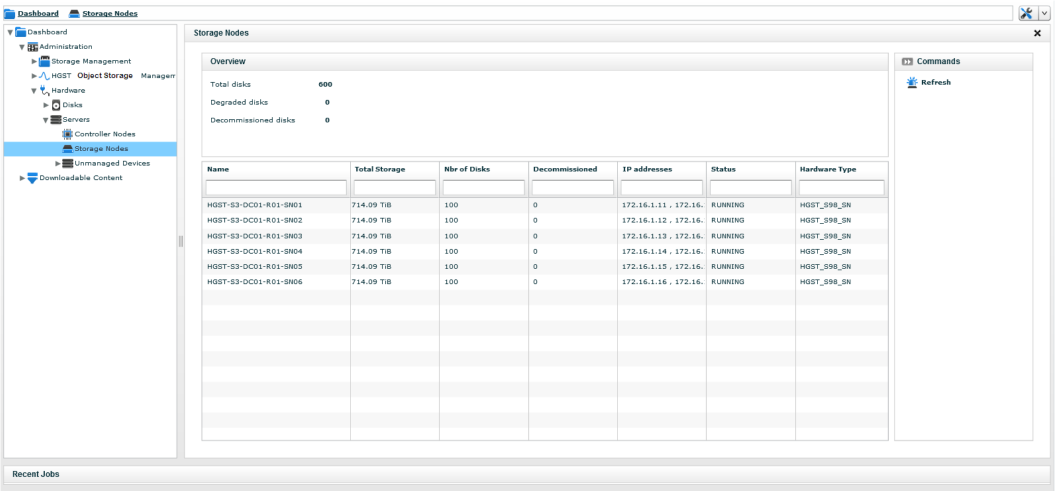
1. Shut down the Storage Node from the CMC.

Caution: Shut down **only** the Storage Node that is paired with the Storage Enclosure Basic containing the FRU.

a) In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Servers** > **Storage Nodes**.

b) Select the desired Storage Node.

Figure 2: A Storage Node Pane in the CMC

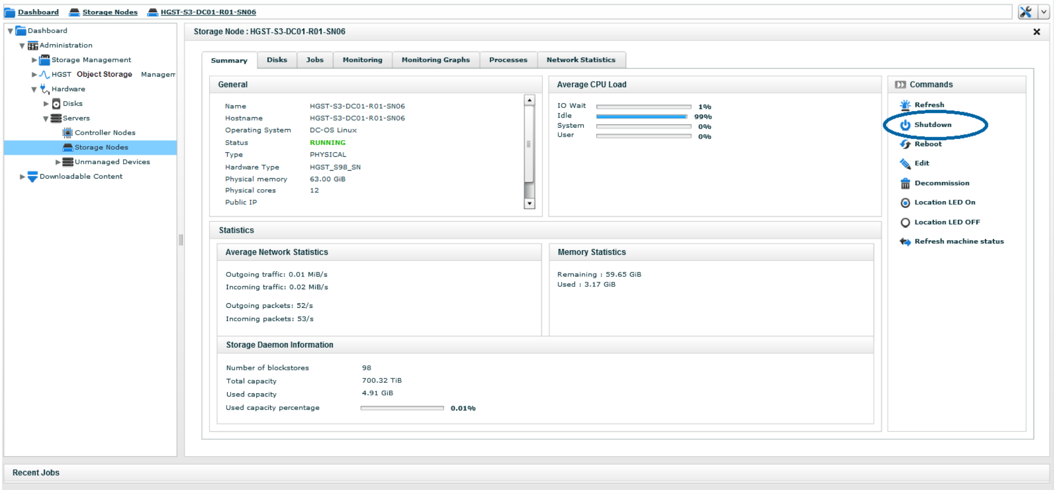


Total disks 600
Degraded disks 0
Decommissioned disks 0

Name	Total Storage	Nbr of Disks	Decommissioned	IP addresses	Status	Hardware Type
HGST-S3-DC01-R01-SN01	714.09 TiB	100	0	172.16.1.11 , 172.16.	RUNNING	HGST_S98_SN
HGST-S3-DC01-R01-SN02	714.09 TiB	100	0	172.16.1.12 , 172.16.	RUNNING	HGST_S98_SN
HGST-S3-DC01-R01-SN03	714.09 TiB	100	0	172.16.1.13 , 172.16.	RUNNING	HGST_S98_SN
HGST-S3-DC01-R01-SN04	714.09 TiB	100	0	172.16.1.14 , 172.16.	RUNNING	HGST_S98_SN
HGST-S3-DC01-R01-SN05	714.09 TiB	100	0	172.16.1.15 , 172.16.	RUNNING	HGST_S98_SN
HGST-S3-DC01-R01-SN06	714.09 TiB	100	0	172.16.1.16 , 172.16.	RUNNING	HGST_S98_SN

c) In the **Commands** pane, click **Shutdown**.

Figure 3: The Shutdown Button in the Commands Pane



Name HGST-S3-DC01-R01-SN06
Hostname HGST-S3-DC01-R01-SN06
Operating System DC-OS Linux
Status RUNNING
Type PHYSICAL
Hardware Type HGST_S98_SN
Physical memory 63.00 GiB
Physical cores 12
Public IP

IO Wait 1%
Idle 99%
System 0%
User 0%

Outgoing traffic: 0.01 MiB/s
Incoming traffic: 0.02 MiB/s

Outgoing packets: 52/s
Incoming packets: 53/s

Remaining : 59.65 GiB
Used : 3.17 GiB

Number of blockstores 98
Total capacity 700.32 TiB
Used capacity 4.91 GiB
Used capacity percentage 0.01%

d) Wait for the **Status** field to change to **DONE**.

Warning: Even if all LEDs are off, you must still wait until the CMC shows **DONE** in the **Status** field.

All I/O to the Storage Enclosure Basic attached to this Storage Node is now quiesced.

2. In the CMC, navigate to: **Dashboard** > **Administration** > **Hardware** > **Disks** > **Degraded**.
3. From the **Degraded disks list**, select all degraded disks.
4. In the **Commands** pane, click **Reset**.
   After the reset is complete, the disks disappear from the degraded disks list.
5. Power on the Storage Node.
6. When the Storage Node is successfully started, log into the CMC, and navigate to: **Dashboard** > **Administration** > **Hardware** > **Disks** > **Degraded**.
7. If the disks reappear in the degraded disks list or when you receive new events about degraded disks, proceed as follows:

   > **Note:** It may take several minutes before the disks are again indicated as degraded or when an event is raised.

   a) Power off the Storage Node again.
   b) Send a support team to check the connected cables on the disks.
   c) Replace and/or reconnect the cables to the degraded disks.
   d) When the intervention is completed, power on the Storage Node again.
   e) Repeat steps 2-5.
8. If the disks still did not disappear from the degraded disks list, proceed as follows:
   a) Power off the Storage Node again.
   b) Send a support team to replace the motherboard and the host adapter.
   c) When the intervention is completed, power on the Storage Node again.
   d) Repeat steps 2-5.

   The disks disappear from the degraded disks list.

6.3.4 Troubleshooting Degraded Disks on a Controller Node

Both Storage Nodes and Controller Nodes may have degraded disks over time. For information on replacing degraded disks, see the *HGST Active Archive System FRU Replacement Guide*.

- If a Controller Node has a degraded disk that is **not a boot disk**, follow the instructions of the following sections.
  1. Troubleshooting a Degraded Disk on a Storage Node on page 49
  2. Performing a Detailed SMART Analysis on page 52
  3. Further Checks on page 54.
- If the **boot disks** of your Controller Node are degraded, continue as follows:
  1. Power off the Controller Node.
  2. Replace the first boot disk with a new disk and verify that the Controller Node can boot from the second original boot disk.

     If replacing the first boot disk does not work, reinstall the first original boot disk and replace the second original boot disk with the new disk.

     Verify that the Controller Node can boot from the first original boot disk.

     If the Controller Node cannot boot from the first disk (with new second boot disk), power off the Controller Node and replace both boot disks with new ones.
  3. Follow the procedure as described in KB Article SCS018 to reinstall the Controller Node.

     > **Warning:** This is a very complex procedure which should only be executed with assistance of HGST Support.

6.3.5 Performing a Detailed SMART Analysis

To perform a Self-Monitoring, Analysis and Reporting Technology (SMART) analysis on a degraded disk, proceed as follows:

1. Establish an SSH session to the node associated with the degraded disk.
   - If the disk is on a Controller Node, open an SSH session to that Controller Node.
   - If the disk is on a Storage Node or a Storage Enclosure Basic, open an SSH session to the Storage Node as follows:

     A. Find the IP address of the Storage Node in the CMC.
     B. Open an SSH session to the Management Node. The OSMI menu appears.
     C. Exit the OSMI menu. The Linux prompt appears.
     D. Open an SSH session to the target Storage Node. The OSMI menu appears.
     E. Exit the OSMI menu. The Linux prompt appears.
2. Run a SMART analysis by issuing the following command:

   ```
   smartctl -x /dev/device_name
   ```

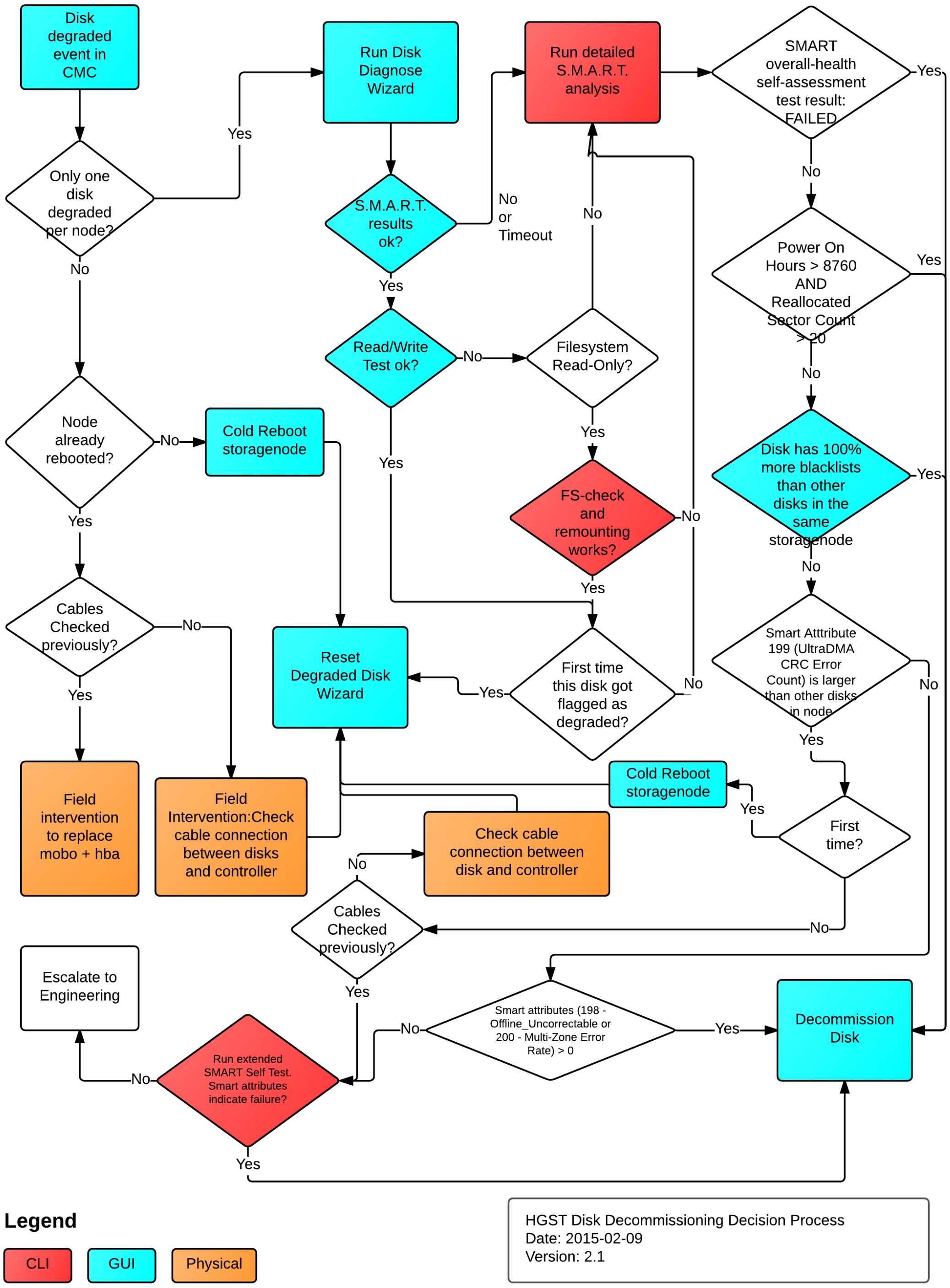
   Caution: Events contain references to partitions and not necessarily to devices (hard disks). For example, `/dev/sdf1` is a partition on the blockdevice `/dev/sdf`.
3. In the resulting console output, check the status of the SMART overall-health self-assessment test result:
   - `Failed`: Decommission the disk.
   - `Passed`: Check the attributes `Reallocated_Sector_Ct` and `Power_On_Hours`:
     - If `Reallocated_Sector_Ct` > 20 *and* `Power_On_Hours` <= 8760, decommission the disk.
     - If these two conditions are not met, check the blacklist of the Storage Node.

   Tip: You can find individual attributes with the following commands:

   - `smartctl -x /dev/sda | grep health`
   - `smartctl -x /dev/sda | grep Reallocated_Sector`
   - `smartctl -x /dev/sda | grep Power_On`

6.3.6 Degraded Disk Troubleshooting Flowchart

Figure 4: Degraded Disk Flowchart


Disk degraded event in CMC
Only one disk degraded per node?
Yes
No
Run Disk Diagnose Wizard
S.M.A.R.T. results ok?
No or Timeout
Yes
Run detailed S.M.A.R.T. analysis
SMART overall-health self-assessment test result: FAILED
Yes
No
Power On Hours > 8760 AND Reallocated Sector Count > 20
Yes
No
Read/Write Test ok?
No
Yes
Filesystem Read-Only?
No
Yes
FS-check and remounting works?
No
Yes
Node already rebooted?
No
Yes
Cold Reboot storagenode
Disk has 100% more blacklists than other disks in the same storagenode
Yes
No
Cables Checked previously?
No
Yes
Reset Degraded Disk Wizard
First time this disk got flagged as degraded?
Yes
No
Smart Atttribute 199 (UltraDMA CRC Error Count) is larger than other disks in node
No
Yes
Cold Reboot storagenode
Yes
First time?
No
Field intervention to replace mobo + hba
Field Intervention:Check cable connection between disks and controller
Check cable connection between disk and controller
No
Cables Checked previously?
Yes
Escalate to Engineering
Smart attributes (198 - Offline_Uncorrectable or 200 - Multi-Zone Error Rate) > 0
No
Yes
Decommission Disk
Run extended SMART Self Test. Smart attributes indicate failure?
No
Yes
Legend
CLI
GUI
Physical
HGST Disk Decommissioning Decision Process
Date: 2015-02-09
Version: 2.1

For instructions on replacing degraded disks, see the *HGST Active Archive System FRU Replacement Guide*.

6.3.7 Reviewing the Number of Blacklists of a Degraded Disk

To review and compare the number of blacklists, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **Storage Management** > **Storage Services**.
2. Click **storagepool**.
3. Click the affected Storage Node.
4. Select a **dssstoragedaemon**.
5. Look at the graphics in the right corner to see the blacklists.
6. To get the exact number of blacklists issued for a specific blockstore, click on a blockstore.
   This shows you the exact amount of blacklists.
7. Click **Physical Details** to see the device name of the disk that hosts the blockstore.
8. Compare the blacklists for the degraded disk with other disks in the same node:
   - If that one disk has 100% more blacklists than any other disk in the same node, decommission the disk.
   - If not, run further checks.

6.3.8 Further Checks

If the previous checks did not give you a conclusion, proceed with these following checks:

> **Caution:** Events contain references to partitions and not necessarily to devices (hard disks). For example, `/dev/sdf1` is a partition on the block device `/dev/sdf`.

1. In the detailed SMART analysis, check the SMART attribute #199 (`UDMA_CRC_Error_Count`), or use the following command:

   ```
   smartctl -x /dev/device_name | grep UDMA
   ```
2. If `UDMA_CRC_Error_Count` is greater than any other disk in the same Storage Node and this is the first time the disk was degraded, click **Reset** in the **Degraded Disk** window in the CMC.
3. If `UDMA_CRC_Error_Count` is greater than any other disk in the same Storage Node and this is not the first time the disk was degraded, proceed as follows:

   A. Shut down the Storage Node through the CMC.
   B. Send a support team to check the cable connection between the disk and the Storage Node or Storage Enclosure Basic.
   C. Replace or reconnect the cable if necessary.
   D. When the intervention is completed, power on the Storage Node.
   E. When the Storage Node has started, connect to the CMC.
   F. In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Disks** > **Degraded**.
   G. In the degraded disks list, select the degraded disk.
   H. In the **Commands** pane, click **Reset**. The disk should disappear from the list.
   I. If the disk does not disappear, proceed with step 6.
4. If `UDMA_CRC_Error_Count` is not greater than other disks, or if the reset did not remove the disk from the degraded disks list, check the following SMART attributes:
   - Attribute #198: `Offline_Uncorrectable`.
   - Attribute #200: `Multi_Zone Error`.

   > **Tip:** You can also use the following commands to get these attributes:

   - ◆ `smartctl -x /dev/device_name | grep Offline_Uncorrectable`
   - ◆ `smartctl -x /dev/device_name | grep Multi_Zone`
5. If any of these attributes are **above 0**, decommission the disk.

6. Otherwise, perform an extended SMART self-test with the following command:

   ```
   smartctl -t long /dev/device_name
   ```

   Note: This command will take a long time to complete.

7. Check the result and the following SMART attributes:

   - Test result must be: `SUCCESS`.
   - Check attribute #9 (`Power_On_Hours`) and attribute #5 (`Reallocated_Sector_Ct`). If the node is on for **8760 hours or below***, the reallocated sectors must be **below 20**.
   - Attribute #199 (`UDMA_CRC_Error_Count`) may **not** be **higher than other disks** in the node.
   - Attribute #198 (`Offline_Uncorrectable`) must be **0**.
   - Attribute #200 (`Multi_Zone Error`) must be **0**.

8. If the test or any of these SMART attributes indicate failure, decommission the disk.
9. If not, contact HGST Support.

6.3.9 Decommissioning a Degraded Disk

The Active Archive System has a policy that can decommission disks automatically . However, you can also decommission disks manually. If you want to manually decommission disks, you have to be careful. It is possible that the automatic decommissioning policy can start after your manual decommission of a disk, possibly leading to data loss. Therefore, you have to take the following precautions when you want to decommission a disk manually.

1. Disable the **Auto decommission disks** policy. For more information about disabling this policy, see Disabling the Auto Decommission Policy on page 68.
2. Verify the disk safety in the CMC and its "Last Update". Validate that the disk safety allows decommissioning a disk and that the relevance of the disk safety is recent enough:

   - If the disk safety is older than 12 hours, wait 12 hours for an updated status of the disk safety or execute a manual real-time **monitor name space** action.

3. Check the number of disks that are currently being decommissioned via the **Auto decommission disks** policy:

   - Navigate to **Dashboard** > **Administration** > **Hardware** > **Disks** > **Degraded**.
   - At the bottom of the **Degraded Disks** pane, click **Decommissioning and Autodecommissioning Disks**.
   - The number of disks that have the status **AUTODECOMMISSIONING** are disks that are currently being repaired. It is possible that the disks in **AUTODECOMMISSIONING** state are not yet reflected in the current disk safety.

     The event OBS-DISK-0020 is sent for each disk that is decommissioned with the auto-decommission policy. This event is sent by the Controller Node that hosts the CMC.

4. Verify that there are no recent manual decommissioning jobs initiated.

When you have taken the precautions, you can start the decommission of the disk as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Disks** > **Degraded**.
2. Select the degraded disk that you want to decommission.
3. In the **Commands** pane, click **Decommission**.
4. Choose whether or not you want to **erase the disk** (secure erase following NIST SP 800-88).

   Note: Erasing the disk can take several hours, depending on the amount of data on the disk.

   The **erase the disk** option erases the disk, using Enhanced Secure Erase on page 56.

   Once the wizard completes, the disk is automatically moved to the **Decommissioned disk** section of the CMC dashboard. When the disk gains the status "decommissioned", a repair crawl starts for all name spaces.

> **Note:** The repair crawl is an iteration of all objects in a name space. The crawl causes repair tasks to be created and put into a queue for the maintenance agents to reserve and execute.
>
> The crawl and repair queue are managed by a storage daemon on a Storage Node. This ensures that the Active Archive System starts repair activities for objects that were affected by the decommission immediately, rather than waiting until the next 24-hour crawl begins.

5. After the manual decommission of a disk, wait until the disk safety is back to normal before enabling the **Auto decommission disks** policy again.

You must replace decommissioned disks. For instructions on replacing disks, see the *HGST Active Archive System FRU Replacement Guide*.

6.3.10 Enhanced Secure Erase

While a normal secure erase overwrites the user data with zeros, the enhanced secure erase writes predetermined data patterns. The patterns are put on all user data areas, including sectors that are no longer in use due to reallocation. The data patterns are defined by the disk manufacturer.

On successful completion of the enhanced secure erase, this command disables security (in other words, returns the device to security state SEC1), and invalidates any existing user password. Any previously valid master password and master password identifier remains valid.

The enhanced secure erase of a disk consists of these two steps:

1. Enable security on the disk.
2. Erase disk using ATA Secure Erase

6.4 Handling a Read-Only File System

A file system can become read-only for several reasons:

- The partition has filled up and the framework has set it to read-only to prevent it from increasing.
- An error has occurred on the file system and the OS set it to read-only.
- There is a faulty disk.

If you see an event indicating that there is a read-only partition, similar to the following message:

```
Read-only partition '/dev/device_name' [machine 'node_hostname'] detected
```

To handle a read-only file system:

1. Retrieve the IP address of the affected node through the CMC.
2. Restart the node and verify that the issue is resolved.

   If the node is the Management Node itself, the CMC becomes unavailable temporarily.

PostgreSQL Disk

If the restart of the node did not resolve the issue:

1. Establish an SSH connection to the node and leave the OSMI menu if applicable.
2. Verify if the disk is used for PostgreSQL, this is mounted on `/mnt/postgresql`: `cat /proc/mounts`
3. If the disk is used by PostgreSQL:
   a) Execute a failover of the Management Node.
   b) Proceed with the next section.
4. If the disk is not used by PostgreSQL, then continue with the next section.

Non-PostgreSQL Disk

If the disk is not used by PostgreSQL, first decommission the disk and then replace the disk. For instructions on replacing the disk, see the *HGST Active Archive System FRU Replacement Guide*.

6.5 Handling Node Failures

A node failure occurs when it has a failed motherboard. If a node fails, you must replace it entirely. You must decommission a node before replacing it.

Note: If the node to be replaced is already halted, follow the replacement instructions for the node in the *HGST Active Archive System FRU Replacement Guide*.

6.5.1 Decommissioning a Node in the CMC

Prerequisites

- If the node to be decommissioned is the Management Node, first perform a failover. For more information, see Executing a Failover on page 58.

  Note: If the node to be replaced is already halted, follow the replacement instructions for the node in the *HGST Active Archive System FRU Replacement Guide*.

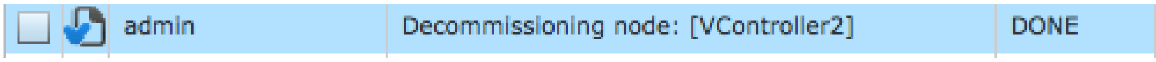
- Make sure that you have run the policy **Backup qpackageserver to other node** successfully. If this action has not succeeded, contact HGST Support.
- Take note of the network information for the node to be replaced, especially if the new node will use the same network information as the original node.
- Take note of any modifications that have been made to any configuration files on the node(for example Arakoon or any BitSpread component). The new node is installed and configured with default settings, so these modifications need to be applied when it is fully operational.

Warning: Having two Controller Nodes decommissioned at the same time could lead to interruption in management functionality and data unavailability until one of the Controller Nodes is replaced. Do not decommission any Controller Node if there are existing decommissioned Controller Nodes that have not yet been completely replaced.

To decommission a node, do the following:

1. Log into the CMC.
2. Navigate to **Dashboard** > **Administration** > **Hardware** > **Servers**.
3. Choose either **Controller Nodes** or **Storage Nodes**, depending on which type of node that you are decommissioning.
4. Click on the node to be decommissioned.
5. In the right pane, click Decommission and confirm. A job progress window appears with the progress of the decommission job.
6. Wait until this decommission job has the status **DONE**.

Figure 5: Decommissioning Status



□		admin	Decommissioning node: [VController2]	DONE

7. Navigate to **Dashboard** > **Administration** > **Hardware** > **Servers**.
8. Check the status of the node under **Controller Nodes** or **Storage Nodes**.
   The status should show **DECOMMISSIONED**.

Follow the replacement instructions for the node in the *HGST Active Archive System FRU Replacement Guide*.

6.5.2 Executing a Failover

If the PostgreSQL partition has failed, or a NIC has failed on the Management Node, fail over the CMC.

Warning: When you are upgrading your setup, do not execute a failover. First complete the upgrade before you start the failover.

To execute a failover, follow the instructions in Executing a Normal Failover on page 58.

6.5.2.1 Executing a Normal Failover

Prerequisites

- Ensure that the **Backup qpackagesserver to Other Node** policy has ran successfully at least once. To verify, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**. Click **Backup qpackagesserver to Other Node** and verify the **Last Run** time. If needed, you can enable the policy and click **Start Now** from the menu on the right.
- Ensure that the framework MetaStore is not degraded. Add more SSDs to the framework MetaStore if necessary.
- Verify the status of the Controller Node:
  - If the Management Node has failed completely, turn off the node and remove it from the environment.
  - If only the SSD failed, you can leave this Controller Node online, for example, if you want to keep the functionality of the node.

Caution: Perform the failover in a `screen` session. For more information about `screen`, see https://www.gnu.org/software/screen/manual/screen.html.

Use this procedure if your Management Node has failed, or if an SSD in your Management Node is faulty, especially if this is the SSD that contains the PostgreSQL partition.

- When your Management Node is offline, the following services are affected:
  - Monitoring
  - Event logging

  Note: Data processing is not affected.
- When an SSD of your Management Node is defective, the following symptoms may occur:
  - Logging into the CMC fails, even though you use the correct credentials.
  - The PostgreSQL mount point is read-only.

The failover procedure takes between 15 and 30 minutes.

To execute a normal failover, proceed as follows:

1. Log in to one of the Controller Nodes which is **not** the Management Node.

   Tip: If you know to which Controller Node the last backup was sent to, log in to that Controller Node.
2. Start a `screen` session.

   ```
   screen -S failover
   ```

   For more information about `screen`, see https://www.gnu.org/software/screen/manual/screen.html.
3. Run the following script:

   ```
    /opt/qbase3/bin/python /opt/qbase3/utils/executeFailoverScript.py
   ```

Note: This script may take up to 5 minutes to respond.

If you did not log into the right Controller Node, you get the following error message:

```
Script executed on wrong node, please execute this script on IP_address
```

Note: If the script failed for other reasons, please contact HGST Support.

4. If the script returned the `Script executed on wrong node` error message, note the IP address in the message, log out of this Controller Node and log into the correct node. Then repeat steps 2 and 3.
5. When the script has finished successfully, log into the CMC.
6. Navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Logging** > **Jobs**.
7. Identify the job **failover management applications** and wait until the failover completes successfully.
8. Navigate to **Dashboard** > **Administration** > **Storage Management** > **MetaStores**.
9. Check if any of the MetaStores are degraded. Add more SSDs to them if necessary.

   In other words, if a MetaStores has less than 3 members, expand it by selecting an available SSD through the CMC.

Note: If you only replaced the SSD of the Management Node and you want to use the affected Controller Node as Management Node again, please contact HGST Support for the failback script.

When you have completed the failover successfully, you can, if wanted, set up the backup policy of the software repository to use a specific Controller Node. This is described in the section Back Up Software Repository Policy on page 66.

6.5.2.2 Replacing a Node

Follow the replacement instructions for the node in the *HGST Active Archive System FRU Replacement Guide*.

6.6 Managing Unmanaged Disks

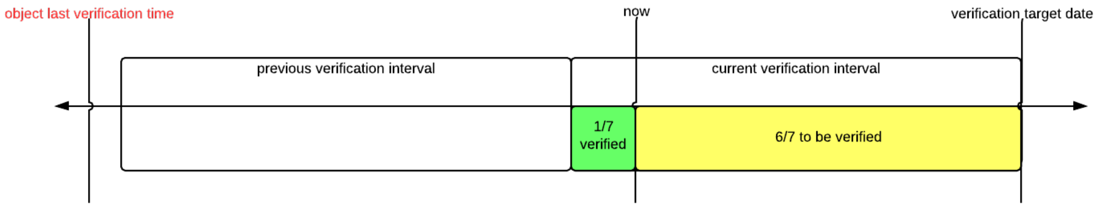
An unmanaged disk is a newly installed disk that the Active Archive System cannot determine a purpose for (in other words, whether it is a replacement disk or really a new disk). The scenario in which this happens is when you accidentally replace the wrong disk in a Controller or Storage Node. For instructions on how to correct this problem, see the *HGST Active Archive System FRU Replacement Guide*.

Warning: Adding disks to the Active Archive System or changing the configuration of any hardware in the Active Archive System is not supported. Please contact HGST Support for more information.

6.7 Handling Unverified Objects

Objects that have the status "unverified", have not been successfully verified at the proper verification interval, as shown in the following image:

Figure 6: Verification Progress: Unverified Objects


object last verification time
now
verification target date
previous verification interval
current verification interval
1/7 verified
6/7 to be verified

Unverified objects are detected through the `monitor name space` action.

Cause	Resolution
Object has at least one blockstore in one one of its spreads that is OFFLINE	Get the blockstore back online.
Checkblock(s) of the object contain CRC32 errors, or entire checkblocks are missing , but the repair keeps on failing	May not occur. Contact HGST Support for an investigation and recovery.
Not all objects can be verified in the set interval, most likely due to heavy load.	The system is under heavy load. Increase the interval to a higher value.

For more information about object verification, see Object Verification on page 180.

Forced Name Space Verification

It is possible to force a name space verification, but it must be done with the highest level of precaution because this action has a negative impact on the performance of the Active Archive System, due to a very high load on it temporarily.

Before forcing a name space verification, you must make an estimate on the minimum verification interval needed to verify all objects.

To force the name space verification, do the following:

1. Open a Q-Shell session on a Controller Node.
   a) Log into a Controller Node using SSH.
   The OSMI menu appears.
   b) Exit the OSMI menu.
   The Linux prompt appears.
   c) Start a Q-Shell session: `/opt/qbase3/qshell`
2. Find the name of the proper name space and its original verification interval.

```
q.dss.manage.listNameSpaces()
```

3. Set the verification interval in seconds.

   For example, 172,800 seconds (2 days).

```
q.dss.manage.setNameSpaceVerificationInterval("name_space_name",172800)
```

4. Set the target date in seconds, relative to current date and time. Ideally, use the same value as the new verification interval.

For example, 172,800 seconds (2 days).

```
q.dss.manage.setNameSpaceVerificationTarget("name_space_name",172800)
```

5. Wait for name space verification to complete.
6. When the name space verification has completed, reset the name space verification interval to its original value and set the new target date with the new interval.

   For example for the default 1 year, this is 31,536,000 seconds.

6.8 Troubleshooting Hardware Issues

This section provides troubleshooting tips for issues you might encounter during when managing the Active Archive System hardware. For more troubleshooting tips, see the *HGST Active Archive System Troubleshooting Guide*.

6.8.1 Field Replaceable Units

Problem	Recommended Action
The PostgreSQL partition has failed, or a NIC has failed on the Management Node.	Fail over the CMC. **Warning:** When you are upgrading your setup, do not execute a failover. First complete the upgrade before you start the failover. To execute a failover, follow the instructions in *Managing Hardware* in the *HGST Active Archive System Administration Guide*.
The wrong disk was replaced.	If you accidentally replace the wrong disk, it shows up in the CMC as an unmanaged disk. An unmanaged disk is a newly installed disk that the Active Archive System cannot determine a purpose for (in other words, whether it is a replacement disk or really a new disk). **Warning:** Adding disks to the Active Archive System or changing the configuration of any hardware in the Active Archive System is not supported. Please contact HGST Support for more information. Correct this problem as follows: **1.** Physically remove the new disk, and replace it with the disk that was accidentally removed. **2.** In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Disks** > **Unmanaged**. **3.** Select the new disk, and in the **Commands** pane, click **Delete**. When you first remove the disk through the CMC, the disk will most likely be added again by the monitoring agent before you can actually remove the disk from the node. If this happens, repeat the steps above to delete the disk again.
You shut down a node in order to replace it or something in it, but when you powered on the new/fixed node, it did not boot or was not detected by the CMC.	Connect a monitor to the node's VGA port, and a keyboard to its USB port. Restart the node. Observe any error messages that it outputs.

6.8.2 System Expansion

Problem	Recommended Action
There is no documentation for how to add drives to an existing Storage Enclosure Basic array.	Do not attempt this. The Active Archive System does not currently support adding more drives to an existing Storage Enclosure Basic array.

6.8.3 General

Problem	Recommended Action

A node is halted or hung, or unreachable after a reboot.	Do a cold reset on the node from IPMI as follows:
	Note: If the Management Node is the node that is hung, perform a failover before executing the procedure below For more information on failing over the Management Node, see *Managing Hardware* in the *HGST Active Archive System Administration Guide*.
	1. In the CMC, browse to **Dashboard** > **Administration** > **Hardware** > **Servers** > **Storage Nodes** or **Dashboard** > **Administration** > **Hardware** > **Servers** > **Controller Nodes**, depending on the type of halted node.
	2. Select the halted node (identified as having the status **HALTED**).
	3. Under the **Summary** tab, in the **General** box, record the IPMI IP address.
	4. Open an SSH session to the Management Node.
	5. Issue the following IPMI command at the Linux prompt, replacing *`IPMI_IP_Address`* with the IPMI IP address recorded above.
	`ipmitool -I lanplus -H IPMI_IP_Address -U ADMIN-P ADMIN chassis`
	` power reset`

Cannot refresh machine status when the Management Node is shut down.	When shutting down a metadata store (in other words, an Arakoon cluster) from the CMC interface, you may not see the current machine status on the last Controller Node to be shut down. This is because an Arakoon cluster requires a minimum of two Controller Nodes in which one is selected as master and reports its status to the CMC.
	You can use either OSMI or the CMC to work around this problem:
	If you are using the OSMI interface, first determine which Controller Node is master for the Arakoon cluster before shutting them down. To determine which Controller Node is master, look for the metadata store's master node using option `1` of the OSMI interface:
	`/opt/qbase3/apps/osmi/osmi`
	`Select 3 for Machines and services`
	`Select 2 for Metastores`
	`Select 1 for list MetaStores`
	`--------------`
	`- MetaStores -`
	`--------------`
	`1) Name: env_metastore (READ/WRITE)`
	`    Master node: node_0_9003`
	`    Node status: {node_0_9003: running}   (DEGRADED)`
	`    Number of keys: 59`
	`2) Name: framework (READ/WRITE)`
	`    Master node: node_0_9001`
	`    Node status: {node_0_9001: running}   (DEGRADED)`
	`    Number of keys: 1740`

Problem	Recommended Action

	`3) Name: userdata (READ/WRITE)`
	`    Master node: node_0_9005`
	`    Node status: {node_0_9005: running}`
	`    Number of keys: 6`
	To avoid this problem when using the CMC, shut down all the Storage Nodes first, then shut down the three Controller Nodes at the same time.

A disk is missing.	Under certain circumstances for a very short window of time, you may notice that a disk is marked as AUTODECOMMISSIONING but does not appear in the list of degraded or decommissioned disks in the CMC. To find a disk that seems to be missing, do the following: **1.** Check the **Live Events** table in the CMC. **2.** Check the **Degraded Disks** page in the CMC. For more information, see *Configuring Maintenance Policies* in the *HGST Active Archive System Administration Guide*.
A disk shows errors.	Run diagnostics on the disk and the Storage Enclosure Basic. For more information, see the *HGST Active Archive System Customer Support Tools*.

SMART data on a decommissioned drive is needed.	To collect SMART data on a decommissioned drive, proceed as follows.
	Warning: Collect this information prior rebooting the Storage Node, as rebooting the Storage Node erases the SMART data for the degraded disks.
	1. Determine the drive's serial number and system IP:
	A. In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Disks** > **Decommissioned**.
	B. Select the drive path for which the SMART data is required.
	C. Write down the serial number presented in the main window. For example, `2EG3RU6J`.
	D. In the same window, click the node's hostname link.
	E. Write down either of the node's private IP address.
	2. Determine the drive's SMART data.
	A. Open an SSH session to the Management Node.
	B. Exit the OSMI menu.
	C. At the Linux prompt, open an SSH session to the node that contains the decommissioned disk using the IP address obtained above.
	D. Execute the following command to determine the correct SMART data file for decommissioned disk.
	`grep serial_number /tmp/smartinfo/*`
	For example,
	`root@HGST-S3-DC01-R01-SN05:~# grep 2EG3RU6J /tmp/smartinfo/*`
	`/tmp/smartinfo/smartctl_scsi-35000cca23b06cb00.txt:Serial`
	` number:                2EG3RU6J`
	E. Look at the file to get additional SMART details.
	`cat filename`

<table>
<tr><th>Problem</th><th>Recommended Action</th></tr>
<tr><td></td><td>

For example,

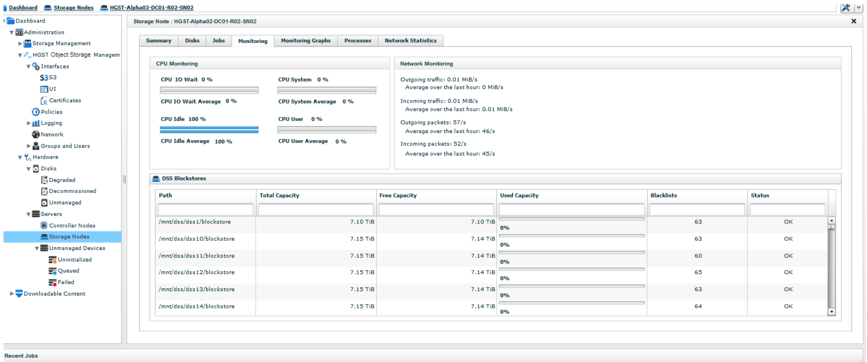
```
root@HGST-S3-DC01-R01-SN05:~# cat /tmp/smartinfo/
smartctl_scsi-35000cca23b06cb00.txt
smartctl 5.41 2011-06-09 r3365 [x86_64-linux-3.11.0-26-generic]
 (local build)
Copyright (C) 2002-11 by Bruce Allen, http://
smartmontools.sourceforge.net

Vendor:               HGST
Product:              HUH728080AL4200
Revision:             a703
User Capacity:        8,001,563,222,016 bytes [8.00 TB]
Logical block size:   4096 bytes
Logical Unit id:      0x5000cca23b06cb00
Serial number:                2EG3RU6J
Device type:          disk
Transport protocol:   SAS
Local Time is:        Wed May 13 13:43:24 2015 PDT
Device supports SMART and is Enabled
Temperature Warning Enabled
SMART Health Status: OK
```

</td></tr>
<tr><td>The hot-swapped disks are being ignored.</td><td>You can replace multiple disks at once, but then you have to install the disks in the same order as you have removed them. For example, if you remove the disks of slot 4, 5, and 8 in that order, you have to install the new disks in the same order, so first slot 4, then 5, and eventually slot 8. If you install the disks in a different order, you have to restart the node.</td></tr>
<tr><td>There are blacklists.</td><td>First check for disk problems, then check for network problems. For more information, see <i>Managing Hardware</i> in the <i>HGST Active Archive System Administration Guide</i>.</td></tr>
<tr><td>Many blacklists appear when a Storage Node is rebooted.</td><td>

When a Storage Node is rebooted, the blockstores are not set to **OFFLINE** status. Therefore, any attempts to write while the system is rebooting result in blacklist operations. Alone, this should not be enough to cause any sort of failure from the perspective of your s3 applications. However, it does manifest in the CMC dashboard's blacklist graph and in the Storage Node's monitoring tab.

Best practice is to reboot the Storage Node using the CMC (**Dashboard** > **Administration** > **Hardware** > **Servers** > **Storage Nodes**, select a node, and in the **Commands** pane, click **Reboot**), which sets it to **HALTED**.

Figure 7: Network Connectivity



Storage Node : HGST-Alpha02-DC01-R02-SN02

CPU Monitoring

CPU IO Wait 0 %
CPU IO Wait Average 0 %
CPU Idle 100 %
CPU Idle Average 100 %
CPU System 0 %
CPU System Average 0 %
CPU User 0 %
CPU User Average 0 %

Network Monitoring

Outgoing traffic: 0.01 MiB/s
Average over the last hour: 0 MiB/s
Incoming traffic: 0.01 MiB/s
Average over the last hour: 0.01 MiB/s
Outgoing packets: 57/s
Average over the last hour: 46/s
Incoming packets: 52/s
Average over the last hour: 45/s

DSS Blockstores

Path	Total Capacity	Free Capacity	Used Capacity	Blacklists	Status
/mnt/dss/dss1/blockstore	7.10 TiB	7.10 TiB	0%	63	OK
/mnt/dss/dss10/blockstore	7.15 TiB	7.14 TiB	0%	63	OK
/mnt/dss/dss11/blockstore	7.15 TiB	7.14 TiB	0%	60	OK
/mnt/dss/dss12/blockstore	7.15 TiB	7.14 TiB	0%	65	OK
/mnt/dss/dss13/blockstore	7.15 TiB	7.14 TiB	0%	63	OK
/mnt/dss/dss14/blockstore	7.15 TiB	7.14 TiB	0%	64	OK

</td></tr>
<tr><td>There are degraded disks.</td><td>First check for disk problems, then check for network problems. For more information, see <i>Managing Hardware</i> in the <i>HGST Active Archive System Administration Guide</i>.</td></tr>
</table>

Problem	Recommended Action
The PostgreSQL mount point is read-only.	Check to see if there is a defective SSD on your Management Node. See *Managing Hardware* in the *HGST Active Archive System Administration Guide*.
Blockstores show different total capacities.	The DSS1 and DSS2 blockstores have an extra 50GB partition named `sandboxtmp`. This partition is used to store logs.
There are memory errors.	If you see an event or receive a notification through SNMP about an ECC memory error, you must replace the DIMM in the node that has the error. For more information about replacing the DIMMs, see the *HGST Active Archive System FRU Replacement Guide*.
There are fan or temperature warnings.	If you see an event or receive a notification through SNMP about a fan or temperature warning, you might need to replace the fan in the node that has the error. For more information about replacing a fan, see the *HGST Active Archive System FRU Replacement Guide*.

7 Configuring Maintenance Policies

Topics:

- Back Up Model Database Policy
- Back Up Software Repository Policy
- Blockstore Automatic Disk Decommissioning Policy
- Collapse MetaStore Transaction Logs Policy
- Defragment MetaStores Policy
- Phone Home Policy
- System Maintenance Policies

These topic describes the various maintenance policies and explains how to customize them.

7.1 Back Up Model Database Policy

The **Back Up Model Database** policy makes a daily backup of all your configuration data, which can be used for disaster recovery.

To enable this policy, proceed as follows:

1. In the CMC go to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**.
2. In the **Policies** list, select **Backup Model Database**.
3. Click **Edit Policy** > **Enable**.
4. Select the desired storage policy that this backup policy will use.

   After you save your configuration, a new name space named `_osis_model_backup` is created.

7.2 Back Up Software Repository Policy

The **Back Up Software Repository** policy backs up the Management Node to one of the other Controller Nodes daily. The Controller Node with the backup becomes the Management Node if the original stops functioning. Running this policy keeps the time needed for a failover to a minimum. The Controller Node for the backup is chosen automatically by default.

After installing your environment, you should activate this policy at once. To activate it, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**.
2. In the **Policies** list, select **Backup qpackagesserver to other node**.
3. In the right pane, click **Start Policy Now**.
4. Choose the desired backup node.
5. (Optional) To change the Controller Node to which the backup is sent to, edit the policy as follows:
   a) In the right pane, click **Edit Policy**.
   b) Choose the desired backup node.

7.3 Blockstore Automatic Disk Decommissioning Policy

The **Auto Decommission Disks** policy can reduce the time between a disk becoming degraded and being repaired. This policy only applies to blockstore disks of a Storage Node.

> **Note:** Disks with RAID partitions are *not* included in automatic disk decommissioning.

When this policy is enabled, any degraded blockstore disk is automatically decommissioned with the following restrictions:

- The blockstore disk must be degraded longer than a configurable amount of time (the *backoff interval*), which is 30 minutes by default.

  When there are multiple degraded disks on a single machine at the same time within a maximum interval of five minutes apart, the backoff interval is multiplied by 24 (*backoff multiplier*) so as to allow the operator to repair a broken connection to a Storage Enclosure Basic. By default, *multiple disks* means "two disks". You can change "multiple disks" to mean something else by editing the policy. For example, if you change "multiple disks" to mean "eight disks" then the *backoff interval* is only multiplied by the *backoff multiplier*when 8 disks are degraded.
- Only one blockstore is decommissioned per node at a time.
- A maximum of 5 blockstores on multiple nodes can be decommissioned at a time.

After decommissioning a disk, there is a repair operation. A new automatic decommission is initiated only after the current repair operation is completed. You can confirm that the blockstore repair operation is done by checking the status of the DSS: when the status is `ABANDONED` in the DSS, the repair operation is done.

7.3.1 About Degraded and Decommissioned Disks

When Does a Drive Become Degraded?

A drive is considered degraded in the following situations.

- The overall SMART health status of a disk is anything other than OK
- There is an indication of an I/O or file system failure
- A mount point becomes read-only: if `touch` *`mount_point`*`/.read_only_test` returns a message stating that the mount point is read-only, then the disk is considered degraded
- The disk is not detected: If a device exists in the internal database but is not viewable using the command `fdisk -l`, then the disk is considered missing and is marked as degraded

What is the Difference Between Degraded and Decommissioned Disks?

A *degraded* disk indicates that the system has observed deviant behavior related to the disk. However, the system does not start any repair activity for the data on the disk. A *decommissioned* disk is one that the system considers bad and has started repair activity on in preparation for immediate or future replacement. The act of decommissioning is non-reversible.

When Does a Disk Become Decommissioned?

A disk becomes decommissioned when:

- An administrator manually decommissions it through the CMC
- The auto-decommission feature determines that it is safe to be decommissioned.

How Does Automatic Decommissioning Occur?

The automatic decommissioning of drives is handled by a management policy called **Auto Decommission Disks**. The policy runs once an hour by default. You can change the settings of this policy through the CMC. For more information on changing the settings, see Editing the Auto Decommission Policy on page 68.

You can also disable this policy altogether; for more information, see Disabling the Auto Decommission Policy on page 68.

What Can Be Automatically Decommissioned?

- Degraded disks
- HDDs that are exclusively used as blockstores, in other words, that hold object content

> **Note:** No automatic decommissioning happens in the following situations:
>
> - The current overall disk safety of the environment is 0
> - The number of disks to be decommissioned would bring the disk safety to 0 or below 0
> - The overall disk safety of the environment could not be determined

What Cannot Be Automatically Decommissioned?

- HDDs that are part of a software RAID or used by a MetaStore for transaction logs
- SSDs

What Events Are Related to Automatic Decommissioning?

See Complete List of Events on page 118.

7.3.2 Viewing Auto Decommissioned Drives

You can find out which disks are currently being auto decommissioned or have completed auto decommissioning in two locations in the CMC: the **Live Events** table, or the **Degraded Disks** page.

1. Log into the CMC.
2. Check the **Live Events** table for disks that are currently being automatically decommissioned:

   Two events are raised when an automatic decommissioning is executed.

   - Automatic decommissioning started for a disk (`OBS-DISK-0020`)
   - Automatic decommissioning completed for disk (`OBS-DISK-0021`)

   For more information on the **Live Events** table, see Event Viewer on page 93.
3. Check the **Degraded Disks** page for disks that have completed the automatic decommissioning process:
   a) Navigate to **Dashboard** > **Administration** > **Hardware** > **Disks** > **Degraded**.
   b) At the bottom of the **Degraded Disks**, click the **Decommissioning and Autodecommissioning Disks** tab. All automatically decommissioned disks are listed.

7.3.3 Enabling the Auto Decommission Policy

1. Log into the CMC.
2. Navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**.
3. In the **Policies** list, select **Auto decommission disks**.
4. In the **Commands** pane, click **Enable Policy**.
5. Click **OK** to confirm.

7.3.4 Disabling the Auto Decommission Policy

1. Log into the CMC.
2. Navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**.
3. In the **Policies** list, select **Auto decommission disks**.
4. In the **Commands** pane, click **Disable Policy**.
5. Click **OK** to confirm.

7.3.5 Editing the Auto Decommission Policy

1. Log into the CMC.

2. Navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**.
3. In the **Policies** list, select **Auto decommission disks**.
4. In the **Commands** pane, click **Edit Policy** and fill out the **Edit Policy** form:
   - **Policy Interval**: how often this policy runs. The default is every 60 minutes.
   - **Backoff Interval**: the number of minutes to wait for a disk in a degraded state before beginning to auto decommission it. The default is 30 minutes. If a disk is degraded for a number of minutes smaller than this value, this policy does not auto decommission it.
   - **Backoff Multiplier**: the multiplier for the backoff interval. The default 24. The **Backoff Interval** is multiplied by this value when the value of **Simultaneous Disk Failures** is exceeded. This allows you to investigate and remediate the issue (check cabling and so on). For example: It is possible that the disk controller has failed rather than the disks themselves.
   - **Simultaneous Disk Failures**: the number of disks on the *same* node that must be degraded within a five minute window before the **Backoff Multiplier** is applied. The default is 2 disks.
   - **Erase Disks After Decommissioning**: specifies whether or not the disk is securely erased after repairs are complete. Secure erase is off by default.
5. Click **Next** and **OK** to apply the changes.

7.4 Collapse MetaStore Transaction Logs Policy

The basic intention of this policy is to limit the number of transaction logs (*tlogs*) per MetaStore. A tlogs is a chronological record of all metadata operations that have not been committed to a backup database file (called `head.db`) that can be used for recovery purposes.

When the MetaStore process terminates in an uncontrolled fashion, the Active Archive System replays the tlogs on top of the backup database file. The number of tlogs determines the amount of time it takes to replay. However, the collapse operation requires that one node of the MetaStore cluster does not operate in the majority while the collapse is ongoing and hence poses an temporary risk of lowered availability.

The **Collapse MetaStore Transaction Logs** policy is enabled by default, but might require modification to match your availability requirements. To modify this policy, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**.
2. In the **Policies** list, select **Collapse MetaStore Transaction Logs**.
3. In the right pane, click **Edit Policy**.
4. In the **Edit Policy** wizard, determine how frequently this policy should run (default is every day). If your system does not generate a lot of tlogs, you may increase this number.
5. Determine the **Tlog threshold** (default is `25`). When your system reaches this number of tlogs, this policy runs.
6. Determine the **Event threshold** (default is `30`). When your system reaches this number of tlogs, it starts triggering events that indicate that the tlog collapsing has stalled or is progressing slower than anticipated.
7. Select the storage policy that the collapsed tlogs will use. Take note that you can only choose this policy once.
   After you update these settings, a new name space named `_metastorebackup` is created.

7.5 Defragment MetaStores Policy

The **Defragment MetaStores** policy defragments the database files of the MetaStores on a weekly basis. It has a default timeout value of 170 minutes. To change this timeout, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**.
2. In the policies list, click **Defragment MetaStores database files**.
3. In the right pane, click **Edit Policy**.
4. In the **Edit Policy** window, set the desired timeout value.

7.6 Phone Home Policy

The **Phone Home** policy generates a report with the following information:

- Active events at the time of the report creation
- All events in the last 24 hours, in reverse chronological order
- Worst case overall disk safety
- MetaStore statuses

To configure the **Phone Home** policy, proceed as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**.
2. In the **Policies** list, select **Send a report of the highest severity events**.
   The **phone_home policy** window appears.
3. In the **Commands** pane, click **Edit Policy**.
   The **Edit Policy** wizard appears.
4. Specify an **e-mail address** to receive real-time event messages.

   Note: Currently, only a single e-mail address can be configured for the `admin` user. If you want multiple e-mail addresses to receive real-time event messages, create a mailing list in your mail infrastructure. Add `HGST_PhoneHome@hgst.com` to this mailing list for additional proactive support.

5. Select the **minimal error condition level** to be included.

   Choose one of the following:

   - **Critical**
   - **Urgent**
   - **Error**
   - **Warning**
   - **Info**
   - **Unknown**

   Caution: The **Unknown** level is not recommended. If you select **Unknown**, the Active Archive System only escalates events that have no level.

6. Click **Next** to accept the changes.
7. In the **phone_home policy** window, in the right pane, click **Start Policy**.

The **Phone Home** policy runs on a daily basis. If you want to generate a report instantly, proceed as follows:

1. Open the **Send a report of the highest severity events** policy.
2. In the right pane, click **Start Policy Now**.

   This action automatically generates a report. The report is sent to the e-mail address specified in the policy.

7.7 System Maintenance Policies

7.7.1 Aggregate Storagepool Info Policy

The **Aggregate Storagepool Info** policy requests the status of every node in the storage pool and updates the CMC dashboard.

To change the frequency of these updates, proceed as follows:

1. In the CMC go to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies**.

2. In the **Policies** list, select **Aggregate Storagepool Info**.
3. Click **Edit Policy** in the right-hand pane.
   The **Edit Policy** wizard appears.
4. In the **Edit Policy** window, you can change the following settings:

   - The number of minutes between each run of the policy
   - The length of the time the policy may take before timing out

   Note: For environments with 16-30 Storage Nodes, this value must be at least 15 minutes. For larger environments, this value must be at least 30 minutes.

7.7.2 Clean Up Old Versions Policy

The **Clean Up Old Versions** policy runs daily. It cleans up all successful jobs on the CMC that are older than seven days, but ignores all failed jobs.

This policy cannot be configured. To view the details of this policy, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies** > **Clean Up Old Versions**.

7.7.3 Monitor Blockstores Policy

The **Monitor Blockstores** policy runs daily. It checks the current status of the blockstores (`DISABLED`: temporarily offline, `HALTED`: offline). If any of the blockstores are either `DISABLED` or `HALTED`, a message is sent to the configured e-mail address with these nodes. If none of the blockstores have this status, no message is sent.

Note: The result of this policy is not visible in the CMC.

This policy cannot be configured. To view the details of this policy, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Policies** > **Monitor Blockstores**.

If you receive an e-mail indicating that there are disabled or halted blockstores, you can check these as follows:

- In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Storage Services**.
- In the **Storage Services** list, click the desired **Storagepool**.
- In the **Storage Nodes** window, click the desired Storage Node (the one hosting the disabled or halted blockstores).
- Click on the desired **dssstoragedaemon**.
- In the **Storage Daemon** window, you can see the blockstores and their status.

8 Logging

Topics:

- Collecting Log Files
- Default Log Behavior
- Normal Log Collection
- IPMI Log Collection
- Emergency Log Collection
- Extracting Log Files
- Other Log Options
- Log File Management
- Client Daemon Log File Statistics
- Arakoon Log File Statistics
- Troubleshooting Logging Issues

8.1 Collecting Log Files

The log collector tool collects the following types of logs. One or many of these log file types can be included in any single run:

Log Type	Description of Process/Log	Location
dssclientdaemons	DSS client daemon logs. These are collected by the process on the Controller Node responsible for erasure coding / responding to external client requests.	`/opt/qbase3/var/log/dss/clientdaemons`
dssstoragedaemons	DSS storage daemon logs. These are collected by the process on the Storage Nodes responsible for writing erasure encoded data to and from disks, also responsible for coordinating repair activity.	`/opt/qbase3/var/log/dss/storagedaemons/`
dssmaintenanceagents	DSS maintenance daemon logs. These are collected by the process on the Storage Nodes responsible for executing repair activity.	`/opt/qbase3/var/log/dss/maintenanceagents`
arakoonclusters	Arakoon cluster logs. These are collected by the process on the Controller Nodes responsible for storing and serving object metadata information.	`/opt/qbase3/var/log/arakoon/`
alldaemons	All of the above processes	`/var/log/syslog, /var/log/messages, /var/log/kern.log`

Log Type	Description of Process/Log	Location
system	System logs (syslog, messages, kernel, daemon, pound, system info)	`/var/log`
pylabs	Pylab logs (monitor agent - application server - workflowengine - postgres - apache - archive); logs related to the management processes	`/opt/qbase3/var/log/pylabslogs`
upgrade	Logs related to Active Archive System upgrades	Upgrade logs in `/opt/qbase3/var/log/`
all	All log types above	
rrds	RRD files (network bandwidth - temperatures - CPU load - ...)	

The log collector tool can collect logs in the following date modes:

- Date (one day log collection)
- Date range (More than one day log collection)

By default the log collector tool collects logs from all running nodes. You can specify one specific node to collect data from it, or exclude some nodes from log collection. Additionally, when logs are collected from the Management Node, the following storage information is collected as well:

- Revision information
- Environment defaults
- Information about storage policies, capacity, and disk safety
- Information about the blockstores (disks), their locations and their status
- Information about storage daemons and their status
- Information about the configured storage policies (spread width, disk safety, maximum superblock size, hierarchy awareness)
- Information about the configured MetaStores

The log collector tool sends log bundles to the Management Node by default.

For more information on the options you can set for the log collector tool, see Normal Log Collection on page 73.

8.2 Default Log Behavior

The log collection tool has the following default behavior:

- The tool collects all logs by default, except any other mode was specified.
- The tool monitors the log collection on environment nodes.
- The tool collects logs for "yesterday" date if you did not specify a date.
- The tool uploads collected log bundles to the Management Node.
- The tool uploads collected log files to HGST Support.

This behavior may change with the selected options in the command.

8.3 Normal Log Collection

For normal collection of logs, run the `log_collector_trigger.py` command **on the Management Node**:

```
/opt/qbase3/apps/log_collector/log_collector_trigger.py -t LOGTYPE... [-m MODE |
-k KEYWORD...] [-I DAEMONID..., [-L]] [-d LOGDATE | -F FROMDATE, -T TODATE]
[-u UPLOAD, [-N NAMESPACE]] [-s] [-n NODE...] [-x EXCLUDE...] [-r RACKNAME...]
```

```
 [-C] [-Q] [-l LOGGING] [-f NUMBEROFFILES] [-y]
```

With the following options:

Option	Description
`-t, --type`	This determines which **log types** are collected. Valid values are: • `dssclientdaemons` • `dssstoragedaemons` • `dssmaintenanceagents` • `arakoonclusters` • `alldaemons` • `system` • `all` • `pylabs` • `rrds` For information about these log types, see Collecting Log Files on page 72.
`-m, --mode`	Use this to determine the **log mode**. Valid values are: • `allogs`: collects all log messages. • `errorlogs`: collects only error logs from all log files. Default. • `keyword`: collects only the log messages containing (a) specified keyword(s).
`-k, --keyword`	Type any keyword(s) to collect from log files. When one or more keywords is specified, the log collection mode is the keyword(s) by default. You do not need to specify logs mode when specifying a keyword(s).
`-I, --id`	Daemon ID for log collection. When using this option, you cannot use more than one log type. This option can only be used with the following **log types**: `dssclientdaemons`, `dsscachedaemons`, `dssstoragedaemons`, `dssmaintenanceagents`, `arakoonclusters`.
`-L, --triggerall`	Trigger log collection on all nodes when daemon ID is specified. This option is only available when using the `-I, --id` option.
`-d ,--date`	Use this to determine the date to collect logs. The default value is: yesterday. Valid values are: • `today` • `yesterday` • and date in *`yyyy-mm-dd`* format
`-F, --from`	Use this to specify a date range to collect logs. The date must be in *`yyyy-mm-dd`* format and must be chronologically prior to the date specified by the `-T, --to` option.
`-T, --to`	Use this to specify a date range to collect logs. The date must be in *`yyyy-mm-dd`* format and must be chronologically after the date specified by the `-F, --from` option.
`-N, --namespace`	Use this option to specify which DSS name space is used to upload log files to. This option should be used when specifying upload option (`-u, --upload`) with `localdss` value. If not, the name space option is ignored.
`-s, --skipcheck`	Use this to skip the tool existence check on environment nodes. Do not use this unless you are certain of the existence of the tool on the environment nodes you will collect the logs from.

Option	Description
-C, --compress	Compress all collected output log files in one .tgz file.
-n, --node	Use this to specify one or more nodes for log collection. In other words, you can use this option to restrict the search/stop to specific node(s).
-x, --exclude	Use this to exclude one or more nodes from the log collection.
-r, --rack	Use this option to specify data center names and rack names instead of individual nodes. Separate names with ":" You can specify multiple rack names.
-l, --logging	Use this to set the logging level. Valid values are INFO and DEBUG. INFO is the default logging level. Specify DEBUG when detailed logs are needed. **Important:** • Use DEBUG level only if necessary, as this level will grow your log files significantly. • In case such a growth is detected, you receive an event similar to [/opt/qbase3/var/log/arakoon/Object2/node_2_9007/node_2_9007.log] Reached [327.562] MB in last hour. • When DEBUG level is used, there is a risk that /mnt/sandboxtmp may become full, which would result in the log collector stopping with an error. In this case, you should receive warning (95% full), error (96% full), and critical (98% full) events.
-f, --filescount	Use this to specify the number of files to collect from.
-Q, --no-monitor	Do not monitor log collection status on environment nodes after log collection triggering. The log collector tool exits after triggering log collection on environment nodes. By default, the log collection status is monitored.
-y,--yes	Always answer with yes.
-A, --running	Use only with the -G and -X options. If used, you will only monitor / stop the running nodes only.

8.4 IPMI Log Collection

To collect IPMI logs, run the following command **on the Management Node**:

```
(ipmitool sel elist ; ipmitool chassis status ; ipmitool bmc selftest ; \
ipmitool mc selftest ; ipmitool sensor list ; ipmitool pef list ; ipmitool sdr ; \
ipmitool -c sdr ; dmidecode -t 15) | tee -a  ipmidmilog.txt
```

8.5 Emergency Log Collection

You cannot trigger the normal log collection on the nodes in your environment if you do not have, or cannot create, a cloudAPI connection. In this case you can use the emergency log collection method. With the emergency collection method, you collect logs without a cloudAPI connection. The logs can:

- Be sent to the Management Node (provide IP address, user name and password)
- Be stored on the node itself, in the following directory /opt/qbase3/var/tmp/log_collector_tar

To collect logs in emergency situations, proceed as follows:

1. Log into the desired node.
2. Go to log collector tool directory (`/opt/qbase3/apps/log_collector`).
3. Run the `log_collector.py` script:

```
/opt/qbase3/bin/python log_collector.py -t LOGTYPE... --no-cloudapi
[-m MODE | -k KEYWORD...] [-I DAEMONID...] [-d LOGDATE | -F FROMDATE, -T TODATE]
[-u UPLOAD, [-N NAMESPACE]] [-l LOGGING] [-f NUMBEROFFILES] [-M MANAGEMENTIP]
[-U USERNAME] [-P PASSWORD]
```

where:

Option	Explanation
`--no-cloudapi`	Use this option to do an emergency collection without cloud api connection. This option is mandatory in emergency collections.
`-M, --management`	IP address of the management node to send the logs to. If no management ip is given, the collected logs are stored locally.
`-U, --user`	Username of the management node. Use this only when a management IP address has been provided.
`-P, --password`	Password of the management node. Use this only when a management IP address has been provided.

Note: Next to the previous mentioned options, all the options of the Normal Log Collection on page 73 are available.

Caution: Storage information cannot be collected in an emergency collection.

8.6 Extracting Log Files

After collecting log files, the tool can extract output log files on the management node.

To use this function, use the following code:

```
log_collector_trigger.py -e -d LOGDATE | -F FROMDATE, -T TODATE [-n NODE...]
[-x EXCLUDE...] [-r RACKNAME...] [-R] [-a FILEPATH...][-l LOGGING]
```

With the following options:

Option	Description
`-e, --extract`	Extract the collected logfiles to the same directory of the log collector tool (/ `opt/qbase3/var/log/log_collector`). You must also specify a log date or date range.
`-d ,?date`	Use this to specify the date of the log files to upload to the evidence environment. Valid values are`today`, `yesterday`, or a date in *yyyy-mm-dd* format. The default value is `yesterday`.
`-F, --from`	Use this to specify a date range of the log files. The date must be in *yyyy-mm-dd* format and must be set prior than the `-T, --to` date.
`-T, --to`	Use this to specify a date range of the log files. The date must be in *yyyy-mm-dd* format and must be set after the `-F, --from` date.
`-n, --node`	Use this to specify the node(s) to upload its log files to HGST Support.

Option	Description
`-x, --exclude`	Use this to specify the node(s) to be excluded from uploading its log files to the HGST Support.
`-r, --rack`	Use this option to specify data center names and rack names instead of individual nodes. Separate data center name and rack name with `":"` You can specify multiple rack names.
`-R, --remove`	Use this option to remove the logfiles from the Management Node after uploading the output log files to HGST Support or other http server.
`-a , --file`	Use this option to upload specific files. Fill in a file path to upload only the files in this location. You can specify multiple file paths.
`-l, --logging`	Use this to set the logging level. Valid values are `info` and `debug`. The default value is `info`. Specify debug when you need detailed logs.

8.7 Other Log Options

Use the `-G` option to get the status of the current log collecting:

```
log_collector_trigger.py -G [-A] [-n NODE...] [-x EXCLUDE...] [-r RACKNAME...]
 [-l LOGGING]
```

Use the `-X` option to stop the current log collecting:

```
log_collector_trigger.py -X [-A] [-n NODE...] [-x EXCLUDE...] [-r RACKNAME...]
 [-l LOGGING]
```

Use the `-O` option to collect the tool logfiles:

```
log_collector_trigger.py -O [-n NODE...] [-x EXCLUDE...] [-r RACKNAME...]
 [-l LOGGING]
```

If you use any of the previous codes, you can use the following options:

Option	Explanation
`-A, --running`	Use only with the -G and -X option. If used, you will only monitor / stop the runnin nodes only.
`-n, --node`	You can use this to restrict the search/stop to specific (a) node(s).
`-x, --exclude`	You can use this to specify the node(s) to be excluded from the search/stop.
`-r, --rack`	Use this option to specify data center names and rack names instead of individual nodes. Separate data center name and rack name with ":" You can specify multiple rack names.
`-l, --logging`	Use this to set the logging level. Valid values are `info` and `debug`. The default value is `info`. Specify `debug` when you need detailed logs.

8.8 Log File Management

Log rotation management is implemented as follows:

- All log files reside on the following drive: `/mnt/sandboxtmp/logging`.
- The log rotation occurs every hour if the file size becomes 10 MB or larger.

- The system keeps the following number of logs, depending on the process:

Process	**Number of Logs**
application server	14
apache	1000
dssclientdaemon	1000
rsync	1000
pylabs	7
ejabberd	1000
arakoon server	1000
postgres	1000
workflow engine	1000
pound	1000
snmp	1000
dhcp	1000
monitoring agent	1000

- Events are triggered in case a log file becomes abnormally large (> 256 MB).
- If less than 10 % of free space is left on the log file partition, the oldest log files are moved to a system name space called `_logging_backup`. Each time logs are moved to the logging name space, log files that older than the parameter `days_to_keep` are removed by an automatic clean-up process.
- The percentage of free space on the partition and the parameter `days_to_keep` can be configured through `monitoring.xml`, in the following section:

```
<partitions>
   <partition>
      <mountpoint>/mnt/sandboxtmp</mountpoint>
      <dir>logging</dir>
      <min_free_percent>10</min_free_percent>
      <backup_until_percent>15</backup_until_percent>
      <days_to_keep>90</days_to_keep>
   </partition>
</partitions>
```

8.8 Applicationserver Log Files

For the application server, there are two log rotation mechanisms, which influence each other:

- System log rotation: OS log rotation
- Active Archive System framework log rotation: creator of the actual applicationserver log files

The Active Archive System framework log rotation creates the actual applicationserver log files. The configuration is stored in `/opt/qbase3/cfg/qconfig/logtargetfs.cfg`. The creation of a new log file is based on the size of the log file. When the file grows beyond 5,000 lines (default), a new log file is created.

The system log rotation creates archive files (.tgz) of these applicationserver log files on a daily basis. By default a maximum of 14 log archives is created. However, by default there is a parameter `size` defined in the system log rotate configuration (`/etc/logrotate.d/applicationserver`). The `daily` parameter is overruled, and an archive is created for each new log file that is found for the applicationserver log. As such the system log rotate creates an archive for each found log file (again, up to 14 archives).

In normal situations, the Active Archive System log file does not grow beyond 5,000 lines and there will only be one such log file per day. This will result in a maximum of 14 days of logging that can be retrieved.

If the system is heavily loaded, it is likely that the Active Archive System log file grows very fast and that there are many log files generated per day. Since the system log rotate creates an archive of each found log file, it is possible that the time frame of available logging becomes very small. For example, when there are seven log files generated each day, then you can find only logging of two day with the default settings.

This small time frame can make it very difficult to investigate issues. Keep this in mind in case that you find very little different log archives.

8.9 Client Daemon Log File Statistics

This section contains the documentation about the log file statistics, which are found in the client daemon log files (`/opt/qbase3/var/log/dss/clientdaemons/<guid>/clientx.log`).

Time duration is always given in seconds, throughput in MiB per second.

8.9 Numbers with Standard and 'other_location' Variants

The standard variants are calculated by taking all transfers into accounts, while the 'other_location' variants only take the transfers to/from a different location (different node) into account.

Variable Name	2nd Part of Variable Name (What is Measured?)	Function Whose Duration/ Throughput is Measured
(other_location_)add_blocks_	duration/throughput	Blockstore writing the wireblocks to the disks
(other_location_)add_full_copy_	duration/throughput	Blockstore writing a full-copy superblock to disk.
(other_location_)get_blocks_	duration/throughput	Blockstore reading the wireblocks from disk.
(other_location_)get_full_copy_	duration/throughput	Blockstore reading a full-copy superblock to disk.

8.9 Other Numbers

Variable Name	2nd Part of Variable Name (What is Measured?)	Function
get_sb_	duration/throughput	Gets a superblock by reading the necessary data from disks and possibly decoding it.
dec_sb_	duration/throughput	Decodes a number of wireblocks to build a superblock.
enc_sb_	duration/throughput	Encodes a superblocks into a number of wireblocks.
wr_sb_	duration/throughput	Sends a superblock to an output channel.
rd_sb_	duration/throughput	Generates a superblock given some input data.
add_obj_md_	duration	Writes the object metadata to disk.
ck_blocks_	duration/throughput	Performs the verification of the wireblocks.
ck_full_copy_	duration/throughput	Performs the verification of a full-copy superblock.

Variable Name	2nd Part of Variable Name (What is Measured?)	Function
delete_blocks_	duration	Deletes the wireblocks from disk.
delete_full_copy_	duration	Deletes the full-copy superblocks from disk, same value as delete_blocks_
spread_gen_normal	(time duration)	Generates a wireblock spread given a certain policy.
spread_gen_custom	(time duration)	Same as above, but with a list of preferred blockstore Id's.
sd_add_blocks_	duration/throughput	Blockstore performing the actual write of the wireblocks to disk.
sd_get_blocks_	duration/throughput	Blockstore performing the actual read of the wireblocks from disk.
sd_ck_blocks_	duration/throughput	Blockstore performing the actual verification on the wireblocks on disk.
sd_delete_blocks_	duration	Blockstore performing the actual deletion of the wireblocks from disk.
sd_add_full_copy_	duration/throughput	Blockstore performing the actual write of the full-copy superblock to disk.
sd_get_full_copy_	duration/throughput	Blockstore performing the actual read of the full-copy superblock from disk.
sd_ck_full_copy_	duration/throughput	Blockstore performing the actual verification of the full-copy superblock on disk.
sd_delete_full_copy_	duration/throughput	Blockstore performing the actual deletion of the full-copy superblock from disk.
sd_add_obj_md_	duration	Blockstore performing the actual write of the object metadata to disk (using the Tlog_registry).
put_	duration/throughput	The full process of the 'put' of an object.
get_	duration/throughput	The full process of a successful 'get' of an object.
failed_get_	duration/throughput	The full process of a failed 'get' of an object.
delete_	duration	The full process of scheduling an object for deletion.
rcc_get_	duration/throughput	'get' of an object from the CacheCluster which was successful.
rcc_missed_get_	duration/throughput	'get' of an object from the CacheCluster where the object was not found

Variable Name	2nd Part of Variable Name (What is Measured?)	Function
rcc_failed_get_	duration/throughput	'get' of an object from the CacheCluster which failed with an exception
rcc_put_	duration/throughput	'put' of an object to the CacheCluster which was successful.
rcc_failed_put_	duration/throughput	'put' of an object to the CacheCluster which failed.
repair_NORMAL_	duration	Normal repair of an object.
repair_DECOMMISSION_	duration	Decommission repair of an object (after the decommissioning of a disk).
repair_REBALANCE_	duration	Rebalance of an object.
repair_CLEAN_	duration	Clean repair of an object.
repair_VERIFY_	duration	Verification of an object.
s3_get_	duration/throughput	Gets an object after a S3 GET request.
s3_put_	duration/throughput	Puts an object after a S3 PUT request.
s3_auth_	duration	Performs the authentication after a S3 request.
s3_md5_	duration/throughput	Calculates the md5sum during a S3 PUT request.
encrypt_	duration/throughput	Performs the encryption during a PUT, if requested.
decrypt_	duration/throughput	Performs the decryption during a GET, if necessary

8.10 Arakoon Log File Statistics

This section explains Arakoon statistics that are saved in Arakoon log files. There is a default frequency of five minutes to start new Arakoon statistics. The values are always calculated for the interval.

The Arakoon statistics can be found in the Arakoon log files as a dictionary "stats". The log files can be found in `/opt/qbase3/var/log/arakoon/<arakoon cluster>/<arakoon node>/<arakoonnode>.log`.

8.10 General Arakoon Statistics

Variable Name	Unit	Meaning
start	epoch	Time stamp of the start of the statistics interval
last	epoch	Time stamp of last operation in the interval
avg_set_size	bytes	Average object size for set operations
avg_get_size	bytes	Average object size for get operations
avg_range_size	n/a	Average number of keys returned with range operations

Variable Name	Unit	Meaning
avg_range_entries_size	n/a	Average number of keys returned by range_entries operations
avg_rev_range_entries_size	n/a	Average number of keys returned by rev_range_entries operations
avg_prefix_size	n/a	Average number of keys returned by prefix_keys operations
avg_del_prefix_size	n/a	Average number of keys deleted by delete_prefix operations
harvest_stats	n/a	• n: number of times a batch of client updates is harvested into one paxos value • min: minimum of the number of updates in one paxos value • max: maximum of the number of updates in one paxos value • avg: average number of updates in one paxos value • dev: variance of the number of updates in one paxos value

Tip: The harvest operation is the retrieval of the master node of a batch of operations, sent by the different Arakoon clients, and to get a consensus with its slave nodes for executing the operations in the batch.

8.10 Arakoon Operation Statistics

The next items provide statistics about the various Arakoon operations. The statistics of each operation consist of:

- **n**: number of operations
- **min**: duration of the fastest execution of the operation, in seconds
- **max**: duration of the slowest execution of the operation, in seconds
- **avg**: average duration of the operation, in seconds
- **dev**: variance on the duration of the operation, in seconds

Variable Name	Operation
set_info	SET operation
get_info	GET operation
del_info	DELETE operation
mget_info	MULTI-GET operation, GET multiple keys at once, with failure if a key has no value
mget_option_info	MULTI-GET OPTION operation, no value returns None
seq_info	SEQUENCE operation, a multi-update operation
tas_info	TEST-and-SET operation
range_info	RANGE operation, retrieve a list of keys by defining begin/end key and max. number of keys
range_entries_info	RANGE-ENTRIES operation, same as RANGE but the result is a list of key/value pairs

Variable Name	Operation
rev_range_entries_info	REV-RANGE-ENTRIES operation, same as RANGE-ENTRIES but result is in reversed order
prefix_info	PREFIX keys operation, return all keys with a given prefix
delete_prefix_info	DELETE PREFIX operation, delete all key/value pairs with a given prefix
ops_info	Statistics of all of the mentioned operations in this table

8.10 Arakoon Process Statistics

The last items in the Arakoon statistics give you an insight on the general Arakoon process.

Variable Name	Unit	Meaning
mem_allocated	KB	Amount of virtual memory used by the OCaml runtime
mem_maxrss	KB	Amount of memory of the this process in the main memory
mem_minor_collections	n/a	Number of minor collections by the garbage collector since the start of Arakoon
mem_major_collections	n/a	Number of major collection cycles completed since the start of Arakoon
mem_compactions	n/a	Number of heap compactions
node_is	n/a	List of the nodes in the cluster and its last witnessed TLog index

Tip: For more information about garbage collectors and heap compactions, consult the Real World OCaml website.

8.11 Troubleshooting Logging Issues

This section provides troubleshooting tips for issues you might encounter during when generating or collecting Active Archive System log files. For more troubleshooting tips, see the *HGST Active Archive System Troubleshooting Guide*.

8.11.1 General

Problem	Recommended Action
The log collector tool fails.	When the output of the log collector tool is similar to:

```
Mon Mar 23 14:03:54 2015
HGST-Alpha02-DC01-R02-CN03 -----> Stopped with errors or warnings
HGST-Alpha02-DC01-R02-CN02 -----> Stopped with errors or warnings
HGST-Alpha02-DC01-R02-CN01 -----> Stopped with errors or warnings
HGST-Alpha02-DC01-R02-SN06 -----> Stopped with errors or warnings
HGST-Alpha02-DC01-R02-SN05 -----> Stopped with errors or warnings
HGST-Alpha02-DC01-R02-SN04 -----> Stopped with errors or warnings
HGST-Alpha02-DC01-R02-SN03 -----> Stopped with errors or warnings
HGST-Alpha02-DC01-R02-SN02 -----> Stopped with errors or warnings
HGST-Alpha02-DC01-R02-SN01 -----> Stopped with errors or warnings
```

Problem	Recommended Action
	`Log collection task completed` this indicates that the tool was run from a Controller Node that is not, or no longer, the Management Node (perhaps due to failover). The log collector tool must be run from the Management Node only. Rerun the tool from the Management Node.

The log files are missing even though the log collector ran with no errors.	When the partition that stores log files is full, the system uploads new logs to an internal name space, named `_logging_backup`, instead.
	To retrieve log files from `_logging_backup`, proceed as follows:
	1. Open `http://`*`public_IP`*`:8080/namespace/_logging_backup`, where *`public_IP`* is an accessible IP address of any Controller Node.
	2. Determine the machine GUID of this particular Controller Node:
	A. Enter the Q-Shell on the Management Mode:
	`/opt/qbase3/qshell`
	B. Execute the following three Q-Shell commands, replacing *`node_name`* with the hostname of the machine for which you want to determine the GUID.
	`nodename = '`*`node_name`*`'`
	`api = i.config.cloudApiConnection.find('main')`
	`api.machine.find(name=nodename)['result'][0]`
	The system displays the GUID of *`node_name`*.
	3. Navigate to the subfolder whose name matches the machine GUID you found in the previous step.
	4. Browse to the desired log type and click to download. No authentication is required.
	Tip: Log file names include two epoch timestamps. For example, `clientx.log.2015_07_09_1436469301.gz.1436472682`. The first timestamp corresponds to the ending time in the log file. The second timestamp corresponds to when the file was uploaded to the `_logging_backup` namespace.

9 Monitoring the Active Archive System

Topics:

- About Monitoring
- About Events
- Monitoring Repair Tasks
- Monitoring of Long Running Jobs
- Monitoring the System Using the CMC Dashboard
- Monitoring the System Using Q-Shell
- Event Viewer
- Configuring Event Escalation
- Checking System Health
- Node Monitoring
- Storage Pool Monitoring
- SNMP Polling
- Dead Man Timer
- Collecting Telemetry

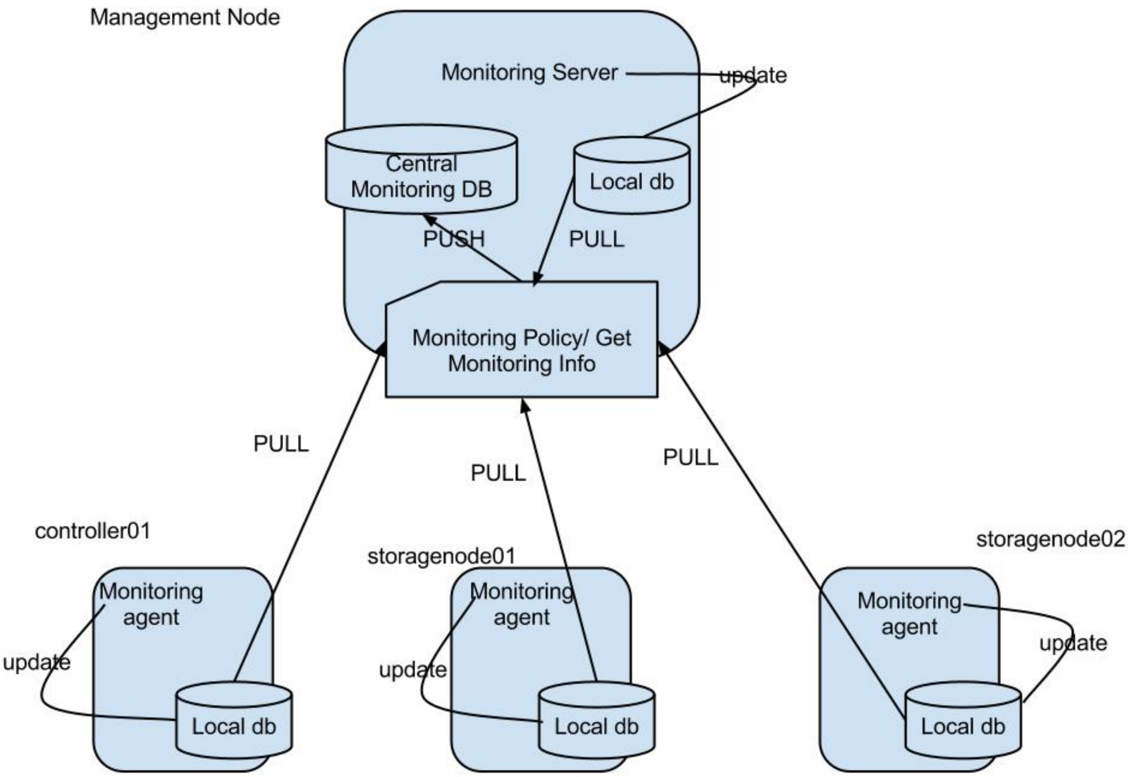
These topics explain how to monitor Active Archive System jobs, events and policies.

The Active Archive System monitoring system is a framework used in an Active Archive System environment.

The purpose of the monitoring system is to collect information about different physical and logical aspects of environment-wide resources.

9.1 About Monitoring

Figure 8: The Active Archive System Monitoring System



Each node in the environment has a *monitoring agent*. This monitoring agent monitors the node once per minute (a *monitoring cycle*).

The resulting monitoring information is updated in the local database. Only the modified objects (compared to the last cycle) are saved in the local database.

The Monitoring policy of the Management Node pulls the monitoring information from all the local databases (including its own) and synchronizes it with the data in the central monitoring database.

The monitoring policy does also the following actions:

- Generates information about the aggregated storage pool usage
- Generates blacklist graphics
- Checks the status of the remote nodes, monitoring agents, and node agents.

The bulk of the time of a monitoring cycle is used for updating the monitoring database, once the data is pulled from a node. Depending upon the load on the node running the CMC, this can take seconds to a minute.

Note: Successful jobs and events are automatically removed when they are older than seven days.

9.2 About Events

Each monitoring agent takes note of events happening on the node and sends those events to the Management Node.

Depending on the configuration of the monitoring policies, the monitoring server collects and escalates those when necessary.

The monitoring policies check the frequency of those events and send the necessary events to the configured e-mail address.

9.3 Monitoring Repair Tasks

A repair task is a task that is executed by a maintenance agent and will repair superblocks that have the status "REPAIR". A superblock gets into the "REPAIR" status when it is not possible to see from the spread that something is wrong with the object, for example when a CRC32 check on a checkblock returns an error.

When a superblock has the status "REPAIR", a repair task is created for a maintenance agent, which will execute a clean repair task on the superblock.

The following table represents the different repair types:

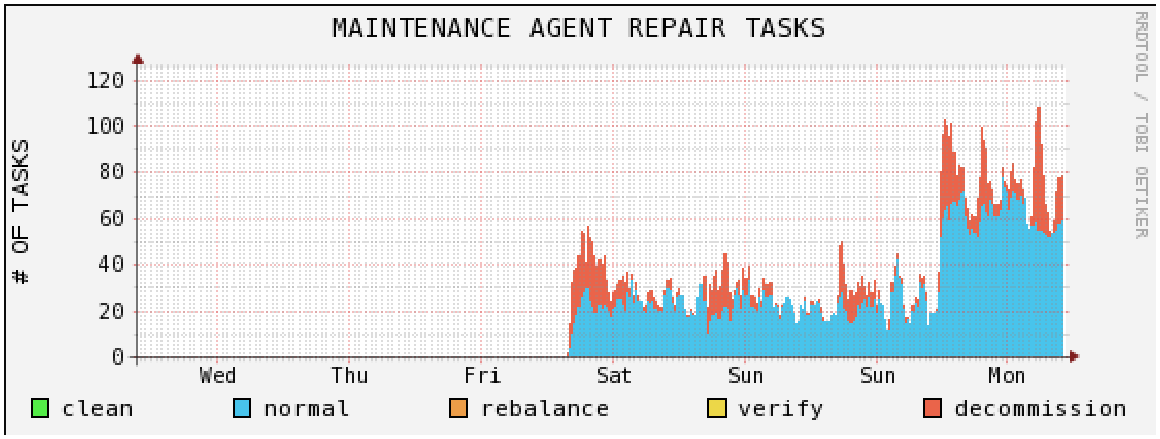
Repair Type	Repair Actions	When
normal	• Don't request from each blockstore the amount of present checkblocks. • Don't check the CRCs of the checkblocks. • Replace missing wireblock sequences on a new blockstore that match the hierarchy rules. • If the result would be an incomplete spread, check how many wireblocks are on all stores to see if the disk safety would not go down. If the disk safety would go down, abort the action.	• Spread contains at least one ABANDONED disk; and • Object status and superblock status of all superblocks in the database are OK; and • Policy of the object and policy of the name space are the same
clean	• New put of the object with the parameters of the new policy. • Storage object version of the new put object is v2.	• Policy of the object is different from the policy of the name space; and • Object or superblock status of at least one superblock is "REPAIR"
rebalance	Rebalance name spaces over the new number of available storage daemons.	Status of *all* superblocks in a storage pool is "REPAIR".
verify	• Request from each blockstore the amount of present checkblocks. • Check CRCs of all checkblocks. • If all wireblock sequences of all stores are OK, update the verification date, if not, execute a 'clean' repair.	• verification time has arrived; and • spread contains no DECOMMISSIONED stores; and • spread contains no ABANDONED stores; and • policy of the object and policy of the name space are the same
decommission	• Request the wireblock sequences from the DECOMMISSIONED blockstores.	• spread contains at least one DECOMMISSIONED store; and • spread contains no ABANDONED stores; and

Repair Type	Repair Actions	When
	• Upload these wireblock sequences to blockstores that match the hierarchy rule.	• policy of the object and policy of the name space are the same; and • object status and superblock status of all superblocks in the database are OK

> **Note:** The verification is completely omitted in a repair crawl when the monitoring agent detects a decommissioned blockstore.

In the CMC, you can find the tasks of the maintenance agent on the dashboard in the right graphic.

Figure 9: Maintenance Agent Repair Tasks



9.4 Monitoring of Long Running Jobs

Some monitoring jobs may take hours to complete, for example crawling a bucket of 300,000,000 objects can take 16 hours. In such a situation, it is useful if an administrator can follow the progress of the job.

The progress cannot be followed in the CMC, but in the storage daemon logs where you can find at a specified interval the following three lines:

```
repair manager: delete crawl: namespace %s: processed %d objects
repair manager: repair crawl: namespace %s: processed %d objects
repair manager: wiping namespace %s: removed %s objects
```

The interval is defined in the storage daemon configuration file by the parameter `crawl_log_progress_interval`.

- Configuration file: `/opt/qbase3/cfg/dss/storagedaemons/`*`guid`*`.cfg`
- Default value: `100 000`.

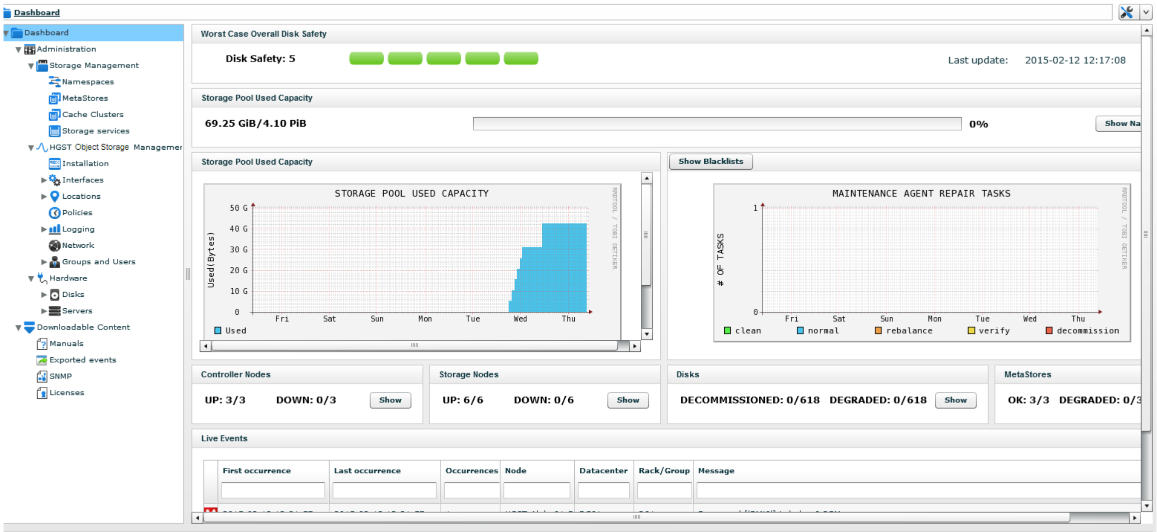
9.5 Monitoring the System Using the CMC Dashboard

The CMC dashboard provides an overview of certain monitoring possibilities:

- The worst case overall Disk Safety.
- The used capacity of the storage pool (in percent and graph).
- The number of read / write blacklists.
- The number of available nodes.

- The number of available disks.
- Number of degraded disks / MetaStores.
- The most recent live events. For more information, see Live Events vs All Events on page 94.

Figure 10: The CMC Dashboard



Disk Safety: 5

Last update: 2015-02-12 12:17:08

69.25 GiB/4.10 PiB 0%

UP: 3/3 DOWN: 0/3

UP: 6/6 DOWN: 0/6

DECOMMISSIONED: 0/618 DEGRADED: 0/618

OK: 3/3 DEGRADED: 0/3

First occurrence	Last occurrence	Occurrences	Node	Datacenter	Rack/Group	Message

Note: On the dashboard, the total number of disks in the environment is not decreased, even if a disk is degraded or decommissioned.

For example, if there are 614/618 disks determined to be OK and there are 2/618 disks degraded, the status of all disks are:

- 614 disks **OK** (**Disks** Pane shows **OK**: 614/618).
- 2 disks degraded (**Disks** Pane shows **DEGRADED**: 2/618).
- 2 disks decommissioned (not shown on dashboard).

9.6 Monitoring the System Using Q-Shell

You can do most of your monitoring of the Active Archive System through the Cloud Management Center (CMC). However, you can also monitor the Active Archive System through the Q-Shell, for example when you do not have a web browser at your disposal.

Monitoring a name space creates a report on the state of the name space. This report contains the number and the capacity taken by the objects, parts (subobject), and superblocks.

9.6 Monitoring a Name Space

The monitoring of a name space via the Q-Shell is done with the command `q.dss.manage.monitorNameSpace('`*`name_space`*`')`.

```
{ 'blockstore_usage' : '' ,
'current_policy_id_stats' :
{ '75e8fd9a109b42232b4e36cdea34871e4' :
   { 'capacity_frontend_delete' : 0 ,
     'capacity_frontend_objects_unverified' : 0 ,
     'capacity_frontend_ok' : 57494180 ,
     'capacity_frontend_repair' : 0 ,
     'capacity_frontend_superblocks_unverified' : 0 ,
     'disksafety_objects' : { 4 : 3 },
     'disksafety_objects_decommissioned' : { 4 : 3 },
     'disksafety_objects_offline' : { 4 : 3 },
     'disksafety_superblocks' : { 4 : 3 },
     'disksafety_superblocks_decommissioned' : { 4 : 3 },
     'disksafety_superblocks_offline' : { 4 : 3 },
     'nr_objects_delete' : 0 ,
     'nr_objects_ok' : 3 ,
     'nr_objects_repair' : 0 ,
     'nr_objects_unverified' : 0 ,
     'nr_superblocks_delete' : 0 ,
     'nr_superblocks_ok' : 3 ,
     'nr_superblocks_repair' : 0 ,
     'nr_superblocks_unverified' : 0 ,
     'policy_stats_hashtbl' :
     { 'capacity' :
        { 'object' :
           { 'all' : 57494180 ,
                 'change policy' : 0 ,
                 'disk safety decommissioned' : { '4' : 57494180 },
                 'disk safety normal' : { '4' : 57494180 },
                 'disk safety offline' : { '4' : 57494180 },
                 'needs conversion' : 0 ,
                 'ok' : 57494180 ,
                 'repair' : 0 ,
                 'unverified' : 0 },
           'part' :
           { 'all' : 57494180 ,
           'change policy' : 0 ,
           'disk safety decommissioned' : { '4' : 57494180 },
           'disk safety normal' : { '4' : 57494180 },
           'disk safety offline' : { '4' : 57494180 },
           'needs conversion' : 0 ,
           'ok' : 57494180 ,
           'repair' : 0 ,
           'unverified' : 0 },
       'superblock' :
       { 'all' : 57494180 ,
       'change policy' : 0 ,
       'disk safety decommissioned' : { '4' : 57494180 },
       'disk safety normal' : { '4' : 57494180 },
       'disk safety offline' : { '4' : 57494180 },
       'needs conversion' : 0 ,
       'ok' : 57494180 ,
       'repair' : 0 ,
       'unverified' : 0 }},
       'nr' :
       { 'object' :
           { 'all' : 3 ,
           'change policy' : 0 ,
           'disk safety decommissioned' : { '4' : 3 },
           'disk safety normal' : { '4' : 3 },
           'disk safety offline' : { '4' : 3 },
           'needs conversion' : 0 ,
           'ok' : 3 ,
           'repair' : 0 ,
           'unverified' : 0 },
           'part' :
           { 'all' : 3 ,
           'change policy' : 0 ,
           'disk safety decommissioned' : { '4' : 3 },
           'disk safety normal' : { '4' : 3 },
           'disk safety offline' : { '4' : 3 },
           'needs conversion' : 0 ,
         'ok' : 3 ,
         'repair' : 0 ,
         'unverified' : 0 },
      'superblock' :
      { 'all' : 3 ,
      'change policy' : 0 ,
      'disk safety decommissioned' : { '4' : 3 },
      'disk safety normal' : { '4' : 3 },
      'disk safety offline' : { '4' : 3 },
      'needs conversion' : 0 ,
      'ok' : 3 ,
      'repair' : 0 ,
      'unverified' : 0 }}}}},
      'full_copy_blockstore_usage' : '' ,
      'last_update' : ( 1396524178.5766261 , 'Apr 3 2014 13:22:58.5766' ),
      'name' : 'myNameSpace' ,
      'object_name_length_stats' : { 'average' :  34.0 ,
      'count' : 3 ,
      'deviation' : 0.0 ,
      'max' : 34.0 ,
      'min' : 34.0 ,
      'name' : 'object_name_length_stats' },
      'old_policy_id_stats' : {},
      'start_date' : ( 1396524178.5640919 , 'Apr 3 2014 13:22:58.5641' ),
'version' : 2 }
```

The data that can be found in the result are:

- `version`: version of the format in which the data is stored. This information is of no further use for the user.
- `name`: name of the monitored name space
- `start_date`: date and time when the last monitor crawl was started
- `last_update`: last update of the cached data
- `current_policy_id_stats`: dict with policy guid as key and value the following information:
    - `nr_objects_ok`: number of healthy objects
    - `nr_superblocks_ok`: number of healthy superblocks
    - `nr_objects_repair`: number of objects that have at least one superblock in repair status
    - `nr_superblocks_repair`: number of superblocks that need repair
    - `nr_objects_delete`: number of objects that need to be deleted
    - `nr_superblocks_delete`: number of superblocks that still need to be deleted
    - `nr_objects_unverified`: number of unverified objects. Every object is verified each 365 days if it is still a healthy object. If this verification would not have taken place within these 365 days, the object is considered as unverified.
    - `nr_superblocks_unverified`: number of unverified superblocks. This is similar as the unverified objects, but for superblocks.
    - `capacity_frontend_ok`: the sum of the sizes of the objects put that are healthy, expressed in bytes
    - `capacity_frontend_repair`: the sum of the sizes of the objects put that need to be repaired, expressed in bytes
    - `capacity_frontend_delete`: the sum of the sizes of the objects put that still need to be deleted, expressed in bytes
    - `capacity_frontend_objects_unverified`: the sum of the sizes of the objects put that is still unverified, expressed in bytes
    - `capacity_frontend_superblocks_unverified`: the sum of the sizes of the superblocks put that is still unverified, expressed in bytes
    - `disksafety_superblocks`: a dict with keys going from 'disksafety - spread width' to disksafety giving for all these values the number of superblocks that have that disk safety, taking into account the ABANDONED blockstores

- ◆ `disksafety_superblocks_decommissioned`: a dict with keys going from 'disksafety - spread width' to disksafety giving for all these values the number of superblocks that have that disk safety, taking into account the ABANDONED and DECOMMISSIONED blockstores
- ◆ `disksafety_superblocks_offline`: a dict with keys going from 'disksafety - spread width' to disksafety giving for all these values the number of superblocks that have that disk safety, taking into account the ABANDONED, DECOMMISSIONED and OFFLINE blockstores
- ◆ `disksafety_objects`: a dict with keys going from 'disksafety - spread width' to disksafety giving for all these values the number of objects that have that disk safety, taking into account the ABANDONED blockstores. For more information about disk safety, see Disk Safety on page 111.
- ◆ `disksafety_objects_decommissioned`: a dict with keys going from 'disksafety - spread width' to disksafety giving for all these values the number of objects that have that disk safety, taking into account the ABANDONED and OFFLINE blockstores
- ◆ `disksafety_objects_offline`: a dict with keys going from 'disksafety - spread width' to disksafety giving for all these values the number of objects that have that disk safety, taking into account the ABANDONED, OFFLINE and DECOMMISSIONED blockstores
- ◆ `policy_stats_hashtbl`: overview with the statistics of all objects, parts, and superblocks with that policy as their target. This is a dict which contains two main sections, "nr" and "capacity", respectively the number of items and the taken disk space of the items. These two main sections show the information by objects, parts, and superblocks.

> **Note:** The Active Archive System parts are not exactly the same as the S3 parts. In S3, there is a flat structure of the parts; in the Active Archive System there is a tree structure, in which the main object (or storage object) consists of parts and each part can consist of child parts.

— `nr`: gives you an overview of the number of items (*objects*, *parts*, *superblocks*)

— `capacity`: gives you an overview of the taken disk space, split in *objects*, *parts*, and *superblocks*. The indicated disk spaces are expressed in bytes.

- – `ok`: all healthy items, which don't have either label REPAIR, CHANGE_POLICY, or UNVERIFIED
- – `all`: all items
- – `repair`: all items which need a repair action
- – `change policy`: items for which a new policy is selected, but which are not yet encoded with the new policy
- – `unverified`: items which are unverified. Every item is verified each 365 days if it is still a healthy object. If this verification would not have taken place within these 365 days, the item is considered as unverified.
- – `disk safety normal`: dict which takes the healthy blockstores into account. The disk safety is the key and the number of items/disk space is the value. This dict can contain multiple key/value pairs.
- – `disk safety offline`: dict which takes the offline blockstores into account.
- – `needs conversion`: remaining items/data volume that need conversion. This is a decreasing value and is ideally 0.
- – `disk safety decommissioned`: dict which takes the decommissioned blockstores into account.

- `old_policy_id_stats`: has been replaced by the parameter `change policy`. It was used to report on the statistics of objects in a name space which were not yet encoded with the target policy. This parameter is still in use for compatibility reasons.
- `object_name_length_stats`: statistics about the length of the object names in the name space.
  - ◆ `average`: average object name length
  - ◆ `count`: number of object to calculate the average length
  - ◆ `variance`: variance of the length
  - ◆ `max`: length of the longest object name
  - ◆ `min`: length of the shortest object name
  - ◆ `name`: name of the statistic

- `blockstore_usage`: dict with the blockstores as key and number of objects and size of each blockstore. To retrieve this information, you have to add the parameter `showBlockstoreUsage=True` in the `monitorNameSpace()` command.

```
{ 'blockstore_usage' : { 0 : { 'count' : 50869 , 'size' : 303687930 },
1 : { 'count' : 50922 , 'size' : 304004340 },
2 : { 'count' : 50799 , 'size' : 303270030 },
3 : { 'count' : 50897 , 'size' : 303855090 },
4 : { 'count' : 50901 , 'size' : 303878970 },
5 : { 'count' : 50923 , 'size' : 304010310 },
6 : { 'count' : 50980 , 'size' : 304350600 },
7 : { 'count' : 50941 , 'size' : 304117770 }}
```

- `full_copy_blockstore_usage`: same as `blockstore_usage` parameter, but this information is only available when small file support is available. This parameter shows the usage of the small file copies. To retrieve this information, you have to add the parameter `showBlockstoreUsage=True` in the `monitorNameSpace()` command.

```
'full_copy_blockstore_usage' : { 0 : { 'count' : 7307 , 'size' : 4669173 },
1 : { 'count' : 7254 , 'size' : 4635306 },
2 : { 'count' : 7377 , 'size' : 4713903 },
3 : { 'count' : 7279 , 'size' : 4651281 },
4 : { 'count' : 7275 , 'size' : 4648725 },
5 : { 'count' : 7253 , 'size' : 4634667 },
6 : { 'count' : 7196 , 'size' : 4598244 },
7 : { 'count' : 7235 , 'size' : 4623165 }},
```

9.6 Monitoring a Storage Pool

The monitoring of a storage pool via the Q-Shell is done with the command `q.dss.manage.monitorStoragePool()`. The result is the aggregated monitoring of all name spaces in that storage pool and shows the same fields as the `monitorNameSpace()`.

9.7 Event Viewer

You can open the events list in two ways:

- In the CMC, go to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Logging** > **Events**
- On the CMC dashboard, at the bottom of the page, you find the **Live Events** table; click **Show more**.

The **Live Events** table shows up to 10,000 events from your environment and the amount of times it occurred.

Note: Successful jobs and events are automatically removed when they are older than seven days.

9.7.1 Event Severity

The events are divided into the following categories of severity:

Symbol	Meaning	Usage
i	INFO	Informational event. No action required.
!	WARNING	An issue requires attention but does not immediately require an intervention. There is no data impact. The issue should be fairly easy to resolve.
☒	ERROR	A component (hardware or software) is failing and needs attention. There is no data impact.
☒	CRITICAL	Issue can cause data loss or service unavailability

For a complete listing of Active Archive System events, see Events on page 118.

9.7.2 Event Details

To open the details of an event, click an event in the event list.

The details window is divided in three sections, providing the following information:

- Event Summary
  - `Message`: user-friendly message of the event.
  - `Severity`: severity of the event. For more information, see Event Severity on page 93.
  - `Source`: the location where the event occurred. This is also a link to that corresponding machine.
  - `Occurrences`: number of times the event occurred.
  - `First Occurrence`: date and time of the first occurrence.
  - `Last Occurrence`: date and time of the latest occurrence.
- Solution
  - If available, the solution to resolve the issue that causes the event.
- Details
  - `Event Type`: internal code of the event.
  - `Event details`: the backtrace of the event, you can copy this information to the clipboard via the button **Copy To Clipboard**.
  - `Tags`: extra metadata, containing system-specific data, for example GUIDs.

Click the X in the top right corner to return to the event list.

9.7.3 Live Events vs All Events

By default, only the events that are currently applicable are visible (live events).

An event remains in the **Live Events** list:

- as long as the event is still applicable.
- for two occurrences of the monitoring interval (see Complete List of Events on page 118), if the event is no longer applicable.
- for 2,000 seconds if the monitoring interval is not specified and the event is no longer applicable.

Click **All Events** to view all events, even if they are no longer applicable on your environment.

9.7.4 Removing Events from the List of Events

To remove one or more events, proceed as follows:

1. Select the checkbox on the left of the desired events.
2. In the right column, click **Delete Events**.

To remove all events in the list, proceed as follows:

1. Select the checkbox the header row of the event list.
2. In the right column, click **Delete Events**.

9.7.5 Exporting Events

It is possible to export your events to a comma-separated file (.csv).

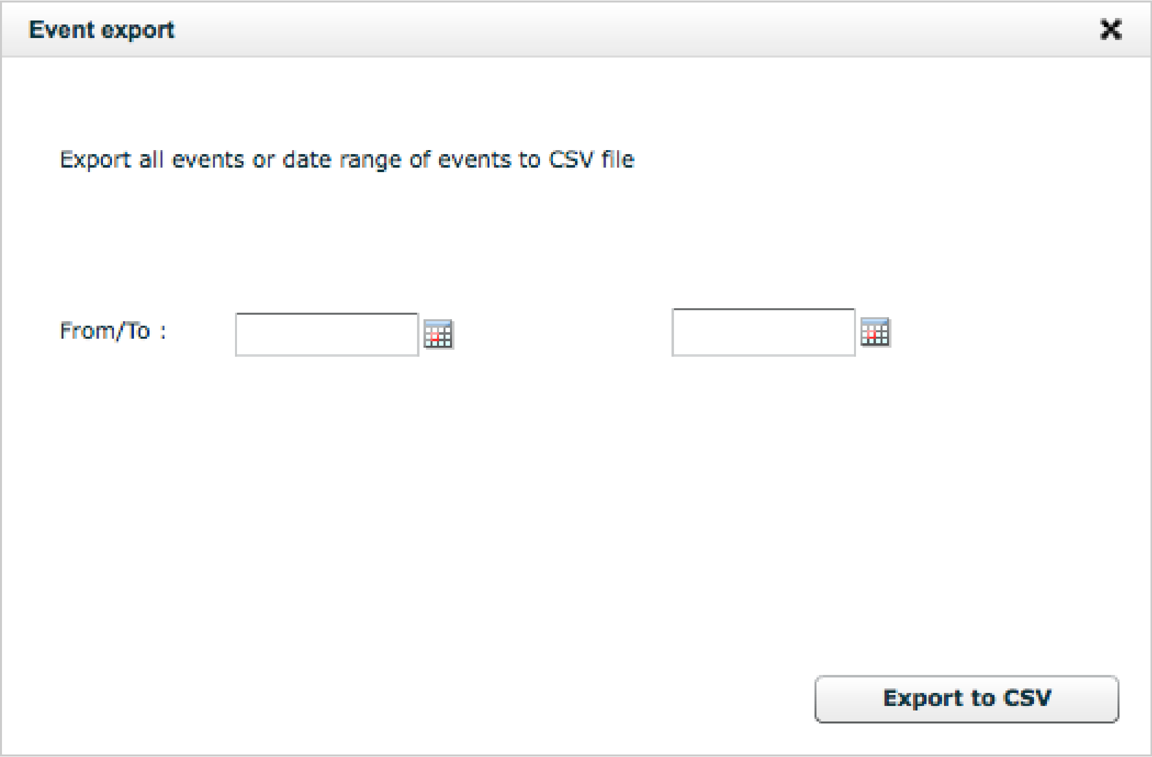
To do so, proceed as follows:

1. Open the list of events window via **Dashboard** > **Administration** > **HGST Object Storage Management** > **Logging** > **Events**.
2. In the right column, click **Export events**.

   The **Event export** appears.

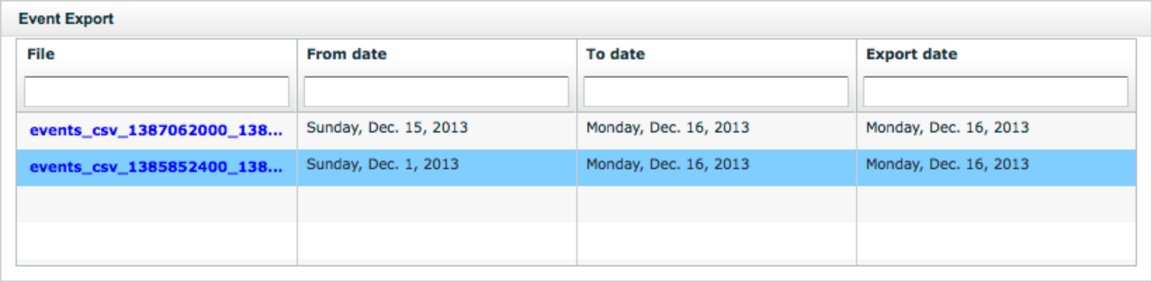
3. Use the calendar icons to select the start and end date of the events that you wish to export.

Figure 11: Exporting Events



4. Click **Export to CSV**.
5. You can download the export via **Dashboard** > **Downloadable Content** > **Exported events**.
6. In the table, click the desired file name to download the file.

Figure 12: Downloadable Content: Exported Events


Event Export

File	From date	To date	Export date
events_csv_1387062000_138...	Sunday, Dec. 15, 2013	Monday, Dec. 16, 2013	Monday, Dec. 16, 2013
events_csv_1385852400_138...	Sunday, Dec. 1, 2013	Monday, Dec. 16, 2013	Monday, Dec. 16, 2013

Note: The export files are automatically removed after one day.

9.8 Configuring Event Escalation

9.8.1 Configuring SNMP Traps

You can configure the Active Archive System to send SNMP traps when events occur. This way the Active Archive System acts as an SNMP client and sends the traps to an SNMP server.

Configure SNMP trapping as follows:

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Installation**.

2. In the right column, click **Configure SNMP**.

   The **Configure SNMP** wizard appears.
3. On the **Traps** tab, specify the **Community String**, **Server**, and **Port** for SNMP trapping.
   - **Community String**: this is similar to a user ID or password to communicate with an SNMP server.
   - **Server**: the IP address or URI of the SNMP server to send the SNMP traps to.
   - **Port**: the IP port via which the SNMP server can be reached.

9.9 Checking System Health

To get an overview of the health of the Active Archive System installation, run the *health checker*. This is an interactive tool, which means that human interaction might be required. For example, when the tool detects that a service is not running, it asks you whether to start the service or not.

You must run the health checker before you upgrade your installation because issues in your installation may block the upgrade. The health checker provides an overview of these issues and makes it easier to tackle them.

To run the health checker, do the following:

1. Log into the Management Node using SSH.
   The OSMI menu appears.
2. Exit the OSMI menu.
   The Linux prompt appears.
3. Start a Q-Shell session:

   ```
   /opt/qbase3/qshell
   ```
4. Start the health checker:
   - If you do not have external public internet connectivity, invoke `q.amplistor.healthCheck(check_public_connectivity=False)`.
   - If you are connected to the external network, invoke `q.amplistor.healthCheck()`.

   Sample output from the health checker:

   ```
   **********************************************************
   * CHECK LOCAL NODE HEALTH STATUS                         *
   **********************************************************
    * Checking local services                                   RUNNING
    *  Checking Ejabberd Engine Running
                                    DONE
    *  Checking DHCP daemon                                     DONE
    *  Checking rsync                                           DONE
    *  Checking TFTP                                            DONE
    *  Checking PostgreSQL                                      DONE
    *  Checking Workflow Engine                                 DONE

    ...

    *  Node stor2: RUNNING                                      DONE
    *  Node cpunode3: RUNNING                                   DONE
    *  Node cpunode2: RUNNING                                   DONE
    *  Node stor1: RUNNING                                      DONE
    *  Node cpunode1: RUNNING                                   DONE
    * Verifying status of nodes                                 FINISHED
   **********************************************************
   * CHECK ENVIRONMENT ACCESSIBILITY AND CONSISTENCY        *
   **********************************************************
    * Verifying environment accessibility                       RUNNING
    *  Creating public cluster                                  RUNNING
   ```

```
 *  Creating public cluster                                  DONE
 *  Ping test to all machines                                DONE

 ...

 *  Verifying model consistency                              DONE
 * Verifying environment consistency                         FINISHED
**********************************************************
* CHECK REMOTE NODES HEALTH STATUS                       *
**********************************************************
 * Checking remote machines                                  RUNNING
 *  Checking health of node stor2                            RUNNING
 *   Checking Remote DSS Storage Daemon                      RUNNING
 *   Checking Remote DSS Storage Daemon                      DONE

 ...

 *  Delete test object from DSS                              DONE
 * Checking remote machines                                  FINISHED
**********************************************************
* JOBS OVERVIEW                                          *
**********************************************************

 0 running jobs

 31 jobs in status ERROR!!

 23793 jobs in status DONE
```

5. Resolve any issues that are found.
   a) Check the output of the health checker for a summary of all historical jobs by status.
   b) If there are failed jobs, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Logging** > **Jobs** in the CMC.
   c) In the **Status** column filter, type `ERROR` and press `Enter`.
   d) Click the individual jobs and failed job steps to see additional details that can be used to further diagnose the root cause.

9.10 Node Monitoring

In the CMC, navigate to **Dashboard** > **Administration** > **Hardware** > **Servers** > **Controller Nodes**.

Click **Controller Nodes** or **Storage Nodes**.

> **Note:** Alternatively, you can also click on the **Show** button of one of the nodes on the CMC dashboard.
>
> In the nodes list, click on one of the nodes to view the detail window.

> **Note:** The detail window retrieves all necessary data. This might take a few seconds.
>
> In the detail window of the node, click on the **Monitoring** tab.

The **Monitoring** tab provides the following monitoring statuses:

- CPU monitoring.
- Network monitoring.
- Blockstore Monitoring (Storage Nodes only).

The **Monitoring Graphs** tab provides the following graphs:

- CPU usage of the last 24 hours.
- CPU load of the last 24 hours.
- The memory usage of the last 24 hours.
- The swap memory usage of the last 24 hours.
- Network usage of the last 24 hours (both sent and received packages).
- The partition usage of the last 24 hours.
- The number of keys used in the MetaStore (Controller Nodes only).

Refreshing the Node Status

If you want to see the current status of the node, proceed as follows:

1. In the nodes list, click on the desired node, to view the detail window.
2. In the detail view window, click **Refresh machine status** in the right-hand pane.

9.11 Storage Pool Monitoring

9.11.1 Monitoring the Storage Pool

There are two ways to monitor the storage pool:

- On the Dashboard page, you can view both the status bar and the aggregated graph of the **storage pool used capacity**.
- You can view the detail window of the storage pool as follows:
  1. In the CMC, go to **Administration** > **HGST Object Storage Management** > **Storage Services** > **Storagepool**.
  2. In the **Storagepool** window, click on the desired Storage Node.
  3. In the **Storage node** window, click on the desired storage daemon to view the detail window.

     Note: The **Detail** window retrieves all necessary data. This might take a few seconds.

Note: While you are putting or getting information from the storage pool, there may be a difference between the storage pool status bar, the aggregated graph, and the detail view graphics.

This is because these three are updated with a different interval:

- The dashboard status bar has the most recent details, as it is updated with the highest frequency.
- The Individual Storage Daemon graphics can have a lag of 5-10 minutes.
- The aggregate storage pool graph on the dashboard can have a lag equal to the interval of the monitoring policy (by default 30 minutes).

9.11.2 Monitoring the Storage Policy

Note: The storage policy is fixed, and is not configurable.

1. Log into the CMC.
2. Navigate to **Dashboard** > **Administration** > **Storage Management** > **Storage policies**.
   The **Storage Policy Management** window shows the following information:
   - The name of the **default policy**.
   - The setting for **full copy** of small files, if small file support is active.
   - The **spread width** and **safety**.
   - The **safety strategy**.
   - The **maximum superblock size** (in bytes, not in MiB).
   - The **lowest disk safety** an object has on the policy.

- The number of **objects with a lower disk safety** than the desired safety (percentage). This is not necessarily the lowest disk safety.

3. Click on the storage policy.
   The **details** dialog displays all the previous mentioned information, and includes the following data:
   - The number of **message blocks**.
   - The **total number of objects** stored via this policy.
   - The **disk safety details**: all objects stored are grouped together with their current safety level.

9.11.3 Disk Safety Details

The disk safety details is the current status of all objects. They are grouped together per disk safety level.

Each safety level has the following information:

- The number of objects that has this safety level (number and percentage).
- A number of dots next to it, corresponding to their Disk Safety level (a disk safety of 4 will generate 4 dots).

If a group of objects has no lowered disk safety, all dots are green.

If a group of objects has a lowered safety, a number of dots are grey, depending on how much lower the current safety is.

If a group of objects has a negative disk safety (data loss!), extra red dots are generated left of the grey ones. This should be avoided!

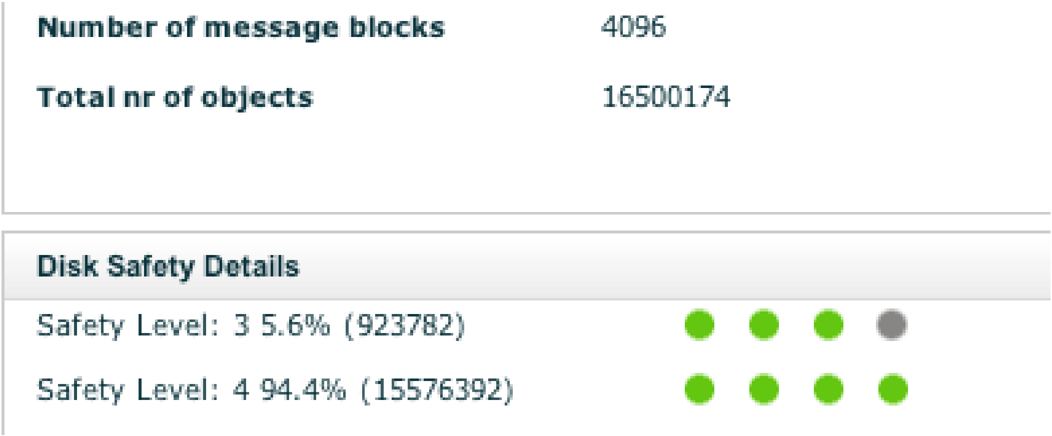
9.11.3 Sample Disk Safety Details

A storage policy has the following details:

- A spread width of 18.
- A safety of 5.
- 16 500 174 objects stored.
- 923 782 objects have a lowered disk safety of 3 instead of 4.

This will result in the following Disk Safety Details window:

Figure 13: Disk Safety Details



Number of message blocks	4096
Total nr of objects	16500174

Disk Safety Details

Safety Level: 3 5.6% (923782) ● ● ● ●

Safety Level: 4 94.4% (15576392) ● ● ● ●

9.12 SNMP Polling

For SNMP Polling, the Active Archive System acts as an SNMP server to provide data to external SNMP clients.

9.12.1 Which Data Can Be Polled?

- Only actual values can be polled. For trending, you will have to login into the Management Node.

- Issues that require (immediate) customer attention are trapped through SNMP. They cannot be polled.
- The highest update frequency of monitoring data is 15 minutes.
- Only globalized stats are exposed, not individual component stats.

9.12.2 Downloading the MIB Variables

Download the MIB variables as follows:

1. In the CMC, go to **Dashboard** > **Downloadable content** > **SNMP**.
2. When the MIB appears, click **Download**.
3. Save the MIB variables and upload them to your SNMP server.

9.12.3 Configuring SNMP Polling

Configure SNMP polling as follows:

1. In the CMC, go to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Installation**.
2. In the right column, click **Configure SNMP**.

   The **Configure SNMP** wizard appears.
3. On the **Polling** tab, specify the **Community string** for SNMP polling.
4. Set this string in your SNMP client, in order to make it communicate with the Active Archive System.

   Tip: When you configure your SNMP client, make sure that it connects to the public virtual IP address of your Cloud Service.

   This avoids the need to reconfigure the SNMP polling when there is a failover of the Management Node.

9.13 Dead Man Timer

When the application that is sending data to a client daemon ceases to send data, the client daemon can get into a situation where all resources are exhausted (because they are reserved for the current ongoing data processing tasks) and no new store or retrieve threads can be spawned. In order to prevent such a situation, a dead man timer has been implemented.

The dead man timer is enabled by default on all client daemons. When enabled, threads that have no activity for two hours are terminated.

You can edit the configuration files of any client daemon to enable/disable/modify the dead man timer as follows.

1. Log into the node using SSH.
   The OSMI menu appears.
2. Open a terminal session.
3. Exit the OSMI menu.
   The Linux prompt appears.
4. Open the client daemon configuration file, `/opt/qbase3/cfg/dss/clientdaemons/`*`daemon_guid`*`.cfg`.
5. In the `HTTP` section of the client daemon configuration file, add/edit the following entries:

   ```
   request_wait_connection_timeout = value_1
   request_process_connection_timeout = value_2
   ```

   where:

   - *`value_1`* is the time an established HTTP connection can remain idle (seconds).
   - *`value_2`* is the time it can take to read from or write to a socket for a single superblock (seconds).
6. Save and exit the configuration file.

7. Restart the client daemon with a Q-Shell command:

```
q.dss.clientdaemons.restartOne(daemon_guid)
```

9.14 Collecting Telemetry

9.14.1 About Telemetry Collection

The telemetry collection feature is installed on all nodes. It runs *telemetry collection agents* on all Controller and Storage Nodes. The *telemetry collection master*, running on the Management Node, aggregates telemetry from the telemetry collection agents, including the agent running on the Management Node, encrypts the data using asymmetric keys, and forwards it to an HGST destination every 24 hours, at 3:00 a.m., using SSL for transport.

What is Collected

The data collected includes:

- Storage Enclosure Basic metrics
- System level information:
  - Rack serial number
  - Hardware inventory data
  - Time series data for system metrics
- IPMI data
- Object storage metrics
- Telemetry agent configuration
- Log file (`/var/log/hawk/callhome.log`)

Where Data is Temporarily Stored

The telemetry collection master stores data from each telemetry collection agent in separate directories, named

`/mnt/hawk/callhome_data/`*`date`*`/`*`node_MAC_address`*`/`. In the HGST destination, the data is stored in separate *`date`*`/`*`node_MAC_address`*`/` directories.

How Long Data is Retained

All collected data is kept on all telemetry collection agents in `/mnt/hawk/callhome_data` for 7 days.

How Failovers are Handled

If a node is down when a telemetry collection agent is supposed to run:

- No data is collected for that node, and no directory for that node is created in the dated directory in the HGST destination.
- The telemetry collection master starts a disaster recovery script on all nodes.

If the Management Node goes down, the Active Archive System automatically initiates a failover to another Controller Node newly designated as the Management Node, and the telemetry collection master fails over to the new Management Node also. Since the virtual IP address for the Management Node stays the same after a failover, and all other nodes use this virtual IP address, nothing else changes.

9.14.2 Displaying Telemetry Collection Categories

To see what categories of data are collected, run the following command from the Linux prompt of the Management Node:

```
/mnt/hawk/callhome/callhome.py --list-metrics category
```

The `callhome.py` command has the following options:

Option	Description
`--list-metrics`	Display the categories of metrics collected. Valid values are: • `all`: display all categories of metrics collected. • `IPMI`: display categories of metrics collected from the Intelligent Platform Management Interface. • `LSHW`: display categories of metrics collected from the Hardware Lister. • `AD`: display categories of metrics collected from Active Archive System logs. • `COLLECTL`: display categories of metrics collected from server component performance statistics. • `JBOD`: display categories of metrics collected from Storage Enclosure Basic status and inventory data. If a value is omitted for this option, the category headings are displayed.

1. **Log into the Management Node using SSH.**
   The OSMI menu appears.
2. **Exit the OSMI menu.**
   The Linux prompt appears.
3. **Run the `callhome.py` script:**

   ```
   /mnt/hawk/callhome/callhome.py --list-metrics category
   ```

   For example:

   ```
   /mnt/hawk/callhome/callhome.py --list-metrics
   ['IPMI','LSHW','AD','COLLECTL','JBOD']
   ```

   ```
   /mnt/hawk/callhome/callhome.py --list-metrics all
   [IPMI]
   Temperature (CPU PCH System Peripheral VcpuVRM VmemABVRM VmemCDVRM) FAN
    status Power status
   Chassis status Disk status Memory status Network status
   [LSHW]
   System Bus Memory Processor Bridge Network Storage Disk Volume Input
    CommunicationDisplay Power
   [AD]
   {'Event Logs': 'Put events Get events',
   'Storage Nodes': 'Agents Daemons Network Partitions System',
   'Controller Nodes': 'Client Daemons MetaStore'}
   [COLLECTL]
   {'Network': 'RxPkt TxPkt RxKB TxKB RxErr RxDrp RxFifo RxFra RxCmp RxMlt
    TxErr TxDrp
   TxFifo TxColl TxCar TxCmp RxErrs TxErrs',
   'Disk': 'Name Reads RMerge RKBytes Writes WMerge WRBytes Request QueLen Wait
    SvcTim Util',
   'NFS': 'ReadsS WritesS MetaS CommitS Udp Tcp TcpConn BadAuth BadClient
    ReadsC WritesC MetaC
   CommitC Retrans AuthRef',
   'CPU': 'Sys User Nice Wait IRQ Soft Steal Idle Totl Intrpt Intrpt/sec Ctx/
   sec Proc/sec
   ProcQue ProcRun L-Avg1 L-Avg5 L-Avg15',
   'Memory': 'Tot Used Free Shared Buf Cached Slab Map Commit SwapTot SwapUsed
    SwapFree SwapIn
   ```

```
SwapOut Dirty Clean Laundry Inactive PageIn PageOut PageFaults PageMajFaults
 HugeTotal
HugeFre HugeRsvd SUnreclaim'}
[JBOD]
Vendor ID Product ID Product revision level Unit serial number Tick counter
 Monitor loop
counter Monitor loop recent latencey Monitor loop maximum latency Offline
 state reason mask
Power state PSU A AC failure counter PSU B AC failure counter PHY reset -
 last ID
PHY reset - event count BIST failure - event count Enclosure status
 Temperature sensors
Voltage sensors Current sensors
```

```
/mnt/hawk/callhome/callhome.py --list-metrics IPMI
Temperature (CPU PCH System Peripheral VcpuVRM VmemABVRM VmemCDVRM) FAN
 status
Power status Chassis status Disk status Memory status Network status
```

```
/mnt/hawk/callhome/callhome.py --list-metrics JBOD
Vendor ID Product ID Product revision level Unit serial number Tick counter
 Monitor loop
counter Monitor loop recent latencey Monitor loop maximum latency Offline
 state reason mask
Power state PSU A AC failure counter PSU B AC failure counter PHY reset -
 last ID
PHY reset - event count BIST failure - event count Enclosure status
 Temperature sensors
Voltage sensors Current sensors
```

```
/mnt/hawk/callhome/callhome.py --list-metrics AD
{'Event Logs': 'Put events Get events',
'Storage Nodes': 'Agents Daemons Network Partitions System',
'Controller Nodes': 'Client Daemons MetaStore'}
```

```
/mnt/hawk/callhome/callhome.py --list-metrics COLLECTL

{'Network': 'RxPkt TxPkt RxKB TxKB RxErr RxDrp RxFifo RxFra RxCmp RxMlt
 TxErr TxDrp
TxFifo TxColl TxCar TxCmp RxErrs TxErrs',
'Disk': 'Name Reads RMerge RKBytes Writes WMerge WRBytes Request QueLen Wait
 SvcTim Util',
'NFS': 'ReadsS WritesS MetaS CommitS Udp Tcp TcpConn BadAuth BadClient
 ReadsC WritesC MetaC
CommitC Retrans AuthRef',
'CPU': 'Sys User Nice Wait IRQ Soft Steal Idle Totl Intrpt Intrpt/sec Ctx/
sec Proc/sec
ProcQue ProcRun L-Avg1 L-Avg5 L-Avg15',
'Memory': 'Tot Used Free Shared Buf Cached Slab Map Commit SwapTot SwapUsed
 SwapFree SwapIn
SwapOut Dirty Clean Laundry Inactive PageIn PageOut PageFaults PageMajFaults
 HugeTotal
HugeFre HugeRsvd SUnreclaim'}
```

```
/mnt/hawk/callhome/callhome.py --list-metrics LSHW
System Bus Memory Processor Bridge Network Storage Disk Volume Input
 CommunicationDisplay Power
```

10 Tuning the Active Archive System

Topics:

- Tuning for Optimal Connection Management
- Tuning for Maximum Throughput
- Tuning for Optimal Repair Performance
- Optimal Performance Vs Maximum Number of Concurrent Users
- Updating the Time Settings

This section describes some configuration options to optimize your Active Archive System installation.

10.1 Tuning for Optimal Connection Management

By default, your client daemon will accept any incoming TCP connection and process it. Our measurements show that there is no real performance benefit in exceeding more than 100 concurrent connections per client daemon. This section shows how to configure your client daemon to prevent that more that 100 concurrent connections are made at the same time:

To tune your Active Archive System for optimal connection management, proceed as follows:

1. Connect to your Controller Node and enter the Q-Shell.
2. Open the configuration files of your client daemon: `/opt/qbase3/cfg/dss/clientdaemons/`*`guid`*`.cfg` .
3. Locate the `http_max_conn` parameter in the `[config]` section. Add it if this parameter is not present.

   This parameter defines the maximum number of parallel streams to this client daemon. The value of `http_max_conn` is empty by default (unlimited).
4. Change the value to `100`.
5. Save and close the configuration file.
6. Restart the updated client daemon in the Q-Shell:

```
q.dss.clientdaemons.restartOne(daemonguid)
```

When a client application attempts to make more than the configured maximum number of connections to the client daemon, we will accept this TCP connection (up to the configured number of TCP connections that can be in the accepted state) but we will not process this connection yet. We will process it as soon as the number of HTTP connections that is being processed by the client daemon, is lower than the configured number.

10.2 Tuning for Maximum Throughput

Note: The Active Archive System is preconfigured for maximum throughput; you do not have to change any of the files described in the following sections.

Client Daemon

Open the client daemon configuration file:

`/opt/qbase3/cfg/dss/clientdaemons/<guid>.cfg` .

Update the following parameters in the `[config]` section:

```
max_node_connections = 6000
max_node_connections_per_blockstore = 1
max_environment_syncstore_connections = 16
max_object_syncstore_connections = 360
max_object_syncstore_connectionsper_per_object_syncstore = 30
max_open_file_descriptors = 8192
```

This gives the client daemon process a capacity for over 1,000 incoming HTTP connections.

Storage Daemon

Open the storage daemon configuration file:

`/opt/qbase3/cfg/dss/storagedaemons/<guid>.cfg .`

Update the following parameters in the `[config]` section:

Add them if they are not yet present:

```
max_node_connections = 512
max_node_connections_per_blockstore = 16
max_environment_syncstore_connections = 4
max_object_syncstore_connections = 48
max_object_syncstore_connectionsper_per_object_syncstore = 4
max_open_file_descriptors = 8192
```

When updating the `max_node_connections` , you have to restart the storage daemons and monitoring agents on all storage nodes in order to become effective.

Maintenance Agent

Open the maintenance agent configuration file:

`/opt/qbase3/cfg/dss/maintenanceagents/<guid>.cfg .`

Update the following parameters in the `[config]` section.

Add them if they are not yet present:

```
max_node_connections = 6000
max_node_connections_per_blockstore = 1
max_environment_syncstore_connections = 4
max_object_syncstore_connections = 48
max_object_syncstore_connectionsper_per_object_syncstore = 4
max_open_file_descriptors = 8192
```

10.3 Tuning for Optimal Repair Performance

When you install the Active Archive System, the settings for repair are optimized for large objects. If you have one or more name spaces that have many small files (millions), make the following change to the storage daemons that are master for those name spaces:

1. Open the OSMI menu.
2. Identify the master storage daemon of a name space in OSMI by listing the name spaces: **OSMI** > **Policies and Namespaces** > **List Namespaces**.
3. In the configuration file of the storage daemon for those name spaces, change `repair_queue_max_size`.

   The default is set to `10,000`. Increase this value by increments of `4,000` and see what the effect is.

10.4 Optimal Performance Vs Maximum Number of Concurrent Users

You can tune the Active Archive System client daemons for two situations:

- Slow upload/download but with maximum number of concurrent users.
- Fast upload/download with a maximum performance.

You have to decide which type of tuning your installation requires.

Typically a Controller Node has four client daemons and 64 GiB of memory. On such a node, 16 GiB is reserved for the operating system, the management framework, and the MetaStores, which means that 48 GiB of memory is available for 4 client daemons, or 12 GiB per client daemon.

However, the settings described in the following sections are calculated for Controller Nodes with 32 GiB memory, of which 24 GiB is assigned to BitSpread, and only one client daemon. As mentioned, in real life, typically a Controller Node has four client daemons, which means then a 6 GiB per client daemon.

10.4.1 Tuning for Maximum Number of Concurrent Users

Up to 256 threads can transfer data to the blockstores, with a spread width of 18. This means that 18 superblocks can be sent simultaneously to the blockstores.

The memory requirement for the backend is calculated as follows: 3 x 18 x 64 = 3.4 GiB (backend memory limit)

- `3`: 1.5 (encoded superblock) + 1.5 (encoded superblock in network buffers)
- `18`: number of simultaneously sent superblocks
- `64`: superblock size in MiB

This means that there is 21 GiB left for the front end memory, able to serve 336 (= 21 GiB / 64 MiB) concurrent slow readers and writers.

To tune the client daemon for this situation:

1. Open the client daemon configuration file `/opt/qbase3/cfg/dss/clientdaemons/`*`guid`*`.cfg` .
2. Update the file with the following parameters in the `config` section. It is possible that the parameters are not available in the file:
   - `superblock_mem_limit`: 22548578304, 3221225472 (21 GiB and 3 GiB, respectively for frontend memory and backend memory limit)

     Note: Update accordingly per client daemon. In most situations, there are four client daemons per Controller Node, so divide the mentioned values by four (or by the number of installed client daemons).

     It is recommended to always provide the two values, but if only one is given, the value is used for both the frontend and the backend memory.

   - `nr_put_threads`: 1
   - `nr_get_threads`: 1
   - `rest_put_mode`: nonblocking
   - `rest_get_mode`: nonblocking
3. Restart the updated client daemon in the Q-Shell:

```
q.dss.clientdaemons.restartOne(daemonguid)
```

10.4.2 Tuning for Optimal Performance

Note: Optimal performance means a higher throughput for single-stream data transports.

For fast uploaders and downloaders, you can allow the superblock size up to 256 MiB.

Up to 256 threads can transfer data to the blockstores, with a spread width of 18. This means that 18 superblocks can be sent simultaneously to the blockstores.

The memory requirement for the back end is calculated as follows: 3 x 18 x 254 = 13.7 GiB.

- `3`: 1.5 (encoded superblock) + 1.5 (encoded superblock in network buffers)
- `18`: number of simultaneously sent superblocks
- `254`: superblock size in MiB

This means that there is 12 GiB left for the front end memory, able to serve 48 (= 12 GiB / 256 MiB) concurrent fast readers and writers.

To tune the client daemon for this situation, proceed as follows:

1. Open the client daemon configuration file: `/opt/qbase3/cfg/dss/clientdaemons/`*`guid`*`.cfg`)
2. Update the file with the following parameters in the `config` section. It is possible that the parameters are not available in the file:
    - `superblock_mem_limit`: 12884901888,12884901888 (both 12 GiB, respectively for frontend memory and backend memory limit)

        > **Note:** Update accordingly per client daemon. In most situations, there are four client daemons per Controller Node, so divide the mentioned values by four (or by the number of installed client daemons).

        It is recommended to always provide the two values, but if only one is given, the value is used for both the frontend and the backend memory.
    - `nr_put_threads`: 8
    - `nr_get_threads`: 8
    - `rest_put_mode`: blocking
    - `rest_get_mode`: blocking
3. Restart the updated client daemon in the Q-Shell.

```
q.dss.clientdaemons.restartOne(daemonguid)
```

10.4.3 In-depth Information

Assume that you have a Controller Node with a frontend connection 2 x 1 GbE (gibibit/s, half-duplex) and 32 GiB memory. With the default settings of 32 MiB superblocks and 1 GiB superblock_mem_limit, the Active Archive System can process 32 streams (1 GiB / 32 MiB). With the 2 x 1 GbE inbound, BitSpread can handle 256 MiB/s, so each stream should at least deliver 8 MiB/s in a best-case scenario (256 MiB/s / 32 streams). However, a best-case scenario is almost impossible to achieve.

In reality, you have to set the superblock_mem_limit to 1/5th of the available memory for the client daemon. In this example, a maximum of 24 GiB of memory can be assigned to BitSpread, the remaining is taken for the Controller Node operating system and other processes. So a 4.8 GiB (24 GiB / 5) can be set as superblock_mem_limit, meaning that each stream should deliver 2 MiB/s.

Users have typically an upload speed of 128 KiB/s or more, so the Active Archive System must be capable of streaming a maximum of 2048 streams in parallel (256 MiB / 128 KiB).

When a superblock is ready to be written on disk, BitSpread needs four times the superblock size:

- 1 time for encryption
- 1.5 times for encoded version
- 1.5 times for encoded version in network buffers

This means that BitSpread needs 128 MiB of memory to encode and upload one superblock.

The upload of data by a user is a slow process, whereas the writing of the superblocks to the blockstores is a very fast process. BitSpread uses two resource managers to manage the memory consumption.

- **source**: for managing the uploads by the user (PUT request) or download requests of the user (GET request).
- **transfer**: for managing the superblocks to write to the blockstores (PUT request by user) or to retrieve from the blockstores (GET request).

10.5 Updating the Time Settings

10.5.1 Date and Time in the Active Archive System

The Management Node is the reference for the date and time of your Active Archive System installation. All nodes synchronize their date and time with this node.

Controller Nodes temporarily set the date and time at installation time. But once the NTP service is set up during the installation, Controller Nodes synchronize with the NTP server you specified at bringup. You must ensure that the Management Node has an active link with an NTP server.

It may occur that the date and time is incorrect on the Management Node, which will be reflected on the Storage Nodes. However, in such a situation the date and time on the nodes are consistent and have very little to no impact on the functionality of your installation. If the time difference is small, then the Controller Node will synchronize with the NTP server again. If the time difference would be too large, then the NTP service on the Controller Node must be restarted.

10.5.2 Selecting an NTP Server

To update the NTP server and the active time zone, proceed as follows:

If the time difference between the current Controller Node time and the new timezone too big, or if you want to synchronize the nodes immediately, restart the NTP service on the Management Node via an SSH session.

```
/etc/init.d/ntp restart
```

Making a time-leap on the Controller Nodes holding a MetaStore, could make them appear as if they are not online. Restarting the MetaStores on those Controller Nodes resolves this situation.

1. In the CMC, navigate to **Dashboard** > **Administration** > **HGST Object Storage Management** > **Installation**.
2. In the **Commands** pane, click **Set datacenter timezone**.
   The **Set datacenter timezone** wizard appears.
3. Select the proper **timezone** and confirm.
4. In the **Commands** pane, click **Configure NTP**.
   The **Configure NTP** wizard appears.
5. In this wizard, type the **URL** or **IP address** of the NTP server, or leave the default.

   Caution: Make sure that the Controller Node can reach the NTP server, by IP address or through DNS.

6. Click **Next** to apply the changes.

A Glossary

Topics:

- Application Server
- Blacklists
- Cache Daemon
- Client Daemon
- Cloud Management Center (CMC)
- Controller Node
- Disk Safety
- HTTPS Proxy Server
- Intelligent Platform Management Interface (IPMI)
- Maintenance Agent
- Management Node
- MetaStore
- Object Store Management Interface (OSMI)
- Repair Spread
- Storage Daemon
- Storage Node
- Superblock

The Active Archive System consists of the following components.

A.1 Application Server

The application server is the component responsible for exposing various services over XML-RPC. The definition and implementation of an application server service is completely separated from the underlying transport. Each node has a running application server, but not all nodes have the same application server services running.

A.2 Blacklists

A blacklist is a list of blockstores that cannot be used for read/write operations. The reasons for a blockstore not being able to handle a read/write operation are:

- Broken disk
- File system error
- File system full
- Network failure
- I/O errors

When a client daemon detects that a blockstore cannot perform a read or write, it adds the blockstore to a blacklist. The blockstores on the blacklist are not contacted for further read/write operations, temporarily, to save time and network bandwidth. Every 30 seconds, the Active Archive System checks whether the blockstores on the blacklist are operational again. Any operational blockstore is removed from the blacklist. Non-operational blockstores remain on the blacklist. For more information on how the Active Archive System checks blockstores on the blacklist, see Handling Blacklists on page 48.

A.3 Cache Daemon

The BitSpread cache daemon is a service/process/daemon running on each Controller Node. There are 4 client daemons per Controller Node.

- The cache daemon caches superblocks (data) for frequently read objects.
- Each cache daemon can read from any physical cache location.
- The actual cache resides on one or more SSDs on one or more Controller Nodes.
- The cache is configurable in size upon creation, but the size cannot be modified afterwards and requires a intervention from HGST Support to remove.
- The client daemon process checks for cached data from the cache daemons before requesting the data from backend storage on object read requests.

A.4 Client Daemon

The BitSpread client daemon is a service/process/daemon running on each Controller Node.

- The client daemon handles REST read, write, update, and delete requests for data.
- The client daemon encodes and decodes data as requested from applications, which are external to the storage system.
- A single Controller Node can have one or more client daemons running.
- Each client daemon has the same view of the underlying storage.
- Each client daemon listens for REST requests on all public networks, but on a unique port to that Controller Node (`7070`, `7071`, `7072`, and `7073`).
- Each client daemon communicates internally on a unique port (`23510`, `23511`, ?).

A.5 Cloud Management Center (CMC)

The Cloud Management Center (CMC) is a web-based interface to the Active Archive System that supports the following administrative tasks:

- Managing Storage
  - Managing buckets
  - Managing MetaStores
  - Managing cache clusters
  - Monitoring the active storage pool
- Managing the Active Archive System
  - Managing the installation in general
  - Managing the rack and data center information
  - Collecting configuration information
  - Monitoring the existing policies
  - Monitoring the logs of jobs, events and policies
  - Managing the LAN networks
  - Managing users and groups
- Managing Hardware
  - Managing the status of all Controller Nodes and Storage Nodes
  - Managing uninitialized, queued, or failed devices
  - Managing degraded or decommissioned disks
  - Managing system health
- Downloading Content

- Exporting a list of executed events (only available for 24 hours)
- Exporting MIB variables (for uploading to your SNMP server)
- Downloading licenses

A.6 Controller Node

Controller Nodes are high performance servers in the Active Archive System. Controller Nodes are prepackaged with the Active Archive System software, MetaStore, and management framework. They provide high performance access over multiple network interfaces, and can serve data over the following network protocols:

- HTTP/REST
- S3

Controller Nodes are equipped with additional ports that are used by the backend storage pool.

A Controller Node operates in a high availability cluster to provide fully shared access to the storage pool, metadata caching in high-performance Solid State Disks (SSDs), and metadata protection.

A Controller Node consists of the following components:

- Client Daemon on page 110
- Cache Daemon on page 110
- MetaStore on page 112

A.7 Disk Safety

The disk safety is the number of blockstores from a specific spread that can be lost while still being able to recover the original data.

If more blockstores are lost than the disk safety allows, the data is no longer recoverable.

A.8 HTTPS Proxy Server

The Active Archive System provides HTTPS functionality, in order to have encrypted data communication to the public interfaces (the Cloud Management Center (CMC) and the S3 interface).

The Active Archive System uses pound, a reverse SSL proxy, for this purpose.

`pound` uses OpenSSL for its encryption/decryption. Therefore, all certificates supported by OpenSSL can be used. For more information about `pound`, see http://www.apsis.ch/pound.

A.9 Intelligent Platform Management Interface (IPMI)

IPMI (Intelligent Platform Management Interface) is a standardized computer system interface, used by system administrators to manage a remote computer system and monitor its operation. Both Controller Nodes and Storage Nodes are equipped with this interface. Controller Nodes have a dedicated physical IPMI network interface card (NIC), while the Storage Nodes have a shared IPMI NIC.

IPMI is used for the following purposes:

- Remote Power Cycle on page 47
- Remote Capture of Screen and Keyboard on page 47
- Viewing Sensor Readings on page 47
- Toggling a Location LED on page 47

A.10 Maintenance Agent

The maintenance agent is a component of the BitSpread technology. It is a service/process/daemon running on each Storage Node, but it can be configured for any node in the environment. There are 14 maintenance agents per Storage Node on SA-7000.

The maintenance agent is responsible for the self-repairing nature of the BitSpread storage backend.

- The maintenance agent polls storage daemons for:
  - Objects to be repaired.
  - Objects to be deleted.
- The maintenance agent instructs the storage daemon responsible for the name space, to conduct the actual deletion process.

Note: The polling for repair work is done every 15 minutes.

The maintenance agent works closely with the storage daemon. It does not need the client daemon to get access to the BitSpread storage backend.

A.11 Management Node

The Management Node is a logical component. It is the designation of the Controller Node 01 in Rack 01 (in other words, `*SystemID*-DC01-R01-CN01`). The Cloud Management Center (CMC) is pre-installed on this node.

The main functions of the Management Node are:

- Storage management
- Job scheduling

A.12 MetaStore

The metadata store, or MetaStore, contains information about metadata of objects, blockstores, superblocks, spreads, policies, and name spaces. The underlying technology is the Arakoon distributed key-value store. A MetaStore is implemented as an Arakoon *cluster* running on all three Controller Nodes. In an Arakoon cluster, one Controller Node is selected as master and reports the cluster (MetaStore) status to the Cloud Management Center (CMC).

- The MetaStore runs on solid state disks (SSD) for high input/output operations per second (IOPS).
- Each MetaStore consists of three SSDs on Controller Nodes for high availability.
- The three SSDs comprising the MetaStore are together considered a *cluster*.
  - Each participating Controller Node is considered a *node*.
  - Each service is considered an *instance*.
- To write and retrieve data from a MetaStore, you need a majority of the participating instances available. For example, on a three-node MetaStore, two instances must be available.
- Each instance has an associated transaction log (*tlog*), which resides on a hard disk drive.
- You only need one intact copy of the database or tlogs to rebuild the entire Arakoon cluster, guaranteeing data safety when multiple components fail.
- Each name space has one associated MetaStore.
- A MetaStore can service one or more name spaces.

For more information about Arakoon, see http://www.arakoon.org/.

A.13 Object Store Management Interface (OSMI)

The Object Store Management Interface (OSMI) is a menu-based management interface that does not require knowledge of Q-Shell or interactive Python.

Using the OSMI you can:

- Execute several tasks in your environment.
- Get information from your environment.

The OSMI is installed by default on each Controller Node.

A.14 Repair Spread

Repair spread is a safety strategy (property of a policy) for objects.

When an upload of a superblock fails, the repair spread attempts a new upload for that superblock, but with a different spread. The upload fails when no valid spread can be found.

The repair spread is the number of blockstores over which your data is distributed upon. For example, a spread width of 18 indicates that all data is spread over 18 different blockstores.

A.15 Storage Daemon

The BitSpread storage daemon is a service/process/daemon running on each Storage Node. There are 6 14 storage daemons per Storage Node on SA-7000.

- The storage daemon receives requests from the client daemon(s) and maintenance agent(s) and acts as a gateway for requests to blockstores.
- Each storage daemon is responsible for a set of blockstores on the node.
- Every storage daemon listens on all network segments, but listens on a unique port for that Storage Node (`23520`, `23521`).
- Each bucket has one storage daemon that acts as an entry point for the information management in that bucket. This is referred to as the *master storage daemon* for that bucket.
- The main responsibilities of a master storage daemon for a bucket are:
  - Providing troubleshooting information
  - Keeping a list of objects that need to be repaired or deleted. It enumerates objects to be repaired every four hours.
- A storage daemon can be the master for multiple name spaces.

A.16 Storage Node

Storage Nodes provide high-density and power-efficient storage for the Active Archive System. Each Storage Node is paired with a Storage Enclosure Basic storage array.

A Storage Node consists of the following components:

- Storage Daemon on page 113
- Maintenance Agent on page 112

A.17 Superblock

A superblock is the logical storage unit for the BitSpread storage backend. This unit is not visible to the end user.

When storing an object in the BitSpread storage backend, the client divides the object into one or more superblocks before encoding the data itself. In a typical BitSpread setup, superblock sizes vary between 1 MiB - 256MiB.

B Ports, Protocols, and Services

Topics:

- General Services
- PXE Services
- NFS (RPC)
- Application Server
- Arakoon
- BitSpread/DSS

The Active Archive System uses the following ports and services by default.

Abbreviations Used in the Tables

- **Dest Port**: destination port, port on the defined destination to access the service
- **Src**: source, application or machine initiating a request
- **Dst**: destination, machine handling an incoming request
- **Mgmt**: Management Node
- **Ctrl**: Controller Node
- **Ctrl DC**: Controller Node in each data center
- **Stor**: Storage Node
- **MGMT**: management network
- **STOR**: storage network
- **PUB**: public network
- **WFE**: workflow engine
- **Random**: port is randomly chosen by the port mapper
- **AMF**: Action Message Format, for sending objects from the CMC to the Active Archive System backend.

B.1 General Services

Service	Dest Port	TCP/UDP	Src	Dst	Network	Process
Apache	80	TCP	Client browser, all nodes	Mgmt	MGMT/STOR/ PUB	httpd
WFE	9876	TCP	localhost	Mgmt	MGMT/STOR	python
rsync	7777	TCP	Mgmt	All nodes	MGMT	rsync
postgres	5432	TCP	localhost	Mgmt	MGMT/STOR	postgres
NTP Server	123	UDP	All nodes	Mgmt	MGMT/ STOR(*)	ntpd
SNMP	161	UDP	Any SNMP poller	Mgmt	PUB	snmpd
Agent controller (ejabberd)	5222, 5223, random	TCP	All nodes	Mgmt	MGMT/STOR	beam.smp
Agent controller (ejabberd)	4369	TCP	All nodes	Mgmt	MGMT/STOR	epmd
Pound (UI)	443 (**)	TCP	Any client sending data to the Active Archive System	Mgmt	PUB	pound
SSH	22	TCP	All nodes	All nodes	MGMT/STOR/ PUB	sshd

Note:

- (*): The public network is used when the Management Node synchronizes with an external NTP server, such as the default NTP server ntp.pool.org
- (**): The port for `Pound` is the default port and can be changed in the CMC

B.2 PXE Services

Service	Dest Port	TCP/UDP	Src	Dst	Network	Process
TFTP	69	UDP	All nodes	Mgmt	MGMT	in.tftpd
DHCP	9991	TCP	All nodes	Mgmt	MGMT	dhcpd
DHCP-helper	67	UDP	All nodes	Ctrl	MGMT	dhcp-helper

Note:

- PXE services are only required during factory installation of the Active Archive System.
- The DHCP helper service runs on one Controller Node per data center in a multi-geo setup. All nodes within that data center can connect to this Controller Node.

B.3 NFS (RPC)

Service	Dest Port	TCP/UDP	Src	Dst	Network	Process
mountd	Random	TCP/UDP	All nodes	Mgmt	MGMT	rpc.mountd
statd	Random	TCP/UDP	All nodes	Mgmt	MGMT	rpc.statd
portmapper	111	TCP/UDP	All nodes	Mgmt	MGMT	rpcbind

B.4 Application Server

Service	Dest Port	TCP/UDP	Src	Dst	Network	Process
XMLRPC	8888	TCP	localhost	localhost	n/a	python (twistd)
REST	8889	TCP	localhost	localhost	n/a	python (twistd)
PXE	8890	TCP	localhost	localhost	n/a	python (twistd)
AMF	8899	TCP	localhost	localhost	n/a	python (twistd)

Note: The XMLRPC, REST, PXE, and AMF services run on the Management Node and can only be accessed through the Apache service, which also runs on the Management Node.

B.5 Arakoon

The *client port* is the port on the Controller Node through which an Arakoon client can access the Arakoon cluster, for example the `framework` cluster. The *messaging port* is a port for communication between the different Arakoon nodes.

B.5 framework

Port Type	Dest Port	TCP/UDP	Src	Dst	Network	Process
Client	9001, 9002, ...	TCP	Ctrl	Ctrl	MGMT/STOR	arakoon
Messaging	9001, 9002, ...	TCP	Ctrl	Ctrl	MGMT/STOR	arakoon

B.5 env_metastore

Port Type	Dest Port	TCP/UDP	Src	Dst	Network	Process
Client	9001, 9002, ...	TCP	All nodes (DSS daemons)	Ctrl	MGMT/STOR	arakoon
Messaging	9001, 9002, ...	TCP	Ctrl	Ctrl	MGMT/STOR	arakoon

B.5 object_metastore

The object_metastore can have any name.

Port Type	Dest Port	TCP/UDP	Src	Dst	Network	Process
Client	9001, 9002, ...	TCP	All nodes (DSS daemons)	Ctrl	MGMT/STOR	arakoon
Messaging	9001, 9002, ...	TCP	Ctrl	Ctrl	MGMT/STOR	arakoon

B.6 BitSpread/DSS

B.6 Client Daemon

Port Type	Dest Port	TCP/UDP	Src	Dst	Network	Process
Internal	23510, 23511, ...	TCP	localhost	localhost	n/a	dss
S3 external	7070,7071,7072,7073	TCP	Client app (no HTTPS)	Ctrl	PUB	dss
S3 external	7070,7071,7072,7073	TCP	Client app (HTTPS)	Ctrl	PRIV	dss
Pound	Next available	TCP	Client app	Ctrl	PUB	pound

Note: The ports for S3 are the default ports and can be changed in the Cloud Management Center.

The `pound` port is to be set manually, by default the next available port is proposed.

B.6 Other BitSpread/DSS Services

Service	Dest Port	TCP/UDP	Src	Dst	Network	Process
Cache Daemon	9999, 10000, ...	TCP	Ctrl	Ctrl	MGMT/STOR	dss
Storage Daemon	23520, 23521, ...	TCP	All nodes (DSS daemons)	Stor	MGMT/STOR	dss
Maintenance Agent	none	TCP	none	Stor	MGMT/STOR	dss

C Events

Topics:

- Events Explained
- Complete List of Events
- Events in Detail

These topics provide a complete list of Active Archive System events.

C.1 Events Explained

Event Property	Explanation
Stored in DRP	This event is stored in the Active Archive System database and retained for 7 days
Trapped over SNMP	When this event occurs, the event is also trapped over SNMP, if configured
Sent through email	When this event occurs, the event is also sent over SMTP, if configured
Monitoring interval	When monitoring state, the monitoring agent will only check for this specific state once in the monitoring interval
Dedupe period	When the same state is detected in the dedupe period, only a single event is issued

C.1 About Event Deduplication

Event deduplication is the act of updating the number of occurrences and time stamp of the last occurrence of an existing event, instead of creating a new event. The deduplication occurs for events which have the following properties:

- Identical event type; and
- Same severity; and
- Same source/machine; and
- Same event message

The events must also occur within the dedupe period. For example, event X has a dedupe period of 15 minutes. When the same event occurs within the dedupe period of 15 minutes, the properties "Occurrences" and "Last occurrence" is updated. If the same event X occurs beyond the dedupe period, then a new event is added to the list of events.

The dedupe period is always counted from the first occurrence of an event.

C.2 Complete List of Events

Tip

- When the Monitoring interval is "-", then the value is variable.
- (*): Depending on the percentage, the severity can increase from warning to critical.(80% - 85%, 90%)
- Take into account while spotting your exact event that any value between the "<" and ">" brackets is a variable.

C.2 Agent Events

Agent events have the following format: OBS-AGENT-*EventID*, for example OBS-AGENT-0006.

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0006	Machine agent is down on <machine_name>.	CRITICAL	yes	no	yes	30 min (monitor policy)	1 h

C.2 Application Events

Application events have the following format: OBS-APPLICATION-*EventID*, for example OBS-APPLICATION-0001

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0001	Cache daemon <daemon_id> is <info_status> but <status>.	WARNING	yes	no	yes	15 min	1 h
	Storage daemon <daemon_id> is <info_status> but <status>.	WARNING	yes	no	yes	10 min	1 h
0006	Machine agent application has status <status>.	ERROR	yes	yes	yes	30 min	1 h
0025	Backup of application <application_name> failed.	CRITICAL	yes	yes	yes	1 day (qpserver backup policy)	1 d
	Osis backup failed.	CRITICAL	yes	yes	yes	1 day (osis backup policy)	1 d
	Osis backup failed. Error: <error>.	CRITICAL	yes	yes	yes	1 day (osis backup policy)	1 d
0047	Too many open files <nr_of_files> for cache daemon <id>.	ERROR	yes	yes	yes	15 min	1 h
	Too many open files <nr_of_files> for client daemon <id>.	ERROR	yes	yes	yes	10 min	1 h
	Too many open files <nr_of_files> for maintenance agent <id>.	ERROR	yes	yes	yes	10 min	1 h
	Too many open files <nr_of_files> for storage daemon <id>.	ERROR	yes	yes	yes	10 min	1 h
	Too many open files <nr_of_files> for node <node_name> on MetaStore <name>.	ERROR	yes	yes	yes	5 min	1 h
0049	Metastore cluster <name> is DOWN.	CRITICAL	yes	yes	yes	5 min	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0050	Failed to create cloudAPI. please check the logs.	ERROR	yes	no	yes	-	-
0051	Duplicate agent sessions found for agent <agent_guid>.	WARNING	yes	yes	yes	30 min	1 h
0052	Too many incoming connections <nr_of_connections> for cache daemon <id>.	ERROR	yes	no	yes	15 min	1 h
	Too many incoming connections <nr_of_connections> for client daemon <id>.	ERROR	yes	no	yes	10 min	1 h
	Too many incoming connections <nr_of_connections> for storage daemon <id>.	ERROR	yes	no	yes	10 min	1 h
0053	SSL certificate <certificate> has expired.	CRITICAL	yes	no	yes	10 min	1 h
	SSL certificate <certificate> does not exist.	ERROR	yes	no	yes	10 min	1 h
	SSL certificate <certificate> will expire in less than 5 days.	WARNING	yes	no	yes	10 min	1 h
0054	Cannot upgrade MetaStore/ MetaStoreclient from version X to version Y	ERROR	yes	yes	yes	n/a	1 h

C.2 MetaStore Events

MetaStore events have the following format: OBS-ARAKOON-`EventID`, for example OBS-ARAKOON-0001.

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0004	Collapsing of transaction logs failed for MetaStore <name>.	ERROR	yes	yes	yes	-	1 h
	Our MetaStore transaction log collapsing policy is not configured. It needs to be configured in order to run this policy.	ERROR	yes	yes	yes	-	1h
0005	The number of keys in the <name> MetaStore exceeds <critical_threshold>%.	CRITICAL	yes	yes	yes	5 min	1 h
	The number of keys in the <name> MetaStore exceeds <error_threshold>%.	ERROR	yes	yes	yes	5 min	1 h
	The number of keys in the <name> MetaStore exceeds <warning_threshold>%.	WARNING	yes	yes	yes	5 min	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0006	Metadata store '<cluster_name>' has no master node: can not re-initialize node '<node_name>'.	CRITICAL	yes	no	yes	-	1 h
	No master node for MetaStore '<name>'.	ERROR	yes	no	yes	-	1 h
0007	MetaStore safety low. Node <node>' missing for cluster <cluster>.	CRITICAL	yes	yes	yes	5 min	15 min
0008	Node <node_name> on MetaStore <cluster_name> is lagging <number_of_keys> keys.	WARNING	yes	yes	yes	5 min	1 h
	Node(s) '<node_name>' in MetaStore '<cluster_name>' are lagging behind.	ERROR	yes	yes	yes	-	1 h
0009	More than %s tlogs found on node %s of MetaStore %s.	CRITICAL	yes	yes	yes	5 min	1 h
0010	Database partition for MetaStore node <metastore_name>::<node_name> is more than <error_threshold>% full.	ERROR	yes	no	yes	5 min	1 h
	Database partition for MetaStore node <metastore_name>::<node_name> is more than <warning_threshold> % full.	WARNING	yes	no	yes	5 min	1 h
0011	Database partition for MetaStore node <metastore_name>::<node_name> is more than <x>% full.	CRITICAL (*)	yes	no	yes	5 min	1 h
0012	Transaction log partition for MetaStore node <cluster_name>::<node_name> is low on space.	ERROR	yes	no	yes	5 min	1 h
	Transaction log partition for MetaStore node <cluster_name>::<node_name> is low on space.	WARNING	yes	no	yes	5 min	1 h
0013	Node <node_name> in Metastore <cluster_name> will be stopped.	CRITICAL	yes	yes	yes	5 min	1 h
	Transaction log partition for MetaStore node <cluster_node>::<node_name> is low on space.	CRITICAL	yes	yes	yes	5 min	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0014	Could not collapse transaction logs for MetaStore '<name>'. Reason: <reason>.	CRITICAL	yes	yes	yes	-	1 h
0015	MetaStore master fail-over triggered on MetaStore '<name>'.	INFO	yes	no	no	-	0 min
0018	MetaStore instance <cluster_name>::<node_name> automatic recovery failed after <number> retries on machine <machine name>.	CRITICIAL	yes	yes	yes	-	0 min
0019	Found MetaStore backup file(s) on machine <machine name>	INFO	yes	no	no	5 min	1 h
0020	The number of keys in the X MetaStore exceeds Y%	CRITICAL	yes	yes	yes	5 min	1 h
0021	MetaStore X node Y usage thresholds fell below critical levels, setting to ACTIVE	INFO	yes	yes	yes	5 min	1 h

C.2 Disk Events

Disk events have the following format: OBS-DISK-*EventID*, for example OBS-DISK-0001.

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0001	Disk <old_diskname> was renamed to <new_diskname> after reboot.	INFO	yes	no	yes	15 min	1 h
0002	Disk X can't be detected on machine Y	CRITICAL	yes	yes	yes	15 min	15 min
0003	Problems found while detecting disks.	WARNING	yes	no	yes	15 min	1 h
0004	Hdparm security attributes are already enabled on disk. Disk <ID> cannot be erased.	ERROR	yes	no	yes	-	1 h
	Hdparm security attributes are frozen on disk. Disk (<uuid>) can not be erased.	ERROR	yes	no	yes	-	1 h
	Hdparm Failed to detect disk (<uuid>) attributes.	ERROR	yes	no	yes	-	1 h
0005	Disk(s) <disk_names> can't be decommissioned on machine <machine_name>.	ERROR	yes	no	yes	-	1 h
0006	Error: decommissioning disks failed. Reason : disks <disk_guids> are in the same raid.	ERROR	yes	no	yes	-	1 h
0007	New empty disk(s) detected.	INFO	yes	no	yes	15 min	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0008	New non-empty disk(s) detected: <disk_ids>.	ERROR	yes	no	yes	15 min	1 h
0009	Replacement disk size (<size> MB) is smaller than the original disk size <size> MB). on node <machine_name>.	ERROR	yes	no	yes	15 min	1 h
0010	Replacement disk type (<type>) is different from the original disk type (<original_type>) on node <machine_name>.	ERROR	yes	no	yes	15 min	1 h
0011	Can not detect physical location of disk [<ID>] on machine [<machine_name>].	ERROR	yes	yes	yes	15 min	1 h
0012	Disk bus location detected.	INFO	yes	no	no	15 min	1 h
0013	Partition '<name>' has too few free inodes (<inodes>).	WARNING	yes	no	yes	15 min	1 h
0014	Mountpoint <name> found in /etc/fstab but not mounted.	WARNING	yes	no	yes	monitoring agent startup	1 h
0015	Disk(s) not detected: <disk_ids>.	ERROR	yes	no	yes	15 min	15 min
0016	No disk found in the model with bus location <bus_location>.	ERROR	yes	no	yes	15 min	0 min
	More than one disk with bus location <bus_location> found.	ERROR	yes	no	yes	15 min	0 min
	Disk with bus location <bus_location> is not decommissioned in the model.	ERROR	yes	no	yes	15 min	0 min
0017	Filesystem sync is running for more than [<timeout>] seconds	ERROR	yes	yes	yes	5 min	default
0018	Decommissioned or autodecommissioning disk detected	INFO	no	no	no	15 min	-
0019	Disk needs urgent replacement	CRITICAL	yes	yes	yes	-	default
0020	Automatic decommissioning started for a disk	INFO	yes	yes	yes	n/a	default
0021	Automatic decommissioning completed for a disk	INFO	yes	yes	yes	n/a	default
0022	New empty disk <disk> on machine <machine> needs repurposing	ERROR	yes	yes	yes	n/a	default

C.2 DSS Events

DSS events have the following formats:

- OBS-BLOCKSTORE-*EventID*, for example OBS-BLOCKSTORE-0026.

- OBS-STORAGEPOOL-*EventID*, for example OBS-STORAGEPOOL-0001.

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0026	Critical threshold (<threshold>%) exceeded for block store <blockstore_path>.	CRITICAL	yes	yes	yes	10 min	15 min
0027	Blockstore <blockstore_path> has released disk space.	INFO	yes	no	no	10 min	1 h
0028	Error checkblock count (<count>) exceeded for blockstore.	ERROR	yes	yes	yes	10 min	1 d
	Warning checkblock count (<count>) exceeded for blockstore <blockstore_path>.	WARNING	yes	yes	yes	10 min	1d
	Error checkblock count (<count>) exceeded for blockstore '<blockstore_path>' with less than <percentage>% full copy files.	ERROR	yes	yes	yes	10 min	1 d
	Warning checkblock count (<count>) exceeded for blockstore '<blockstore_path>' with less than <percentage> full copy files.	ERROR	yes	yes	yes	10 min	1 d
0029	Critical checkblock count (<count>) exceeded for block store <blockstore_path>.	CRITICAL	yes	yes	yes	10 min	15 min
	Critical checkblock count (<count>) exceeded for block store '<blockstore_path>' with less than <percentage>% full copy files.	CRITICAL	yes	yes	yes	10 min	15 min
0030	Blockstore <blockstore_path> is OFFLINE.	WARNING	yes	yes	yes	10 min	1 h
	You have <offline_blockstores_count> blockstore(s) that is(are) temporarily offline. Blockstores are only in this status when some of your Storage Nodes are down.	WARNING	yes	yes	yes	-	1 h
	You have <offline_blockstores_count> blockstore(s) that is(are) offline. Offlined blockstores will only be used again when they are put online through an operator action	WARNING	yes	yes	yes	-	1 h
	You have <count> blockstore(s) that is(are) temporarily offline. Blockstores are only in this status when some of your Storage Nodes are down. You have <count> blockstore(s) that is(are) offline.	WARNING	yes	yes	yes	-	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
	Offlined blockstores will only be used again when they are put online through an operator action.						
0031	Blockstore <blockstore_path> has status DECOMMISSIONED for more than <nr> days.	ERROR	yes	no	yes	10 min	1 h
	Blockstore <blockstore_path> has status DECOMMISSIONED for more than <nr> days.	WARNING	yes	no	yes	10 min	1 h
0032	Block store IDs do not match for <blockstore_path> (<ID> - <ID>).	ERROR	yes	no	yes	10 min	1 h
0033	OFFLINE blockstores detected	WARNING	yes	no	no	10 min	1 h
0034	Could not access blockstore partition partition_path	CRITICAL	yes	yes	yes	10 min	1h
0001	Storage pool statistics not updated in the past day. Last update: <last_update>.	INFO	yes	yes	yes	1 h	1 day
0005	Found namespaces with a deprecated codec version.	ERROR	yes	no	yes	1 h	1 h
0025	Storage pool is more than 70% full.	WARNING	yes	yes	yes	30 min (monitor policy)	1 h
0026	Storage pool is more than 80% full.	ERROR	yes	yes	yes	30 min (monitor policy)	1 h
0027	Storage pool is more than 90% full.	CRITICAL	yes	yes	yes	30 min (monitor policy)	1 h
0028	Storage policies detected that are affected by a known issue, ...	CRITICAL	yes	no	yes	-	1h

C.2 Environment Events

Environment events have the following format: OBS-ENVIRONMENT-`EventID`, for example OBS-ENVIRONMENT-0001.

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0001	Policy '<policy_name>' is disabled	CRITICAL	yes	yes	yes	1 d	1 d
	Policy '<policy_name>' did not run in more than <interval> hours	CRITICAL	yes	yes	yes	1 d	1 d
	Policy '<policy_name>' never ran before	CRITICAL	yes	yes	yes	1 d	1 d
0002	Duplicate rack or data center IDs found	ERROR	yes	yes	yes	-	1 d

C.2 Generic Events

Generic events have the following format: OBS-GENERIC-*EventID*, for example OBS-GENERIC-0001.

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0001	Machine <machine_name> is HALTED.	CRITICAL	yes	yes	yes	30 min (monitor policy)	30 min
	Machine <machine_name> is down [status: <machine_status>].	CRITICAL	yes	yes	yes	30 min (monitor policy)	30 min
0002	Found a partially installed patch: <patch_name>.	WARNING	yes	yes	yes	1 h	30 min
0003	Failed to save job <action_name>.	WARNING	yes	no	yes	-	1 h
0004	Failed to initialize node <machine_name>.	CRITICAL	yes	no	yes	-	1 h
	Job '<name>' failed on machine <machine_name>.	CRITICAL	yes	no	yes	-	1 h
	Policy '<name>' failed on machine <machine_name>.	CRITICAL	yes	no	yes	-	1 h
0071	Core dump files found in /var/crash/ Filename: <name> (<type> – <date>).	WARNING	yes	yes	yes	1 d	1 h
0072	Too many instances running for policy <policy_name>, skipping execution.	CRITICAL	yes	no	yes	-	1 h
0073	Failed login attempt with username <username>.	WARNING	yes	no	yes	-	15 min
0074	Too many failed login (<threshold>) attempts in the last (<seconds>) seconds.	ERROR	yes	no	yes	-	1 h
0075	Failure while executing event handling logic for event type: <type ID>	CRITICAL	yes	yes	yes	-	1 h
0076	Errors while updating machine(s) configurations during failover	ERROR	yes	yes	yes	-	1 h

C.2 Network Events

Network events have the following format: OBS-NETWORK-*EventID*, for example OBS-NETWORK-0001.

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0001	NTP daemon is down.	WARNING	yes	no	yes	1 h	1 h
	HTTPS proxy is enabled but not running.	ERROR	yes	no	yes	10 min	1 h
	HTTPS proxy <proxy_type> port <port> is unavailable.	ERROR	yes	no	yes	10 min	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
	Application <applicationname> is HALTED.	ERROR	yes	no	yes	10 min	1 h
	Cache daemon <ID> has status <status>.	ERROR	yes	no	yes	15 min	1 h
	Unable to connect to cache daemon <ID>:<port>.	ERROR	yes	no	yes	15 min	1 h
	Client daemon <GUID>' has status <status>.	ERROR	yes	no	yes	10 min	1 h
	Unable to connect to client daemon <ID>:<port>.	ERROR	yes	no	yes	10 min	1 h
	Maintenance agent <ID> has status <status>.	ERROR	yes	no	yes	10 min	1 h
	Storage daemon <ID>' has status <status>.	ERROR	yes	no	yes	10 min	1 h
	Unable to connect to storage daemon <IP>:<port> (<exception>).	ERROR	yes	no	yes	10 min	1 h
	Metastore node <metastore_name>::<metastore_node> is DOWN.	ERROR	yes	no	yes	5 min	1 h
	Application monitoringagent down on node <machine_name> [Auto restart].	CRITICAL	yes	no	yes	30 min (monitor policy)	1 h
0003	NIC <name> (<mac_address>) in half duplex mode.	WARNING	yes	yes	yes	5 min	1 h
0004	NIC <name> (<mac_address>) has speed <speed> below threshold.	WARNING	yes	yes	yes	5 min	1 h
0005	HTTPS proxy is disabled but running as PID <pid>.	WARNING	yes	no	no	10 min	1 h
	Application <application_name> is RUNNING.	INFO	yes	no	no	15 min	1 h
	Cache daemon <ID> is UP.	INFO	yes	no	no	15 min	1 h
	Client daemon <ID> is UP.	INFO	yes	no	no	10 min	1 h
	Maintenance agent Cache <ID> is UP.	INFO	yes	no	no	10 min	1 h
	Storage daemon <ID> is UP.	INFO	yes	no	no	10 min	1 h
	Metastore node <metastore_name>:<node_name> is UP.	INFO	yes	no	no	5 min	1 h
0006	SMTP is not configured correctly.	WARNING	yes	no	no	-	1 h
0007	Cache cluster <ID> has status <status>.	ERROR	yes	no	no	15 min	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0008	Interface <dev> (<name>) is UP but not configured in /etc/network/interfaces.	WARNING	yes	no	no	monitoring agent startup	1 h
0009	NIC [<name>] renamed from '<old_name>' to '<new_name>'.	WARNING	yes	no	no	5 min	1 h
0010	NIC <name> (<mac_address>) has no IP <IP> configured.	ERROR	yes	no	no	5 min	1 h

C.2 Physical Machine Events

Physical machine events have the following format: OBS-PMACHINE-*EventID*, for example OBS-PMACHINE-0001.

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0001	<mountpoint> is more than <x>% full.	CRITICAL	yes	yes	yes	5 min	1 h
	<mountpoint> is more than <x>% full.	ERROR	yes	yes	yes	5 min	1 h
	<mountpoint> is more than <x>% full.	WARNING	yes	yes	yes	5 min	1 h
0002	Swap is not available.	ERROR	yes	yes	yes	15 min	1 h
	Swap usage is over <x>%.	WARNING	yes	yes	yes	15 min	1 h
0005	Cache daemon [<ID>] memory threshold exceeded (<x>MB).	WARNING	yes	yes	yes	15 min	1 h
	Cache daemon [<ID>] shared memory threshold exceeded (<x>MB).	WARNING	yes	yes	yes	15 min	1 h
	Client daemon [<ID>] memory threshold exceeded (<x>MB).	WARNING	yes	yes	yes	10 min	1 h
	Client daemon [<ID>] shared memory threshold exceeded (<x>MB).	WARNING	yes	yes	yes	10 min	1 h
	Maintenance agent [<ID>] memory threshold exceeded (<x>MB).	WARNING	yes	yes	yes	10 min	1 h
	Maintenance agent [<ID>] shared memory threshold exceeded (<x>MB).	WARNING	yes	yes	yes	10 min	1 h
	Storage daemon [<ID>] memory threshold exceeded (<x>MB).	WARNING	yes	yes	yes	10 min	1 h
	Storage daemon [<ID>] shared memory threshold exceeded (<x>MB).	WARNING	yes	yes	yes	10 min	1 h
	MetaStore node [<metastore_name>::<node_name>]	WARNING	yes	yes	yes	5 min	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
	memory threshold exceeded (<x>MB).						
	MetaStore node [<metastore_name>::<node_name>] shared memory threshold exceeded (<x>MB).	WARNING	yes	yes	yes	5 min	1 h
	Less than <x>% memory free: <x>%.	WARNING	yes	yes	yes	5 min	1 h
0006	Interface <name> (<mac_address>) has <x> transmission errors.	WARNING	yes	yes	yes	5 min	1 h
0007	Interface <name> is overloaded for <x>%.	WARNING	yes	yes	yes	5 min	1 h
0012	Disk <ID> has SMART failures [Overall Status 'UNKNOWN'].	CRITICAL	yes	yes	yes	15 min	1 h
0013	I/O errors on disk (<disk_name>).	CRITICAL	yes	yes	yes	30 min	1 h
	Filesystem errors on disk <disk_name>, partition <partition_name>.	CRITICAL	yes	yes	yes	30 min	1 h
0017	Software RAID array <device_path> has status <status>.	WARNING	yes	yes	yes	15 min	1 h
0019	Kernel dmesg errors detected.	ERROR	yes	yes	yes	30 min	1 h
0020	Mountpoint <mountpoint> is read-only.	CRITICAL	yes	yes	yes	15 min	1 h
0023	Unknown lantype.	CRITICAL	yes	no	no	-	1 h
	DHCP method not supported for <device> (<name>).	WARNING	yes	no	no	5 min	1 h
0024	Macaddress <macaddress> is already in use.	CRITICAL	yes	no	no	-	1 h
0025	Machine with guid <machineguid> has already a NIC with name <interfacename>.	ERROR	yes	no	yes	-	1 h
0026	Invalid IP address <ipaddress> specified.	ERROR	yes	no	yes	-	1 h
	Invalid IP address <ipaddress>.	ERROR	yes	no	yes	-	1 h
	Invalid IP address specified.	ERROR	yes	no	yes	-	1 h
0027	Failed to Retrieve application agent. Reason: could not find application called "cloudapi".	ERROR	yes	no	yes	-	1 h
0028	Unable to retrieve nic with macaddress <macaddress> or name <interfacename>.	ERROR	yes	no	no	-	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0033	Unable to stop application <applicationname>. Error: <exception>.	ERROR	yes	no	no	-	1 h
0040	Invalid nic configuration. Expected exactly one match.	ERROR	yes	no	no	-	1 h
0042	Unable to start application <name>. Error: <exception>.	ERROR	yes	yes	yes	-	1 h
	Unable to restart application <name>. Error: <exception>.	ERROR	yes	yes	yes	-	1 h
0045	Machine is not running.	ERROR	yes	no	no	-	30 min
0046	Agent not running, restarting agent.	WARNING	yes	no	no	-	1 h
0047	Restarting agent failed.	ERROR	yes	yes	yes	-	1 h
0049	Machine object has no agent guid.	ERROR	yes	no	no	-	1 h
0054	DNS resolving fails.	ERROR	yes	yes	yes	monagent startup	15 min
0054	Interface '<dev>' (<name>) is DOWN.	ERROR	yes	yes	yes	5 min	15 min
0054	Unable to contact default gateway <gateway>. Reason: <error>.	ERROR	yes	yes	yes	5 min	15 min
0055	IP address <ipaddressguid> already in use for lan with guid <languid>.	ERROR	yes	no	no	-	1 h
0061	Template name is not uniquely defined.	ERROR	yes	no	no	-	1 h
0067	Gateway <gateway> does not belong to any of the configured networks.	CRITICAL	yes	no	no	-	1 h
0104	Load average over the last 15 minutes is high (<load>).	WARNING	yes	yes	yes	15 min	1 h
0105	The SMART control on disk <ID> is disabled [Overall status: DISABLED].	CRITICAL	yes	yes	yes	15 min	1 h
0106	Machine was rebooted.	INFO	yes	yes	yes	monagent startup	1 h
0107	NTP cannot adjust time on node '<machinename>', time difference is greater than 1000 seconds, NTP needs to be restarted.	CRITICAL	yes	no	no	1 h	1 h
0108	Cannot shutdown Machine(s) <machine_names> and they will be skipped during shutdown process.	WARNING	yes	no	yes	-	1 h

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0109	Machine(s) <machine_name> failed to be shutdown.	WARNING	yes	no	yes	-	1 h
0110	Number of processes exceeds threshold: <nr> > <threshold>	WARNING	yes	no	yes	15 min	1 h
0111	Processes found with state dead / zombie <process_names>	ERROR	yes	no	no	15 min	1 h
0113	Machine <machine_name> is in status: STOPPING	WARNING	yes	no	no	30 min (monitoring policy)	-
0114	Fan speed ['<name>'] is below <speed> RPM.	CRITICAL	yes	no	yes	15 m	1 h
	Fan speed ['<name>'] is below <speed> RPM.	ERROR	yes	no	yes	15 m	1 h
	Fan speed ['<name>'] is below <speed> RPM.	WARNING	yes	no	yes	15 m	1 h
0115	PSU <name> failed.	ERROR	yes	no	yes	15 m	1 h
0116	Abnormal log file size detected for <file name>. Reached <size> MB in last hour	ERROR	yes	yes	yes	1 h	1 h

C.2 Storage Events

Storage events have the following format: OBS-STORAGE-`EventID`, for example OBS-STORAGE-0004.

Event ID	Event Message	Severity	Stored in DRP	SNMP Trap	Sent via Email	Monitoring Interval	Dedupe Period
0004	Objects found with low disk safeties.	CRITICAL	yes	yes	yes	1 h	15 min
0005	Current MetaStore configuration for storage daemon / client daemon / maintenance agent with ID X does not match desired configuration.	CRITICAL	yes	no	yes	At monitoring agent startup	30 min
0006	Unverified objects found	CRITICAL	yes	yes	yes	1 day	-

C.3 Events in Detail

C.3.1 Application Events

C.3.1 OBS-APPLICATION-0001

Details	Description
Event Message	One of the following: `Cache daemon daemon_id is info_status but status.`

Details	Description
	Storage daemon *daemon_id* is *info_status* but*status*.
Severity	WARNING
Solution	No action needed, the system will set the corresponding application to HALTED.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	10 or 15 minutes
Dedupe period	1 hour

C.3.1 OBS-APPLICATION-0006

Details	Description
Event Message	Machine agent application has status *status*.
Severity	ERROR
Solution	Investigate why the agent has this status by checking the agent logfile on the physical machine: `/opt/qbase3/var/log/applicationserver.log`
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	30 minutes
Dedupe period	1 hour

C.3.1 OBS-APPLICATION-0025

Details	Description
Event Message	One of the following: Backup of application *application_name* failed. Osis backup failed. Osis backup failed. Error: *error*
Severity	CRITICAL
Solution	Active Archive System Management Framework Internal Database Backup failed. Check the job log to figure out why this backup failed
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes

Details	Description
Monitoring interval	Set in either the qpserver or osis backup policy (default is 1 day).
Dedupe period	1 day

C.3.1 OBS-APPLICATION-0047

Details	Description

Event Message	One of the following:
	`Too many open files` *`nr_of_files`* `for cache daemon` *`id`*`.`
	`Too many open files` *`nr_of_files`* `for client daemon` *`id`*`.`
	`Too many open files` *`nr_of_files`* `for maintenance agent` *`id`*`.`
	`Too many open files` *`nr_of_files`* `for storage daemon` *`id`*`.`
	`Too many open files` *`nr_of_files`* `for node` *`node_name`* `on MetaStore` *`name`*`.`

Severity	ERROR
Solution	Investigate the cause of the high amount of file descriptors by executing the following command: `lsof -p` *`process_id_of_the_application`*
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Varying between 5-15 minutes.
Dedupe period	1 hour

C.3.1 OBS-APPLICATION-0049

Details	Description
Event Message	`Metastore cluster` *`name`* `is DOWN.`
Severity	CRITICAL
Solution	Consult the logfile and verify the cause of the failure. Consult HGST Support, if needed.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.1 OBS-APPLICATION-0050

Details	**Description**
Event Message	`Failed to create cloudAPI. Please check the logs.`
Severity	CRITICAL
Solution	Investigate the logs on the machine monitoring agent for more details.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	N.A.
Dedupe period	0

C.3.1 OBS-APPLICATION-0051

Details	**Description**
Event Message	`Duplicate agent sessions found for agent` *`agent_guid`*`.`
Severity	WARNING
Solution	Automatic recovery will take place, no manual intervention required.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	30 minutes
Dedupe period	1 hour

C.3.1 OBS-APPLICATION-0052

Details	**Description**

Event Message	One of the following:
	`Too many incoming connections` *`nr_of_connections`* `for cache daemon` *`id`*`.`
	`Too many incoming connections` *`nr_of_connections`* `for client daemon` *`id`*`.`
	`Too many incoming connections` *`nr_of_connections`* `for storage daemon` *`id`*`.`

Severity	ERROR
Solution	Restart the application or increase the maximum allowed connections configuration parameter.
Stored in DRP	Yes
Trapped over SNMP	No

Details	Description
Sent through e-mail	Yes
Monitoring interval	10 or 15 minutes
Dedupe period	1 hour

C.3.1 OBS-APPLICATION-0053

Details	Description

Event Messages	One of the following:
	`SSL certificate` *`certificate`* `will expire in less than 5 days.`
	`SSL certificate` *`certificate`* `does not exist.`
	`SSL certificate` *`certificate`* `has expired.`

Severity	Respectively: • WARNING • ERROR • CRITICAL
Solution	-
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	10 minutes
Dedupe period	1 hour

C.3.1 OBS-APPLICATION-0054

Details	Description
Event Messages	`Cannot upgrade MetaStore/ MetaStoreclient from version` *`X`* `to version` *`Y`*`.`
Severity	ERROR
Solution	-
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	N/A
Dedupe period	1 hour

C.3.2 Disk Events

C.3.2 OBS-DISK-0001

Details	Description
Event Message	`Disk` *`old_diskname`* `was renamed to` *`new_diskname`* `after reboot.`
Severity	INFO
Solution	Automatic correction is started.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.2 OBS-DISK-0002

Details	Description
Event Message	`Disk` *`X`* `can't be detected on machine` *`Y`*
Severity	CRITICAL
Solution	Operator intervention required to check why disk is not detected.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	15 minutes

C.3.2 OBS-DISK-0003

Details	Description
Event Message	`Problems found while detecting disks.`
Severity	Solution
WARNING	Operator intervention required. Check the output of the `lshw` command for disks.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.2 OBS-DISK-0004

Details	Description
Event Message	One of the following:

Details	Description

	`hdparm security attributes are already enabled on disk. Disk` *`ID`* `can not be erased.`
	`hdparm security attributes are frozen on disk. Disk (`*`uuid`*`) can not be erased.`
	`hdparm Failed to detect disk (`*`uuid`*`) attributes.`

Severity	ERROR
Solution	Investigate the root cause by executing the `hdparm` command manually.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.2 OBS-DISK-0005

Details	Description
Event Message	`Disk(s)` *`disk_names`* `can't be decommissioned on machine` *`machine_name`*`.`
Severity	ERROR
Solution	Manual intervention is required.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.2 OBS-DISK-0006

Details	Description
Event Message	`Error: decommissioning disks failed. Reason : disks` *`disk_guids`* `are in the same raid.`
Severity	N/A
Solution	Manual intervention is required.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.2 OBS-DISK-0007

Details	Description
Event Message	`New empty disk(s) detected.`
Severity	INFO
Solution	The new disk is checked to see if it has been added for replacement of an old disk.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.2 OBS-DISK-0008

Details	Description
Event Message	`New non-empty disk(s) detected:` *`disk_ids`*`.`
Severity	ERROR
Solution	No replacement operation starts on the new disk. Manual intervention is needed to empty the disk and insert it again.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.2 OBS-DISK-0009

Details Description	
Event Message	`Replacement disk size (`*`size`*` MB) is smaller than the original disk size (`*`size`*` MB) on node `*`machine_name`*`.`
Severity	ERROR
Solution	No replacement operation starts on the new disk. Manual intervention is needed to add new disk with size greater than or equal to original disk size.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.2 OBS-DISK-0010

Details	Description
Event Message	`Replacement disk type (`*`type`*`) is different from the original disk type (`*`original_type`*`) on node `*`machine_name`*`.`
Severity	ERROR
Solution	No replacement operation starts on the new disk. Manual intervention is needed to add a new disk of the same type as the original disk.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.2 OBS-DISK-0011

Details	Description
Event Message	`Can not detect physical location of disk [`*`ID`*`] on machine [`*`machine_name`*`].`
Severity	ERROR
Solution	A machine reboot is needed if USB key was unplugged from the running system.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.2 OBS-DISK-0012

Details	Description
Event Message	`Disk bus location detected.`
Severity	INFO
Solution	No action required.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.2 OBS-DISK-0013

Details	Description
Event Message	`Partition '`*`name`*`' has too few free inodes (`*`inodes`*`).`
Severity	WARNING
Solution	N/A
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.2 OBS-DISK-0014

Details	Description
Event Message	`Mountpoint `*`name`*` found in /etc/fstab but not mounted.`
Severity	WARNING
Solution	Please verify mounted partitions and `fstab` information.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Monitoring agent startup
Dedupe period	1 hour

C.3.2 OBS-DISK-0015

Details	Description
Event Message	`Disk(s) not detected: `*`disk_ids`*`.`
Severity	ERROR
Solution	N/A
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	15 minutes

C.3.2 OBS-DISK-0016

Details	Description

Event Message	One of the following:
	`No disk found in the model with bus location `*`bus_location`*`.`
	`More than one disk with bus location `*`bus_location`*` found.`

Details	Description
	`Disk with bus location` *`bus_location`* `is not decommissioned in the model.`
Severity	ERROR
Solution	N/A
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	0 minutes

C.3.2 OBS-DISK-0017

Details	Description
Event Message	`Filesystem sync is running for more than [`*`timeout`*`]`
Severity	ERROR
Solution	-
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	*`default`*

C.3.2 OBS-DISK-0018

Details	Description
Event Message	`Decommissioned and autodecommissioning disk detected`
Severity	INFO
Solution	-
Stored in DRP	No
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	15 minutes
Dedupe period	-

C.3.2 OBS-DISK-0019

Details	Description
Event Message	`Disk needs urgent replacement`
Severity	CRITICAL
Solution	-

Details	Description
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	-
Dedupe period	*`default`*

C.3.2 OBS-DISK-0020

Details	Description
Event Message	`Automatic decommissioning started for a disk`
Severity	Info
Solution	N/A
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	N/A
Dedupe period	*`default`*

C.3.2 OBS-DISK-0021

Details	Description
Event Message	`Automatic decommissioning completed for a disk`
Severity	Info
Solution	N/A
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	N/A
Dedupe period	*`default`*

C.3.2 OBS-DISK-0022

Details	Description
Event Message	`New empty disk` *`disk`* `on machine` *`machine`* `needs repurposing`
Severity	ERROR
Solution	A new empty disk detected that is not suitable for automatic replacement. Disk needs manual repurposing.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes

Details	Description
Monitoring interval	N/A
Dedupe period	*default*

C.3.3 DSS Events

C.3.3 OBS-DSS-BLOCKSTORE-0026

Details	Description
Event Message	Critical threshold (*threshold_value* %) exceeded for block store *blockstore_path*. *threshold_value* can be: • 90 % • 96 % • 98 %
Severity	Respectively: • WARNING • ERROR • CRITICAL
Solution	No action needed.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	10 minutes
Dedupe period	15 minutes

C.3.3 OBS-DSS-BLOCKSTORE-0027

Details	Description
Event Message	Blockstore *blockstore_path* has released disk space.
Severity	INFO
Solution	This is an informational message. No immediate action required.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	10 minutes
Dedupe period	1 hour

C.3.3 OBS-DSS-BLOCKSTORE-0028

Details	Description
Event Message	One of the following:

Details	Description
	Error checkblock count (*count*) exceeded for block store *blockstore_path*. Error checkblock count (*count*) exceeded for block store '*blockstore_path*' with less than *percentage*% full copy files. **Threshold: 12,000,000 counts** Warning checkblock count (*count*) exceeded for block store *blockstore_path*. Warning checkblock count (*count*) exceeded for block store '*blockstore_path*' with less than *percentage* full copy files. **Threshold: 10,000,000 counts**
Severity	Respectively: • ERROR • WARNING
Solution	No immediate action required.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	10 minutes
Dedupe period	1 day

C.3.3 OBS-DSS-BLOCKSTORE-0029

Details	Description
Event Message	One of the following: Critical checkblock count (*count*) exceeded for block store*blockstore_path* Critical checkblock count (*count*) exceeded for block store '*blockstore_path*' with less than *percentage*% full copy files **Threshold: 15,000,000 counts**
Severity	CRITICAL
Solution	Blockstore is automatically set to read only.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes

Details	Description
Monitoring interval	10 minutes
Dedupe period	15 minutest

C.3.3 OBS-DSS-BLOCKSTORE-0030

Details	Description

Event Message	One of the following:
	`Blockstore` *`blockstore_path`* `is OFFLINE.`
	`You have` *`disabled_blockstores_count`* `blockstore(s) that is(are) temporarily offline. Blockstores are only in this status when some of your Storage Nodes are down.`
	`You have` *`offline_blockstores_count`* `blockstore(s) that is(are) offline. Offlined blockstores will only be used again when they are put online through an operator action.`
	`You have` *`count`* `blockstore(s) that is(are) temporarily offline. Blockstores are only in this status when some of your storage nodes are down. You have` *`count`* `blockstore(s) that is(are) offline. Offlined blockstores will only be used again when they are put online through an operator action.`

Severity	WARNING
Solution	No immediate action required.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.3 OBS-DSS-BLOCKSTORE-0031

Details	Description
Event Message	`Blockstore` *`blockstore_path`* `has status DECOMMISSIONED for more than` *`nr`* `days.`
Severity	WARNING or ERROR
Solution	Check the dss log files and repair statistics to identify why this blockstore has not been automatically changed to the ABANDONED status
Stored in DRP	Yes
Trapped over SNMP	No

Details	Description
Sent through e-mail	Yes
Monitoring interval	10 minutes
Dedupe period	1 hour

C.3.3 OBS-DSS-BLOCKSTORE-0032

Details	Description
Event Message	`Block store IDs do not match for` *`blockstore_path`* `(`*`ID - ID`*`).`
Severity	ERROR
Solution	Please contact HGST Support.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	10 minutes
Dedupe period	1 hour

C.3.3 OBS-DSS-BLOCKSTORE-0033

Details	Description
Event Message	`OFFLINE blockstores detected`
Severity	WARNING
Solution	No immediate action required. This situation is corrected automatically.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	10 minutes
Dedupe period	1 hour

C.3.3 OBS-DSS-BLOCKSTORE-0034

Details	Description
Event Message	`Could not access blockstore partition` *`partition_path`*
Severity	CRITICAL
Solution	Check disk for errors and make sure the partition is mounted correctly
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	10 minutes
Dedupe period	1 hour

C.3.3 OBS-DSS-STORAGEPOOL-0001

Details	Description
Event Message	`Storage pool statistics not updated in the past day. Last update: `*`last_update`*`.`
Severity	INFO
Solution	Validate why the storage pool monitoring data is not up to date. Operator intervention may be required.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	1 hour
Dedupe period	1 day

Note: Name space monitoring does not run when a repair is ongoing for a specific name space. Repair activities can be validated through the CMC or OSMI.

Name space monitoring can be initiated manually through OSMI. However, a manual run takes resources from the MetaStore and might slow down other operations (repair, ingest, outgest).

C.3.3 OBS-DSS-STORAGEPOOL-0005

Details	Description
Event Message	`Found name spaces with a deprecated codec version.`
Severity	ERROR
Solution	Change the name space to use a new codec version.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	1 hour
Dedupe period	1 hour

C.3.3 OBS-DSS-STORAGEPOOL-0025

Details	Description
Event Message	`Storage pool is more than 70% full.`
Severity	WARNING
Solution	Validate these figures against the design targets of the environment and match them to your capacity planning.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Defined in monitor policy (default: 30 minutes)
Dedupe period	1 hour

C.3.3 OBS-DSS-STORAGEPOOL-0026

Details	Description
Event Message	`Storage pool is more than 80% full.`
Severity	ERROR
Solution	Validate these figures against the design targets of the environment and match them to your capacity planning.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Defined in monitor policy (default: 30 minutes)
Dedupe period	1 hour

C.3.3 OBS-DSS-STORAGEPOOL-0027

Details	Description
Event Message	`Storage pool is more than 90% full.`
Severity	CRITICAL
Solution	Validate these figures against the design targets of the environment and match them to your capacity planning.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Defined in monitor policy (default: 30 minutes)
Dedupe period	1 hour

C.3.3 OBS-DSS-STORAGEPOOL-0028

Details	Description
Event Message	`Storage policies detected that are affected by a known issue that might result in incorrect hierarchical data spreading. The data store in S3 buckets or AXR name spaces using these policies may not be stored with the expected node, rack or data center failure protection.`
Severity	CRITICAL
Solution	Consult knowledge base article BSP044.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	not applicable
Dedupe period	1 hour

C.3.4 Physical Machine Events

C.3.4 OBS-PMACHINE-0001

Details	Description
Event Message	`mountpoint is more than percentage% full.` Threshold: • 95 % • 96 % • 98 %
Severity	Respectively: • WARNING • ERROR • CRITICAL
Solution	The mentioned machine has a mount point which is using too much disk space. Investigate why this mount point is using that so much disk space.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0002

Details	Description

Event Message	One of the following:
	`Swap usage is over x%.`
	`Swap is not available.`

Severity	Respectively WARNING or ERROR
Solution	The machine is using a part of its swap space. In normal circumstances, the system should not be using swap space. Verify what processes are using the swap space. Contact HGST Support if needed.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0005

Details	**Description**

Event Message	One of the following:
	`Cache daemon [ID] memory threshold exceeded (xMB)`
	`Cache daemon [ID] shared memory threshold exceeded (xMB)`
	`Client daemon [ID] memory threshold exceeded (xMB)`
	`Client daemon [ID] shared memory threshold exceeded (xMB)`
	`Maintenance agent [ID] memory threshold exceeded (xMB)`
	`Maintenance agent [ID] shared memory threshold exceeded (xMB)`
	`Storage daemon [ID] memory threshold exceeded (xMB)`
	`Storage daemon [ID] shared memory threshold exceeded (xMB)`
	`MetaStore node [metastore_name::node_name] memory threshold exceeded (xMB)`
	`MetaStore node [metastore_name::node_name] shared memory threshold exceeded (xMB)`
	`Less than x% memory free: x%`

Severity	WARNING
Solution	The specified machine has one or more processes which are using a lot of memory. Identify the process(es), and, if abnormal, try to resolve them. If this is expected behavior, consider updating the thresholds. Contact HGST Support for this action.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Various
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0006

Details	**Description**
Event Message	`Interface name (mac_address) has x transmission errors.`
Severity	WARNING
Solution	The mentioned machine has on one or more of its network interface cards which is having too many packet errors.

Details	Description
	Identify the reason for these errors.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0007

Details	Description
Event Message	`Interface` *`name`* `is overloaded for` *`x`*`%.`
Severity	WARNING
Solution	The mentioned machine has one or more of its network interface cards which has to handle too much traffic. Identify the reason for this traffic.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0012

Details	Description
Event Message	`Disk` *`ID`* `has SMART failures [Overall Status 'UNKNOWN'].`
Severity	CRITICAL
Solution	On the mentioned machine, for one or more disks SMART failures have been detected. This may mean that one or more disks are broken or are nearly broken. A replacement of the disk might be necessary.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0013

Details	Description
Event Message	One of the following: `I/O errors on disk (`*`disk_name`*`).`

Details	Description
	`File system errors on disk` *`disk_name`*`, partition` *`partition_name`*`.`
Severity	CRITICAL
Solution	On the mentioned machine, for one or more partition failures have been detected. This may mean that one or more partitions are broken or are nearly broken. A replacement of the disk might be necessary.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	30 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0017

Details	Description
Event Message	`Software RAID array` *`device_path`* `has status` *`status`*`.`
Severity	WARNING
Solution	Operator intervention required.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0019

Details	Description
Event Message	`Kernel dmesg errors detected.`
Severity	ERROR
Solution	Investigate the root cause of the kernel `dmesg` errors
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	30 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0020

Details	Description
Event Message	`Mount point` *`mountpoint`* `is read-only.`

Details	Description
Severity	CRITICAL
Solution	A read-only file system has been detected on the machine. This indicates an issue with one of the partitions or the file system.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0023

Details	Description

Event Message	One of the following:
	`Unknown LAN type.`
	`DHCP method not supported for device (`*`name`*`).`

Severity	Respectively CRITICAL or WARNING.
Solution	Should not happen; please open a case with HGST Support.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0024

Details	Description
Event Message	`Macaddress` *`macaddress`* `is already in use.`
Severity	CRITICAL
Solution	This event is thrown if a NIC is added to a machine using a MAC address that is already used on another NIC in the DRP. This should not happen. Please contact your vendor and ask to file a bug report.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0025

Details	Description
Event Message	`Machine with guid` *`machineguid`* `has already a NIC with name` *`interfacename`*`.`
Severity	ERROR
Solution	A NIC gets added to a machine with the same order as an already existing NIC. This should not happen; please open a case with HGST Support.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0026

Details	Description

Event Message	One of the following:
	`Invalid IP address` *`ipaddress`* `specified.`
	`Invalid IP address` *`ipaddress`*`.`
	`Invalid IP address specified.`

Severity	ERROR
Solution	While adding an IP address to a new NIC, the workflow discovers that the IP address does not exists in the database. This IP address should have been added in a different workflow. This means there might be an issue with the workflow to add the IP address, the workflow engine, or there is an issue with the database. Run a complete health check on the environment and open a case with HGST Support.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0027

Details	Description
Event Message	`Failed to Retrieve application agent. Reason: could not find application called "cloudapi".`

Details	Description
Severity	ERROR
Solution	The cloudapi application was not found in the database while the workflow attempted to retrieve the appliance agent GUID. Run a full health check and contact HGST Support.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0028

Details	Description
Event Message	`Unable to retrieve nic with macaddress` *`macaddress`* `or name` *`interfacename`*`.`
Severity	ERROR
Solution	This can occur when a workflow attempts to modify a NIC of a vmachine. In this case the database does not contain the mac address or NIC number that was provided to the workflow. This can be caused by a database inconsistency. Run a full health check on the SSO and contact HGST Support.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0033

Details	Description
Event Message	`Unable to stop application` *`applicationname`*`. Error:` *`exception`*`.`
Severity	ERROR
Solution	This error can occur when a workflow attempts to stop an application, but the status remains active. Manually stop the application on the pmachine. If this also fails, check if the application is still visible in the process list and kill it. If it is not running, check for application PID files in `/opt/qbase3/var/pid/` and clean them up.\r\n.

Details	Description
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0040

Details	Description
Event Message	`Invalid nic configuration. Expected exactly one match.`
Severity	ERROR
Solution	While removing IP addresses from a NIC, the workflow discovers that the MAC address does not exist or exists multiple times in the database. This means there might be an issue with the database or applications. Run a complete health check on the environment and contact HGST Support.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0042

Details	Description

Event Message	One of the following:
	`Unable to start application` *`name`*`. Error:` *`exception`*`.`
	`Unable to restart application` *`name`*`. Error:` *`exception`*`.`

Severity	CRITICAL
Solution	A software component failed to (re)start. Run a health check to try and pinpoint the problem. Please open a case with HGST Support if the problem cannot be pinpointed.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0045

Details	Description
Event Message	`Machine is not running.`
Severity	CRITICAL
Solution	Investigate why the machine is not running.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	30 minutes

C.3.4 OBS-PMACHINE-0046

Details	Description
Event Message	`Agent not running, restarting agent.`
Severity	WARNING
Solution	No immediate action necessary.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0047

Details	Description
Event Message	`Restarting agent failed.`
Severity	ERROR
Solution	Open a logger and try to restart the agent again to try and pinpoint the problem.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0049

Details	Description
Event Message	`Machine object has no agent guid.`
Severity	ERROR
Solution	Should not happen; please open a case with HGST Support.
Stored in DRP	Yes

Details	Description
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0054

Details	Description

Event Message	One of the following:
	`DNS resolving fails`
	`Interface '`*`dev`*`' (`*`name`*`) is DOWN.`
	`Unable to contact default gateway `*`gateway`*`. Reason: `*`error`*`.`

Severity	CRITICAL
Solution	An interface on the mentioned machine is down. Identify the reason for this error.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Respectively monitoring agent startup or 5 minutes.
Dedupe period	15 minutes

C.3.4 OBS-PMACHINE-0055

Details	Description
Event Message	`IP address `*`ipaddressguid`*` already in use for lan with guid `*`languid`*`.`
Severity	ERROR
Solution	Choose a different ip address or delete the ip address from the specified lan.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0061

Details	Description
Event Message	`Template name is not uniquely defined.`
Severity	ERROR
Solution	This should not happen; please open a case with HGST Support.

Details	Description
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0067

Details	Description
Event Message	`Gateway` *`gateway`* `does not belong to any of the configured networks.`
Severity	CRITICAL
Solution	This should not happen; please open a case with HGST Support.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0104

Details	Description
Event Message	`Load average over the last 15 minutes is high (`*`load`*`).`
Severity	WARNING **Background Information** The load average on a Storage Node is determined by: • the number of parallel streams • the size of the operations • the policy parameters • the type of operation (read, write, delete or update).
Solution	Your monitoring should be refined so that events do not trigger while the normal load is being put on your environment. Update the thresholds if necessary. Only if an exceptional load is being put in the environment, these events may trigger. They will indicate the following: • If the event applies to a single machine: There might be a hardware issue with this machine. Investigate the cause. • If the event applies to all Storage Nodes: There is a change in the I/O patterns that could lead to performance impact.

Details	Description
	Investigate what has changed. If the change is expected and the performance impact acceptable, update the load triggers.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0105

Details	Description
Event Message	`The SMART control on disk` *`ID`* `is disabled [Overall status: DISABLED].`
Severity	WARNING
Solution	SMART control is not enabled on the specified node for the specified disk. To enable it, run the following command at the Linux prompt on the specified node: `smartctl -s /dev/sdX`
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0106

Details	Description
Event Message	`machine was rebooted.`
Severity	INFO
Solution	Operator intervention required to check the cause of reboot if this was an unplanned reboot.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	monagent startup
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0107

Details	Description
Event Message	`NTP cannot adjust time on node '`*`machinename`*`', time difference is`

Details	Description
	`greater than 1000 seconds, NTP needs to be restarted.`
Severity	CRITICAL
Solution	NTP is auto restarted.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	1 hour
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0108

Details	Description
Event Message	`Cannot shutdown Machine(s)` *`machine_names`* `and they will be skipped during shutdown process.`
Severity	N/A
Solution	Turn off the machine manually.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0109

Details	Description
Event Message	`Machine(s)` *`machine_names`* `failed to be shutdown.`
Severity	N/A
Solution	Check the shutdown job details for more information.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0110

Details	Description
Event Message	`Number of processes exceeds threshold:` *`nr > threshold`*
Severity	WARNING

Details	Description
Solution	Investigate the root cause
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0111

Details	Description
Event Message	`Processes found with state dead / zombie` *`process_names`*`.`
Severity	ERROR
Solution	Investigate the root cause
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS–PMACHINE-0113

Details	Description
Event Message	`Machine` *`machine_name`* `is in status STOPPING`
Severity	WARNING
Solution	Verify why machine is in status STOPPING.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	30 min (monitoring policy)
Dedupe period	-

C.3.4 OBS-PMACHINE-0114

Details	Description
Event Message	`Fan speed ['`*`name`*`'] is below` *`speed`* `RPM.`
Severity	• WARNING: Between 3000 - 2501 RPMs • ERROR: Between 2500 - 1 RPMs • CRITICAL: When the fan is not spinning at all, 0 RPMS

Details	Description
Solution	Check the fan. Replace it if necessary. For instructions on fan replacement, see the *HGST Active Archive System FRU Replacement Guide*.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0115

Details	Description
Event Message	`PSU` *`name`* `failed.`
Severity	ERROR
Solution	Check the power supply unit (PSU). Replace it if necessary. For instructions on PSU replacement, see the *HGST Active Archive System FRU Replacement Guide*.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	15 minutes
Dedupe period	1 hour

C.3.4 OBS-PMACHINE-0116

Details	Description
Event Message	Abnormal log file size detected for *`filename`*. Reached *`size`* MB in last hour.
Severity	ERROR
Solution	-
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	1 hour
Dedupe period	1 hour

C.3.5 MetaStore Events

C.3.5 OBS-ARAKOON-0004

Details	Description
Event Message	One of the following: `Collapsing of transaction logs failed for MetaStore` *`name`*`.`

Details	Description
	Your MetaStore transaction log collapsing policy is not configured. It needs to be configured in order to run this policy.
Severity	ERROR
Solution	Failure to run the internal MetaStore maintenance can cause performance loss or introduce delays upon failure. Investigate the root cause and contact HGST Support.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hours

C.3.5 OBS-ARAKOON-0005

Details	Description
Event Message	One of the following: • The number of keys in the *name* MetaStore exceeds *critical_threshold* %. Threshold: ◆ 98 % ◆ 96 % • The number of keys in the *name* MetaStore exceeds *warning_threshold*%. Threshold: ◆ 95 %
Severity	Respectively: • CRITICAL • ERROR • WARNING
Solution	Validate these figures against the design targets of the environment and match them to your capacity planning. Consider adding an additional MetaStore.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.5 OBS-ARAKOON-0006

Details	Description

Event Message	One of the following:
	`Metadata store '`*`cluster_name`*`' has no master node: can not re-initialize node '`*`node_name`*`'.`
	`No master node for MetaStore '`*`name`*`'.`

Severity	Respectively: • CRITICAL • ERROR
Solution	Investigate why the MetaStore has no master node. Operator intervention is required.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.5 OBS-ARAKOON-0007

Details	Description
Event Message	`MetaStore safety low. Node `*`node`*` missing for cluster `*`cluster`*`.`
Severity	CRITICAL
Solution	Investigate why the MetaStore has no master node. Operator intervention is required.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	15 minutes

C.3.5 OBS-ARAKOON-0008

Details	Description

Event Message	One of the following:
	`Node `*`node_name`*` on MetaStore `*`cluster_name`*` is lagging `*`number_of_keys`*` keys on machine `*`machine_name`*` (*)`
	`Node(s) '`*`node_name`*`' in MetaStore '`*`cluster_name`*`' are lagging behind.`

Severity	Respectively: • WARNING

Details	Description
	• ERROR
Solution	Investigate why the MetaStore has nodes that are lagging behind. Operator intervention is required.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

Note: (*) The event message is correct regarding the node name and the MetaStore, but the machine name is incorrect.

C.3.5 OBS-ARAKOON-0009

Details	Description
Event Message	More than *num_logs* tlogs found on node *node_name* of MetaStore *cluster_name*
Severity	CRITICAL
Solution	Investigate why the MetaStore has nodes that are lagging behind. Operator intervention is required.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.5 OBS-ARAKOON-0010

Details	Description
Event Message	One of the following: • Database partition for MetaStore node *cluster_name*::*node_name* is more than *error_threshold*% full. Threshold: 96 % • Database partition for MetaStore node *cluster_name*::*node_name* is more than *warning_threshold*% full. Threshold: 95 %
Severity	Respectively: • ERROR

Details	Description
	• WARNING
Solution	Mark the MetaStore as `FULL`.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.5 OBS-ARAKOON-0011

Details	Description
Event Message	`Database partition for MetaStore node` *`metastore_name`*`::`*`node_name`* `is more than` *`x`*`% full.` Threshold: 98 %
Severity	CRITICAL
Solution	Mark the MetaStore as `FULL`.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.5 OBS-ARAKOON-0012

Details	Description
Event Message	`Transaction log partition for MetaStore node` *`metastore_name`*`::`*`node_name`* `is low on space.` Threshold: • 1.5 free space left • 1.3 free space left
Severity	Respectively: • WARNING • ERROR
Solution	Mark the MetaStore as `FULL`.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.5 OBS-ARAKOON-0013

Details	Description
Event Message	`Node` *`node_name`* `in Metastore` *`cluster_name`* `will be stopped. Transaction log partition for MetaStore node` *`metastore_name`*`::`*`node_name`* `is low on space.` Threshold: 1.2 free space left
Severity	CRITICAL
Solution	Mark the MetaStore as `FULL`.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.5 OBS-ARAKOON-0014

Details	Description
Event Message	`Could not collapse transaction logs for MetaStore '`*`name`*`'. Reason:` *`reason`*`.`
Severity	CRITICAL
Solution	-
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.5 OBS-ARAKOON-0015

Details	Description
Event Message	`MetaStore master fail-over triggered on MetaStore '`*`name`*`'.`
Severity	INFO
Solution	-
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	0

C.3.5 OBS-ARAKOON-0018

Details	Description
Event Message	`MetaStore instance` *`cluster name`*`::`*`node name`* `automatic recovery failed after` *`number`* `retries on machine` *`machine_name`*`.`
Severity	CRITICAL
Solution	-
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	-
Dedupe period	0 min

C.3.5 OBS-ARAKOON-0019

Details	Description
Event Message	`Found MetaStore backup file(s) on machine` *`machine_name`*
Severity	INFO
Solution	-
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	5 min
Dedupe period	1 h

C.3.5 OBS-ARAKOON-0020

Details	Description
Event Message	`The number of keys in the` *`X`* `MetaStore exceeds` *`Y`*`%`
Severity	CRITICAL
Solution	MetaStore is automatically set to `FULL`
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 min
Dedupe period	1 h

C.3.5 OBS-ARAKOON-0021

Details	Description
Event Message	`MetaStore` *`X`* `node` *`Y`* `usage thresholds fell below critical levels, setting to ACTIVE`

Details	Description
Severity	INFO
Solution	MetaStore is checked if eligible to be automatically set to `ACTIVE`
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 min
Dedupe period	1 h

C.3.6 Network Events

C.3.6 OBS-NETWORK-0001

Details	Description

Event Message	One of the following:
	`NTP daemon is down`
	`HTTPS proxy is enabled but not running`
	`HTTPS proxy` *`proxy_type`* `port` *`port`* `is unavailable`
	`Application` *`applicationname`* `is HALTED`
	`Cache daemon` *`ID`* `has status` *`status`*
	`Unable to connect to cache daemon` *`ID`*`:`*`port`*
	`Client daemon` *`GUID`* `has status` *`status`*
	`Unable to connect to client daemon` *`ID`*`:`*`port`*
	`Maintenance agent` *`ID`* `has status` *`status`*
	`Storage daemon` *`ID`* `has status` *`status`*
	`Unable to connect to storage daemon` *`IP`*`:`*`port`* `(`*`exception`*`)`
	`MetaStore node` *`metastore_name`*`::`*`metastore_node`* `is DOWN`
	`Application monitoring agent down on node` *`machine_name`* `[Auto restart]`

Severity	• WARNING: message 1 • ERROR: messages 2 - 12 • CRITICAL: message 13
Solution	Validate why the the event prompts (applicationserver/ port down, network issues,...).
Stored in DRP	Yes
Trapped over SNMP	No

Details	Description
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.6 OBS-NETWORK-0003

Details	Description
Event Message	`NIC` *`name`* `(`*`mac_address`*`) in half duplex mode.`
Severity	ERROR
Solution	Validate why the network interface has been switched back to half duplex. The network interface or the switch the interface is connected to might be broken.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.6 OBS-NETWORK-0004

Details	Description
Event Message	`NIC` *`name`* `(`*`mac_address`*`) has speed` *`speed`* `below threshold.`
Severity	WARNING
Solution	Validate if the related network interface is working correctly.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.6 OBS-NETWORK-0005

Details	Description

Event Message	One of the following:
	`HTTPS proxy is disabled, but running as PID` *`pid`*`.`
	`Application` *`application_name`* `is RUNNING.`
	`Cache daemon` *`ID`* `is UP.`
	`Client daemon` *`ID`* `is UP.`
	`Maintenance agent` *`ID`* `is UP.`

Details	Description

	`Storage daemon` *`ID`* `is UP.`
	`Metastore node` *`metastore_name`*`::`*`node_name`* `is UP.`

Severity	• WARNING: message 1 • INFO: messages 2 - 7
Solution	This event clears former events, no solution needed.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.6 OBS-NETWORK-0006

Details	Description
Event Message	`SMTP is not configured correctly.`
Severity	WARNING
Solution	Manual intervention required.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.6 OBS-NETWORK-0007

Details	Description
Event Message	`Cache cluster` *`ID`* `has status` *`status`*`.`
Severity	ERROR
Solution	N/A
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	15 Dedupe
minutes period	1 hour

C.3.6 OBS-NETWORK-0008

Details	Description
Event Message	`Interface` *`dev`* `(`*`name`*`) is UP but not configured in /etc/network/interfaces.`
Severity	WARNING

Details	Description
Solution	N/A
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Monitoring agent startup
Dedupe period	1 hour

C.3.6 OBS-NETWORK-0009

Details	Description
Event Message	`NIC [`*`name`*`] renamed from '`*`old_name`*`' to '`*`new_name`*`'.`
Severity	WARNING
Solution	N/A
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.6 OBS-NETWORK-0010

Details	Description
Event Message	`NIC `*`name`*` (`*`mac_address`*`) has no IP `*`IP`*` configured.`
Severity	ERROR
Solution	N/A
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	5 minutes
Dedupe period	1 hour

C.3.7 Generic Events

C.3.7 OBS-GENERIC-0001

Details	Description

Event Message	One of the following:
	`Machine `*`machine_name`*` is HALTED.`
	`Machine `*`machine_name`*` is down [status: `*`status`*`].`

Severity	CRITICAL

Details	Description
Solution	Validate why the machine is down / powered off.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	Determined by monitor policy (default is 30 minutes)
Dedupe period	30 minutes

C.3.7 OBS-GENERIC-0002

Details	Description
Event Message	`Found a partially installed patch:` *`name_patch_version`*
Severity	WARNING
Solution	Try to resume the failed upgrades.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	1 hour
Dedupe period	30 minutes

C.3.7 OBS-GENERIC-0003

Details	Description
Event Message	`Failed to save job` *`action_name`*
Severity	WARNING
Solution	Check the logs on the system and try to fix the root cause.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.7 OBS-GENERIC-0004

Details	Description

Event Message	One of the following:
	`Failed to initialize node` *`machine_name`*`.`
	`Job` *`job_description`* `failed on machine` *`machine_name`*`.`
	`Policy` *`policy_description`* `failed on machine` *`machine_name`*`.`

Severity	ERROR

Details	Description
Solution	Check the logs of the failed job and try to remove the root cause of the failure. For the logs, navigate to: **Dashboard** > **Administration** > **HGST Object Storage Management** > **Logging** > **Jobs** > **Policies**.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.7 OBS-GENERIC-0071

Details	Description
Event Message	`Core dump files found in` *`/var/crash/`* `Filename:` *`name`* `(type -` *`date`*`).`
Severity	WARNING
Solution	Retrieve the core files, provide them to HGST Support for further diagnosis and remove them from the machine to clear the event
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	1 day
Dedupe period	1 hour

C.3.7 OBS-GENERIC-0072

Details	Description
Event Message	`Too many instances running for policy` *`policyname`*`, skipping execution.`
Severity	CRITICAL
Solution	Investigate why the previous instance of this policy is still running.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	No
Monitoring interval	Variable
Dedupe period	1 hour

C.3.7 OBS-GENERIC-0073

Details	Description
Event Message	`Failed login attempt with username` *`username`*`.`
Severity	Warning
Solution	-
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	15 minutes

C.3.7 OBS-GENERIC-0074

Details	Event
Description Message	`Too many failed login (`*`threshold`*`) attempts in the last (`*`seconds`*`) seconds.`
Severity	ERROR
Solution	-
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Variable
Dedupe period	1 hour

C.3.7 OBS-GENERIC-0075

Details	Description
Event Message	`Failure while executing event handling logic for event type` *`type ID`*
Severity	CRITICAL
Solution	-
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	-
Dedupe period	1 hour

C.3.7 OBS-GENERIC-0076

Details	Description
Event Message	`Errors while updating machine(s) configurations during failover`
Severity	ERROR

Details	Description
Solution	Manual intervention is required
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	No monitoring interval is configured, its only raised during failover
Dedupe period	1 hour

C.3.8 Other Events

C.3.8 Agent Events

C.3.8 OBS-AGENT-0006

Details	Description
Event Message	`Machine agent is down on` *`machine_name`*`.`
Severity	CRITICAL
Solution	Please check the machine agent.
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Determined by monitor policy (default is 30 minutes)
Dedupe period	1 hour

C.3.8 Environment Events

C.3.8 OBS-ENVIRONMENT-0001

Details	Description

Event Message	`Policy '`*`policy_name`*`' is disabled`
	`Policy '`*`policy_name`*`' did not run in more than` *`interval`* `hours`
	`Policy '`*`policy_name`*`' never ran before`

Severity	CRITICAL
Solution	Make sure 'Backup model database' policy is enabled and executing properly.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	1 d
Dedupe period	1 d

C.3.8 OBS-ENVIRONMENT-0002

Details	Description
Event Message	`Duplicate rack or data center IDs found.`
Severity	ERROR
Solution	Please contact HGST Support.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	-
Dedupe period	1 d

C.3.8 Storage Events

C.3.8 OBS-STORAGE-0004

Details	Description
Event Message	`Objects found with low disk safeties.` Environment Statistics Environment Statistics Degraded disks: x / y Decommissioned disks: x / y Offline disks: x / y Healthy disks: x / y Halted systems: x / y Decommissioned systems: x / y
Severity	CRITICAL
Solution	One or more disks have failed and repair has not been able to correct this. Validate that repair is ongoing and moving in the right direction, this means leading to an increased disksafety.
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	1 hour
Dedupe period	15 minutes

C.3.8 OBS-STORAGE-0005

Details	Description
Event Message	`Current MetaStore configuration for storage daemon/client daemon/ maintenance agent with ID X does not match desired configuration.`
Severity	WARNING
Solution	No immediate action required.

Details	Description
Stored in DRP	Yes
Trapped over SNMP	No
Sent through e-mail	Yes
Monitoring interval	Once during monitoring agent startup
Dedupe period	30 minutes

C.3.8 OBS-STORAGE-0006

Details	Description
Event Message	`Unverified objects found`
Severity	CRITICAL
Solution	-
Stored in DRP	Yes
Trapped over SNMP	Yes
Sent through e-mail	Yes
Monitoring interval	1 day
Dedupe period	-

D Object Verification

Topics:

- About Object Verification
- Object Verification in Detail

D.1 About Object Verification

The goal of object verification is to make sure that all objects, stored in the Active Archive System, are correctly stored and in case of issues, that the object is repaired.

The process of object verification consists of:

- Checking if the object exists
- Checking if all superblocks of the object exist
  - For each superblock, check if each blockstore contains enough checkblocks
  - For each checkblock, check if the CRC32 is correct

The Active Archive System software has a default object verification process, running as a background process. The goal of this process is to verify all objects in a name space within a certain time interval, by default 1 year.

D.1 Extra Metadata

The following metadata is added to run the verfication process:

- On a name space:
  - `NsVi`: name space verfication interval, the fixed interval in which all objects of the name space must be verified.
  - `NsVt`: name space verification target date, the target date by which all objects in a name space must have been verified.
- On an object:
  - `ObjVd`: object verification date, the date and time when the object verification has completed successfully

D.1 Basic Verification Flow

In a running Active Archive System setup without errors, all objects have a verification date in the previous or current verification interval.

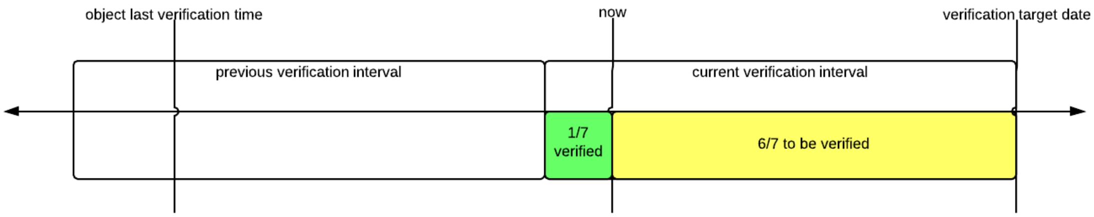
When an object is successfully verified, its verification date (`ObjVd`), is set to the date and time when it has been verified.

The verification of objects is done in iterations because it is not possible to verify all objects at the beginning of the verification interval.

The number of objects in a name space that will be verified per iteration, is calculated by using the target date, current date, and verification interval. For more information, see Object Verification in Detail on page 181.

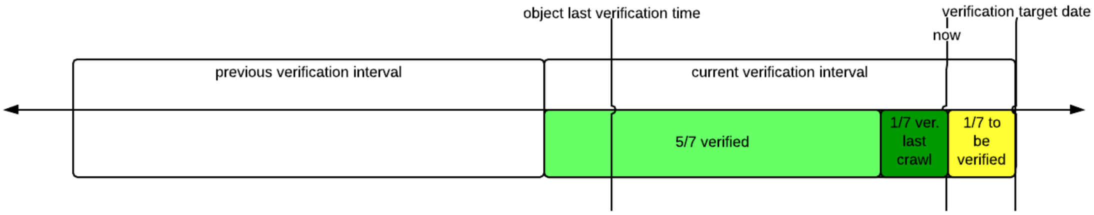
The image below shows that after one iteration, there is 1/7th of the object verified, and thus that it takes seven iterations before all objects in the name space to be verified.

Figure 14: Verification Progress: First Iteration


object last verification time
now
verification target date
previous verification interval
current verification interval
1/7 verified
6/7 to be verified

In the following image, you see the progress when the sixth iteration has completed.

Figure 15: Verification Progress: Sixth Iteration


object last verification time
verification target date
now
previous verification interval
current verification interval
5/7 verified
1/7 ver. last crawl
1/7 to be verified

D.2 Object Verification in Detail

The verification of objects is executed over several iterations. The number of objects that are verified in an iteration is determined during a repair crawl (iteration) and is calculated as follows:

```
\[Vf = 1 - (NsVt - now)/NsVi\]
```

where:

- **Vf**: number of objects that will be verified during the iteration
- **NsVt**: name space verification target date, in epoch
- **now**: current date and time, in epoch
- **NsVi**: name space verification interval, in seconds, default 1 year or 31,536,000 seconds

The following counters are used during a repair crawl:

- **Ov**: number of objects that have passed the repair crawl and that are already vefified during the current interval; this is the number of objects with *ObjVd >= (NsVt - NsVi)*
- **Oa**: number of objects that have passed the repair crawl and that are marked for verification
- **Ot**: total number of objects that have passed the repair crawl

D.2 Marking an Object for Verification

An object is marked for verification when the following conditions are met:

- object verification date is outside the current verification interval: *ObjVd < NsVt - NsVi*
- *(Ov + Oa)/Ot < Vf*
- object is not marked for repair, rebalance, or change policy
- object does not contain an offline blockstore

For the objects that meet these conditions, the repair crawl process creates a verification task.

D.2 Verification Task

The verification task is executed by a maintenance agent. The task consists of:

- check if the object exists
- check if all superblocks of the object exist
    - for each superblock, check if each blockstore contains enough checkblocks
    - for each checkblock, check if the CRC32 is correct

Follow-up actions depend on the result of the verification task:

- object verified as faultless: *ObjVd = now*, the object gets a new verification date
- object verified as erroneous: repair task is started for the object, a new verification is executed in a next repair crawl
- objects with a storage layout from 3.x or earlier, are always repaired. As a result, the entire storage
- will be gradually upgraded to the new storage layout in the first verification interval.

Note

An offline or unreachable blockstore does not trigger a repair task but the verification task will fail. In this case the verification date is not updated.

Index

A

accounts
- adding 13, 14, 14, 14
- modifying 13, 13, 15
- removing 14, 14, 15, 15

analytics 12, 101
Apache 115
application server 22, 109
Arakoon
- cluster 112
- interval 81
- log files 81
- statistics 81
- updates 36

ATA Secure Erase 56
automatic disk decommissioning 67

B

backoff interval 67, 67, 68
backoff multiplier 67, 67, 68
blacklists
- graphics 86

blockstore 48, 54, 56, 62, 67, 68, 68, 68, 109, 112
blockstores
- disabled 71
- halted 71

boot disk 51
boot log 22
bucket operations
- DELETE Bucket 31

buckets
- deleting 29, 29

C

cache 110
cache clusters 72, 110, 170
client machine
- Linux 27

cloudAPI 75, 118, 131, 149
cold reset 62
console 12, 22
conventions 9, 9, 10
copyright
- notice 2

Cyberduck 36

D

daemon
- cache 16, 110, 117, 131, 149, 170
- client
  - configuration file 31, 31
- storage 55, 88, 113, 117

database
- backup 69
- defragmentation 69

dead man timer 100
DHCP helper 116
disk
- blockstore 67
- decommissioned 62, 67, 67
- degraded
  - list 48, 49
  - notification 48, 49
- missing 62
- repair 67
- SMART data 52, 62, 67
- unmanaged 59, 61
- wrong 61

disk safety 22, 48, 55, 70, 99, 111
documentation 10
domain name 27
DSS 22

E

ejabberd 115
encryption 33
epoch time 32
EULA 101
event
- logging 58

events 48, 86, 93, 93, 94, 94, 94, 94, 95, 118, 118, 118, 131, 131, 136, 143, 149, 163, 170, 173, 177
executeFailoverScript.py 58

F

failover 42, 58, 58, 61, 62, 66, 83, 100
fan
- failure 149
- warnings 62

field replaceable unit (FRU) 61
file system
- read only 56

G

graphical user interface 12, 88, 110
GUI 110

H

hardware
  disk
    decommissioned 62
    degraded
      notification 48, 49
    missing 62
    SMART data 62
  ECC 62
  fan errors 62
  management 46
  memory errors 62
  temperature warnings 62
health checker 96
HTTPS 33, 111

I

IPMI 46, 47, 47, 47, 47, 62, 75, 111

L

log bundles 72, 73
log collector 83
log files 72, 72, 73, 73, 75, 76, 77
logrotate 33

M

maintenance agent 22, 55, 112
management 12
management framework 35
Management Node
  virtual IP address 12, 18, 27
metadata 69
metadata store 62, 62
MetaStore
  batch processing 36
  data 22
  defragmentation 69
  full
    automatic 34
    reactivating 34
  head.db 35, 35
  management and monitoring 22
  reactivating 35
  read only 35
  recovery 35, 35, 35, 35
  replay 35, 35
  status 70
  system
    env_metastore 22, 116
    framework 22, 116
  tlog collapse 35, 35, 69
  transaction logs (tlogs) 35, 35, 35, 36, 69
  updates 36
metering 32
MIB 100
monitoring 58, 85, 86, 88, 97, 98, 98, 98, 99
monitoring agent 22, 35, 86
monitoring cycle 86
multi-geo 116

N

name spaces
  _metering_info 32, 33
  deleting 29
network
  firewall rules 42
  IP addresses 42
  NTP server 42
  primary private network 42
  public 41
  secondary private network 42
  SFP+ 1G module 42
  ssh 42
  status 403 response 42
  virtual IP address 42
  wrong port 42
NIC 61
node
  controller
    SSDs 110
  decommissioning 57
  management
    virtual IP address 12, 27
  replacement 57
  storage 113
NTP server 108, 108, 115

O

objects
  deleting 30, 30
  unverified 60
  verification 180
OpenSSL 111
OSMI 16, 16, 26, 113

P

passwords
- master password 56
- master password identifier 56

policies 37, 112
policy
- aggregate storagepool info 70
- auto decommission disks 55, 68, 68, 68
- backup model database 66
- backup software repository 66
- clean up old versions 71
- collapse MetaStore transaction logs 69
- defragment MetaStores 69
- events 70
- jobs
  - failed 71
  - successful 71
- maintenance 66, 70
- monitor blockstores 71
- monitoring 86
- phone home 70

port mapper 115
ports
- pound 117
- S3 117

postgres 115, 115
PostgreSQL
- mount point 58
- partition 61

pound 33, 111, 115, 117
power down 20, 21, 21
power supply unit (PSU)
- failure 149

power up 20, 20, 20, 20
PXE 116

Q

Q-Shell 12, 15, 16, 22, 26, 29, 30, 60, 89, 96, 100, 104, 106

R

rack
- serial number 101

RAID 67, 67, 118, 136, 149
recovery 35, 35, 35, 35, 66, 69
repair crawl 55
repair operation 67
repair spread 113
RSA ciphers 33
rsync 115

S

S3 29
s3cmd
- installing 27

S3 domain name 27
sandboxtmp 62
scope 11
secure erase
- enhanced 56
- normal 56

security state SEC1 56
services 115
shutdown
- unclean 35

small file support 37
SNMP 95, 99, 100, 115
startup 20, 20, 20, 20
startup log 22
storage
- MetaStore
  - batch processing 36
  - full
    - automatic 34
    - reactivating 34
  - head.db 35, 35
  - reactivating 35
  - read only 35
  - recovery 35, 35, 35
  - replay 35, 35
  - tlog collapse 35, 35
  - transaction logs (tlogs) 35, 35, 35, 36
- S3
  - buckets
    - adding 28
  - client
    - adding 27
  - metering
    - disabling 32
    - enabling 32
    - logs 32
  - multipart 31
  - parameters 31
  - users
    - adding 26, 26, 26, 27

superblock 110, 113
system expansion 62
system health 96

T

telemetry 101
time zone 108
troubleshooting 18, 22, 22, 36, 41, 48, 49, 49, 49, 51, 52, 53, 54, 54, 55, 59, 61, 68, 83
typographical 9

V

VGA 12
virtual IP address 18, 42, 101

W

WFE 115

X

XML-RPC 109