



# View for ClusterStor™ Administration Guide

**(1.2.0)**

**S-3026**

---

# Contents

1 About View for ClusterStor™ Administration Guide .....	4
2 Command Prompt Conventions.....	6
3 About Restarting the View for ClusterStor Service.....	8
4 Log in to View for ClusterStor.....	9
4.1 Log Out from View for ClusterStor.....	10
5 System Overview.....	12
5.1 Other GUI Elements.....	17
6 User Account Configuration.....	25
7 Alarms Overview.....	28
7.1 View Current Alarm States.....	28
7.2 Cray Defined Alarms.....	32
7.3 Site Local Alarms Fields.....	35
7.4 Override Cray Defined Alarms.....	38
7.5 Create New Alarm Definition.....	39
7.6 Configure Notification Email.....	41
8 Jobs Table.....	42
9 Create and Modify Dashboards.....	46
10 Dashboard Metrics.....	48
10.1 Lustre Metrics.....	49
10.2 Health Metrics.....	52
10.3 InfiniBand Metrics.....	52
10.4 Jobstats Metrics.....	54
11 Predefined Dashboards.....	58
11.1 ClusterStor Server Metrics Dashboard.....	61
11.2 ClusterStor Storage Overview Dashboard.....	62
11.3 InfiniBand Errors Dashboard.....	63
11.4 InfiniBand Performance Dashboard.....	67
11.5 Job Details Dashboard.....	69
11.6 Job Details Compare Dashboard.....	71
11.7 Lustre File System Capacity Dashboard.....	73
11.8 Lustre File System Metadata Dashboard.....	76
11.9 Lustre File System Performance Dashboard.....	77
12 InfiniBand Features.....	79
12.1 Reset InfiniBand Topology Baseline.....	80
12.2 Modify or Rebuild the InfiniBand Name Map File.....	80

---

13 Annotations.....	83
14 Sample Workflows.....	86
14.1 Dashboard Navigation.....	86
14.2 Jobs Workflow.....	87
14.2.1 Job Details Dashboards.....	89
14.2.2 OSTs Table.....	92
14.3 OST Workflow.....	94
14.4 MDT Workflow.....	95
14.5 System Performance Workflow.....	96
14.6 InfiniBand Workflow.....	97
14.6.1 InfiniBand Logs.....	98
14.7 ClusterStor Logs.....	99
14.8 Health Events.....	100
15 Update Retention Policies for View for ClusterStor.....	102
16 Collect Current Infrastructure State.....	104

# 1 About View for ClusterStor™ Administration Guide

## Scope and Audience

View for ClusterStor™ is a monitoring and metrics software package created by Cray®, which collects and persists performance and job metrics specific to the Cray ClusterStor storage system. View for ClusterStor collects Lustre® performance metrics, jobs metrics, and system events specific to each storage system that is being monitored. Additionally, system logs, system metrics, and system events from each storage system can be configured to be monitored and View for ClusterStor will collect and persist ibstats metrics from the InfiniBand fabric if connected to the ClusterStor high speed InfiniBand network.

The *View for ClusterStor Administration Guide* covers procedures for using View for ClusterStor's graphical user interface (GUI), which provides a visual representation, via data dashboards and workflows, of performance and jobs metrics and other information that has been collected at different points of time since system monitoring began. This guide assumes the reader is familiar with ClusterStor storage systems.

## Release Information

The *View for ClusterStor Administration Guide* supports View for ClusterStor 1.2.0.

## Record of Revision

Revision	Date	Content Information
View for ClusterStor Administration Guide 1.2.0	12/7/2018	Release 1.2.0
View for ClusterStor Administration Guide 1.1.0	08/13/2018	Release 1.1.0
View for ClusterStor Administration Guide 1.0.1	05/16/2018	Release 1.0.1
View for ClusterStor Administration Guide	03/14/2018	GA release

## Typographic Conventions

*Monospace*

Indicates program code, reserved words, library functions, command-line prompts, screen output, file/path names, key strokes (e.g., `Enter` and `Alt-Ctrl-F`), and other software constructs.

**Monospaced Bold**

Indicates commands that must be entered on a command line or in response to an interactive prompt.

*Oblique or Italics*

Indicates user-supplied values in commands or syntax definitions.

<b>Proportional Bold</b>	Indicates a graphical user interface window or element.
\ (backslash)	At the end of a command line, indicates the Linux® shell line continuation character (lines joined by a backslash are parsed as a single line). Do not type anything after the backslash or the continuation feature will not work correctly.
<code>smaller font size</code>	Some screenshot and code examples require more characters than are able to fit on a line of a PDF file, resulting in the code wrapping to a new line. To prevent wrapping, some examples are displayed with a smaller font to preserve the file formatting.

## Other Conventions

Sample commands and command output used throughout this publication are shown with a generic file system name of `cls12345`.

## Trademarks

The following are trademarks of Cray Inc. and are registered in the United States and other countries: CRAY and design, SONEXION, URIKA, and YARCDATA. The following are trademarks of Cray Inc.: APPRENTICE2, CHAPEL, CLUSTER CONNECT, ClusterStor, CRAYDOC, CRAYPAT, CRAYPORT, DATAWARP, ECOPHLEX, LIBSCI, NODEKARE. The following system family marks, and associated model number marks, are trademarks of Cray Inc.: CS, CX, XC, XE, XK, XMT, and XT. The registered trademark LINUX is used pursuant to a sublicense from LMI, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis. Other trademarks used in this document are the property of their respective owners.

## 2 Command Prompt Conventions

### Host Name and Account in Command Prompts

The host name in a command prompt indicates where the command must be run. The account that must run the command is also indicated in the prompt.

- The `root` or super-user account always has the `#` character at the end of the prompt.
- Any non-`root` account is indicated with `account@hostname>`. A user account that is neither `root` nor `crayadm` is referred to as `user`.

<code>smw#</code>	Run the command on the SMW as <code>root</code> .
<code>sdb#</code>	Run the command on the SDB node as <code>root</code> .
<code>boot#</code>	Run the command on the boot node as <code>root</code> .
<code>login#</code>	Run the command on any login node as <code>root</code> .
<code>hostname#</code>	Run the command on the View for ClusterStor system as <code>root</code> .
<code>user@hostname&gt;</code>	Run the command on the specified system as any non- <code>root</code> user.
<code>[node]\$</code>	Run the command on the specified ClusterStor node as the <code>admin</code> user.
<code>[node]#</code>	Run the command on the specified ClusterStor node as <code>root</code> .

### Directory Path in Command Prompt

Example prompts do not include the directory path, because long paths can reduce the clarity of examples. Most of the time, the command can be executed from any directory. When it matters which directory the command is invoked within, the `cd` command is used to change into the directory, and the directory is referenced with a period (`.`) to indicate the current directory.

For example, here are actual prompts as they appear on the system:

```
smw:~ # cd /etc
smw:/etc# cd /var/tmp
smw:/var/tmp# ls ./file
smw:/var/tmp# su - crayadm
crayadm@smw:~> cd /usr/bin
crayadm@smw:/usr/bin> ./command
```

And here are the same prompts as they appear in this publication:

```
smw# cd /etc
smw# cd /var/tmp
smw# ls ./file
smw# su - crayadm
```

```
crayadm@smw> cd /usr/bin  
crayadm@smw> ./command
```

## 3 About Restarting the View for ClusterStor Service

---

A number of procedures in this guide include a step requiring that the View for ClusterStor service be restarted. Before restarting the service, it is recommended that the user log out of any open View for ClusterStor Graphical User Interface (GUI) sessions. Restarting the service without closing open GUI sessions can cause display problems or other unexpected behavior in the GUI.

The proper procedure to follow when it is necessary to restart View for ClusterStor is as follows:

1. Log out from View for ClusterStor, as described in [Log Out from View for ClusterStor](#) on page 10.
2. Restart the View for ClusterStor service on the View server.

```
hostname# systemctl restart sma
```

3. Restart the View for ClusterStor GUI, if desired, after the service has restarted:
  - a. Reload the browser tab that was previously used to work with the GUI, or open a new browser tab and navigate to the GUI URL.
  - b. Log in to the new GUI session.

## 4 Log in to View for ClusterStor

---

### Prerequisites

- View for ClusterStor™ has been installed and configured.
- The user account has been created by the system administrator.
- The site-specific URL for View for ClusterStor has been obtained from the site's system administrator.

### About this task

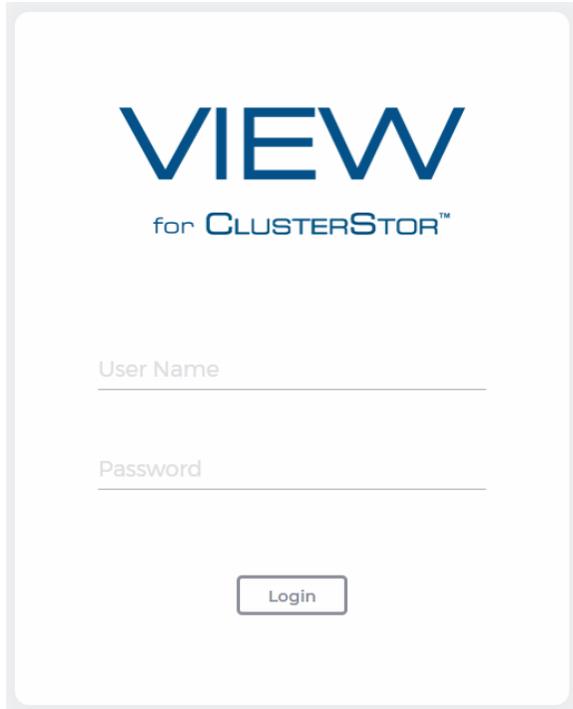
Follow this procedure to log in to the View for ClusterStor graphical user interface (GUI) to begin monitoring ClusterStor systems metrics, job events, and health information.

Note that View for ClusterStor uses the Grafana Global Users feature for authenticating users when they log in to View for ClusterStor. View for ClusterStor administrators configure login information directly from the Grafana interface. See [User Account Configuration](#) on page 25.

### Procedure

1. Open a browser on a client machine.
2. Navigate to the site-specific URL for the View for ClusterStor server.

The View for ClusterStor **Log In** window will appear.



VIEW  
for CLUSTERSTOR™

User Name

Password

Login

3. Enter the **User Name** and **Password**.

4. Click **Login**.

The **System Overview** will open. See: [System Overview](#) on page 12.

## 4.1 Log Out from View for ClusterStor

### Prerequisites

- The user account has been created by the system administrator.
- The user is logged in to the View for ClusterStor™ GUI via the site-specific URL.

### About this task

Follow this procedure to log out from the View for ClusterStor graphical user interface (GUI).

### Procedure

1. Click the **User** icon or the user ID at the upper right of the View for ClusterStor GUI.



2. Select **Logout**.

The user is successfully and securely logged out from View for ClusterStor.

View for ClusterStor coordinates security with Grafana in a parent/child relationship. It is recommended that users log out via the View for ClusterStor logout functionality, as described above, to securely log out. It is possible that a user might log out directly from Grafana, however, this method is not considered secure. When logging out via Grafana, the user must clear their browser cache of all Grafana and View for ClusterStor cookies to ensure they've been logged out completely and securely.

## 5 System Overview

The **System Overview** displays in the content pane of the View for ClusterStor graphical user interface (GUI). The GUI appears as shown in the following figures. Refer to the table that follows for explanations of the GUI elements.

The **System Overview** is displayed as a collection of rectangular tiles, as shown in [View for ClusterStor GUI with Tiled System Overview](#) on page 12. Each tile presents information for a specific ClusterStor system.

When the number of tiles for monitored ClusterStor systems exceeds the display area of the browser, the tiles change into a row-based format, as shown in [View for ClusterStor GUI with System Overview Tiles in Row Format](#) on page 13.

Figure 1. View for ClusterStor GUI with Tiled System Overview

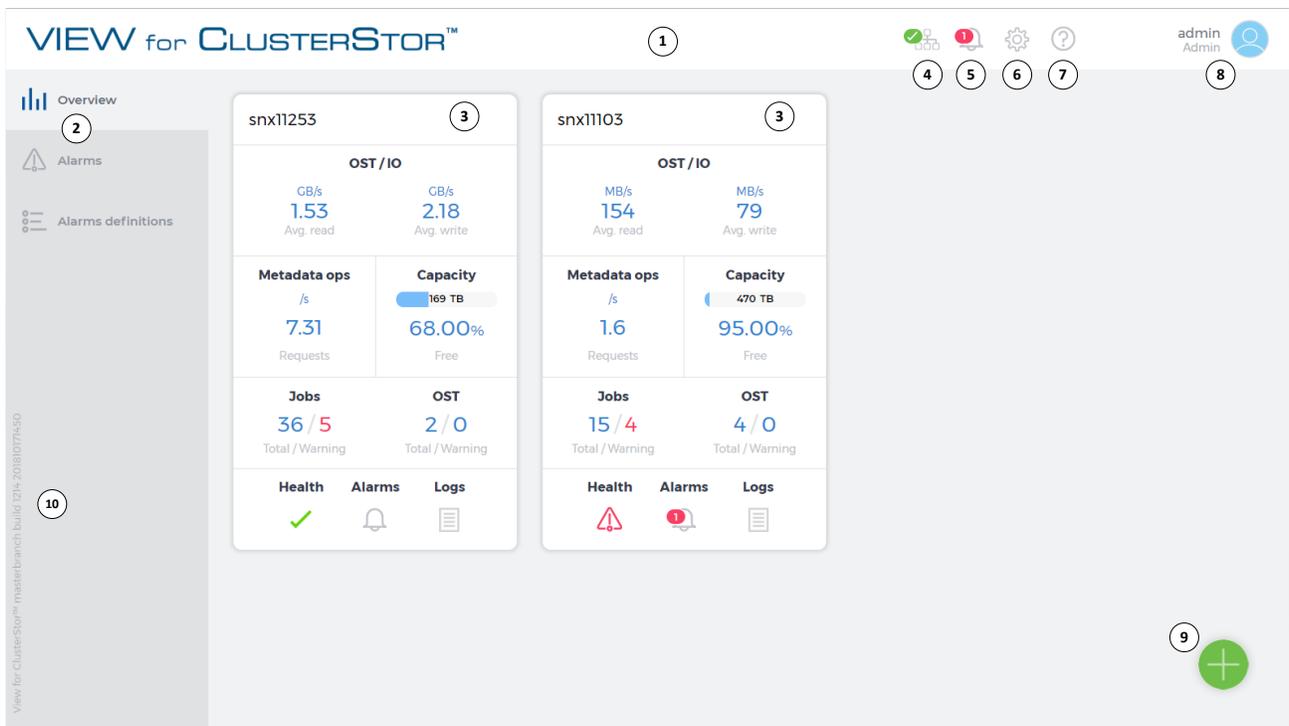


Figure 2. View for ClusterStor GUI with System Overview Tiles in Row Format

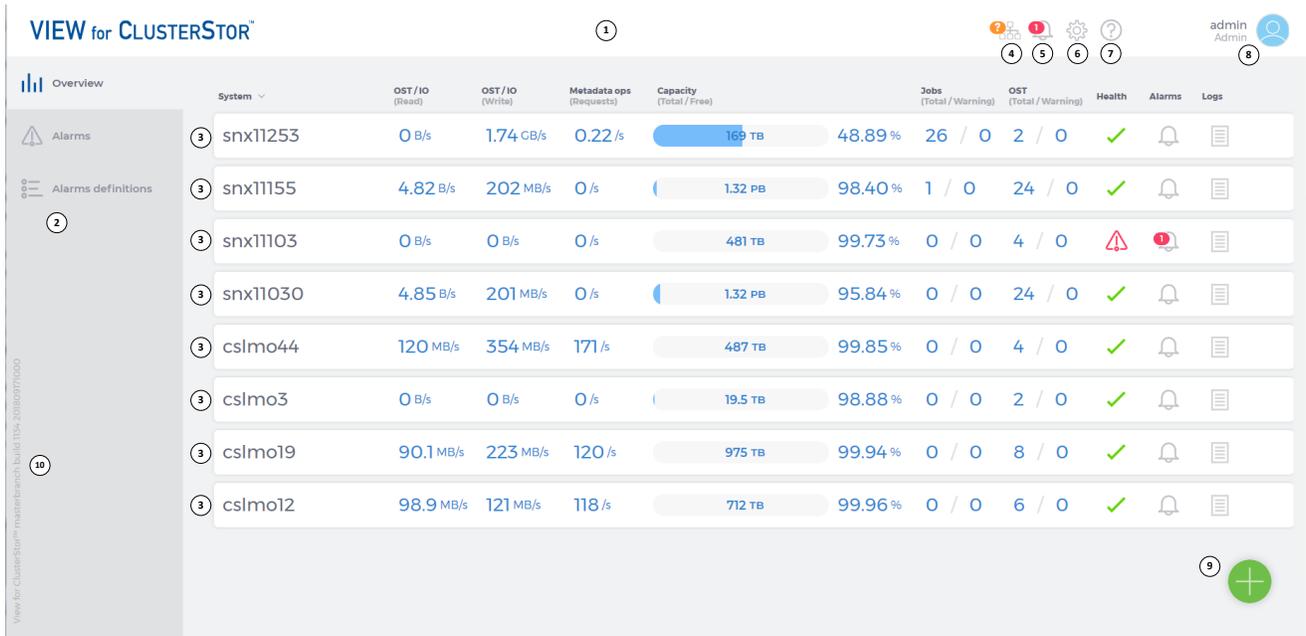


Table 1. GUI Elements

Ref. No.	Description								
1	Header - Displays the View for ClusterStor logo and the icons described in the rows below.								
2	Menu - Select to display the <b>System Overview</b> , <b>Alarms Table</b> , or <b>Alarms Definitions Table</b> in the content pane. See <a href="#">General GUI Layout</a> on page 18 for reference information about the content pane.								
3	Monitored system tiles.								
4	<p><b>IB (InfiniBand) icon</b> - Select to open the <b>IB Page</b> in Kibana. The icon will indicate the state of the IB network with the colored icons:</p> <table border="1"> <thead> <tr> <th>Icon</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td></td> <td>IB OK.</td> </tr> <tr> <td></td> <td>IB warning.</td> </tr> <tr> <td></td> <td>IB undetermined.</td> </tr> </tbody> </table>	Icon	Description		IB OK.		IB warning.		IB undetermined.
Icon	Description								
	IB OK.								
	IB warning.								
	IB undetermined.								
5	<b>Alarms icon</b> - Select to open the <b>Alarms Table</b> to display all alarms for all monitored systems. The icon will indicate:								

Ref. No.	Description	
	<b>Icon</b>	<b>Description</b>
		Number of alarms with Alarm state.
		Number of alarms with Undetermined state.
6	<b>Settings</b> icon - Select to add new users, set permissions, change passwords.	
7	<b>Help</b> icon - Select to launch online help.	
8	<b>User</b> icon - displays currently logged in user name. Select to log out.	
9	<b>Add</b> icon - Use to add a ClusterStor system for monitoring.	
10	Hover to display the build information for the installed version of View for ClusterStor in a larger, darker font. Select to copy the build information.	

## System Tiles

The **System Overview** provides a high-level status of each ClusterStor system being monitored. Each information block of a system tile provides navigation to more detailed metrics and information about the associated ClusterStor system. Values displayed in the **System Overview** tiles represent the last 15 minute time period at the point when the was last refreshed. The **System Overview** does not refresh automatically. To refresh it manually, reload the browser page.

Please refer to the following diagrams and the table that follows for explanations of the information blocks and the navigation action that occurs when selecting the block.

Figure 3. System Overview Tile

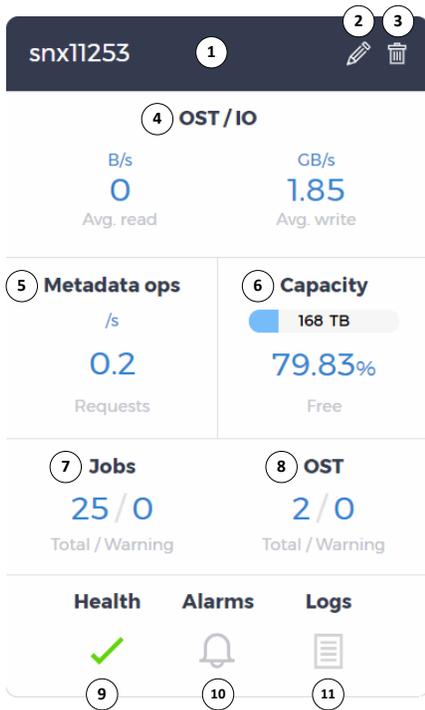


Figure 4. System Overview Tile (Row Format)

System	OST / IO (Read)	OST / IO (Write)	Metadata ops (Requests)	Capacity (Total / Free)	Jobs (Total / Warning)	OST (Total / Warning)	Health	Alarms	Logs	
snx11253	0 B/s	1.74 GB/s	0.22 /s	169 TB	48.89 %	26 / 0	2 / 0	✓	🔔	📄
snx11155	4.82 B/s	202 MB/s	0 /s	1.32 PB	98.40 %	1 / 0	24 / 0	✓	🔔	📄
snx11103 (1)	0 B/s (4)	0 B/s (5)	0 /s	481 TB (6)	99.73 %	0 / 0 (7)	4 / 0 (8)	⚠️ (9)	🔔 (10)	📄 (11) ✎️ (2) 🗑️ (3)
snx11030	4.85 B/s	201 MB/s	0 /s	1.32 PB	95.84 %	0 / 0	24 / 0	✓	🔔	📄
cslmo44	120 MB/s	354 MB/s	171 /s	487 TB	99.85 %	0 / 0	4 / 0	✓	🔔	📄
cslmo3	0 B/s	0 B/s	0 /s	19.5 TB	98.88 %	0 / 0	2 / 0	✓	🔔	📄

Table 2. Explanation of Tile GUI Elements

Ref. No.	Description	Navigates to:
1	Name of the monitored system. Hovering over this in a tile turns the tile header black and displays the <b>Pencil</b> and <b>Trash</b> icons. Hovering over this when the tiles appear in the row format	<b>ClusterStor Storage Overview Dashboard</b> (Grafana)

Ref. No.	Description	Navigates to:						
	highlights the entire tile and displays the <b>Pencil</b> and <b>Trash</b> icons.							
2	<b>Pencil</b> icon - Select to change the configuration of the monitored system. The <b>Pencil</b> icon is only displayed when admin users are logged into View for ClusterStor.	<b>Edit</b> window						
3	<b>Trash</b> icon - Select to remove the monitored system from View for ClusterStor.	Opens a confirmation window.						
4	<b>OST I/O</b> information block: <ul style="list-style-type: none"> <li>• Average read rate per second</li> <li>• Average write rate per second</li> </ul>	<b>Lustre File System Performance Dashboard</b> (Grafana) Displays information specific to the selected ClusterStor system.						
5	<b>Metadata ops</b> information block: Average ops requests per second for all MDTs.	<b>Lustre File System Metadata Dashboard</b> (Grafana) Displays information specific to the selected ClusterStor system.						
6	<b>Capacity</b> information block: Percentage of free file space capacity across all OSTs. The capacity bar graph will show used space in colors that depict capacity status. See <a href="#">Capacity Color Legend</a> .	<b>Lustre File System Capacity Dashboard</b> (Grafana) Displays information specific to the selected ClusterStor system.						
7	<b>Jobs</b> information block: Displays total jobs and number of jobs with warnings.	<b>Jobs Table</b> Displays information specific to the selected ClusterStor system.						
8	<b>OST</b> information block: Displays total number of OSTs and number of OSTs with warnings.	<b>OSTs Table</b> Displays information specific to the selected ClusterStor system.						
9	<b>Health</b> information block: Provides overall system health at a glance. <table border="1" data-bbox="298 1509 875 1740"> <thead> <tr> <th>Icon</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td></td> <td>A green check mark indicates no warnings.</td> </tr> <tr> <td></td> <td>A red warning triangle indicates there are warnings.</td> </tr> </tbody> </table>	Icon	Description		A green check mark indicates no warnings.		A red warning triangle indicates there are warnings.	<b>Health Page</b> (Kibana) Displays information specific to the selected ClusterStor system.
Icon	Description							
	A green check mark indicates no warnings.							
	A red warning triangle indicates there are warnings.							

Ref. No.	Description	Navigates to:						
	<table border="1"> <thead> <tr> <th>Icon</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td></td> <td>An orange question mark indicates health is undetermined or information is missing.</td> </tr> </tbody> </table>	Icon	Description		An orange question mark indicates health is undetermined or information is missing.			
Icon	Description							
	An orange question mark indicates health is undetermined or information is missing.							
10	<p><b>Alarms</b> information block: Indicates number of alarms whose state is Alarm. These alarm indicators are specific to the system being monitored by the tile.</p> <table border="1"> <thead> <tr> <th>Icon</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td></td> <td>Number of alarms with Alarm state for the monitored ClusterStor system.</td> </tr> <tr> <td></td> <td>Number of alarms with Undetermined state for the monitored ClusterStor system.</td> </tr> </tbody> </table>	Icon	Description		Number of alarms with Alarm state for the monitored ClusterStor system.		Number of alarms with Undetermined state for the monitored ClusterStor system.	<p><b>Alarms Table</b></p> <p>Displays information specific to the selected ClusterStor system. To view alarms for all monitored systems, select the <b>Alarms</b> option from the main menu, or select the <b>Alarms</b> icon from the GUI header.</p>
Icon	Description							
	Number of alarms with Alarm state for the monitored ClusterStor system.							
	Number of alarms with Undetermined state for the monitored ClusterStor system.							
11	<p><b>Logs</b> information block: Select to access collected ClusterStor system logs.</p>	<p><b>Logs Page (Kibana)</b></p> <p>Displays information specific to the selected ClusterStor system.</p>						

Table 3. Capacity Color Legend

COLOR	DESCRIPTION
Light Gray 	Normal state; capacity mostly unused.
Light Blue 	Normal state.
Orange 	Warning state; capacity available below threshold.
Red 	Critical state; capacity almost full.

## 5.1 Other GUI Elements

In addition to the graphical user interface (GUI) elements described in [System Overview](#) on page 12, this section describes many of the other general GUI elements.

## General GUI Layout

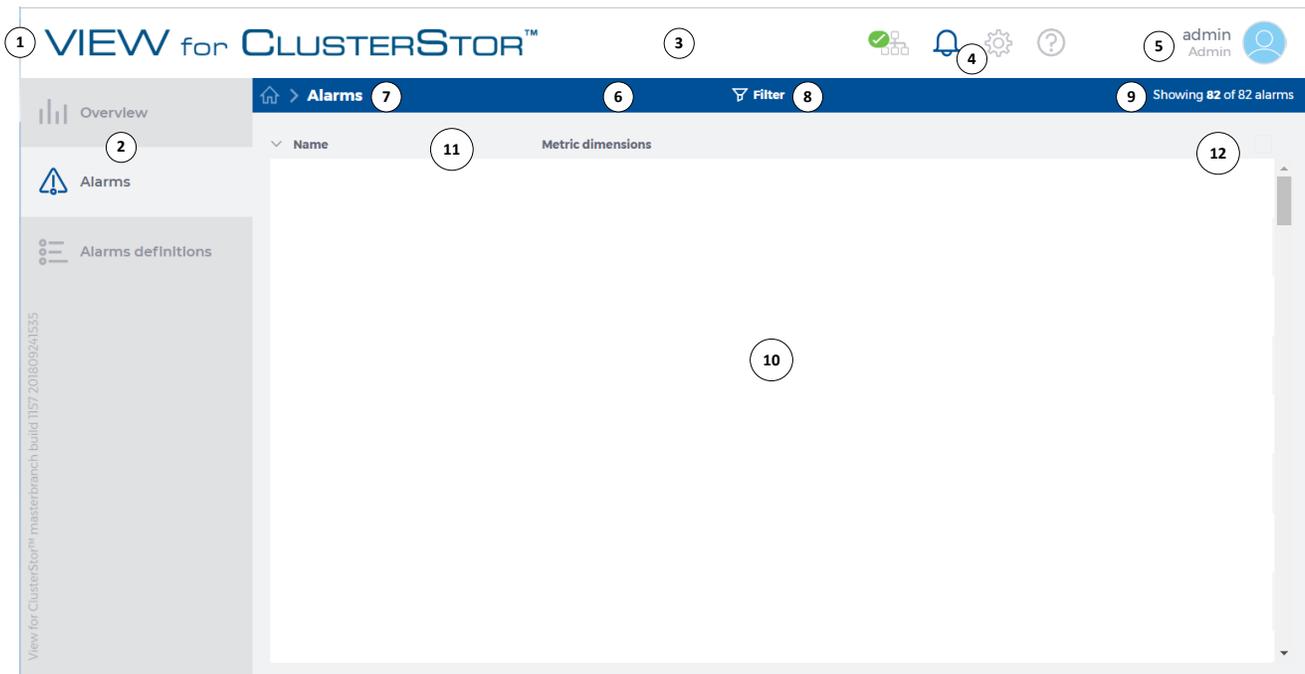


Table 4. General GUI Elements

Ref. No.	Description
1	View for ClusterStor™ logo.
2	Menu - Use to select what displays in the Content Pane.
3	General GUI Header - Displays the View for ClusterStor logo and general icons.
4	General icons - Use to access the <b>IB Page</b> , <b>Alarms Table</b> , <b>Settings</b> , and <b>Help</b> . See <a href="#">Icons Reference</a> on page 19.
5	<b>User</b> icon. See <a href="#">Icons Reference</a> on page 19.
6	Content pane banner. See <a href="#">Content Pane Banner</a> on page 22.
7	Breadcrumbs. See <a href="#">Content Pane Banner</a> on page 22.
8	Actions. See <a href="#">Content Pane Banner</a> on page 22.
9	Information. See <a href="#">Content Pane Banner</a> on page 22.
10	Content pane - Display area for the <b>Alarms Table</b> , <b>Alarms Definitions Table</b> , <b>Jobs Table</b> , <b>OSTs Table</b> , Grafana Dashboards, and Kibana Pages.
11	Content pane column headings.

Ref. No.	Description						
12	Content pane actions icons, displayed on the right side of the column headings row for the <b>Alarms Table</b> and <b>Jobs Table</b> . Numbers shown indicate the number of selected table rows.						
	<table border="1"> <thead> <tr> <th>Table</th> <th>Icons</th> </tr> </thead> <tbody> <tr> <td><b>Alarms Table</b></td> <td>3  </td> </tr> <tr> <td><b>Jobs Table</b></td> <td>2   </td> </tr> </tbody> </table>	Table	Icons	<b>Alarms Table</b>	3  	<b>Jobs Table</b>	2   
Table	Icons						
<b>Alarms Table</b>	3  						
<b>Jobs Table</b>	2   						
	See <a href="#">Icons Reference</a> on page 19.						

## Icons Reference

Icon	Description	Used In
	<b>Add icon</b>	<b>System Overview</b> Note that the ClusterStor system being added must be properly configured to enable Cstream metrics, SNMP, and log forwarding. Also, the name used when adding the ClusterStor system for monitoring must match the file system name used on the ClusterStor system.  Please see the <i>View for ClusterStor Installation and Configuration Guide</i> for details.
	<b>Alarms Definitions icon</b>	Menu Grayed out when not selected.
	<b>Alarms icon</b>	Menu Grayed out when not selected.
	<b>Alarms icon</b>	GUI header, displays in blue when hovered over, remains blue once icon is selected and <b>Alarms Table</b> is opened. <b>System Overview - Health</b>
	<b>Alarms icon with warning</b>	GUI header, displays in blue when hovered over, remains blue once icon is selected and <b>Alarms Table</b> is opened. <b>System Overview - Health</b>
	<b>Alarms icon with undetermined</b>	GUI header, displays in blue when hovered over, remains blue once icon is selected and <b>Alarms Table</b> is opened. <b>System Overview - Health</b>

Icon	Description	Used In
	<b>Alarms</b> icon with warning and undetermined	GUI header, displays in blue when hovered over, remains blue once icon is selected and <b>Alarms Table</b> is opened. Hovering over the icon clearly shows the number of alarms with warnings and with undetermined state <b>System Overview</b> - Health
	<b>Compare</b> icon	<b>Jobs Table</b> Displays in blue when hovered over.
	<b>Copy</b> icon	<b>Alarms Dimensions</b> window.
	<b>Filter</b> icon <sup>1</sup>	Content pane banner
	<b>Filter</b> icon (small) <sup>1</sup>	Table column headings, displays in blue when hovered over
	<b>Grafana Admin Menu</b> icon	Dashboards
	<b>Grafana Menu</b> icon	Dashboards
	<b>Health OK</b> icon	<b>System Overview</b> - Health
	<b>Health Warning</b> icon	<b>System Overview</b> - Health
	<b>Health Unknown</b> icon	<b>System Overview</b> - Health
	<b>Help</b> icon	GUI header, displays in blue when hovered over
	<b>Home</b> icon	Content pane banner
	<b>IB</b> icon	GUI header, displays in blue when hovered over
	<b>IB OK</b> icon	GUI header, displays in blue when hovered over, remains blue once icon is selected and <b>IB Page</b> is opened in Kibana.
	<b>IB Unknown</b>	GUI header, displays in blue when hovered over, remains blue once icon is selected and <b>IB Page</b> is opened in Kibana.

Icon	Description	Used In
	<b>IB Warning</b> icon	GUI header, displays in blue when hovered over, remains blue once icon is selected and <b>IB Page</b> is opened in Kibana.
	<b>Details</b> icon	<b>Jobs Table</b> Displays in blue when hovered over.
	<b>Logs</b> icon	<b>System Overview</b>
	<b>Pencil</b> icon	Content pane banner The <b>Pencil</b> icon is only displayed when admin users are logged into View for ClusterStor.
	<b>Settings</b> icon	GUI header, displays in blue when hovered over, remains blue once icon is selected and Grafana <b>Users Page</b> is opened. The <b>Settings</b> icon is only displayed when admin users are logged into View for ClusterStor.
	<b>System Overview</b> icon	Menu Grayed out when not selected.
	<b>Trash</b> icon	<b>System Overview, Alarms Table</b> Displays in red when hovered over.
	<b>Deselect</b> icon	<b>Alarms Table, Jobs Table</b>
	<b>User</b> icon	GUI header

### <sup>1</sup>About Rounding and Filtering

Data values displayed on some View for ClusterStor tables is rounded to facilitate visual presentation. This affects which data values will be included in the display when applying a range filter, which uses a visual slider to set a minimum and maximum value for the range. Consider the following example of the **Write Bandwidth (Is)** column of the **OSTS Table**. There are currently four jobs where the write bandwidth is displayed as 350MB before applying a range filter of 0B to 350MB.

Job	Device	Write Bandwidth (/s)	Read Bandwidth (/s)
2498173	OST0000	350MB	
2498177	OST0000	350MB	N/A
2498187	OST0000	350MB	N/A
2498170	OST0000	350MB	N/A
2498180	OST0000	346MB	N/A
2498183	OST0000	245MB	N/A

After applying the range filter, only two jobs showing a write bandwidth of 350MB remain, as seen in the next figure. Jobs 2498173 and 2498177 have been filtered out. This is because the actual write bandwidth for these two jobs had been rounded down to 350MB for display purposes. Their unrounded values may have been 350.4MB and 350.2MB, respectively, so they were dropped from the filtered display.

Job	Device	Write Bandwidth (/s)	Read Bandwidth (/s)
2498187	OST0000	350MB	N/A
2498170	OST0000	350MB	N/A
2498180	OST0000	346MB	N/A
2498183	OST0000	245MB	N/A
2498175	OST0000	187MB	N/A
2498178	OST0001	147MB	N/A

## Content Pane Banner

Notice the blue content pane banner shown in the figure in [General GUI Layout](#) on page 18. With the exception of the **System Overview**, the banner appears over any tables, dashboards, or pages that appear in the content pane of the View for ClusterStor graphical user interface (GUI). The banner includes:

Where in Banner	Displays	Description
Left side	Breadcrumbs	<p>Depicts the current location within the application's navigation path. This is the only item displayed in the content pane banner for Grafana dashboards or Kibana pages. Note that the <b>Home</b> icon is always a live link in the displayed navigation path.</p> <p>Example:</p>

Where in Banner	Displays	Description
		Some other navigation paths will have additional live links. For example, in the path <code>home &gt; Alarms &gt; Alarm History</code> , <b>Alarms</b> is a live link.
Center	Actions	<p>Displays any actions the user may take to change the display. For example, the <b>Filter</b> icon on the <b>Alarms</b> or <b>Jobs</b> tables, or time range on the <b>OSTs Table</b>.</p> <p>Examples:</p>  <p>Note that after applying a filter, the number of applied filters will display next to the <b>Filter</b> icon. Select the small <b>X</b> icon to remove the applied filters.</p> 
Right side	Information	<p>Displays information specific to the contents being displayed. For example:</p> <ul style="list-style-type: none"> <li> <b>Alarms Table:</b> Number of alarms displayed and total number of alarms.            </li> <li> <b>Alarms Definitions Table:</b> Email address where notifications are sent.            </li> <li> <b>Jobs Table:</b> Number of jobs loaded, the time it took to load them, and a button to refresh information in the table.            </li> <li> <b>OSTs Table:</b> Number of OSTs displayed and total number of OSTs.            </li> </ul>

## Grafana Header

This section describes the main Grafana header, which appears above every View for ClusterStor dashboard.



Ref. No.	Description
1	Grafana Menu - Configure data sources, users, organizations, server settings.
2	Dashboard Selector - Select from a list of saved dashboards.  <b>NOTE:</b> When navigating directly from a View for ClusterStor tile to a Grafana dashboard, the charts and metrics shown are specific to the ClusterStor system represented in that tile. Once in Grafana, however, using the <b>Dashboard Selector</b> to navigate to a different dashboard will change the scope of the data displayed, to all ClusterStor systems being monitored in View for ClusterStor. To maintain scope to a specific ClusterStor system when navigating within Grafana, use the links that appear when selecting the <b>Details</b> icons at the upper left of any panel in a dashboard.
3	Star Dashboard - Mark or unmark a favorite dashboard. Favorite dashboards appear in the dashboard selector with a star next to the name. The list of saved dashboards can be filtered to show only the user's favorites.
4	Share Dashboard - Share current dashboard by creating a link or snapshot.
5	Settings - Dashboard settings. Make a dashboard editable, display a list of keyboard shortcuts, save a dashboard with a different name.
6	Shift Time Backwards - Change the time range of displayed metrics and graphs to an earlier period.
7	Zoom Out Time Range - Increase the time range used to display metrics and graphs.
8	Shift Time Forwards - Change the time range of displayed metrics and graphs to a later period.
9	Quick Ranges - Select from pre-defined time ranges to apply to the displayed metrics and graphs.
10	Refresh - Manually refresh the metrics and charts displayed.

## 6 User Account Configuration

### Prerequisites

View for ClusterStor™ has been installed.

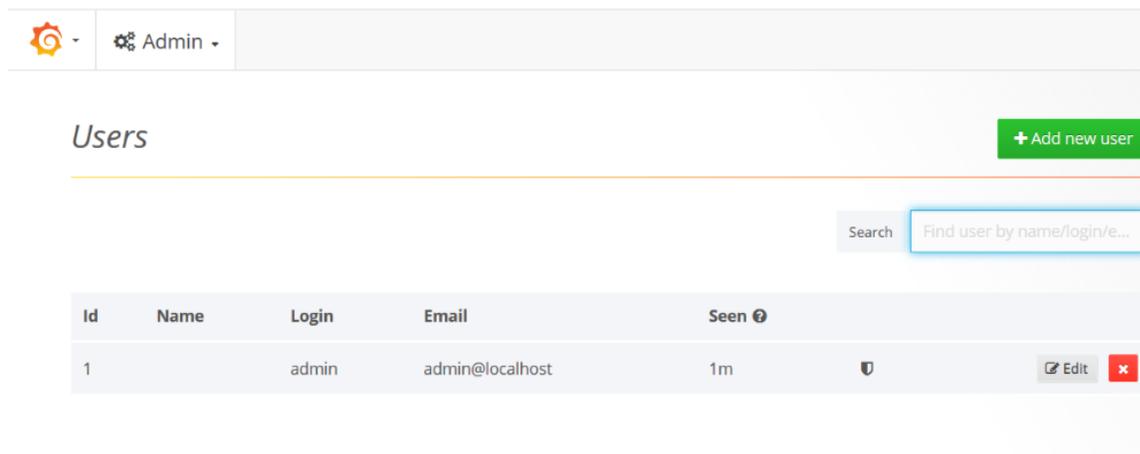
### About this task

View for ClusterStor uses the Grafana Global Users feature for authenticating users when they log in to View for ClusterStor. View for ClusterStor administrators configure login information directly from the Grafana interface. Initial configuration must be done using the Grafana administrator login, which is created during installation. An administrator can create new users and change passwords.

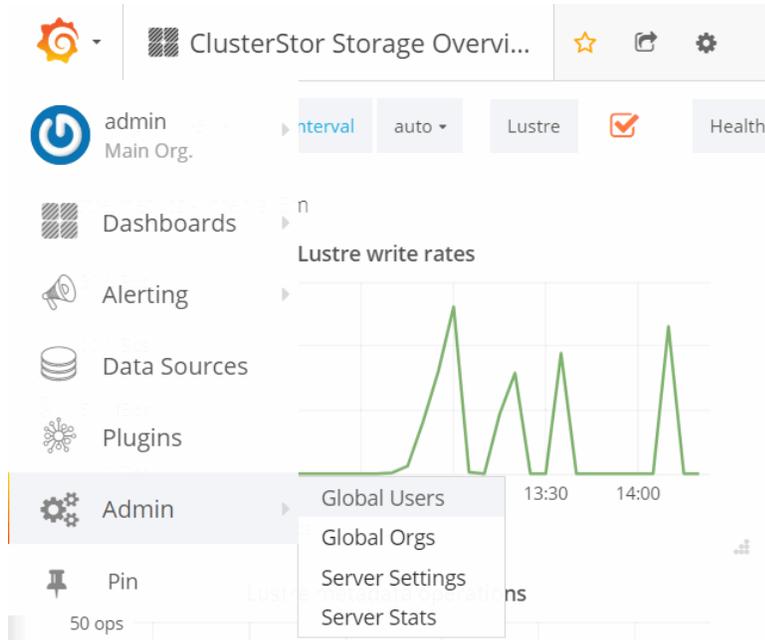
### Procedure

1. Sign in to View for ClusterStor with administrator privileges.
2. Select the **Settings** icon from the View for ClusterStor header, and then select **Users**.

The **Users** page opens in Grafana.



Alternatively, if already in Grafana, open the **Users** page by selecting the **Grafana Menu** icon in the upper left corner, select the **Grafana Admin Menu** icon from the drop down menu, and then select **Global Users**.



From the **Users** page, perform one or more of these actions:

- Add a new user (Step 3 on page 26)
- Modify the information for an existing user (Step 4 on page 26)
- Delete a user (Step 5 on page 27)

3. Select **Add new user** to create a new user. Provide the requested information on the **Add new user** page, and then select **Create**.

Add new user	
Name	<input type="text"/>
Email	<input type="text"/>
Username	<input type="text"/>
Password	<input type="password"/>
<input type="button" value="Create"/>	

4. Select **edit** to change an existing user's information, permissions, or password on the **Edit User** page. Select the appropriate **Update** button after making any changes.

## Edit User

---

Name	
Email	admin@localhost
Username	admin

Update

### Change password

New password

Update

### Permissions

Grafana Admin

Update

Note that users without administrative privileges will be able to change passwords in the same way, but will not be able to edit any other user's information or grant privileges.

5. Select the red **X** icon to the right of an existing user entry to delete that user.

For more information on how to use Grafana, refer to the [Grafana documentation](#) directly.

## 7 Alarms Overview

Alarms and notifications allow the site administrator to monitor the state of the storage system by receiving automated messages about the metrics monitored by the View for ClusterStor™ server. An alarm records the state of the system as indicated by the values in specified metric series.

Alarms are created automatically by View for ClusterStor based on alarm definitions. As new values are added to a metric series, the state of the alarm updates automatically to reflect the most recent values.

An alarm is defined by a set of one or more inequality comparisons of a monitoring subsystem metric value with a threshold value. An alarm definition specifies the type of metric series that the alarm will monitor, the threshold values for determining what the state of the alarm should be, and the frequency for checking the metric series for changes to the metric values.

View for ClusterStor installs several pre-defined alarm definitions. The threshold values for these alarms can be modified by the site administrator. Additional site-local alarm definitions can also be created.

An alarm state may be one of the following:

Alarm State	Description
OK	Indicates that the alarm threshold has not been reached.
Alarm	Indicates that the alarm threshold has been exceeded.
Undetermined	Indicates that the metric series did not receive a new value when expected.

A notification email is a message sent automatically when the state of an alarm changes. The alarm definition specifies which notifications will be sent when the state of an alarm changes.

To disable notifications, the email address can be changed to root@localhost. For more information, see [Configure Notification Email](#) on page 41.

### 7.1 View Current Alarm States

The current state of all alarms can be seen by selecting **Alarms** from the main menu at the left side of the application, or by clicking the **Alarms** icon in the application's header. Selecting an **Alarms** icon from a system-specific tile on the **System Overview** will display all alarms for that system only. Any of these actions will open the **Alarms Table**.

Figure 5. Alarms Table

Name	Alarm ID	Metric name	Metric dimensions
SMA_ClusterStor_service_alerts	baa6e79a-9ce6-4b...	cray_storage.snx_h...	["system_name":"snx1103","product":"ClusterStor","hostname":"snx1103n000","component":"snmpwalk","servi...
IB_symbol_errors	81ea3152-ad7d-4b6...	cray_ib.symbol_err...	["lid":"30","node":"snx1103_Rack_1_Odd_TOR","dest":"caribou01_mlx4_0","disc_port":"mlx4_0-1","guid":"f452140...
IB_symbol_errors	19eb9e6c-5a1d-4b7...	cray_ib.symbol_err...	["lid":"6","node":"snx1103_Rack_1_Even_TOR","dest":"caribou09_mlx4_0","disc_port":"mlx4_0-1","guid":"f452140...
IB_symbol_errors	1fc248ac-a869-4cfa...	cray_ib.symbol_err...	["lid":"6","node":"snx1103_Rack_1_Even_TOR","dest":"snx1103n006_HCA_1","disc_port":"mlx4_0-1","guid":"f4521...
IB_symbol_errors	279d2c26-b259-414...	cray_ib.symbol_err...	["lid":"30","node":"snx1103_Rack_1_Odd_TOR","dest":"nid00118_HCA_1","disc_port":"mlx4_0-1","guid":"f4521403...
SMA_OST_Free_Files	1c10ebc2-1c30-4be...	cray_storage.free_fi...	["system_name":"snx11030","product":"ClusterStor","service":"storage","hostname":"snx1155n010","component"...
SMA_OST_Free_Files	1ed5ecc8-d05d-4e...	cray_storage.free_fi...	["system_name":"snx11030","product":"ClusterStor","service":"storage","hostname":"snx1155n008","component"...
SMA_OST_Free_Files	0239fc7a-5e03-49...	cray_storage.free_fi...	["system_name":"snx11030","product":"ClusterStor","service":"storage","hostname":"snx1155n025","component"...
SMA_OST_Free_Space	2377149d-5945-436...	cray_storage.free_s...	["system_name":"snx11030","product":"ClusterStor","service":"storage","hostname":"unknown","component":"lus...
SMA_OST_Free_Space	259ecb70-924a-4e...	cray_storage.free_s...	["system_name":"cslmo19","product":"ClusterStor","service":"storage","hostname":"cslmo1905","component":"lu...
SMA_OST_Free_Files	25c81ef9-f576-4a6...	cray_storage.free_fi...	["system_name":"cslmo3","product":"ClusterStor","service":"storage","hostname":"cslmo304","component":"lustr...

## Alarms Table Column Descriptions

Column Heading	Description
n/a	Alarm state is indicated in the table's first column by the color of the cell. <ul style="list-style-type: none"> <li>Red: Alarm</li> <li>Orange: Undetermined</li> <li>Transparent: OK</li> </ul>
<b>Name</b>	The name of the alarm definition that was used to create the alarm.
<b>Alarm ID</b>	A globally unique identifier that was assigned to the alarm when it was created.
<b>Metric Name</b>	The name of the metric type that the alarm monitors.
<b>Metric Dimensions</b>	Parameters that distinguish different metric series of the same metric type.

## Filter and Sort the Alarms Table

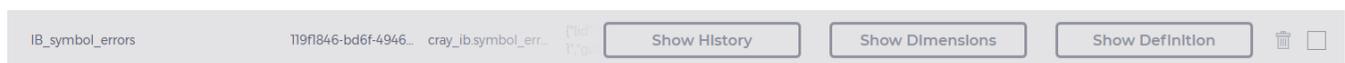
The **Alarms Table** can be filtered and sorted in the following ways:

Method	Location	Description and Use
Filter icon	Content pane banner	Select the <b>Filter</b> icon to open the <b>Filter Alarms</b> window. Specify the desired options for filtering information in the table:

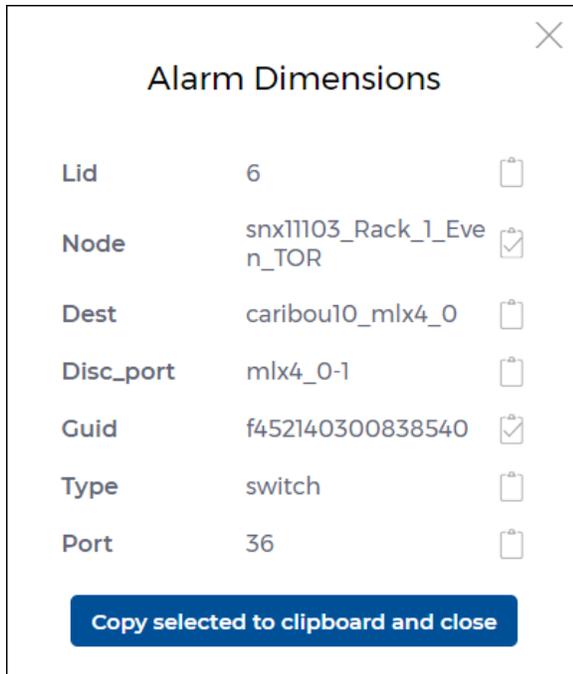
Method	Location	Description and Use
		<ul style="list-style-type: none"> <li>Enter one or more values for alarm ID, alarm name, or metric dimensions.</li> <li>Select one or more of the listed alarm severities.</li> <li>Select one or more of the listed alarm states.</li> </ul> <p>Click <b>Submit</b> to apply the filtering options. The number of applied filters will display next to the <b>Filter</b> icon. Select the small <b>X</b> icon to remove the applied filters.</p>
Column filter icon	Column heading	<p>Create a filter for specific columns in the <b>Alarms Table</b>. Hover over a column heading to display the small <b>Filter</b> icon, and then click the icon.</p> <ul style="list-style-type: none"> <li>Enter some text to search for alarm names, alarm IDs, or metric dimensions. Press <b>Enter</b> to apply the filter.</li> <li>One or more metric names may also be selected from the list of displayed names. Enter text to reduce the number of items shown in the list. Filter options are applied as they are selected. Click the light gray <b>X</b> in the upper right to close the filter options.</li> </ul> <p>The number of applied filters will display next to the <b>Filter</b> icon. Select the small <b>X</b> icon to remove the applied filters.</p>
Column sort	Column heading	Click on a column heading to sort the table by the values in that column. Click a second time to reverse the direction of the sort.
Secondary Column Sort	Column heading	A secondary column can be added to the primary column sort. Click and hold the <b>Alt</b> key while clicking on the second column. To sort the second column in the reverse direction, repeat the procedure.

## Available Alarms Actions

When hovering over a row in the **Alarms Table**, some additional actions will be displayed at the right side of the row.



- **Show History** - Opens the **Alarm History Table**, which lists all of the alarm's state transitions.
- **Show Dimensions** - Opens the **Alarm Dimensions** window, displaying the alarm's metric dimensions.



Specific dimensions can be copied to the clipboard, making it easy to paste the values into the **Metric Dimensions** filter. Select the specific dimensions by clicking the **Copy** icon next to the dimension name. When all desired dimensions have been selected, click **Copy selected to clipboard and close**.

- **Show Definition** - Opens the **Alarm Definition** window, displaying the alarm definition on which the alarm is based.
- **Trash** icon - Use to delete the instantiated alarm that is created from the alarm definition. This icon displays only when an admin user is logged in to View for ClusterStor.
- **Select** checkbox - Use to select multiple alarm rows, up to a maximum of five. Once rows are selected, click the **Trash** icon that appears at the right of the table heading row.

When an alarm is deleted, all history for that alarm is also deleted. When a storage system is removed from View for ClusterStor™, if a node is removed from a storage system, or if any other components are no longer being monitored, an alarm should be deleted. However, if the metric series for a deleted alarm is still active, with new values still being added to its metric series, then the alarm will be re-created automatically when the next value is received.

## Alarm History Table

Time	Transition	Reason
2018-10-17 13:20:01	Ok → Alarm	Thresholds were exceeded for the sub-alarms: last(cray_storage.snx_health) < 3.0 with the values: [1.0], last(cray_storage.snx_health) > 0.0 ...
2018-10-17 12:53:54	Alarm → Ok	The alarm threshold(s) have not been exceeded for the sub-alarms: last(cray_storage.snx_health) > 0.0 with the values: [0.0]
2018-10-16 00:44:12	Ok → Alarm	Thresholds were exceeded for the sub-alarms: last(cray_storage.snx_health) < 3.0 with the values: [1.0], last(cray_storage.snx_health) > 0.0 ...
2018-10-15 20:50:51	Alarm → Ok	The alarm threshold(s) have not been exceeded for the sub-alarms: last(cray_storage.snx_health) > 0.0 with the values: [0.0]
2018-10-15 10:58:15	Ok → Alarm	Thresholds were exceeded for the sub-alarms: last(cray_storage.snx_health) < 3.0 with the values: [1.0], last(cray_storage.snx_health) > 0.0 ...
2018-10-05 14:47:26	Undetermined → Ok	The alarm threshold(s) have not been exceeded for the sub-alarms: last(cray_storage.snx_health) > 0.0 with the values: []

The **Alarm History** Table lists all of the alarm's state transitions. Select **Show Dimensions** to display the **Alarm Dimensions** window.

## 7.2 Cray Defined Alarms

A list of Cray® defined alarm definitions can be viewed by selecting **Alarms Definitions** from the main menu. This opens the **Alarms Definitions Table**. To see alarm definition details, hover over the related table row and then select **Show Alarm Definition**, which will open the **Alarm Definition** window.

Alarms Definitions Notifications are sent to `root@localhost`

Name ^	Description	Notifications
IB_symbol_errors	IB SymbolErrors exceeds 120/s	On
IB_topology_change	IB Topology Change Detection	On
Loss_of_Job_Event_Daemon_Heartbeat	There has been a loss of heartbeats from the Job Event Daemon	On
Root_Partition_Disk_Space_Usage	The disk space of the root partition is over 95%.	On
Root_Partition_inode_Usage	The use of the configured inodes of the root partition is over 95%.	On
SMA_ClusterStor_Metric_Health	Check for Incoming metrics from ClusterStor	On
SMA_ClusterStor_service_alerts	There are one or more service alerts on a ClusterStor	On
SMA_ClusterStor_snmpwalk_timeout	There is a timeout with SNMP communications in the attempt to gain health information from a ClusterStor	On
SMA_OST_Free_Files	The free inode space of a specific OST/MDT is under 5%	On
SMA_OST_Free_Space	The free disk space of a specific OST/MDT is under 5%	On
SSD_Partition_Disk_Space_Usage	The disk space of the SSD partition is over 95%.	On
SSD_Partition_inode_Usage	The use of the configured inodes of the SSD partition is over 95%.	On

Following is a list of Cray defined alarms and their descriptions.

Name	Description
IB_symbol_errors	A View for ClusterStor check on the InfiniBand SymbolErrors metric. When the SymbolErrors exceed 120 errors per second, the alarm will transition to an Alarm state.
IB_topology_change	A View for ClusterStor™ check on the InfiniBand topology. Whenever the InfiniBand fabric topology changes the alarm will transition to an Alarm state and a notification will be generated.
Loss_of_Job_Event_Daemon_Heartbeat	This defines a check of the Job Event Daemon's health. A loss of the heartbeat will transition to an Alarm state.
Root_Partition_Disk_Space_Usage	A View for ClusterStor server system check on the root partition disk space in use. This will issue an alarm when the root partition is over 95%. It should be possible to partially clone this alarm to create an alarm to monitor the amount of space in use for an SSD partition containing the Influx database and the Elasticsearch database.
Root_Partition_inode_Usage	The use of the configured inodes of the root partition is over 95%.

Name	Description
SMA_ClusterStor_Metric_Health	An overall measure of the flow of data from the ClusterStor system to View for ClusterStor.
SMA_ClusterStor_service_alerts	An overall check on the health of the ClusterStor storage cluster by checking the service alerts present on the storage system.
SMA_ClusterStor_snmpwalk_timeout	A timeout with communications during an attempt to gain health information from a ClusterStor system.
SMA_OST_Free_files	A check on inode free space. This alarm will be issued when the percentage amount of free inodes on an OST or MDT device drops below a threshold. The default threshold is 5%.
SMA_OST_Free_Space	A check on free space. This alarm will be issued when the percentage amount of free space on an OST or MDT device drops below a threshold. The default threshold is 5%.
SSD_Partition_Disk_Space_Usage	The disk space usage of the SSD file system partition is over 95%.
SSD_Partition_inode_Usage	The use of the configured inodes of the SSD partition is over 95%.

## Filter and Sort the Alarms Definitions Table

The **Alarms Definitions Table** can be filtered and sorted in the following ways:

Method	Location	Description and Use
Column filter icon	Column heading	<p>Create a filter for specific columns in the <b>Alarms Definitions Table</b>. Hover over a column heading to display the small <b>Filter</b> icon, and then click the icon.</p> <ul style="list-style-type: none"> <li>Enter some text to search for alarm definition names and descriptions. Press <b>Enter</b> to apply the filter.</li> <li>To filter the table by whether notifications are on or off, click the <b>Filter</b> icon and then select <b>On</b> or <b>Off</b> to apply the filter.</li> </ul> <p>The number of applied filters will display next to the <b>Filter</b> icon. Select the small <b>X</b> icon to remove the applied filters.</p>
Column sort	Column heading	Click on a column heading to sort the table by the values in that column. Click a second time to reverse the direction of the sort.
Secondary Column Sort	Column heading	A secondary column can be added to the primary column sort. Click and hold the <b>Alt</b> key while clicking on the second column. To sort the second column in the reverse direction, repeat the procedure.

## 7.3 Site Local Alarms Fields

Site local alarm definitions can be viewed by selecting **Alarms Definitions** from the main menu. This opens the **Alarms Definitions Table**. To see alarm definition details, hover over the related table row and then select **Show Alarm Definition**, which will open the **Alarm Definition** window. The procedure to create site local alarms is described in [Create New Alarm Definition](#) on page 39.

The following tables list fields and commands that are used when defining site local alarms.

Name	Description
actions_enabled	A boolean that indicates whether or not the alarm definition is active. Disabled alarm definitions will not cause email notifications to be issued. Disable an alarm definition for testing or an excessive number of notifications. This can be set to either True or False.
alarm_actions	Indicates the notifications to be assigned to the alarm in case the alarm goes into an Alarm state. If an alarm has an <code>alarm_action</code> field and it transitions into the Alarm state, it will generate an email notification when the alarm transitions into an Alarm state. Alarms created to track the presence or absence of metric data may never transition to an Alarm state and may not have an <code>alarm_actions</code> field. Currently the only supported <code>alarm_actions</code> value is "default".
cause	Describes in non-technical terms the cause or source of the problem. The cause field is attached to email notifications.
description	Describes the alarm. It is displayed in the GUI when viewing an individual alarm.
expression	Describes how the underlying software changes the state of the alarm using both a threshold value and a period value.
ignoreund	A Boolean value that is either True or False and indicates whether the Undetermined state should be ignored for the initial transition of an alarm. Most alarms initially go into the Undetermined state when data becomes present and then transitions into the OK state and generates an email notification. When this field is set True, the first transition into the OK state from an Undetermined state will not cause an email notification to be generated. This field is used to reduce the amount of emails generated when the View for ClusterStor™ service is started or when the software is updated.
match_by	Describes a list of metric dimensions that make an alarm unique across multiple ClusterStor systems and/or devices.
name	Defines a unique identifier for the alarm definition. The name field must be unique across all alarm definitions.
ok_actions	Indicates a list of notifications to be assigned to the alarm in case the alarm goes into the OK state. When an alarm has gone into an Alarm state and then the problem gets resolved, the email notification that the alarm has gone back into an OK state will be crucial for data center personnel to note that the originating problem has been resolved. Currently the only supported <code>ok_action</code> is "default".
severity	Describes the severity of the alarm. The severity level is included in the subject line of the email notification.

Name	Description
solution	Describes in non-technical terms the recommended actions that should be taken by operations or administrative staff when an alarm has transitioned to the Alarm state. The solution field is attached to email notifications with the line that begins with "Recommended action," to ensure alarms can be acted upon.
undalarm	A Boolean value that is either True or False and it indicates whether an Undetermined alarm state is to generate a notification. When monitoring the metric subsystem, the lack of incoming data is a concern.
undcause	Like the cause field, but undcause is inserted into an email notification when an alarm goes into Undetermined state. The email notification will describe what probably happened to cause the metric data to be missing.
undetermined_actions	Indicates a list of notifications to be assigned to the alarm in case the alarm goes into an Undetermined state. The Undetermined state is an indicator that there is a problem receiving the metric data from the source and that no determination of an OK or Alarm state can be made. Most alarm definitions will not contain an undetermined_actions field in the alarm definition.
undsolution	The undsolution field defines text that gets inserted into the email notification after a "Recommended action:" in the email text when an alarm is in an Undetermined state. It describes what the administrative or operations staff should do in response to the alarm being in an Undetermined state.

Following is an example of a new alarm named `Sim_SMA_metric_health`. Site local alarms fields are shown in bold.

```
- name: "Sim_SMA_metric_health"
  description: "Check for Incoming simulated metrics"
  expression: "last(cray_test.metric_health{device=MDT0000}) < 0"
  match_by:
    - "system_name"
  severity: "CRITICAL"
  undetermined_actions: *notification_list
  ok_actions: *notification_list
  cause: "n/a"
  solution: "n/a"
  undcause: "There is an interruption in the flow of metric data from the
ClusterStor storage system."
  undsolution: "The management console of the ClusterStor storage cluster should
be checked to see if it is operational and if there are alerts on its service
console."
  ignoredund: False
  undalarm: True
```

## Command Fields

Command fields extend the possibilities of alarm maintenance by allowing users to delete or modify Cray® defined alarms. If an alarm does not have a command field, it reverts to "default," which is a `create` for its first run, and then `update` on any runs thereafter. Command fields only apply to a single alarm.

The supported values for these fields are:

- "yes"

- yes
- True

Both yes and True are boolean values. "yes" is a string.

Name	Description
adjust_requested	<p>This command only performs a patch on an alarm definition. The only supported safe operation of the <code>patch</code> command is to disable or re-enable an alarm definition. A disabled alarm definition does not cause emails to be generated, but it isn't deleted and can easily be re-enabled. Here is the simplest form of disabling an alarm:</p> <pre data-bbox="500 604 899 688">- name: "Sample_Alarm"   adjust_requested: "yes"   actions_enabled: False</pre> <p>Here is the simplest form of enabling an alarm:</p> <pre data-bbox="500 779 899 863">- name: "Sample_Alarm"   adjust_requested: "yes"   actions_enabled: True</pre>
delete_requested	<p>This deletes an alarm definition. This functionality was introduced in release 1.0.1. There are two use cases: the first is to just delete an alarm definition that a site does not want. The second is to delete a Cray defined alarm and replace it with a site replacement alarm. Here is what an alarm deletion looks like:</p> <pre data-bbox="500 1066 899 1119">- name: "Sample_Alarm"   delete_requested: "yes"</pre>
override_requested	<p>Indicates that this is an override definition. Only the <code>expression</code>, <code>severity</code>, and <code>description</code> fields can be overridden. Here is what an override looks like in the <code>site-local-alarms</code> file:</p> <pre data-bbox="500 1297 1279 1434">- name: "Sample_Alarm_2"   override_requested: "yes"   expression: "avg(cray_test.metric, 500) &lt; 10"   description: "Metric monitored every 5 minutes"   severity: "MEDIUM"</pre>
refresh_requested	<p>This command refreshes the alarm definition without changing the expression or any other field of the alarm definition. In release 1.0.0, an override was implemented with the <code>alarm-definition-patch</code> command, which would sometimes leave the alarm definition in a broken state. The <code>refresh_requested</code> command will fix alarm definitions in that state. Because overrides are implemented differently in 1.0.1, it is not anticipated that repairing alarm definitions will be required. Here is what an alarm refresh looks like:</p> <pre data-bbox="500 1745 915 1797">- name: "Sample_Alarm"   refresh_requested: "yes"</pre>

The following example uses two site local alarm fields and one command field to disable an existing alarm named `Sample_Loss_of_Heartbeat`.

```
- name: "Sample_Loss_of_Heartbeat"  
  adjust_requested: "yes"  
  actions_enabled: False
```

## 7.4 Override Cray Defined Alarms

### About this task

Cray® defined alarms can be overridden by providing new expressions, severities, and/or descriptions. The expression field's threshold value and period can both be changed. Please see [Site Local Alarms Fields](#) on page 35 for additional details about the expression, description, and severity fields.

For descriptions and examples of the types of changes that can be made to alarms, please refer to the following files, which can be found in the `/etc/sma-data/etc` directory on the View for ClusterStor™ server.

- `site-override.alarms.yaml.example`
- `site-local-alarms.yml.j2.example`

The following procedure explains how to override Cray defined alarm definitions.

### Procedure

1. Log in to View for ClusterStor as root.
2. Copy the `site-override.alarms.yaml.example` file to a new file name in the same directory.

```
hostname# cp /etc/sma-data/etc/site-override-alarms.yaml.example \  
/etc/sma-data/etc/site-override-alarms.yaml
```

The file includes override examples for:

- `SMA_OST_Free_Space`: Change the threshold value from 5.0 to 10.0, plus change the severity and description.
- `SMA_OST_Free_Files`: Change the threshold value from 5.0 to 10.0, change the period from 15 minutes to 20 minutes, plus change the severity and description.
- `Root_Partition_Disk_Space_Usage`: Change the threshold value from 95 to 90 and also change the description.
- `Root_Partition_inode_Usage`: Change the threshold value from 95 to 98, change the period from 2 minutes to 5 minutes, and change the description.
- `SSD_Partition_Disk_Space_Usage`: Change the threshold value from 95 to 97.
- `SSD_Partition_inode_Usage`: Change the threshold value from 95 to 96.

3. Edit the new file as needed:
  - a. Open the file in an editor.

```
hostname# vi /etc/sma-data/etc/site-override-alarms.yaml
```

- b. Remove any of the override examples not needed or comment them out.
- c. Change thresholds, period, expressions, severities, and descriptions as needed on any override examples being kept
- d. Save the modified file.

Following is an example override file where only the `SMA_OST_Free_Space` alarm threshold, severity, and description have been changed:

```
# Site-override alarms definition file
#
alarm_definitions:
  # Change SMA_OST_Free_Space: threshold value from 5.0 to 10.0
  # also, change severity and description
  - name: "SMA_OST_Free_Space"
    expression: "avg(cray_storage.free_space_perc, 900) < 10.0"
    severity: "MEDIUM"
    description: "The free disk space of a specific OST/MDT is under 10%."
```

#### 4. Apply the overrides.

```
hostname# cd /etc/sma-data/etc
hostname# docker-compose run --rm alarms remove_all_alarms.sh
hostname# docker-compose run --rm alarms /start.sh
```

This step executes two Docker Compose commands:

- The first removes all currently defined alarms.
- The second sets up all Cray defined alarms and all of the site local overrides.



**CAUTION:** All alarm history will be lost when removing alarms.

New alarm definitions can be viewed via the **Alarms Definitions Table**, which is described in [Cray Defined Alarms](#) on page 32. New alarms may take a while to appear.

## 7.5 Create New Alarm Definition

### About this task

Administrators can create new alarm definitions. For descriptions and examples of the types of changes that can be made to alarms, please refer to the following files, which can be found in the `/etc/sma-data/etc` directory on the View for ClusterStor™ server.

- `site-override.alarms.yaml.example`
- `site-local-alarms.yml.j2.example`

Creating site-specific local alarms involves copying the `site-local-alarms.yml.j2.example` file to a new file, and then making edits as desired. Fields used to define these new alarms are described in detail in [Site Local Alarms Fields](#) on page 35.

The following procedure explains how to create new alarms.

## Procedure

1. Log in to View for ClusterStor as root.
2. Copy the `site-local-alarms.yml.j2.example` file to a new file name in the same directory.

```
hostname# cp /etc/sma-data/etc/site-local-alarms.yml.j2.example \
/etc/sma-data/etc/site-local-alarms.yml.j2
```

3. Edit the new file to define the desired new alarm.

- a. Open the file in an editor.

```
hostname# vi /etc/sma-data/etc/site-local-alarms.yml.j2
```

- b. Remove any of the examples that are not needed or comment them out.
- c. Modify the selected example to define the new alarm. Use the fields described in [Site Local Alarms Fields](#) on page 35.
- d. Save the modified file.

Following is an example of a new alarm named `Sim_SMA_metric_health`. Site local alarms fields are shown in bold:

```
- name: "Sim_SMA_metric_health"
  description: "Check for Incoming simulated metrics"
  expression: "last(cray_test.metric_health{device=MDT0000}) < 0"
  match_by:
    - "system_name"
  severity: "CRITICAL"
  undetermined_actions: *notification_list
  ok_actions: *notification_list
  cause: "n/a"
  solution: "n/a"
  undcause: "There is an interruption in the flow of metric data from the
ClusterStor storage system."
  undsolution: "The management console of the ClusterStor storage cluster
should be checked to see if it is operational and if there are alerts on its
service console."
  ignoredund: False
  undalarm: True
```

4. Create the new alarm(s).

```
hostname# cd /etc/sma-data/etc
hostname# docker-compose run --rm alarms remove_all_alarms.sh
hostname# docker-compose run --rm alarms /start.sh
```

This step executes two Docker Compose commands:

- The first removes all currently defined alarms.
- The second sets up all Cray defined alarms and all of the site local overrides.



**CAUTION:** All alarm history will be lost when removing alarms.

New alarm definitions can be viewed in the **Alarms Definitions Table**, which is described in [Cray Defined Alarms](#) on page 32. New alarms may take a while to appear.

## 7.6 Configure Notification Email

### Prerequisites

View for ClusterStor™ has been installed and configured.

### About this task

View for ClusterStor creates alarms that record the state of the system as indicated by the values in specified metric series. A notification email is a message sent automatically when the state of an alarm changes. The following procedure configures the email address the system uses when sending alarm notifications.

### Procedure

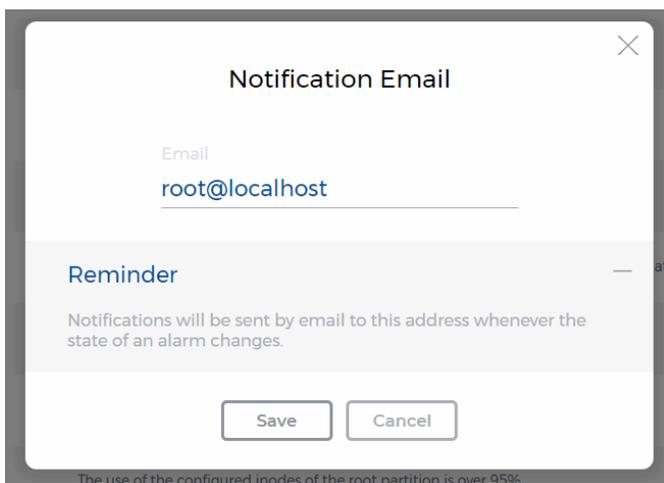
1. Select **Alarms Definitions** from the main menu of the View for ClusterStor graphical user interface (GUI). The **Alarms Definitions Table** will open.
2. Select the **Pencil** icon located at the right side of the content pane banner.



Name	Description	Notifications
SSD_Partition_inode_Usage	The use of the configured inodes of the SSD partition is over 95%.	On

The **Notification Email** window will open.

3. Enter the email address of the intended recipient of alarm notifications.



**Notification Email**

Email

**Reminder**

Notifications will be sent by email to this address whenever the state of an alarm changes.

4. Click **Save**.

## 8 Jobs Table

The **Jobs Table** provides a broad view of the behavior of the jobs that are currently running. The metrics in this table indicate if a job might be behaving in a way that can adversely affect the performance of the ClusterStor™ system. Select the **Jobs** information block on any ClusterStor tile in the **System Overview** to open the **Jobs Table**.

Figure 6. Jobs Table

Host Name	Job ID	Apid	User ID	Application	Start Time	End Time	Duration	Avg. I/O Size	Metadata Ops
pollux-smw	2522680	1356	growfiles_mpi	2018-10-17 09:39:47	2018-10-17 09:43:57	4m 10s	2.57MB	2.02m	
pollux-smw	2522679	1356	growfiles_mpi	2018-10-17 09:39:45	2018-10-17 09:43:48	4m 3s	3.72MB	19.4m	
pollux-smw	2522730	1356	IOR	2018-10-17 10:37:14	2018-10-17 10:40:12	2m 58s	4.19MB	1.25k	
pollux-smw	2522644	16912	serbalt	2018-10-17 08:58:19	2018-10-17 08:58:47	28s	4.19MB	16	
pollux-smw	2522647	1356	IOR	2018-10-17 08:59:03	2018-10-17 09:19:13	20m 10s	1.68MB	4.07k	
pollux-smw	2522649	22569	nemod	2018-10-17 09:00:03	2018-10-17 09:20:13	20m 10s	2.1MB	646k	
pollux-smw	2522650	1356	IOR	2018-10-17 09:01:48	2018-10-17 09:03:15	1m 27s	1.41MB	1.36k	
pollux-smw	2522654	16912	serbalt	2018-10-17 09:07:04	2018-10-17 09:07:36	32s	4.18MB	16	
pollux-smw	2522655	1356	IOR	2018-10-17 09:07:04	2018-10-17 09:14:08	7m 4s	1.34MB	4.89k	
pollux-smw	2522656	7862	jrebsnorth	2018-10-17 09:07:36	2018-10-17 09:10:50	3m 14s	2.09MB	5.01k	
pollux-smw	2522658	16912	serbalt	2018-10-17 09:10:50	2018-10-17 09:11:25	35s	4.18MB	16	
pollux-smw	2522659	7862	jrebsnorth	2018-10-17 09:11:25	2018-10-17 09:14:24	2m 59s	2.09MB	5.01k	
pollux-smw	2522662	16912	serbalt	2018-10-17 09:14:24	2018-10-17 09:14:52	28s	4.19MB	16	

### Jobs Table Column Descriptions

Column Heading	Description
n/a	A job warning is indicated in the table's first column by the color of the cell: <ul style="list-style-type: none"> <li>Red: Warning</li> <li>Transparent: OK</li> </ul>
Host Name	The SMW or Cluster host name from which the job was launched.
Job ID	The job ID which is reported by the workload manager for job metric values.

Column Heading	Description
<b>Apid</b>	The application ID which is reported by the workload manager for job metric values
<b>User ID</b>	The ID of the user who owns the job, as it was reported by the workload manager when the job was started
<b>Application</b>	The name of the job application, as it was reported by the workload manager when the job was started
<b>Start Time</b>	The time that the application started, as reported by the workload manager
<b>End Time</b>	The time that the job ended, as reported by the workload manager
<b>Duration</b>	The calculated amount of time the job has run
<b>Avg. I/O Size</b>	<p>The Average I/O Size is calculated as the total number of bytes read and written since the given job's inception, divided by the total number of read and write requests since the job's inception.</p> <p>The displayed value will appear in red text when it is &lt; 64000 bytes per read/write. Small I/O sizes cause the file system to perform poorly, indicating poor job performance.</p>
<b>Metadata Ops</b>	<p>Metadata Ops is calculated as the sum of all the metadata operations since the given job's inception, from each MDT.</p> <p>The displayed value will appear in red text when it is &gt; 1,000,000 total operations. High metadata operations cause the file system to perform poorly, indicating poor job performance.</p>

## Filter and Sort the Jobs Table

The **Jobs Table** can be filtered and sorted in the following ways:

Method	Location	Description and Use
Filter icon	Content pane banner	<p>Select the <b>Filter</b> icon to open the <b>Filter Jobs</b> window. Specify the desired options for filtering information in the table:</p> <ul style="list-style-type: none"> <li>• Enter one or more partial or full values for host name, job ID, user ID, apid, or app name. Multiple comma-separated values may be entered.</li> <li>• Adjust the sliders to set minimum and maximum values for job duration, average I/O size, or metadata operations. Alternatively, double-click the values above the slider to enter minimum and maximum values manually. Press <b>Enter</b> after typing a value.</li> <li>• Select one of the provided time ranges or select <b>Set custom time range</b> to set the start and end date and time. If a value is not set, the time range will default to the last 15 minutes.</li> </ul> <p>Click <b>Submit</b> to apply the filtering options. The time range and the number of other applied filters will display next to the <b>Filter</b> icon. Select the small <b>X</b> icon to remove the applied filters.</p>

Method	Location	Description and Use
Column filter icon	Column heading	<p>Create a filter for specific columns in the <b>Jobs Table</b>. Hover over a column heading to display the small <b>Filter</b> icon, and then click the icon.</p> <ul style="list-style-type: none"> <li>Enter some text to search for host name, job ID, user ID, apid, or application name. Press <b>Enter</b> to apply the filter. Multiple comma-separated values may be entered.</li> <li>Adjust the sliders to set minimum and maximum values for job duration, average I/O size, or metadata operations. Alternatively, double-click the values above the slider to enter minimum and maximum values manually. Press <b>Enter</b> after typing a value. Click the check mark above a slider to apply the selection, or click the <b>X</b> to cancel.</li> </ul> <p>The number of applied filters will display next to the <b>Filter</b> icon. Select the small <b>X</b> icon to remove the applied filters.</p>
Column sort	Column heading	Click on a column heading to sort the table by the values in that column. Click a second time to reverse the direction of the sort.
Secondary Column Sort	Column heading	A secondary column can be added to the primary column sort. Click and hold the <b>Alt</b> key while clicking on the second column. To sort the second column in the reverse direction, repeat the procedure.

Figure 7. Filter Jobs Window

The screenshot shows a window titled "Filter Jobs" with a close button (X) in the top right corner. The window contains several filter sections:

- Filter by host name:** A text input field with a blue underline.
- Filter by job ID:** A text input field.
- Filter by user ID:** A text input field.
- Filter by apid:** A text input field.
- Filter by app name:** A text input field.
- Filter by job duration:** A slider ranging from 0s to 14d.
- Filter by avg. I/O size:** A slider ranging from 0B to 10GB.
- Filter by metadata ops:** A slider ranging from 0 to 10m.
- Limit search to time range:** A section with buttons for 5m, Last 15m (selected), 30m, 1h, and 3h. Below these buttons is a text input field labeled "Set custom time range".

At the bottom of the window are two buttons: "Submit" and "Cancel".

## Jobs Table Actions

When hovering over a row in the **Jobs Table**, some additional actions will be displayed at the right side of the row.

Job ID	Apid	User ID	Application	Start Time	End Time	Avg. I/O Size	Metadata Ops	2	Sort	Info	Minus
2526184		16912	serbalt	2018-10-19 13:26:17	2018-10-19 13:26:47	4.18MB	16				<input type="checkbox"/>
2526185		7862	jrebsnorth	2018-10-19 13:26:47	2018-10-19 13:29:36	2.09MB	5.01k			Info	<input checked="" type="checkbox"/>
2526186		16912	serbalt	2018-10-19 13:29:37	2018-10-19 13:30:07	4.19MB	16				<input type="checkbox"/>

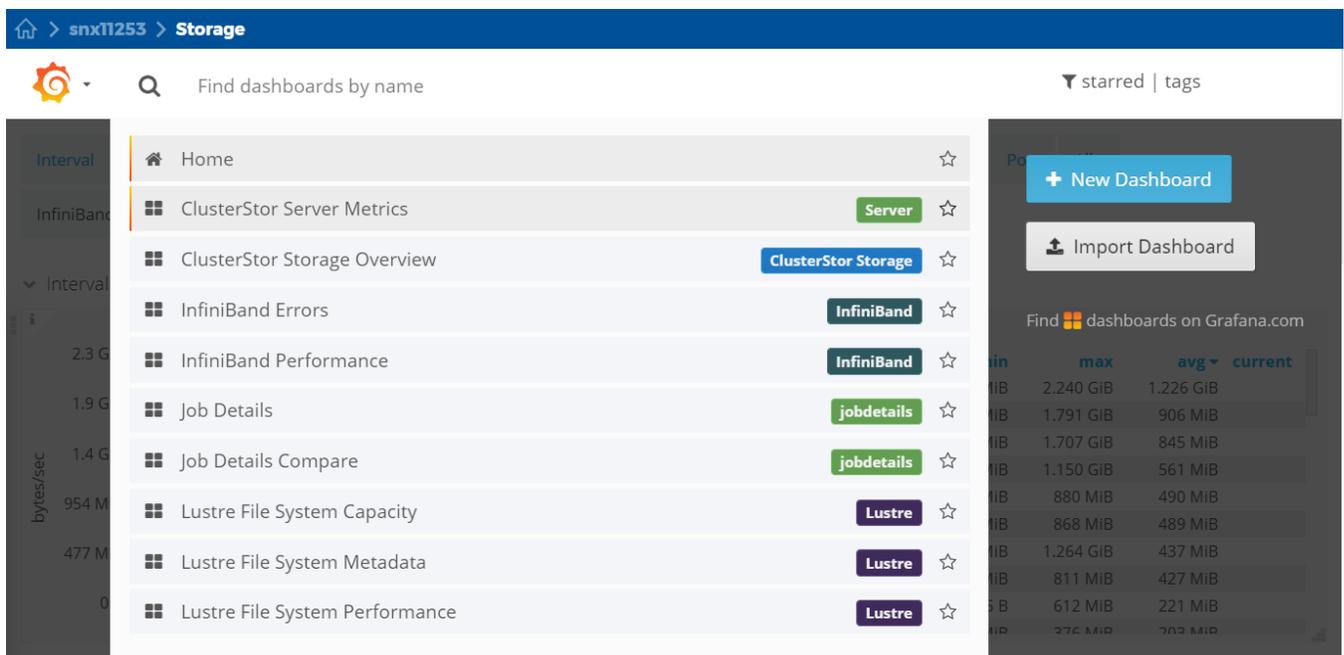
- **Details** icon - Opens the **Job Details Dashboard**, which displays the details of the individual job.
- **Select** checkbox - Use to select multiple jobs rows, up to a maximum of six. Once rows are selected, click one of the following options shown at the right side of the **Jobs Table** column headings row:
  - **Details** icon - Opens the **Job Details Dashboard**, which displays the combined details of the selected jobs.
  - **Compare** icon - When one or two jobs are selected, opens the **Job Details Compare Dashboard**, to view both jobs synced by start times.
  - **Deselect** icon - Clears all of the selected checkboxes in the table.

## 9 Create and Modify Dashboards

### About this task

New View for ClusterStor™ dashboards can be created in Grafana at any time. The basic procedure is as follows:

1. Click the Grafana **Dashboard Selector**.
2. Select **+ New Dashboard**:



3. Choose panels to add to the new dashboard and configure the panels.
4. Save the new dashboard.

For in-depth details about creating and customizing dashboards, please refer to the [Grafana documentation](#) directly.

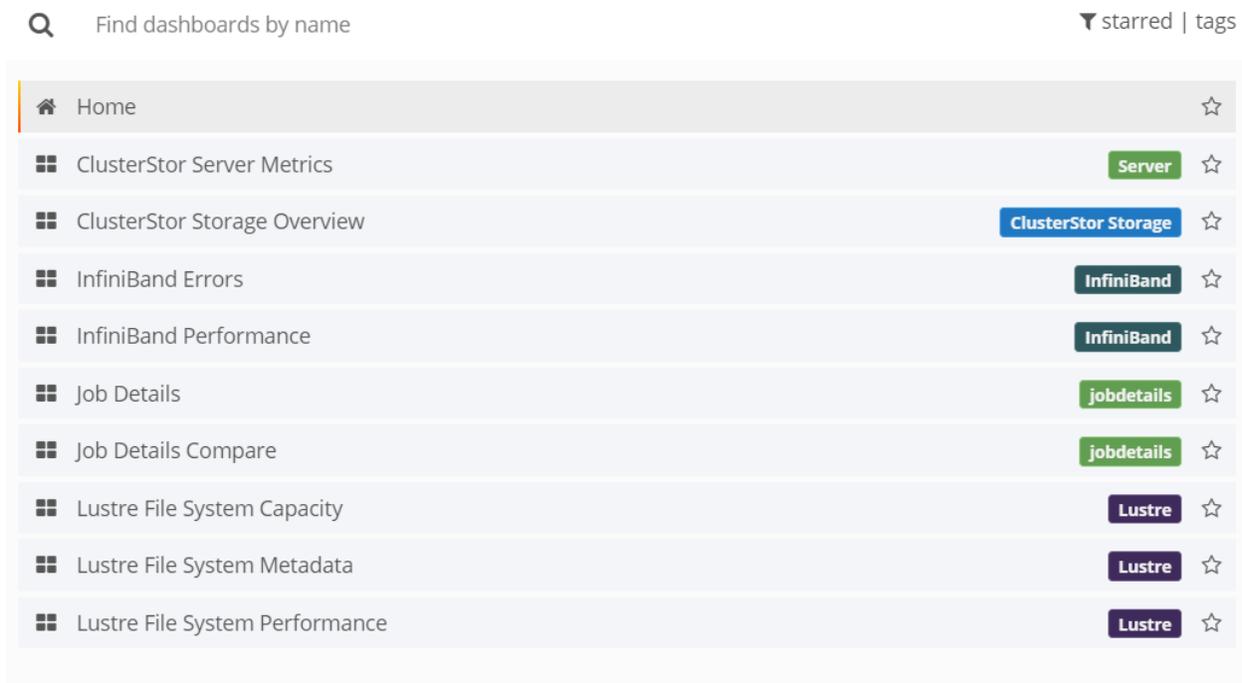
### Modify Predefined Dashboards

View for ClusterStor comes with a number of predefined dashboards that can be modified in Grafana. For more information about the predefined dashboards, see [Predefined Dashboards](#) on page 58.

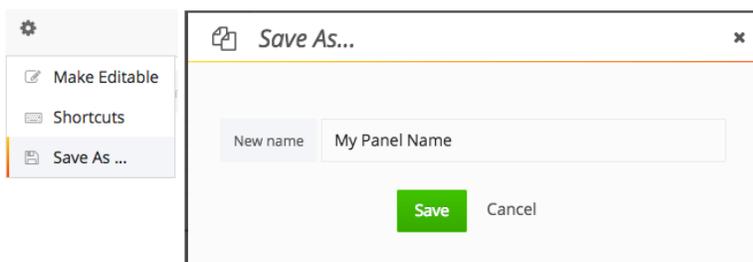
To save modifications that are made to a predefined dashboard, the dashboard must be saved using a new name. Otherwise, updating or restarting View for ClusterStor will overwrite all modifications.

## Procedure

1. Click the Grafana **Dashboard Selector**.
2. Select the predefined dashboard that is to be modified.



3. Select Grafana **Settings** (gear icon), and then select **Save As** from the drop down menu. Rename the dashboard.



The screen will refresh and the dashboard is ready for any modifications.

4. Make any desired changes to the new dashboard.
5. Save the modified dashboard when the changes are complete.

The modified dashboard will now persist through updates and any restarts.

## 10 Dashboard Metrics

The View for ClusterStor™ server collects and stores performance metrics for ClusterStor systems and the InfiniBand network. These metrics make up the performance data that is displayed. The data model for InfluxDB is based on dimensions. This section describes the possible dimensions and the metrics being persisted.

### Dimensions

The following table defines the dimensions described in the sections that follow.

Name	Description	Example values
apid	Application ID	apid, job_id, procid
apname	Application name	The name of the application
component	The component name within the service from which the metric comes	For ClusterStor metrics, this would be <code>Lustre</code>
dest	The opposite endpoint of the IB connection. For example, the <i>dest</i> for a ClusterStor port will always be a Local Data Network (LDN) switch port.	caribou07_mlx4_0, caribou10_mlx4_0
device	The device name	OST000 [0-N], MDT000 [0-N]
device_type	The device type	mds, mdt, oss, ost
disc_port	Local View IB port	metrics was discovered on   mlx4_0, mlx4_1
guid	Global Unique Identifier	64-bit definitions for all the elements within a subnet: chassis, HCAs, switches, routers, and ports
hostname	The FQDN of the host being measured; describes the relevant origin node name	cls12345n002, cls12345n004, etc
job_id	The Lustre job identification number	apid, job_id, procid
lid	Local Identifier	16-bit address assigned by the Subnet Manager
node	Processor node	cls12345_Rack_1_Odd_TOR 1, cls12345_Rack_1_Odd_TOR 2, etc
port	Interface	Integer value

Name	Description	Example values
product	The product name	For View for ClusterStor, this would be ClusterStor
service	The service name that owns the metric	For View for ClusterStor, this would be storage
system_name	The name (serial number) of the storage cluster	ClusterStor Serial Number
type	Component type	switch or ca
userid	Unique identifier of user	UID of user

## 10.1 Lustre Metrics

Lustre® Metrics are those metrics retrieved from the ClusterStor™ system CStream API. These metrics consist of Lustre-specific and system node statistics, and are stored in the `cray_storage` series heading.

### Aggregated Metrics

Metric	Description	Sample Rate
calculated_metadata_ops	Total system aggregated metadata operations across all MDTs	5s
calculated_read_bytes	Total read byte rates aggregated across all OSTs	5s
calculated_write_bytes	Total write byte rates aggregated across all OSTs	5s

The above metrics also have dimensions supported for custom queries:

- component
- device\_type
- product
- service
- system\_name

### OST Metrics

Metric	Description	Sample Rate
bytesavail	The free space that is not reserved by client's write cache or file system reserve	60s
bytesfree	The total amount of unused space in the file system	60s
bytestotal	The total amount of storage in the file system	60s

Metric	Description	Sample Rate
filesfree	The number of unallocated files (free inodes)	60s
filestotal	The number of possible files in a system (total inodes)	60s
free_files_perc	A derived metric calculating the percentage of free inodes still available.	60s
free_space_perc	A derived metric calculating the percentage of free space still available.	60s
read_bytes_rate	Average number of bytes read per second during last interval	5s
write_bytes_rate	Average number of bytes written per second during last interval	5s

The above metrics also have dimensions supported for custom queries:

- component
- device
- device\_type
- hostname
- product
- service
- system\_name

## MDT Metrics

Metric	Description	Sample rate
close_rate	The file close requests per second	5s
connect_rate	The client connection requests per second	5s
create_rate	The pool create requests per second	5s
destroy_rate	The pool destroy requests per second	5s
disconnect_rate	The client disconnect requests per second	5s
free_files_perc	A derived metric calculating the percentage of free inodes still available.	60s
free_space_perc	A derived metric calculating the percentage of free space still available.	60s
getattr_rate	The getattr requests per second	5s
getxattr_rate	The getxattr requests per second	5s
link_rate	The link requests per second	5s
llog_init_rate	The llog_init requests per second	5s

Metric	Description	Sample rate
mkdir_rate	The mkdir requests per second	5s
mknod_rate	The mknod requests per second	5s
notify_rate	The notify requests per second	5s
open_rate	The file open requests per second	5s
process_config_rate	The process config requests per second	5s
quotactl_rate	The quotactl requests per second	5s
reconnect_rate	The client reconnect requests per second	5s
rename_rate	The rename requests per second	5s
rmdir_rate	The rmdir requests per second	5s
setattr_rate	The setattr requests per second	5s
statfs_rate	The statfs requests per second	5s
unlink_rate	The unlink requests per second	5s

The above metrics also have dimensions supported for custom queries:

- component
- device
- device\_type
- hostname
- product
- service
- system\_name

## OSS Linux Metrics

Metric	Description	Sample Rate
cpu_idle_perc	The percentage of time the CPU is idle	15s
cpu_iowait_perc	The amount of time a thread is blocked waiting for I/O to complete	15s
cpu_utilization_perc	The percentage of time the CPU spends processing the instructions of a program or O/S	15s
memory_utilization_perc	The percentage of memory used by processes and the O/S	15s

The above metrics also have dimensions supported for custom queries:

- component

- device
- device\_type
- hostname
- product
- service
- system\_name

## 10.2 Health Metrics

Health metrics are retrieved from polling of the ClusterStor™ system's SNMP walk interface.

Metric	Description	Sample Rate
snx_health	Heartbeat of the ClusterStor health measured by periodically polling the SNMP walk interface	5m

The above metric also has dimensions supported for custom queries:

- component
- hostname
- product
- service
- system\_name

## 10.3 InfiniBand Metrics

InfiniBand Metrics are retrieved from the InfiniBand network and are stored in the `cray_ib` series heading. These metrics are sampled at 30 second or 60 second intervals but only persist a metric value if the delta from the previous value is greater than zero.

Metric	Definition	Default Sample Rate
excessive_buffer_overrun_errors_sec	The number of times that OverrunErrors consecutive flow control update periods occurred per second, each having at least one overrun error.	60s
link_downed_sec	The total number of times per second the Port Training state machine has failed the link error recovery process and downed the link.	60s

Metric	Definition	Default Sample Rate
link_error_recoveries_sec	The total number of times per second the Port Training state machine has successfully completed the link error recovery process.	60s
local_link_integrity_errors_sec	The number of times that the count of local physical errors per second exceeded the threshold specified by LocalPhyErrors.	60s
port_rcv_bytes_sec	The total number of data bytes received per second.	30s
port_rcv_constraint_errors_sec	The total number of packets received on the switch physical port that were discarded per second.	60s
port_rcv_errors_sec	The total number of packets containing an error that were received on the port per second.	60s
port_rcv_pkts_sec	The total number of packets received per second.	30s
port_rcv_remote_phys_errors_sec	The total number of packets marked with the EBP (End of Bad Packet) delimiter received on the port per second. This is typically due to a physical error that was detected and marked by an upstream port.	60s
port_rcv_switch_relay_errors_sec	The total number of packets received on the port per second that were discarded because they could not be forwarded by the switch relay.	60s
port_total_errors_sec	The aggregate of all errors on the port per second (excluding SymbolErrors).	30s
port_xmit_bytes_sec	The total number of data bytes transmitted per second.	30s
port_xmit_constraint_errors_sec	The total number of packets not transmitted from the switch physical port per second.	60s
port_xmit_discards_sec	The total number of packets dropped because the port is down or congested per second.	60s
port_xmit_pkts_sec	The total number of packets transmitted per second.	30s
port_xmit_wait_sec	The number of packets per second that had to wait before being transmitted per second.	30s
symbol_errors_sec	The total number of minor link errors detected per second on one or more physical lanes. This includes 8B/10B coding	60s

Metric	Definition	Default Sample Rate
	violations and is typically an indication of a bit error on the line.	
v115_dropped_sec	The number of incoming VL15 packets dropped per second due to resource limitations (for example, lack of buffers) in the port.	60s

The above metrics also have dimensions supported for custom queries:

- dest
- guid
- lid
- node
- port
- type
- disc\_port

Metric	Description	Default Sample rate
topology	IB topology compare per second across all fabrics. <ul style="list-style-type: none"> <li>• 0 = no difference</li> <li>• 1 = difference</li> </ul>	900s

The above metric also has a dimension supported for custom queries:

- type

## 10.4 Jobstats Metrics

Jobstats metrics display the breakout of Lustre® statistics on a per-job basis. These metrics can determine which jobs are contributing to degraded system performance, and track a job's performance from run to run. They are stored in the `cray_job` series heading.

Certain metrics are reported from each of the nodes in a storage cluster. Since each node only reports when there is updated data since its last report, only summary calculations can be made when a node reports. Summary data will consequently appear on an irregular basis. Per report metrics are calculated every time nodes send data. Summary metrics are calculated every other time a node sends data. Scoring metrics are representative of metrics across the lifetime of a given job.

Metric	Description	Sample Rate
calculated_metadata_ops	The total number of metadata ops across all MDTs	Summary calculation

Metric	Description	Sample Rate
<code>calculated_metadata_ops_sec</code>	The total number of metadata ops across all MDTs, divided by the best estimate of elapsed time in seconds during which the operations occurred.	Summary calculation
<code>calculated_metadata_ops_sum</code>	The total number of metadata ops across all MDTs since the beginning of a job.	Summary calculation
<code>calculated_read_bytes_req</code>	The sum of all bytes read during the last cycle of reporting across all OSTs, divided by the total number of requests across all OSTs.	Summary calculation
<code>calculated_read_bytes_sec</code>	The sum of all bytes read during the last cycle of reporting across all OSTs, divided by the best estimate of elapsed time in seconds during which the bytes were read or written.	Summary calculation
<code>calculated_write_bytes_req</code>	The sum of all bytes written during the last cycle of reporting across all OSTs, divided by the total number of requests across all OSTs.	Summary calculation
<code>calculated_write_bytes_sec</code>	The sum of all bytes written during the last cycle of reporting across all OSTs, divided by the best estimate of elapsed time in seconds during which the bytes were read or written.	Summary calculation

The above metrics have dimensions supported for custom queries:

- `component`
- `device_type`
- `job_id`
- `product`
- `service`
- `system_name`

Metric	Description	Sample Rate
<code>d_close</code>	The number of <code>close</code> requests that have occurred since last calculated.	30s
<code>d_getattr</code>	The number of <code>getattr</code> requests that have occurred since last calculated.	30s
<code>d_getxattr</code>	The number of <code>getxattr</code> requests that have occurred since last calculated.	30s
<code>d_mkdir</code>	The number of <code>mkdir</code> requests that have occurred since last calculated.	30s
<code>d_open</code>	The number of <code>open</code> requests that have occurred since last calculated.	30s
<code>d_quotactl</code>	The number of <code>quotactl</code> requests that have occurred since last calculated.	30s

Metric	Description	Sample Rate
d_rename	The number of <code>rename</code> requests that have occurred since last calculated.	30s
d_rmdir	The number of <code>rmdir</code> requests that have occurred since last calculated.	30s
d_setattr	The number of <code>setattr</code> requests that have occurred since last calculated.	30s
d_unlink	The number of <code>unlink</code> requests that have occurred since last calculated.	30s
io_size	The total of all the read and write bytes since the beginning of a job across all OSTs, divided by the total number of read/write requests since the beginning of the job.	Summary calculation
job_cnt	The number of jobs currently being tracked.	Summary calculation
metadata_ops_sec	The total number of metadata ops on a given MDT, divided by the best estimate of elapsed time in seconds during which the operations occurred.	Summary calculation
metadata_ratio	The total of all read and write bytes since the beginning of a job across all OSTs, divided by the total number of metadata ops across all MDTs since the beginning of the job.	Summary calculation
read_bytes_sec	The total number of bytes read on a given OST (since the last reported value) divided by the best estimate of elapsed time in seconds during which the operations occurred.	Summary calculation
write_bytes_sec	The total number of bytes written on a given OST (since the last reported value) divided by the best estimate of elapsed time in seconds during which the operations occurred.	Summary calculation

The above metrics also have dimensions supported for custom queries:

- component
- device
- device\_type
- hostname
- job\_id
- product
- service
- system\_name

Metric	Description	Sample Rate
status	The start/stop and heartbeat from the SMW providing information on the job scheduler.	Varies

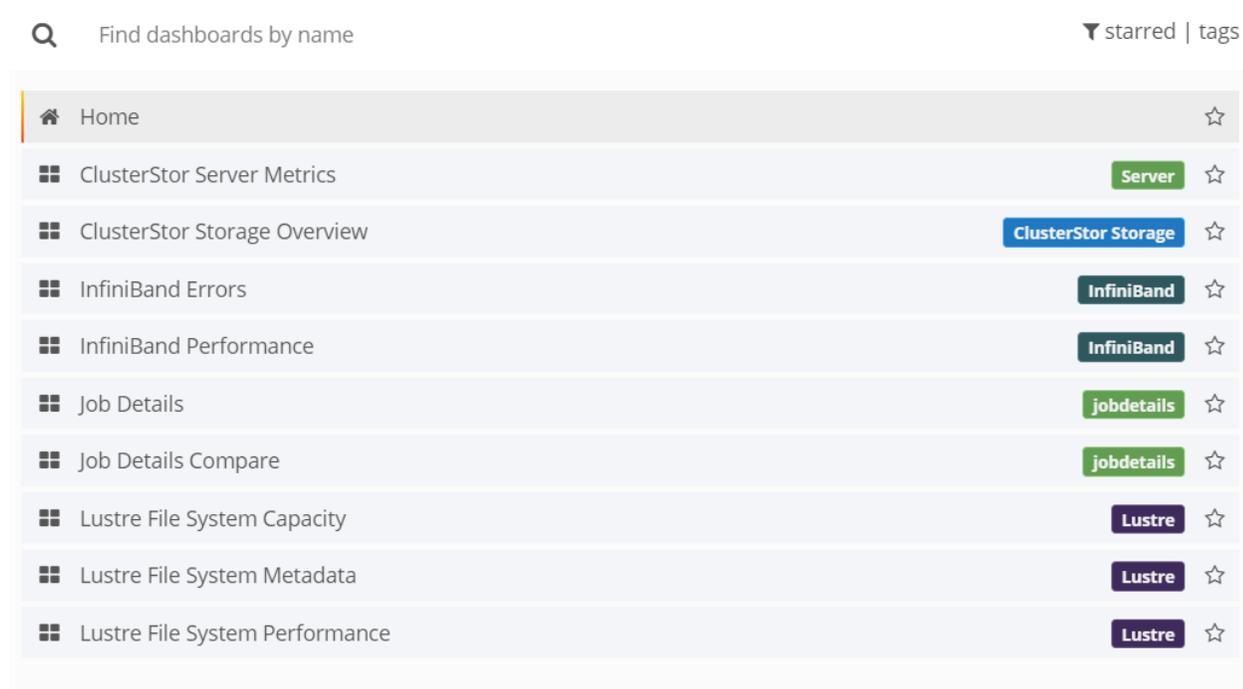
The above metric also has dimensions supported for custom queries:

- apid
- apname
- hostname
- userid

## 11 Predefined Dashboards

View for ClusterStor™ provides out-of-the-box custom Grafana dashboards that may be used to display ClusterStor and InfiniBand metrics. Some of these predefined dashboards can be accessed from the **System Overview** or the **Jobs Table**. All of the predefined dashboards are accessible from the Grafana **Dashboard Selector**.

Figure 8. Grafana Dashboard Selector



**NOTE:** When navigating directly from a View for ClusterStor tile to a Grafana dashboard, the charts and metrics shown are specific to the ClusterStor system represented in that tile. Once in Grafana, however, using the **Dashboard Selector** to navigate to a different dashboard will change the scope of the data displayed, to all ClusterStor systems being monitored in View for ClusterStor. To maintain scope to a specific ClusterStor system when navigating within Grafana, use the links that appear when selecting the **Details** icons at the upper left of any panel in a dashboard.

### General Dashboard Features

- Home Link - Use the **Home** icon in the content pane banner to navigate back to the **System Overview**.
- Templates - Use the following template variables, which appear below the Grafana header, to modify the scope of what each dashboard displays. Select template variable values from the drop-down menu to the right of the variable name. The displayed graphs update immediately to reflect metric values for the selected template settings.

Template Variable	Description	Dashboards Where Used
<b>Server</b>	Select <b>All</b> or one specific ClusterStor server to include in the charts.	ClusterStor Server Metrics, ClusterStor Storage Overview, Job Details, Job Details Compare, Lustre File System Capacity, Lustre File System Metadata, Lustre File System Performance
<b>Interval</b>	Select the interval at which the dashboard graphs are scaled, based on the selected time range. The mean value of the metrics are grouped by time based on the chosen interval setting. This interval setting is automatically calculated from the selected time frame.	ClusterStor Server Metrics, ClusterStor Storage Overview, InfiniBand Errors, InfiniBand Performance, Job Details, Job Details Compare, Lustre File System Capacity, Lustre File System Metadata, Lustre File System Performance
<b>Local IB Port</b>	Select <b>All</b> or specific local InfiniBand ports to include in charts.	InfiniBand Performance
<b>Node Type</b>	InfiniBand node type. Select <b>All</b> , <b>ca</b> , or <b>switch</b> to include in charts.	InfiniBand Errors, InfiniBand Performance
<b>Node Name</b>	Select <b>All</b> or specific node names to include in charts.	InfiniBand Errors, InfiniBand Performance
<b>LID</b>	Select <b>All</b> or specific LIDs to include in charts.	InfiniBand Errors, InfiniBand Performance
<b>Port</b>	Select <b>All</b> or specific port numbers to include in charts.	InfiniBand Errors, InfiniBand Performance
<b>Job</b>	Enter a specific job id or application id (apid) to include in charts.	Job Details, Job Details Compare
<b>MDT</b>	Select <b>All</b> or specific MDTs as the focus in the charts	Job Details, Job Details Compare, Lustre File System Capacity, Lustre File System Metadata
<b>OST</b>	Select <b>All</b> or specific OSTs as the focus in the charts.	Job Details, Job Details Compare, Lustre File System Capacity, Lustre File System Performance
<b>offset</b>	A time offset applied to the charts. Determined automatically when the dashboard is opened from the <b>Jobs Table</b> . Offset value can be changed.	Job Details Compare

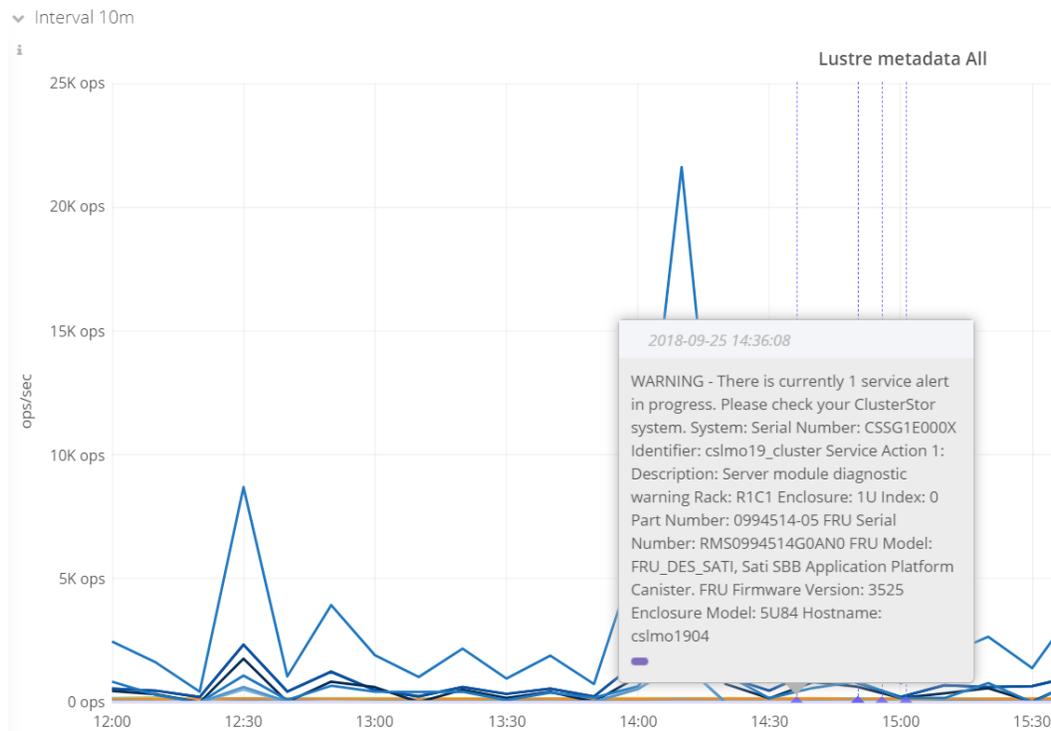
- Annotations - Vertical lines displayed on dashboard charts highlight where major events have occurred. Toggle these annotations on/off by checking/unchecking the options that appear under the Grafana header:



- Lustre** - Critical Lustre errors noted in the ClusterStor logs.

- **Health** - SNMP hardware events noted from a polling of SNMP.
- **InfiniBand** - Changes in InfiniBand topology.

The following figure shows an example of annotations on the **Lustre File System Metadata Dashboard**:

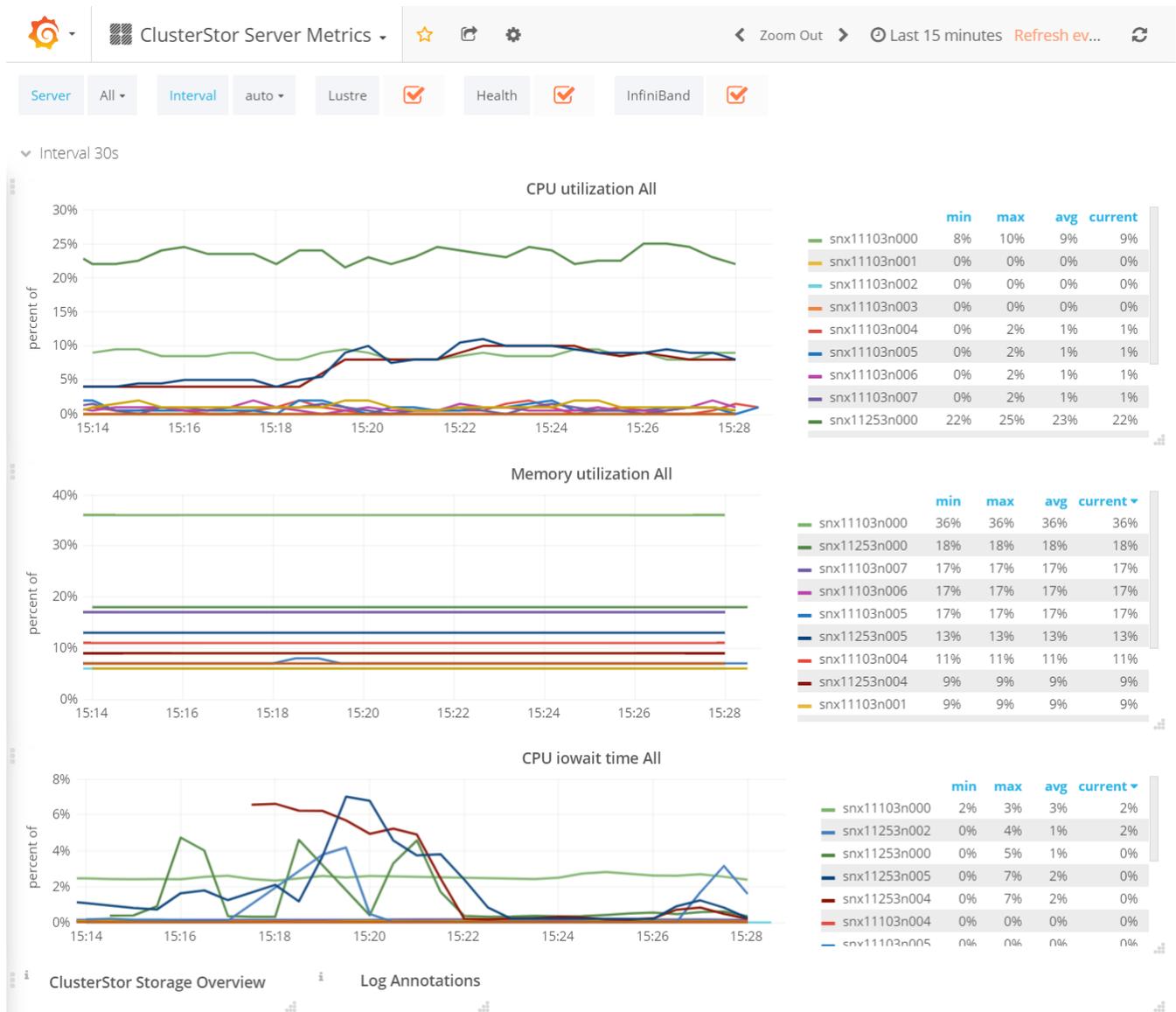


For more information, see [Annotations](#) on page 83.

- **Navigation** - Links at the bottom of each dashboard and the upper left corner of each dashboard panel, aid in navigating common workflows. Click the small **Details** icon to open the dialog containing the link(s).

The screenshot shows the "ClusterStor Storage Overview" dashboard. At the top, there's a navigation bar with a home icon, "snx11253 > Storage", and a "ClusterStor Storage Overvi..." title. Below this are icons for a gear, star, and refresh, along with "Zoom Out", "Last 15 minutes", "Refresh ev...", and a refresh icon. The main content area has tabs for "Server", "All", "Interval", and "auto". There are also checkboxes for "Lustre", "Health", and "InfiniBand". Below the tabs, there's a "Lustre metrics - Interval 30s" section. To the left, there's an "InfiniBand metrics" section with a sub-panel for "IB errors (vL15 drops, symbol, port emit discards)". To the right, there are two gauge charts: "Free data space" at 93% and "Free inode space" at 97%. Two callout boxes are present: one for "ClusterStor Server Metrics" and another for "Lustre File System Capacity".

## 11.1 ClusterStor Server Metrics Dashboard

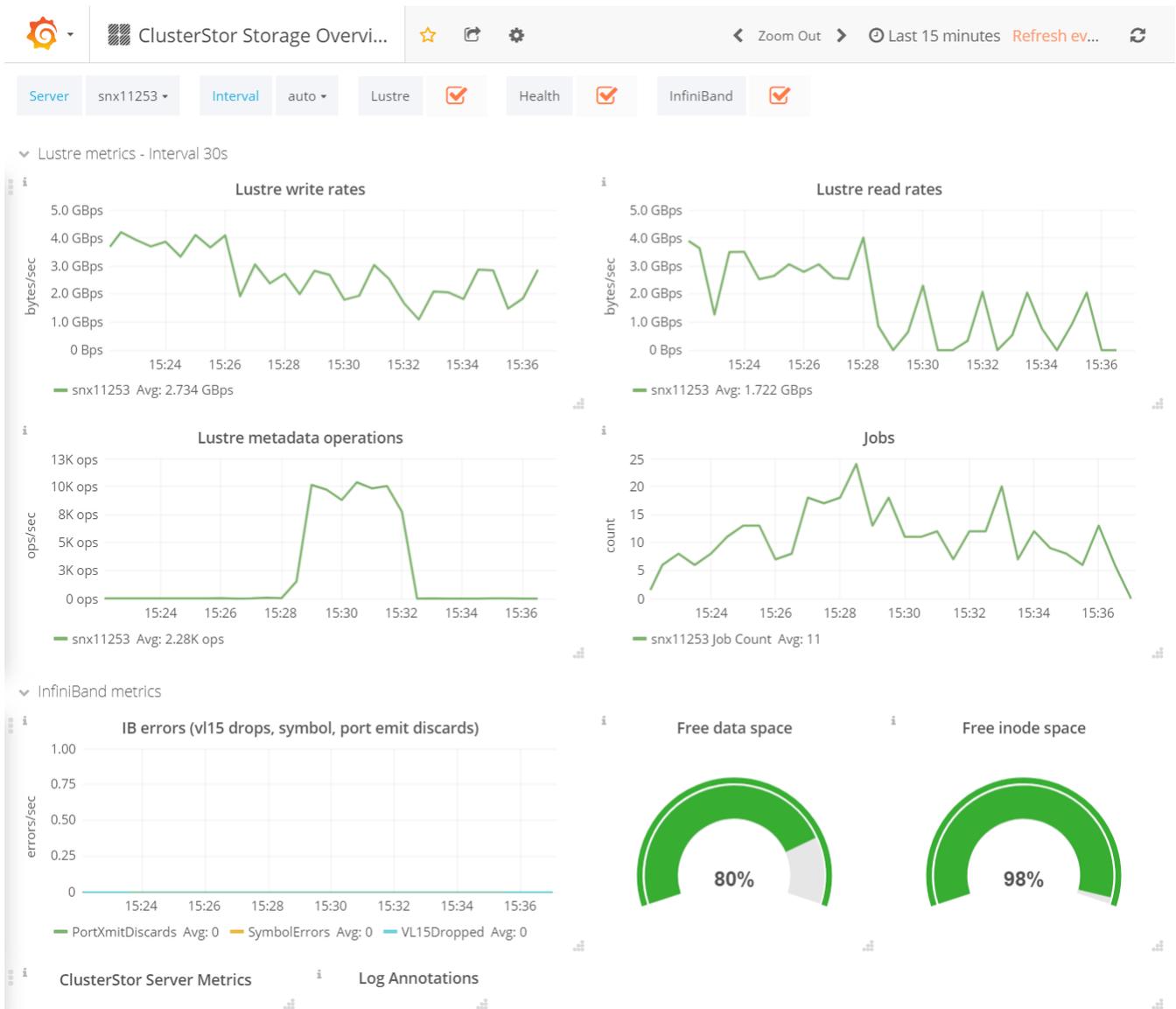


This dashboard has three panels.

Panel	Displayed Info	Metric
CPU utilization	CPU usage for all system nodes	cray_storage.cpu_utilization_perc Represents the percentage of time the CPU spends processing the instructions of a program or O/S.
Memory utilization	Memory usage for all system nodes	cray_storage.memory_utilization_perc Represents the percentage of memory used by processes and the O/S

Panel	Displayed Info	Metric
CPU iowait time	CPU iowait time for all system nodes	cray_storage.cpu_iowait_perc  Represents the amount of time a thread is blocked waiting for I/O to complete.

## 11.2 ClusterStor Storage Overview Dashboard

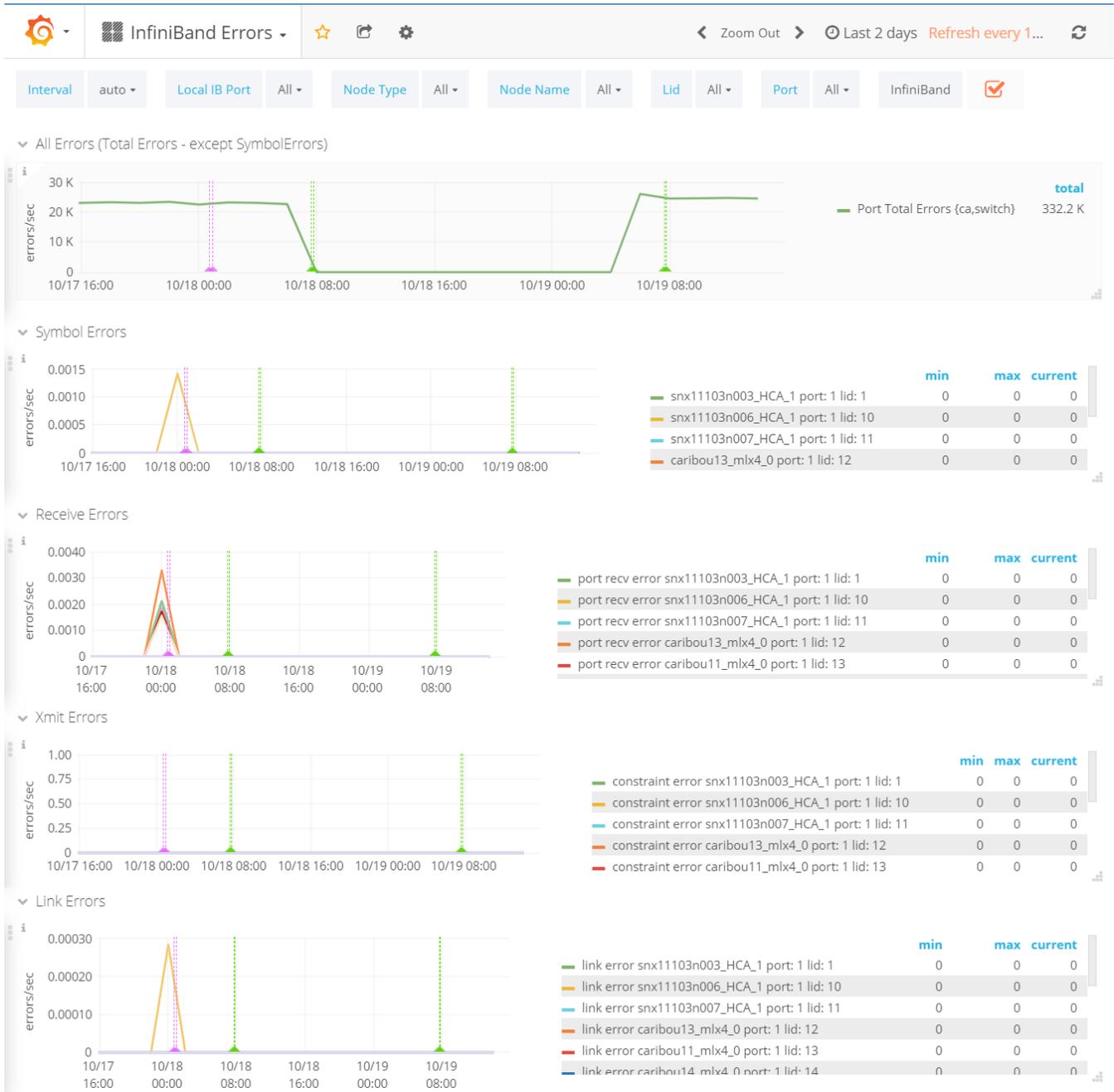


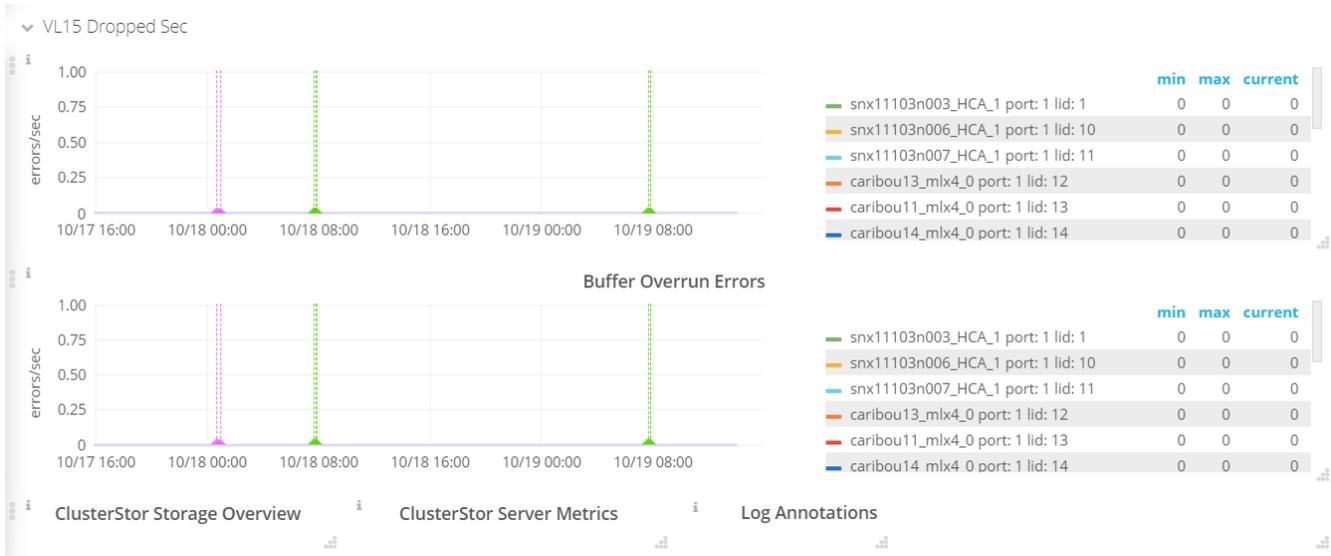
The **ClusterStor Storage Overview Dashboard** is the default home dashboard. Navigate to this dashboard by clicking the ClusterStor™ system name in a tile on the **System Overview**.

The dashboard has seven panels arranged in two rows. The rows can be collapsed or expanded. The dashboard charts can be refined to switch the ClusterStor system focus or to show aggregated data for multiple ClusterStor systems.

Panel	Displayed Info	Metric
Lustre write rates	Aggregated Lustre write rates	<code>cray_storage.calculated_write_bytes</code> Represents the total write byte rates aggregated across all OSTs.
Lustre read rates	Aggregated Lustre read rates	<code>cray_storage.calculated_read_bytes</code> Represents the total read byte rates aggregated across all OSTs.
Lustre metadata operations	Aggregated Lustre metadata rates	<code>cray_storage.calculated_metadata_ops</code> Represents the total metadata operation rates aggregated across all MDTs.
Jobs	Number of jobs that are currently being tracked.	n/a
Free data space	Percentage of free data space for all the OSTs	<code>cray_storage.free_space_perc</code> Represents the percentage of aggregated data space available across the ClusterStor system.
Free inode space	Percentage of free inode space for all the OSTs	<code>cray_storage.free_files_perc</code> Represents the percentage of aggregated inodes available for creating new files across the ClusterStor system.
InfiniBand error counters	Average of three discrete metrics	<ul style="list-style-type: none"> <li><code>cray_ib.port_xmit_discards_sec</code> The total rate per/sec of packets dropped because the port was down or congested.</li> <li><code>cray_ib.symbol_errors_sec</code> The total rate per/sec of component link errors detected on one or more physical lanes. This includes 8B/10B coding violations and is typically an indication of a bit error on the line.</li> <li><code>cray_ib.vl15_dropped_sec</code> The rate per/sec of incoming VL15 packets dropped due to resource limitations (for example, lack of buffers) in the port.</li> </ul>

# 11.3 InfiniBand Errors Dashboard





This dashboard has seven panels tracking InfiniBand errors in errors/sec. The error counter descriptions were taken from the following sources:

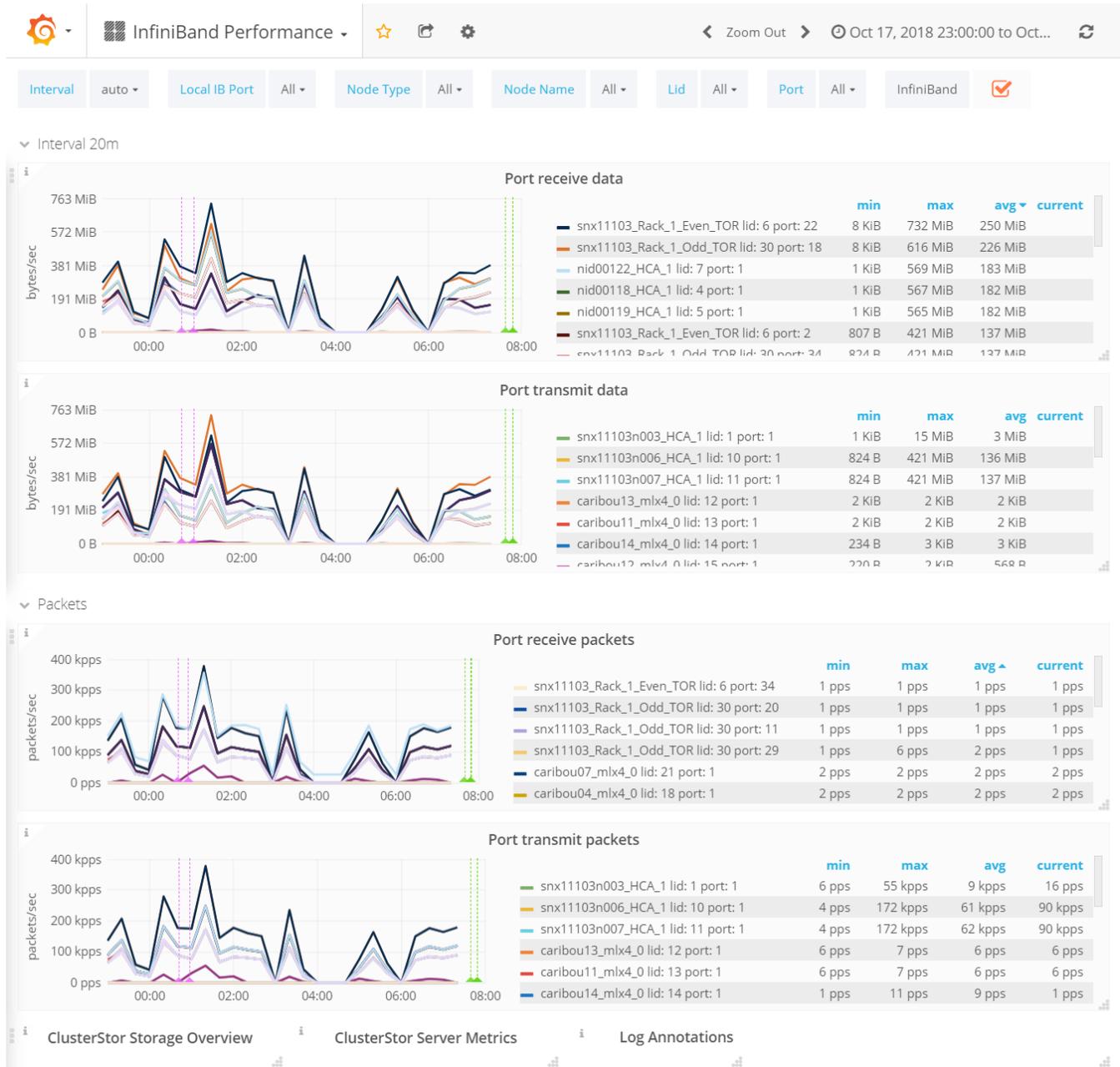
- OpenFabrics Alliance Software Wiki. "Overview of Error Counters". [https://www.openfabrics.org/mediawiki/index.php/Overview\\_of\\_Error\\_Counters](https://www.openfabrics.org/mediawiki/index.php/Overview_of_Error_Counters)
- Mellanox Technologies. "Understanding mlx5 Linux Counters and Status Parameters". <https://community.mellanox.com/docs/DOC-2572>

Panel	Displayed Info	Additional Info
Total Errors (except Symbol Errors)	Overview of all IB errors, per second, except for SymbolErrors.	<p>See the other panels in this table for definitions of the individual counters.</p> <ul style="list-style-type: none"> <li>• LinkErrorRecoveryCounter</li> <li>• LinkDownedCounter</li> <li>• PortRcvErrors</li> <li>• PortRcvRemotePhysicalErrors</li> <li>• PortRcvSwitchRelayErrors</li> <li>• PortXmitDiscards</li> <li>• PortXmitConstraintErrors</li> <li>• PortRcvConstraintErrors</li> <li>• LocalLinkIntegrityErrors</li> <li>• ExcessiveBufferOverrunErrors</li> <li>• VL15Dropped</li> </ul>
Symbol Errors	<p>Number of symbol errors per second.</p> <p>The interpretation of symbols within the packet is done on the HCA/CA. If the translation or interpretation fails, it creates a</p>	<p>Most !SymbolErrors are hardware related. If the counts are small, they can be ignored. If the numbers are large and/or the same CA is reporting this error regularly, it should be looked into. On a node, the HCA and/or cable should be reseated. If the reseal is unsuccessful, it should be replaced. On a switch, reseal the cable or replace the cable.</p>

Panel	Displayed Info	Additional Info
	minor event called a symbol error  Chart legend displays Min, Max, and Current rate.	
Recv Errors	Number of errors per second received on the port. This panel aggregates the following: <ul style="list-style-type: none"> <li>● PortRcvConstraintErrors</li> <li>● PortRcvSwitchRelayErrors</li> <li>● PortRcvErrors</li> <li>● PortRcvRemotePhysicalErrors</li> </ul> Chart legend displays Min, Max, and Current rate.	<ul style="list-style-type: none"> <li>● PortRcvConstraintErrors - This is the number of packets [ received and discarded on   not transmitted by ] a port in the fabric.  There are 2 general reasons for this: The filter for raw packets [ inbound   outbound ] is turned on and these are raw packets. The partition key or IP version check has failed.</li> <li>● PortRcvSwitchRelayErrors - This field counts the number of packets that could not be forwarded by the switch.  The reasons for this include: Virtual lane (VL) mapping errors; Looping; input port and output port are the same; DLID errors; It is a Multicast DLID (0xC000 to 0xFFFE) not configured for this CA, or DLID is outside the LFTS range or greater than the LinearFDBTop, or Port associated with this DLID in the LFTS file does not exist. Usually this is due to the poor implementation of multicast on IB and therefore can be ignored.</li> <li>● PortRcvErrors - These errors can be due to local physical errors, local buffer overruns, or receiving a malformed packet.  If a malformed packet is received, this indicates a problem somewhere else on the fabric. There are bad messages on the wire.</li> <li>● PortRcvRemotePhysicalErrors - The end bad packet (EBP) flag is set. This is usually a problem between the physical and logical layers.</li> </ul>
Xmit Errors	Number of errors per second transmitted on the port. This panel aggregates the following: <ul style="list-style-type: none"> <li>● PortXmitConstraintErrors</li> <li>● PortXmitDiscards</li> </ul> Chart legend displays Min, Max, and Current rate.	<ul style="list-style-type: none"> <li>● PortXmitConstraintErrors - This is the number of packets [ received and discarded on   not transmitted by ] a port in the fabric.  There are two general reasons for this: The filter for raw packets [ inbound   outbound ] is turned on and these are raw packets, or the partition key or IP version check has failed.</li> <li>● PortXmitDiscards - This counter tracks packets that were discarded instead of transmitted.</li> </ul>

Panel	Displayed Info	Additional Info
		<p>This usually indicates congestion in the fabric. The CA this packet was supposed to be sent to cannot accept it. After so many retries and/or too many incoming packets, the packet to be transmitted gets dropped. If the fabric is being routed well, without deadlocks or credit loops, these should be transient.</p>
Link Errors	<p>Number of link errors per second on the port. This panel aggregates the following:</p> <ul style="list-style-type: none"> <li>• LocalLinkIntegrityErrors</li> <li>• LinkErrorRecoveryCounter</li> <li>• LinkDownedCounter</li> </ul> <p>Chart legend displays Min, Max, and Current rate.</p>	<ul style="list-style-type: none"> <li>• LocalLinkIntegrityErrors - The number of times that the count of local physical errors exceeded the threshold specified by LocalPhyErrors.</li> <li>• LinkErrorRecoveryCounter - Total number of times the Port Training state machine has successfully completed the link error recovery process.</li> <li>• LinkDownedCounter - Counts links downed. This is usually associated with a node reboot. If not associated with a reboot, could be a failing connection.</li> </ul>
Buffer Overrun Errors	<p>Number of buffer overrun errors per second</p> <p>Chart legend displays Min, Max, and Current rate.</p>	<p>ExcessiveBufferOverrunErrors - The number of times that OverrunErrors consecutive flow control update periods occurred, each having at least one overrun error.</p>
VL15 Dropped Sec	<p>Number of dropped management packets per second on VL15.</p> <p>Chart legend displays Min, Max, and Current rate.</p>	<p>VL15 is the default virtual lane for management packets. They are the first to be dropped when there are resource limitations on the port. This is usually related to not enough space in the buffers. In many instances, these errors can be ignored.</p> <p>There have been instances when these messages were very closely correlated to user problems in time and fabric space. If they are being dropped, the buffers are being kept very busy with other data and this could indicate congestion.</p>

## 11.4 InfiniBand Performance Dashboard

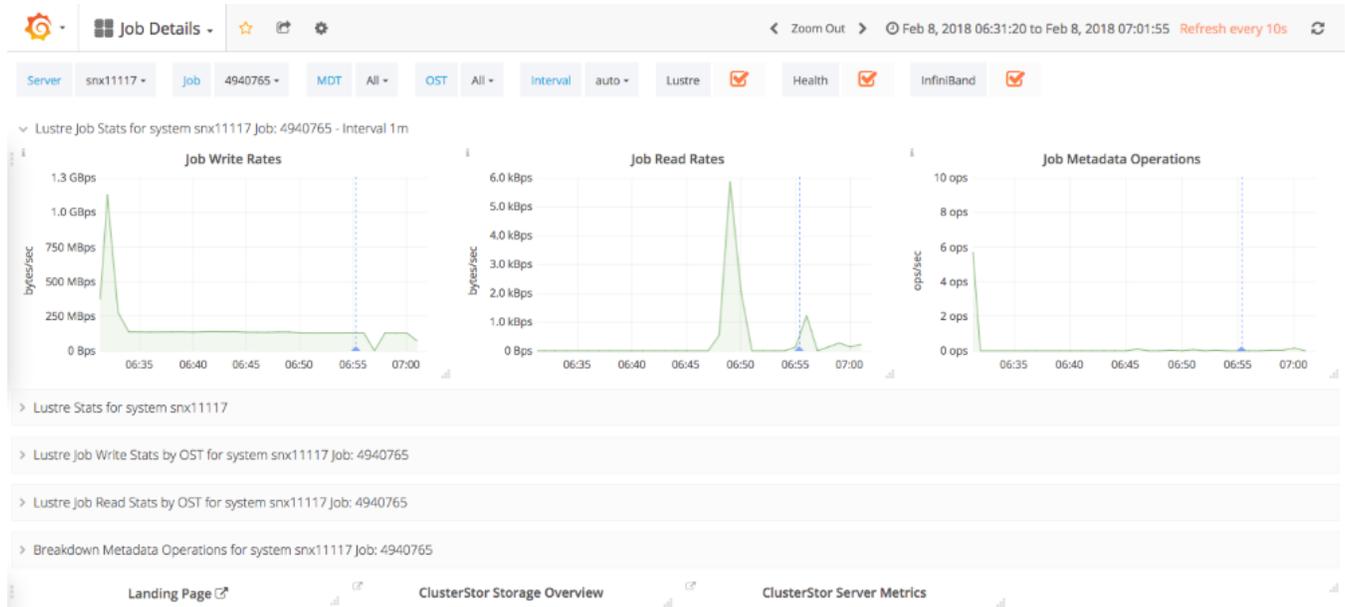


This dashboard has four panels.

Panel	Displayed Info	Metric
Port receive data	IB data received rate, in bytes/sec. Chart legend displays Min, Max, Average, and Current rate.	cray_ib.port_recv_bytes_sec Represents the total rate per/sec of data octets, divided by four, received on all virtual lanes (VL) from the port.

Panel	Displayed Info	Metric
Port transmit data	IB data transmitted rate, in bytes/sec. Chart legend displays Min, Max, Average, and Current rate.	<code>cray_ib.port_xmit_bytes_sec</code> Represents the total rate per/sec of data octets, divided by four, transmitted on all VLs from the port.
Port receive packets	IB data packets received rate, in bytes/sec. Chart legend displays Min, Max, Average, and Current rate.	<code>cray_ib.port_recv_pkts_sec</code> Represents the total rate per/sec of packets received on all VLs from the port.
Port transmit packet	IB data packets transmitted rate, in bytes/sec. Chart legend displays Min, Max, Average, and Current rate.	<code>cray_ib.port_xmit_pkts_sec</code> Represents the total rate per/sec of packets transmitted on all VLs from the port.

## 11.5 Job Details Dashboard



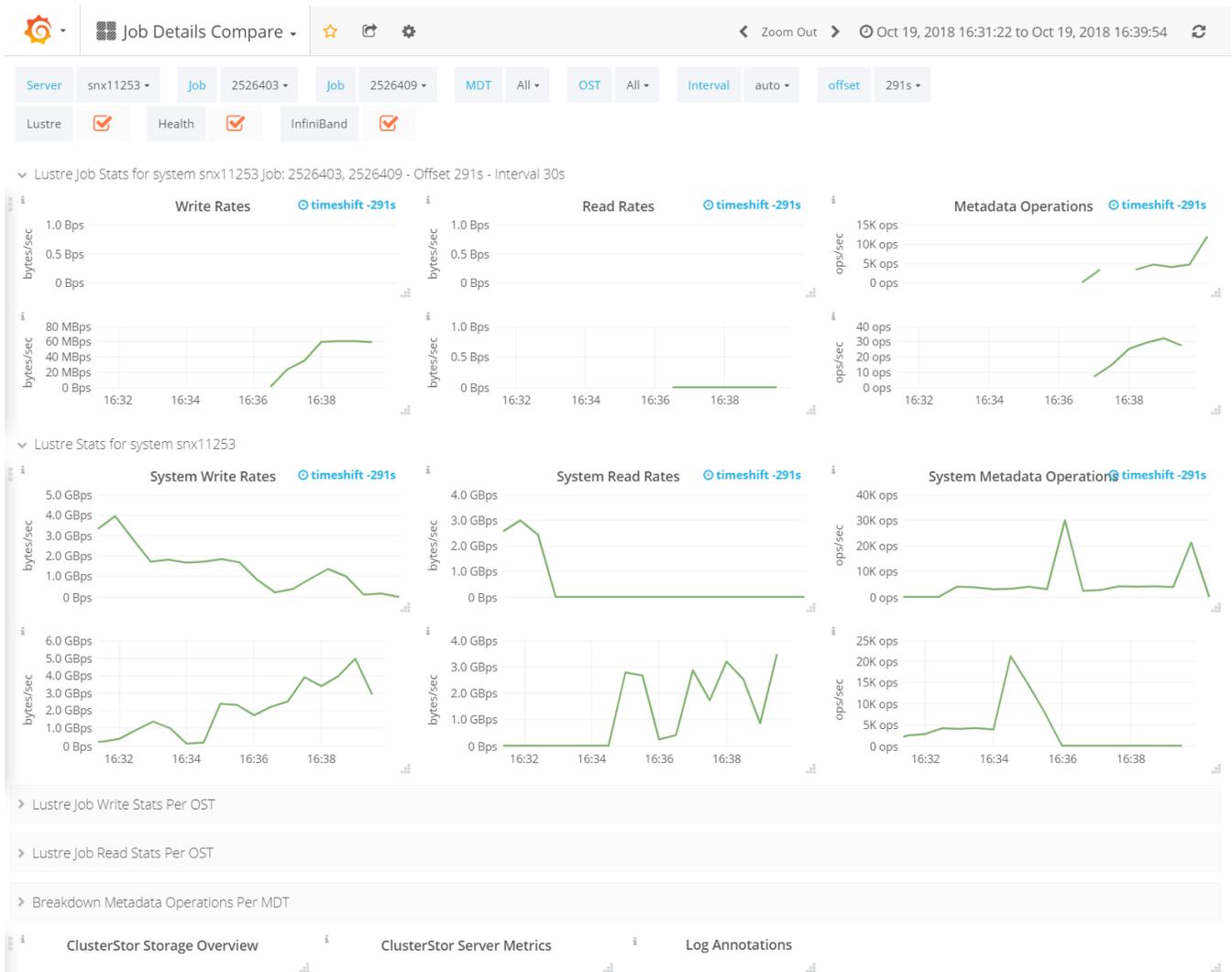
This dashboard has nine panels in five rows.

Panel (Row)	Displayed Info	Metric
Job Write Rates (Lustrre Job Stats)	Aggregate write rates for a job.	<code>cray_job.write_bytes_sec</code> Represents the total write byte rates initially aggregated across all OSTs.
Job Read Rates (Lustrre Job Stats)	Aggregate read rates for a job.	<code>cray_job.read_bytes_sec</code> Represents the total read byte rates initially aggregated across all OSTs.

Panel (Row)	Displayed Info	Metric
Job Metadata Operations (Lustre Job Stats)	Aggregate metadata operations for a job.	<code>cray_job.metadata_ops_sec</code> Represents the total metadata operation rates initially aggregated across all MDTs.
System Write Rates (Lustre Stats)	Aggregate write rates for the system.	<code>cray_storage.write_bytes_rate</code> Represents the total write byte rates aggregated across all OSTs. Stacking this row below the Lustre Job Stats row allows a direct comparison of a job's performance with the system performance.
System Read Rates (Lustre Stats)	Aggregate read rates for the system.	<code>cray_storage.read_bytes_rate</code> Represent the total read byte rates aggregated across all OSTs. Stacking this row below the Lustre Job Stats row allows a direct comparison of a job's performance with the system performance.
System Metadata Operations (Lustre Stats)	Aggregate MDT's metadata operations on the system.	<code>cray_storage.calculated_metadata_ops</code> Represents the total metadata operation rates aggregated across all MDTs.
OST Writes (Lustre Job Write Stats)	A job's write statistics.	<code>cray_job.write_bytes_sec</code> Represents the total write bytes rate aggregated across all OSTs, grouped by OST.
OST Reads (Lustre Job Read Stats)	A job's read statistics.	<code>cray_job.read_bytes_sec</code> Represents the total read bytes rate aggregated across all OSTs, grouped by OST.
Metadata Operations (Breakdown of Metadata Operations)	A job's metadata operations (per MDT) that a job has done.	Represented through multiple metrics, listed below. Not all possible metadata operations are listed. <ul style="list-style-type: none"> <li>• <code>cray_job.close</code> The number of <code>close</code> requests.</li> <li>• <code>cray_job.open</code> The number of <code>open</code> requests.</li> <li>• <code>cray_job.getatt</code> The number of <code>getattr</code> requests.</li> <li>• <code>cray_job.rename</code> The number of <code>rename</code> requests.</li> <li>• <code>cray_job.setattr</code> The number of <code>open setattr</code> requests.</li> <li>• <code>cray_job.rmdir</code> The number of <code>rmdir</code> requests.</li> </ul>

Panel (Row)	Displayed Info	Metric
		<ul style="list-style-type: none"> <li>cray_job.mkdir The number of mkdir requests.</li> <li>cray_job.unlink The number of unlink requests.</li> <li>cray_job.quotctl The number of quotctl requests.</li> </ul>

## 11.6 Job Details Compare Dashboard

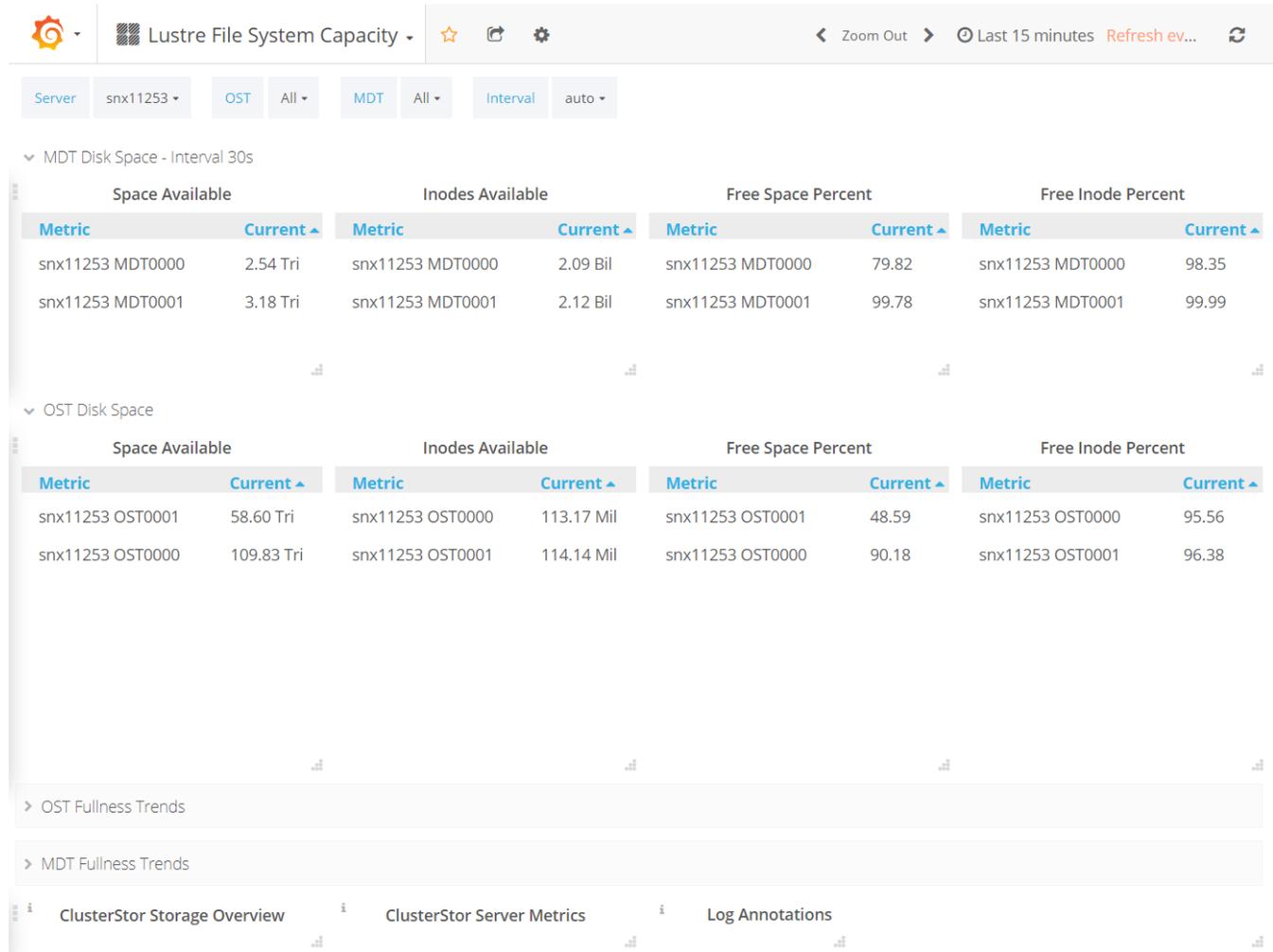


This dashboard has nine panels in five rows.

Panel (Row)	Displayed Info	Metric
Job Write Rates (Lustre Job Stats)	Aggregate write rates for a job.  The panel is stacked to show the job with the earliest start time and is time-shifted to align with the job with the latest start time.	<code>cray_job.write_bytes_sec</code>  Represents the total write byte rates initially aggregated across all OSTs.
Job Read Rates (Lustre Job Stats)	Aggregate read rates for a job.  The panel is stacked to show the job with the earliest start time and is time-shifted to align with the job with the latest start time.	<code>cray_job.read_bytes_sec</code>  Represents the total read byte rates initially aggregated across all OSTs.
Job Metadata Operations (Lustre Job Stats)	Aggregate metadata operations for a job.  The panel is stacked to show the job with the earliest start time and is time-shifted to align with the job with the latest start time.	<code>cray_job.metadata_ops_sec</code>  Represents the total metadata operation rates initially aggregated across all MDTs.
System Write Rates (Lustre Stats)	Aggregate write rates for the system.  The panel is stacked to show the system performance during the job with the earliest start time and is time-shifted to align with the system performance during the job with the latest start time.	<code>cray_storage.write_bytes_rate</code>  Represents the total write byte rates aggregated across all OSTs. Stacking this row below the Lustre Job Stats row allows a direct comparison of a job's performance with the system performance.
System Read Rates (Lustre Stats)	Aggregate read rates for the system.  The panel is stacked to show the system performance during the job with the earliest start time and is time-shifted to align with the system performance during the job with the latest start time.	<code>cray_storage.read_bytes_rate</code>  Represent the total read byte rates aggregated across all OSTs. Stacking this row below the Lustre Job Stats row allows a direct comparison of a job's performance with the system performance.
System Metadata Operations (Lustre Stats)	Aggregate MDT's metadata operations on the system.	<code>cray_storage.calculated_metadata_ops</code>  Represents the total metadata operation rates aggregated across all MDTs.
OST Writes (Lustre Job Write Stats)	A job's write statistics.	<code>cray_job.write_bytes_sec</code>  Represents the total write bytes rate aggregated across all OSTs, grouped by OST.

Panel (Row)	Displayed Info	Metric
	The panel is stacked to show the job with the earliest start time and is time-shifted to align with the job with the latest start time.	
OST Reads (Lustre Job Read Stats)	A job's read statistics.  The panel is stacked to show the job with the earliest start time and is time-shifted to align with the job with the latest start time.	<code>cray_job.read_bytes_sec</code>  Represents the total read bytes rate aggregated across all OSTs, grouped by OST.
Metadata Operations (Breakdown of Metadata Operations)	A job's metadata operations (per MDT) that a job has done.  The panel is stacked to show the job with the earliest start time and is time-shifted to align with the job with the latest start time.	Represented through multiple metrics, listed below. Not all possible metadata operations are listed. <ul style="list-style-type: none"> <li>• <code>cray_job.close</code> The number of <code>close</code> requests.</li> <li>• <code>cray_job.open</code> The number of <code>open</code> requests.</li> <li>• <code>cray_job.getattr</code> The number of <code>getattr</code> requests.</li> <li>• <code>cray_job.rename</code> The number of <code>rename</code> requests.</li> <li>• <code>cray_job.setattr</code> The number of <code>open setattr</code> requests.</li> <li>• <code>cray_job.rmdir</code> The number of <code>rmdir</code> requests.</li> <li>• <code>cray_job.mkdir</code> The number of <code>mkdir</code> requests.</li> <li>• <code>cray_job.unlink</code> The number of <code>unlink</code> requests.</li> <li>• <code>cray_job.quotctl</code> The number of <code>quotctl</code> requests.</li> </ul>

## 11.7 Lustre File System Capacity Dashboard



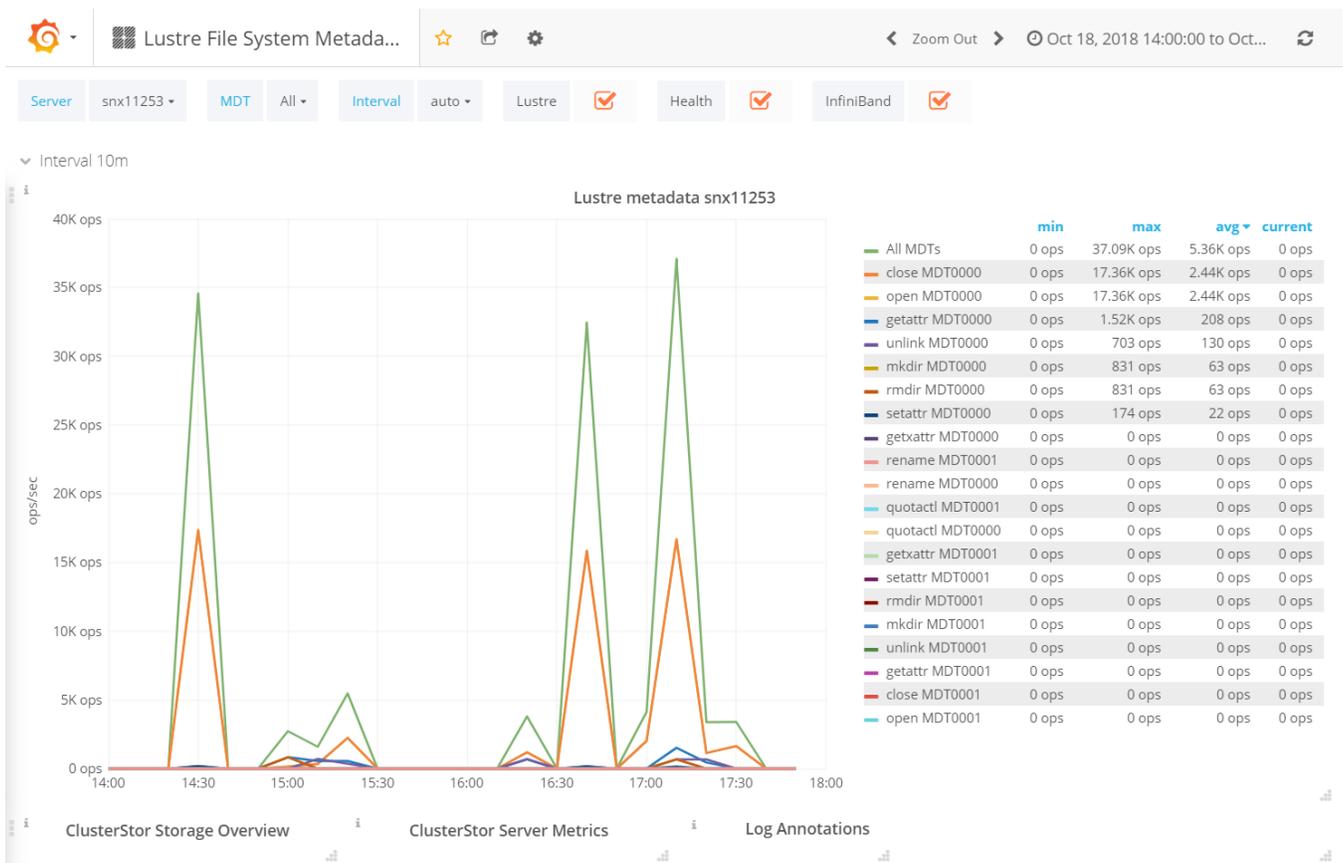
This dashboard has 16 panels in 4 rows.

Panel (Row)	Displayed Info	Metric
Space Available (MDT Disk Space)	MDT free space available. Displayed as tabular data, no chart.	<code>cray_storage.bytesavail</code> Represents the free space that is not reserved by a client's write cache or file system reserve.
Inodes Available (MDT Disk Space)	MDT free inodes available. Displayed as tabular data, no chart.	<code>cray_storage.filesfree</code> Represents the number of unallocated files (free inodes).
Free Space Percent (MDT Disk Space)	MDT free space percentage. Displayed as tabular data, no chart.	<code>cray_storage.free_space_perc</code> Represents the percentage of free space that is not reserved by a client's write cache or file system reserve.

Panel (Row)	Displayed Info	Metric
Free Inode Percent (MDT Disk Space)	MDT free inodes available percentage. Displayed as tabular data, no chart.	<code>cray_storage.free_files_perc</code> Represents the percentage of unallocated files still available for use.
Space Available (OST Disk Space)	OST free space available. Displayed as tabular data, no chart.	<code>cray_storage.bytesavail</code> Represents the free space that is not reserved by a client's write cache or file system reserve.
Inodes Available (OST Disk Space)	OST free inodes available. Displayed as tabular data, no chart.	<code>cray_storage.filesfree</code> Represents the number of unallocated files (free inodes).
Free Space Percent (OST Disk Space)	OST free space percentage. Displayed as tabular data, no chart.	<code>cray_storage.free_space_perc</code> Represents the percentage of free space that is not reserved by a client's write cache or file system reserve.
Free Inode Percent (OST Disk Space)	OST free inodes available percentage. Displayed as tabular data, no chart.	<code>cray_storage.free_files_perc</code> Represents the percentage of unallocated files still available for use.
Space Available (OST Fullness Trends)	OST free space available trend.	<code>cray_storage.bytesavail</code> Represents the free space that is not reserved by a client's write cache or file system reserve.
Inodes Available (OST Fullness Trends)	OST free inodes available trend.	<code>cray_storage.filesfree</code> Represents the number of unallocated files (free inodes).
Free Space Percent (OST Fullness Trends)	OST free space percentage trend.	<code>cray_storage.free_space_perc</code> Represents the percentage of free space that is not reserved by a client's write cache or file system reserve.
Free Inode Percent (OST Fullness Trends)	OST free inodes available percentage trend.	<code>cray_storage.free_files_perc</code> Represents the percentage of unallocated files still available for use.
Space Available (MDT Fullness Trends)	MDT free space available trend.	<code>cray_storage.bytesavail</code> Represents the free space that is not reserved by a client's write cache or file system reserve.
Inodes Available (MDT Fullness Trends)	MDT free inodes available trend.	<code>cray_storage.filesfree</code> Represents the number of unallocated files (free inodes).

Panel (Row)	Displayed Info	Metric
Free Space Percent (MDT Fullness Trends)	MDT free space percentage trend.	<code>cray_storage.free_space_perc</code> Represents the percentage of free space that is not reserved by a client's write cache or file system reserve.
Free Inode Percent (MDT Fullness Trends)	MDT free inodes available percentage trend.	<code>metric cray_storage.free_files_perc</code> Represents the percentage of unallocated files still available for use.

## 11.8 Lustre File System Metadata Dashboard

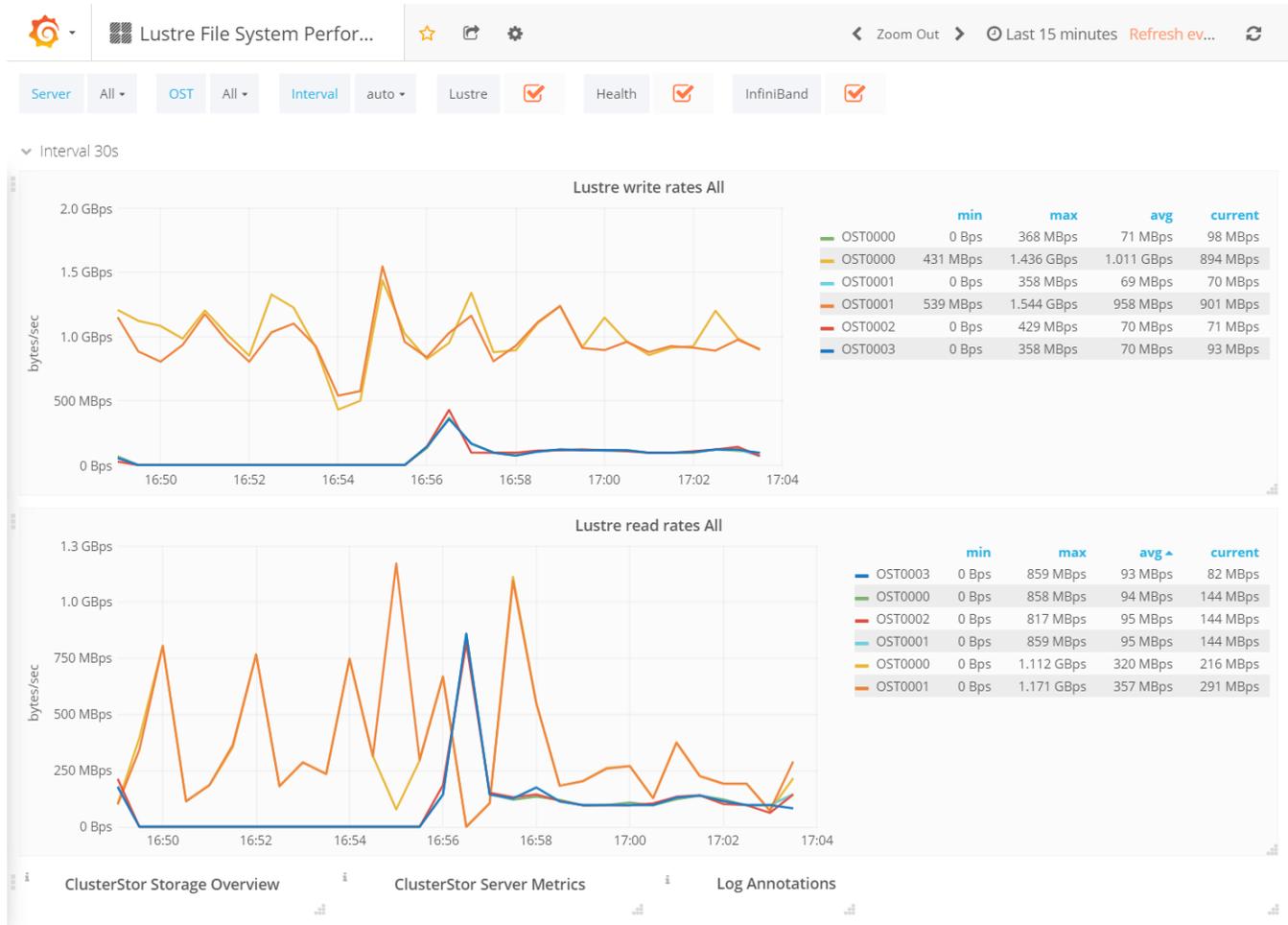


This dashboard has one panel.

Panel	Displayed Info	Metric
Lustre Metadata	MDT's metadata operations (per MDT).	Represented through multiple metrics, listed below. Not all possible metadata operations are listed. <ul style="list-style-type: none"> <li><code>cray_job.close</code></li> </ul>

Panel	Displayed Info	Metric
	Each display can be refined to show a specific MDT(s).	<p>The number of <code>close</code> requests</p> <ul style="list-style-type: none"><li>• <code>cray_job.open</code></li></ul> <p>The number of <code>open</code> requests</p> <ul style="list-style-type: none"><li>• <code>cray_job.getattr</code></li></ul> <p>The number of <code>getattr</code> requests</p> <ul style="list-style-type: none"><li>• <code>cray_job.getxatt</code></li></ul> <p>The number of <code>getxattr</code> requests</p> <ul style="list-style-type: none"><li>• <code>cray_job.rename</code></li></ul> <p>The number of <code>rename</code> requests</p> <ul style="list-style-type: none"><li>• <code>cray_job.quotctl</code></li></ul> <p>The number of <code>quotctl</code> requests</p> <ul style="list-style-type: none"><li>• <code>cray_job.setattr</code></li></ul> <p>The number of <code>open setattr</code> requests</p> <ul style="list-style-type: none"><li>• <code>cray_job.rmdir</code></li></ul> <p>The number of <code>rmdir</code> requests</p> <ul style="list-style-type: none"><li>• <code>cray_job.mkdir</code></li></ul> <p>The number of <code>mkdir</code> requests</p> <ul style="list-style-type: none"><li>• <code>cray_job.unlink</code></li></ul> <p>The number of <code>unlink</code> requests</p>

## 11.9 Lustre File System Performance Dashboard



This dashboard has two panels. The dashboard can be refined to show specific OSTs.

Panel	Displayed Info	Metric
Lustre write rates	Lustre write rates	<code>cray_storage.write_bytes_rate</code> Represents the write byte rates for each OST.
Lustre read rates	Lustre read rates	<code>cray_storage.read_bytes_rate</code> Represents the read byte rates for each OST.

---

## 12 InfiniBand Features

---

InfiniBand (IB) performance, configuration, and errors are collected for any attached InfiniBand networks.

### IB Topology Baseline

When the monitoring system is first initialized, the InfiniBand configuration is detected and a baseline configuration is saved on the View for ClusterStor™ server for future comparison. This initial configuration is considered to be the IB topology baseline. In subsequent discovery, comparisons will be made against the baseline to determine if the topology has changed.

The IB topology is checked by default every 900 seconds (15 minutes), and then compared to the saved IB topology baseline. If there's a divergence:

- A topology change event is logged in the View for ClusterStor logs and in the metric database. The log details of the topology changes can be retrieved via the **IB Page** in Kibana.
- The topology change is posted to the time-series database and an annotation will be created.
- Alarms are triggered and notifications sent to the configured email address.

If the configuration returns to a state equivalent to the baseline, a notification that the configuration has returned to normal is sent.

The saved IB baseline configuration can be updated using the procedure described in [Reset InfiniBand Topology Baseline](#) on page 80.

### Error and Performance Metrics

In addition to collecting and storing the IB topology baseline, error and performance metrics are gathered. For a list of the InfiniBand metrics gathered, see [InfiniBand Metrics](#) on page 52. The error metrics are gathered every 60 seconds. The extended metrics for transmits and receives are gathered every 30 seconds.

### Modify Sample Rates

IB metric intervals can be customized in the `/etc/sma-data/etc/.env` file. Collection intervals of three classes of IB metrics can be controlled:

- Extended performance metrics
- Error metrics
- IB topology updates

Default values, in seconds, are shown below:

```
EXTENDED_METRICS_INTERVAL=30
ERROR_METRICS_INTERVAL=60
IBTOPOLOGY_INTERVAL=900
```

Note that the collector agent polls on a 15 second interval, which places a lower limit on the sample intervals of the three classes of metrics.

Changes made to the `/etc/sma-data/etc/.env` file will not take place until View for ClusterStor is restarted.

## 12.1 Reset InfiniBand Topology Baseline

### About this task

When View for ClusterStor™ is first initialized, the InfiniBand (IB) configuration is detected and a baseline configuration is saved on the View for ClusterStor™ server for future comparison. This initial configuration is considered to be the IB topology baseline. In subsequent discovery, comparisons will be made against the baseline to determine if the topology has changed.

If the IB fabric is reconfigured in any way (e.g., adding or removing IB devices), a new baseline needs to be established. Perform the following procedure on the View for ClusterStor server to create a new baseline after the IB fabric has been reconfigured:

### Procedure

1. Log in to View for ClusterStor as root.
2. Remove the IB topology baseline files from the base host.

```
hostname# rm -f /etc/sma-data/ibperf/ibnetdiscover_cache.dat*
```

3. Log out of any View for ClusterStor GUI sessions that are open in a browser.
4. Restart the View for ClusterStor service.

```
hostname# systemctl restart sma
```

A new IB topology baseline has been established.

5. Restart the View for ClusterStor GUI, if desired, after the service has restarted:
  - a. Reload the browser tab that was previously used to work with the GUI, or open a new browser tab and navigate to the GUI URL.
  - b. Log in to the new GUI session.

## 12.2 Modify or Rebuild the InfiniBand Name Map File

By default, ClusterStor top-of-rack InfiniBand (IB) switches are not given a unique name (e.g., `SwitchX - Mellanox Technologies`). This makes it difficult to identify ClusterStor IB switches. To make it easier to identify these switches in View for ClusterStor™, an IB name map file is created, if it does not already exist, when the View for ClusterStor service is started. The map file includes the switch GUID and defines a corresponding name. The names defined in the file will appear in any IB-related View for ClusterStor dashboards.

The IB name map file, named `ib_name_map.txt`, is located in the `/etc/sma-data/ibperf` directory on the View server. A typical IB name map file looks similar to this:

```
hostname# cat /etc/sma-data/ibperf/ib_name_map.txt
0xf452140300838540 snx11103_Rack_1_Even_TOR
0x0002c903006fe4b0 snx11033_Rack_1_Odd_TOR
0x0002c90300676c30 snx11033_Rack_2_Even_TOR
0xf452140300838340 snx11103_Rack_2_Odd_TOR
```

## Modify the IB Name Map File

The IB name map file is initially populated with only the IB switch GUIDs and fabricated names that are based on the attached ClusterStor devices. However, the file can be edited to map both the IB switches and IB channel adapters (CA).

1. Log in to the View server as `root`.
2. Navigate to the `/etc/sma-data/ibperf` directory.

```
hostname# cd /etc/sma-data/ibperf
```

3. Edit the file, following the example name map file shown above.

```
hostname# vi ib_name_map.txt
```

4. Save the modified file.

## Rebuild the IB Name Map File

To rebuild the IB name map file:

1. Log in to the View server as `root`.
2. Navigate to the `/etc/sma-data/ibperf` directory.

```
hostname# cd /etc/sma-data/ibperf
```

3. Remove the existing name map file.

```
hostname# rm ib_name_map.txt
```

4. Log out of any View for ClusterStor GUI sessions that are open in a browser.
5. Restart the View for ClusterStor service.

```
hostname# systemctl restart sma
```

6. Restart the View for ClusterStor GUI, if desired, after the service has restarted:
  - a. Reload the browser tab that was previously used to work with the GUI, or open a new browser tab and navigate to the GUI URL.
  - b. Log in to the new GUI session.

The map can also be recreated without restarting the View for ClusterStor service, by running the following command:

```
hostname# docker exec -it sma_infiniband_1 python /usr/local/lib/python2.7/dist-packages/cray_ibperf/discovery.py --create_switch_map_file --force
```

---

## Delete Alarms that Have Old Names

Changing switch names as described in the above procedure may result in more than the expected number of `ib_symbol_errors` appearing in the **Alarms Table** because the alarms will display with both the old and new switch names. To ensure this does not happen, remove all the alarms and establish the new alarms using the following steps.

1. Log in to the View for ClusterStor server as `root`.
2. Navigate to the `/etc/sma-data/etc` directory.

```
hostname# cd /etc/sma-data/etc
```

3. Run the following commands:

```
hostname# docker-compose run --rm alarms remove_all_alarms.sh
hostname# docker-compose run --rm alarms /start.sh
```



**CAUTION:** All alarm history will be lost when removing alarms.

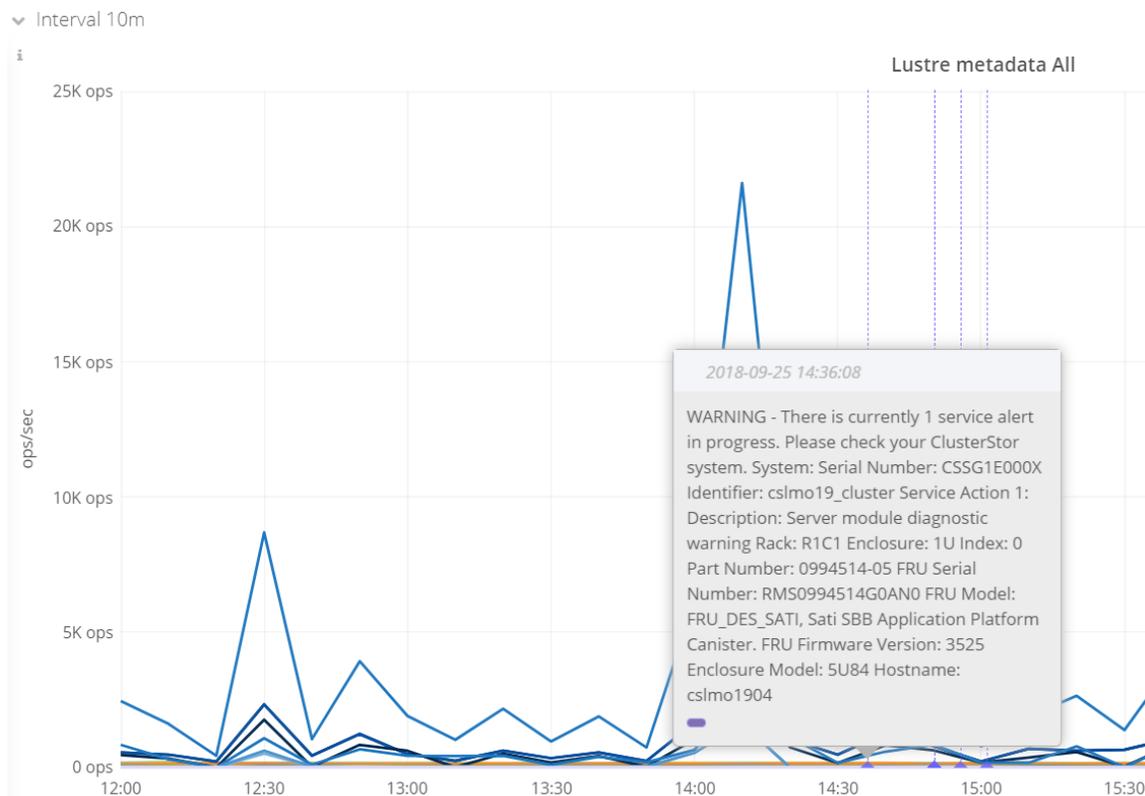
## 13 Annotations

The View for ClusterStor™ server includes an annotation facility for recording and retrieving important events that happen in either the InfiniBand network or the monitored ClusterStor systems. These events are recorded to Elasticsearch, much like logs. A number of these events in the system log fields are displayed in View for ClusterStor as annotations or are highlighted with warning icons in the **System Overview** tiles.

There are currently three classifications of event types:

- **Lustre** - Critical Lustre errors noted in the ClusterStor logs.
- **Health** - SNMP hardware events noted from a polling of SNMP.
- **InfiniBand** - Changes in InfiniBand topology.

The following figure shows an example of annotations on the **Lustre File System Metadata Dashboard**:



### Lustre Event Type

View for ClusterStor highlights when Lustre has issued a critical error on the ClusterStor system. The critical Lustre alerts will be written to the Elasticsearch `srx-logs_*` indices. These are displayed in the graphs as Lustre annotations.

## InfiniBand Event Type

View for ClusterStor tracks and issues alerts when changes to the IB topology occur. It compares the current topology to a saved baseline configuration every 900 seconds. The baseline is the normal operating IB topology.

When View for ClusterStor detects an IB topology transition to or from the baseline, it will send an alarm notification email and log the error in Elasticsearch. The ClusterStor system service alerts will be written to the Elasticsearch `~event-ibtopology_*` indices. These are displayed in the graphs as InfiniBand annotations.

## Health Event Type

Health events are a collection of zero or greater service alerts that are present in the ClusterStor system storage cluster. View for ClusterStor gathers health events by parsing through ClusterStor system log files. The composition of the ClusterStor log files varies depending on the software release that is running on the ClusterStor system. As a result, the health events that appear as annotations in View for ClusterStor are heavily dependent on the ClusterStor release. For example, the following table illustrates the success in detecting OST and MDT failover/failback events across 3 different ClusterStor system updates (SU):

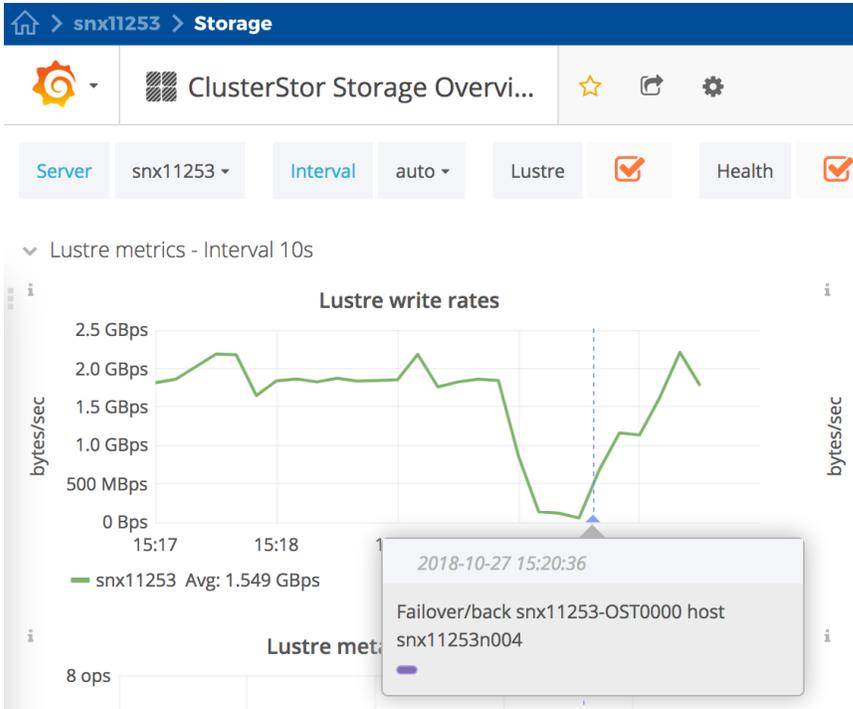
ClusterStor Release	OST Failover/Back	MDT Failover/Back
3.0.0 SU-011	High	High
3.0.0 SU-010	High	Low
2.0.0 SU-028	Medium	Low

Please note that View for ClusterStor annotations are intended to provide system administrators with insight into the ClusterStor system's performance characteristics that are being displayed in Grafana dashboards, but are not intended to provide a definitive picture.

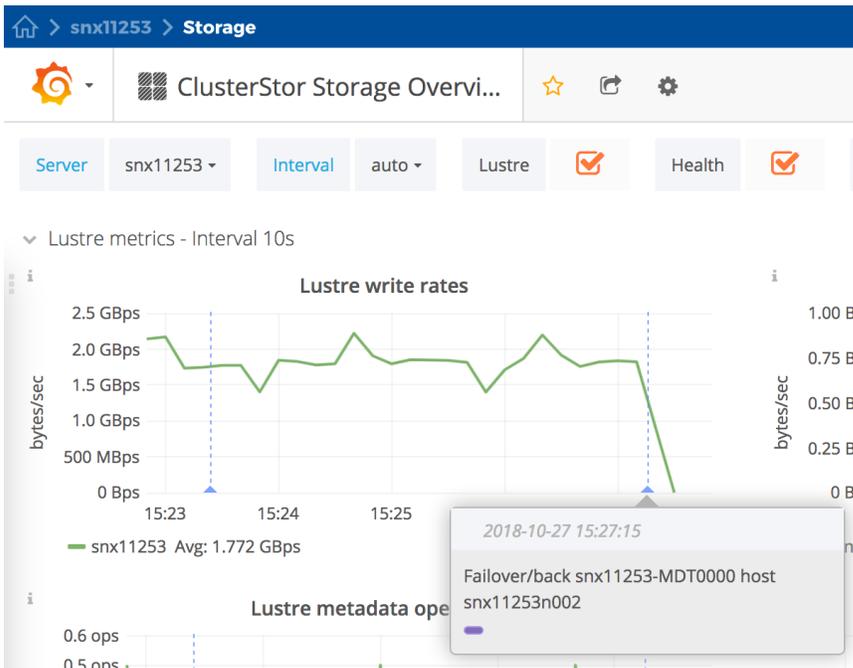
The ClusterStor system service alerts are written to the Elasticsearch `event-*` indices and are displayed in the graphs as Health annotations.

Some examples of these events may include:

- FRU Fan Status (`event-hdwefan_*`)
- Raid Resync start/complete (`event-mdraid_*`)
- Slow drive events (`event-hdslowdrive_*`)
- Indications of ClusterStor OST/MDT failover (`event-stgfailover_*`)
  - OST failover/back event example:



- o MDT failover/back event example:



## 14 Sample Workflows

---

The basic workflows described in this section illustrate the potential uses of View for ClusterStor™, and how to best utilize the variety of information panels.

Each workflow demonstrates a singular use of View for ClusterStor. For most tables, dashboards, or pages discussed in the workflows, selectable GUI elements will open additional information views.

The remainder of this section will discuss the following in more detail:

- Dashboard navigation
- Jobs workflow
- OST workflow
- MDT workflow
- System performance workflow
- InfiniBand workflow
- ClusterStor Logs
- ClusterStor Health Events

Each of these workflows start from the View for ClusterStor **System Overview**.

### 14.1 Dashboard Navigation

#### About this task

View for ClusterStor™ consists of custom created dashboards integrated with predefined Grafana dashboards, as well as Kibana pages. All general navigation begins from the **System Overview** tiles.

**NOTE:** When navigating directly from a View for ClusterStor tile to a Grafana dashboard, the charts and metrics shown are specific to the ClusterStor system represented in that tile. Once in Grafana, however, using the **Dashboard Selector** to navigate to a different dashboard will change the scope of the data displayed, to all ClusterStor systems being monitored in View for ClusterStor. To maintain scope to a specific ClusterStor system when navigating within Grafana, use the links that appear when selecting the **Details** icons at the upper left of any panel in a dashboard.

#### Procedure

1. Select the desired information block topic from one of the ClusterStor system tiles on the screen.  
Some information blocks, when selected, will navigate directly to a predefined Grafana dashboard.

See [System Tiles](#) on page 14 for details about the different information blocks of a tile, and the table, dashboard, or page where the system will navigate when the information block is selected.

The remainder of these steps describes workflow within any Grafana dashboard. For more information about the header that is displayed above each Grafana dashboard, see [Grafana Header](#) on page 23.

2. Perform inter-panel navigation by selecting the **Details** icon on a panel.

Each dashboard panel has a **Details** icon in the upper left corner. Selecting this icon will provide either more information about the panel or navigation links to other dashboards or panels.

3. Use navigation options to access additional details.

Relevant navigation options vary per panel.

4. Click on the browser's **Back** button/arrow to return to the previous screen.

5. Modify time frames and refresh rates by selecting the time ribbon in the upper right corner of the dashboard header. For example:

Last 15 minutes Refresh every 10s

6. Zoom in on a specific time frame by selecting a section on the screen. For example:



7. Select the **Home** icon, located on the left side of the content pane banner, to return to the **System Overview**.

## 14.2 Jobs Workflow

To analyze information about jobs, select the **Jobs** information block from a tile on the **System Overview**, which opens the **Jobs Table**. To analyze information about OSTs, select the **OST** information block from a tile on the **System Overview**, which opens the **OSTs Table**.

Figure 9. Jobs Table

snx11253 > Jobs									
Filter Last 3 hours									
156 jobs loaded in 0.376 seconds a few secs ago									
Host Name	Job ID	Apid	User ID	Application	Start Time	End Time	Duration	Avg. I/O Size	Metadata Ops
pollux-smw	2522680	1356	growfiles_mpi	2018-10-17 09:39:47	2018-10-17 09:43:57	4m 10s	2.57MB	2.02m	
pollux-smw	2522679	1356	growfiles_mpi	2018-10-17 09:39:45	2018-10-17 09:43:48	4m 3s	3.72MB	19.4m	
pollux-smw	2522750	1356	IOR	2018-10-17 10:37:14	2018-10-17 10:40:12	2m 58s	4.19MB	1.25k	
pollux-smw	2522644	16912	serbalt	2018-10-17 08:58:19	2018-10-17 08:58:47	28s	4.19MB	16	
pollux-smw	2522647	1356	IOR	2018-10-17 08:59:03	2018-10-17 09:19:13	20m 10s	1.68MB	4.07k	
pollux-smw	2522649	22569	nemod	2018-10-17 09:00:03	2018-10-17 09:20:13	20m 10s	2.1MB	646k	
pollux-smw	2522650	1356	IOR	2018-10-17 09:01:48	2018-10-17 09:03:15	1m 27s	1.41MB	1.36k	
pollux-smw	2522654	16912	serbalt	2018-10-17 09:07:04	2018-10-17 09:07:36	32s	4.18MB	16	
pollux-smw	2522655	1356	IOR	2018-10-17 09:07:04	2018-10-17 09:14:08	7m 4s	1.34MB	4.89k	
pollux-smw	2522656	7862	jrebsnorth	2018-10-17 09:07:36	2018-10-17 09:10:50	3m 14s	2.09MB	5.01k	
pollux-smw	2522658	16912	serbalt	2018-10-17 09:10:50	2018-10-17 09:11:25	35s	4.18MB	16	
pollux-smw	2522659	7862	jrebsnorth	2018-10-17 09:11:25	2018-10-17 09:14:24	2m 59s	2.09MB	5.01k	
pollux-smw	2522662	16912	serbalt	2018-10-17 09:14:24	2018-10-17 09:14:52	28s	4.19MB	16	

The **Jobs Table** will show a view of all jobs currently doing I/O or Metadata operations. Information from the SMW is integrated with jobstat metrics. The jobstat metrics are flagged to highlight jobs of interest.

- **Avg. I/O Size** - Average I/O Size is calculated by View for ClusterStor as the total number of bytes read and written since the given job's inception, divided by the total number of read and write requests since the job's inception.
- **Metadata Ops** - Metadata Ops is calculated by View for ClusterStor as the sum of all the metadata operations since the given job's inception, from each MDT.

For complete details about the **Jobs Table** columns, see [Jobs Table](#) on page 42.

Information in the **Jobs Table** can be re-sorted by clicking on a column header.

## Search for Jobs

If the job ID or apid was running outside the window of time displayed in the **Jobs Table**, select the **Filter** icon on the content pane banner to open the **Filter Jobs** window.

Specify an **apid** and a **time range**, and then select **Submit**.

Optionally, select **Set custom time range** and select an ending date/time. Do not specify a start time. This will result in a search of the entire database.

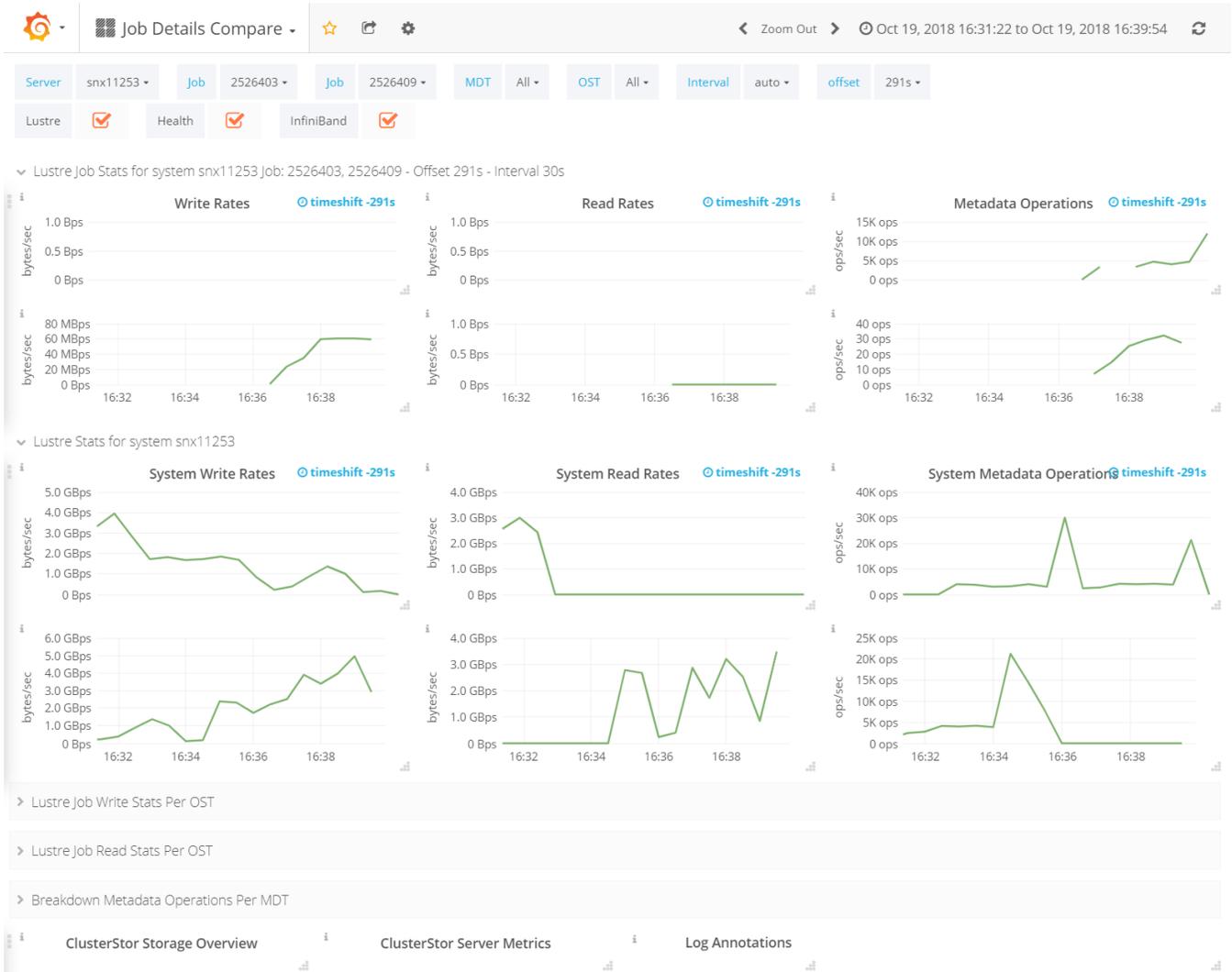
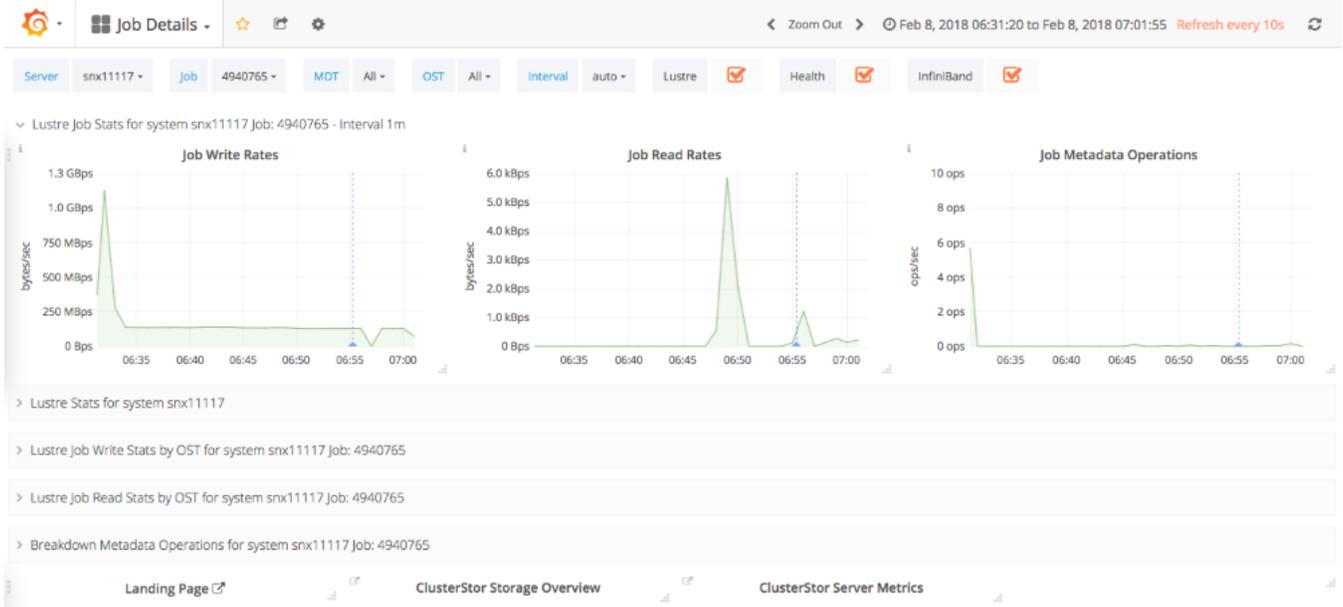
### 14.2.1 Job Details Dashboards

On the **Jobs Table**, hover over the job whose details are to be viewed, and then select the **Details** icon at the right side of the job's row. Alternatively, select multiple jobs by hovering over the job rows and marking the selection checkboxes, and then select the **Details** icon that appears at the right side of the **Jobs Table** column headings.

These actions will open the **Job Details Dashboard**, which provides details about the selected job(s).

To compare two jobs, select the jobs on the **Jobs Table**, and then select the **Compare** icon that appears at the right side of the **Jobs Table** column headings. This will open the **Job Details Compare Dashboard**, which details the two jobs synced by start time.

These two custom dashboards provide additional information for the selected job(s).



Detailed job information presented in the dashboards' panels includes:

- History of job write rates
- History of job read rates
- Historic rate of metadata operations

Open the **Lustre Stats** row to compare how the specific job's performance compares to the overall system performance.

Open the **Lustre Read Stats** or **Lustre Writes Stats** rows to see per-OST details of a job. The last row on the dashboard provides a breakdown of metadata operations (per MDT) for the job.

Note that if multiple jobs were selected, the system displays aggregated information for the selected jobs.

## Direct Job Access

A user can also look up details of a specific job id or apid by navigating directly to one of these dashboards and modifying the template variables, which appear below the Grafana header:

1. Enter the job ID or apid into the **Job** variable.
2. Select a time range that covers the period during which the job ran.

See [Job Details Dashboard](#) on page 69 and [Job Details Compare Dashboard](#) on page 71 for additional information about these dashboards.

## View Job Details for PBS Pro Jobs

The workflow for jobs run with PBS Pro is essentially the same as the workflow for other jobs, with two exceptions:

- Latency - PBS Pro job completion records may become available in View for ClusterStor only after some delay once the job has completed. All of the data in the record is accurate, including job completion time, but the record may not be up to date until some period after the job is completed.

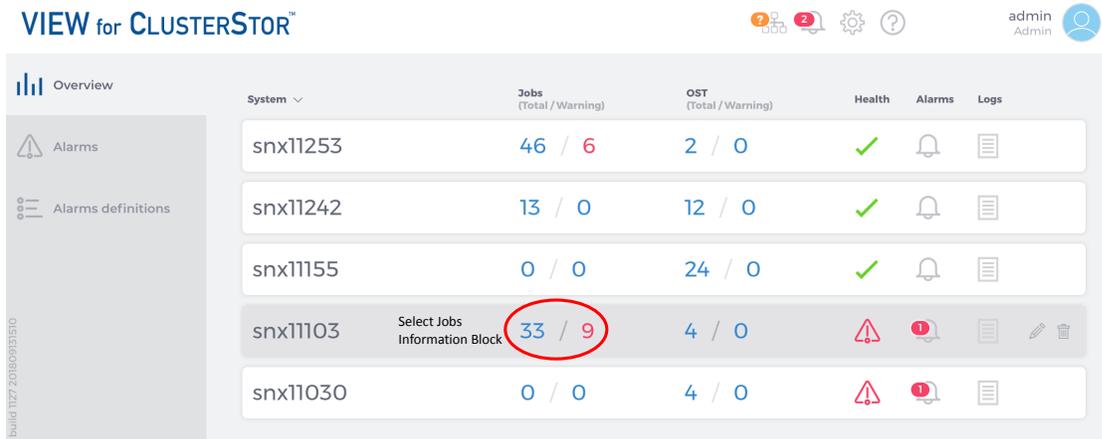
This latency can be reduced by setting the `pbs_epilog_timer` configuration option to a value other than the default of 60 seconds. This option specifies the time interval for the `pbs_epilog` function to scan for and find completed jobs. Setting a shorter interval will result in the View server updating more quickly, but will increase overhead on the PBS Pro server node.

For more information about setting the `pbs_epilog_timer` configuration option, see the *View for ClusterStor Installation and Configuration Guide (S-3025)*.

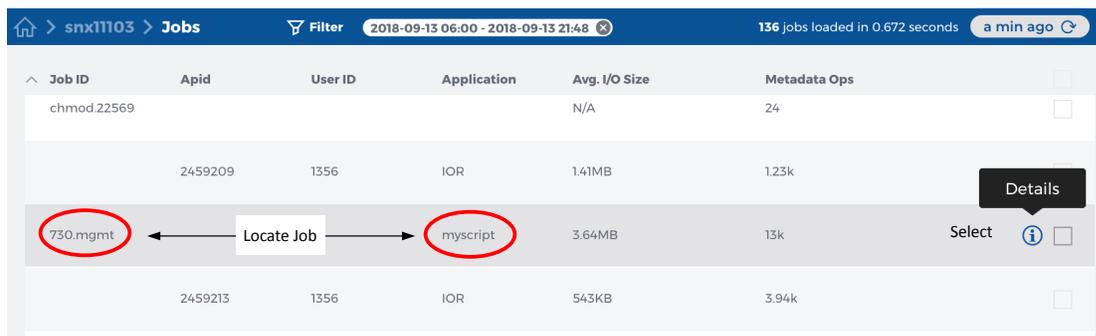
- All PBS jobs will appear in the View for ClusterStor graphical user interface (GUI) with a null apid. Jobs that run on PBS clusters without Cray Alps, do not have apids associated with constituent applications, thus none are presented in View.

1. Select the Jobs information block from a tile on the **System Overview**.

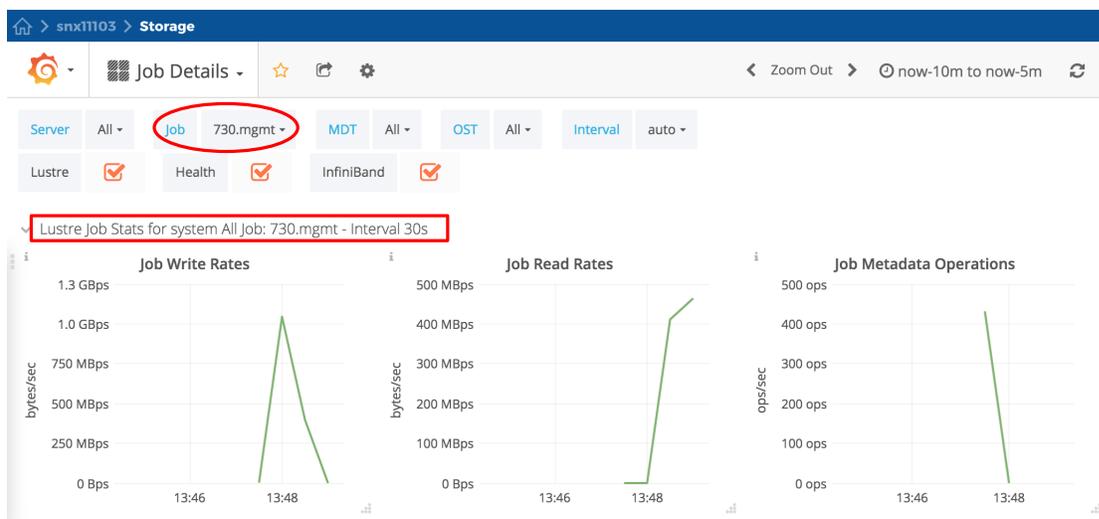
Note that job data is tied to a particular storage device, even if the job uses multiple devices. In the following example, the job was tied to ClusterStor system **snx11103**, the Job ID was **730.mgmt**, and the application name was **myscript**.



- Locate the PBS Pro job in the **Jobs Table** and then select the **Details** icon to open the **Job Details Dashboard**.



- Review the job's detailed information.



## 14.2.2 OSTs Table

To determine what jobs are running on what OSTs, select the **OST** information block from a tile on the **System Overview**. This will open the **OSTs Table**.

Job	Device	Write Bandwidth (/s)	Read Bandwidth (/s)
2494610	OST0000	350MB	N/A
2494609	OST0000	547MB	N/A
2494609	OST0001	449MB	N/A
2494608	OST0001	14.2MB	N/A
2494607	OST0000	350MB	N/A
2494606	OST0001	28.3MB	N/A
2494605	OST0000	228MB	N/A
2494604	OST0001	118MB	N/A
2494603	OST0000	44.1MB	N/A
2494602	OST0001	61.6MB	N/A
2494601	OST0000	350MB	N/A
2494600	OST0001	7.1MB	N/A

Select an alternative time range in the content pane banner to change the number of jobs displayed in the **OSTs Table**.

### OSTs Table Column Descriptions

Column Heading	Description
<b>Job</b>	The job id or apid of the job that ran or is running during the selected time range.
<b>Device</b>	Name of an OST that the job ran on. There may be multiple rows for the same job id or apid with different OSTs shown
<b>Write Bandwidth (/s)</b>	The write bandwidth, per second, of the OST on which the job is running.
<b>Read Bandwidth (/s)</b>	The read bandwidth, per second, of the OST on which the job is running.

### Filter and Sort the OSTs Table

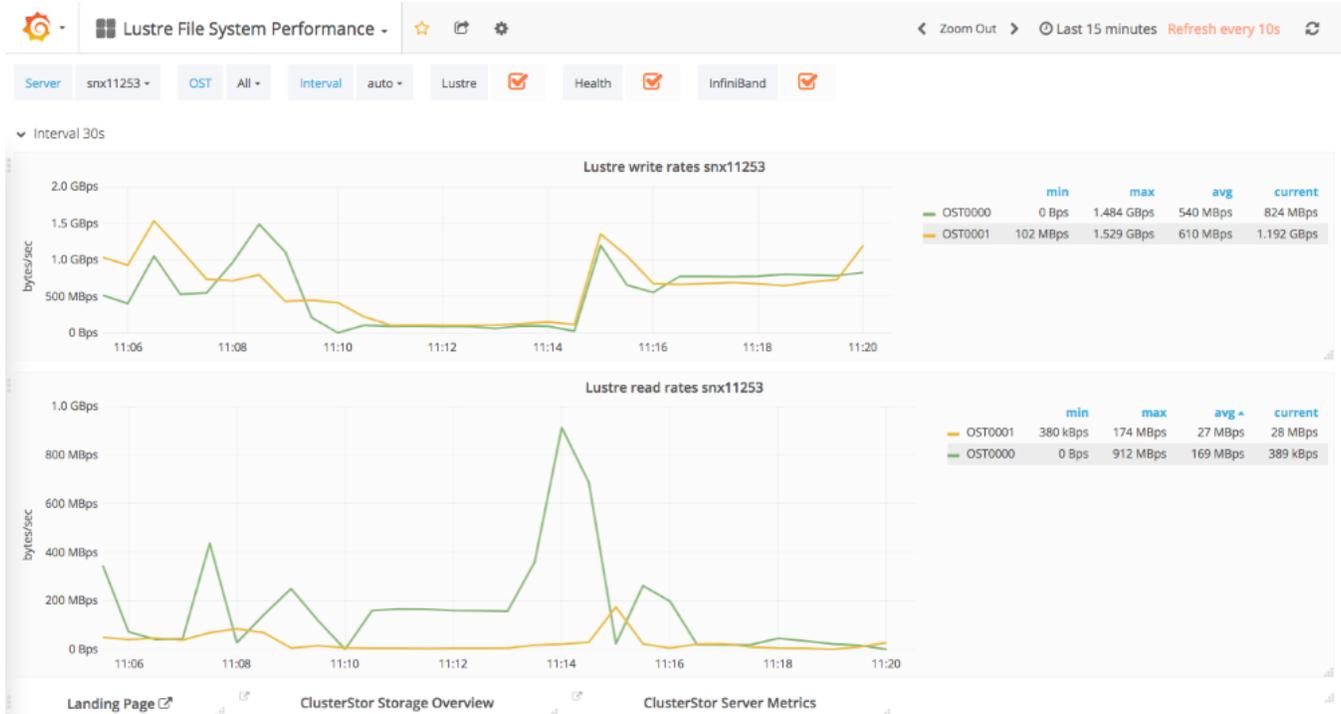
The **OSTs Table** can be filtered and sorted in the following ways:

Method	Location	Description and Use
Column filter icon	Column heading	Create a filter for specific columns in the <b>OSTs Table</b> . Hover over a column heading to display the small <b>Filter</b> icon, and then click the icon.

Method	Location	Description and Use
		<ul style="list-style-type: none"> <li>Enter some text to search for job IDs, or OST devices. Press <b>Enter</b> to apply the filter.</li> <li>One or more OST devices may also be selected from the list of displayed devices. Enter text to reduce the number of items shown in the list. Filter options are applied as they are selected. Click the light gray <b>X</b> in the upper right to close the filter options.</li> <li>Adjust the sliders to set minimum and maximum values for write bandwidth or read bandwidth per second. Alternatively, double-click the values above the slider to enter minimum and maximum values manually. Press <b>Enter</b> after typing a value. Click the check mark above a slider to apply the selection, or click the <b>X</b> to cancel.</li> </ul> <p>The number of applied filters will display next to the <b>Filter</b> icon. Select the small <b>X</b> icon to remove the applied filters.</p>
Column sort	Column heading	Click on a column heading to sort the table by the values in that column. Click a second time to reverse the direction of the sort.
Secondary Column Sort	Column heading	A secondary column can be added to the primary column sort. Click and hold the <b>Alt</b> key while clicking on the second column. To sort the second column in the reverse direction, repeat the procedure.

## 14.3 OST Workflow

Select the **OST I/O** information block from a tile on the **System Overview**, which opens the **Lustre File System Performance Dashboard**.

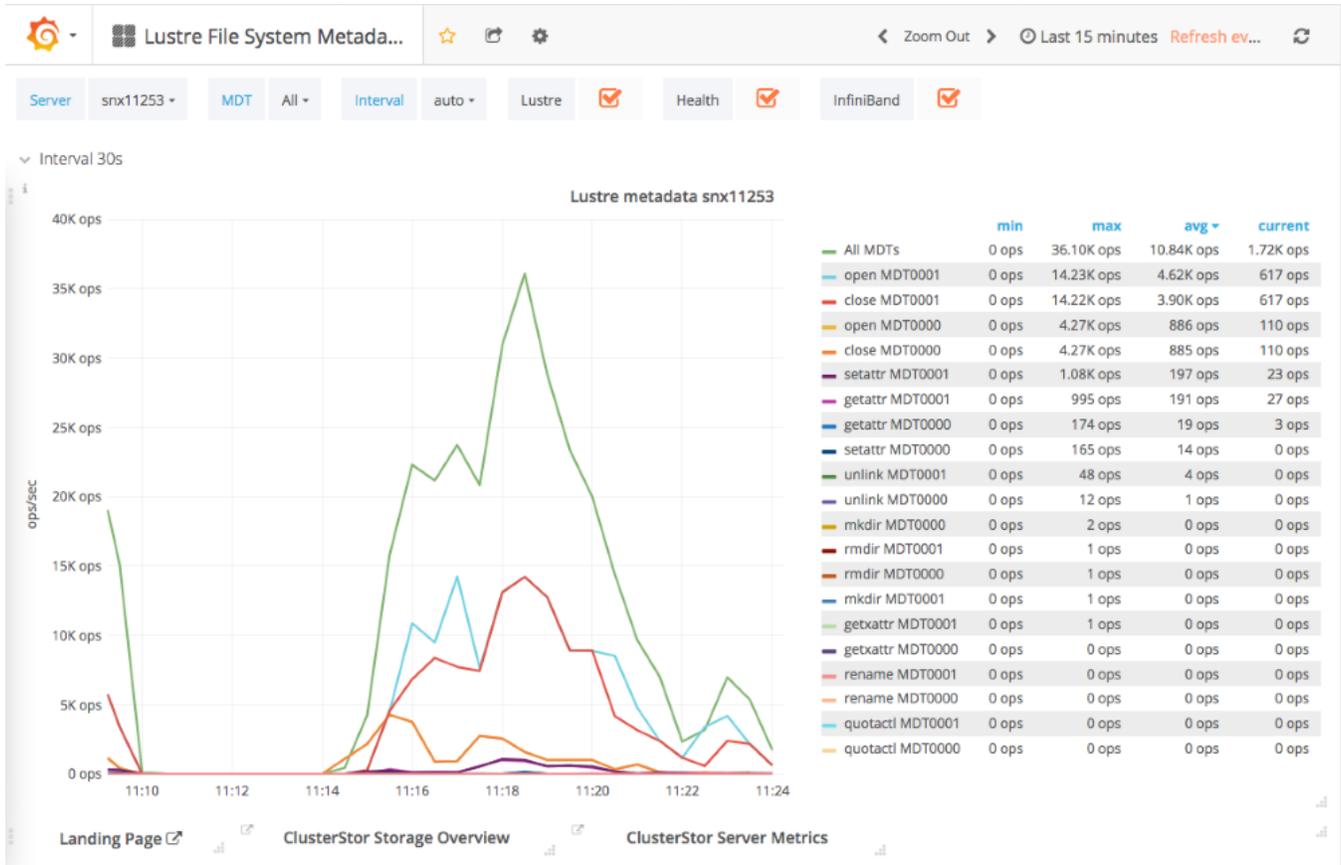


The per-OST write rates and the per-OST read rates are displayed in the dashboard's panels.

See [Lustre File System Performance Dashboard](#) on page 77 for additional information about the dashboard.

## 14.4 MDT Workflow

Select the **Metadata ops** information block from a tile on the **System Overview**, which opens the **Lustre File System Metadata Dashboard**.

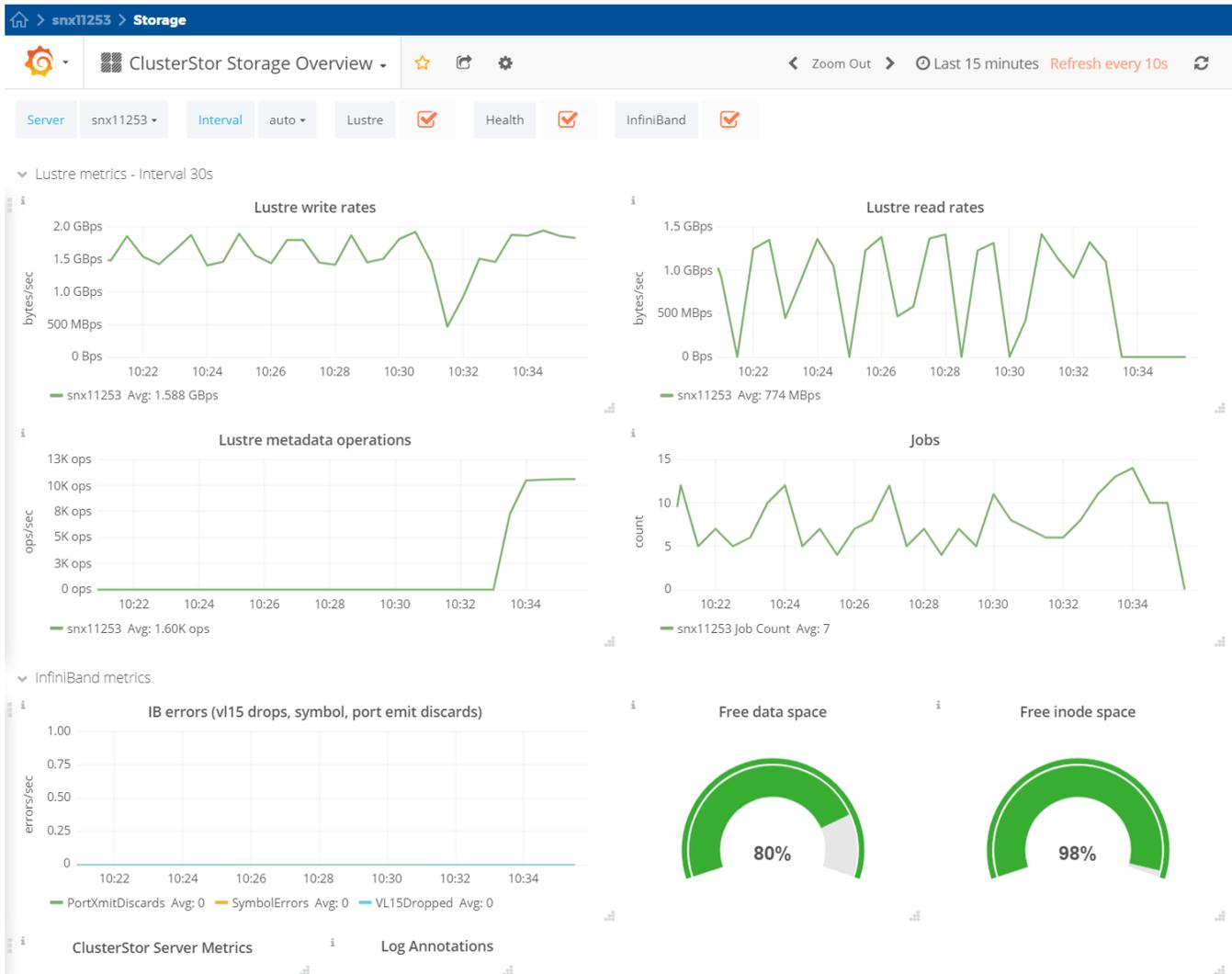


The per-MDT breakdowns of metadata operations are displayed in the dashboard's panels. All metadata operations are collected, but not all possible metadata operations are displayed. The operations listed are those most likely to be of interest.

See [Lustre File System Metadata Dashboard](#) on page 76 for additional information about the dashboard.

## 14.5 System Performance Workflow

Select the name of a monitored ClusterStor system from a tile on the **System Overview**, which opens the **ClusterStor Storage Overview Dashboard**.



Key information about how the overall system is performing is displayed in the dashboard's panels.

See [ClusterStor Storage Overview Dashboard](#) on page 62 for additional information about the dashboard.

## 14.6 InfiniBand Workflow

Select the name of a monitored ClusterStor system from a tile on the **System Overview**, which opens the **ClusterStor Storage Overview Dashboard**. This dashboard displays a summary chart of IB Errors.

See [ClusterStor Storage Overview Dashboard](#) on page 62 for additional information about the dashboard.

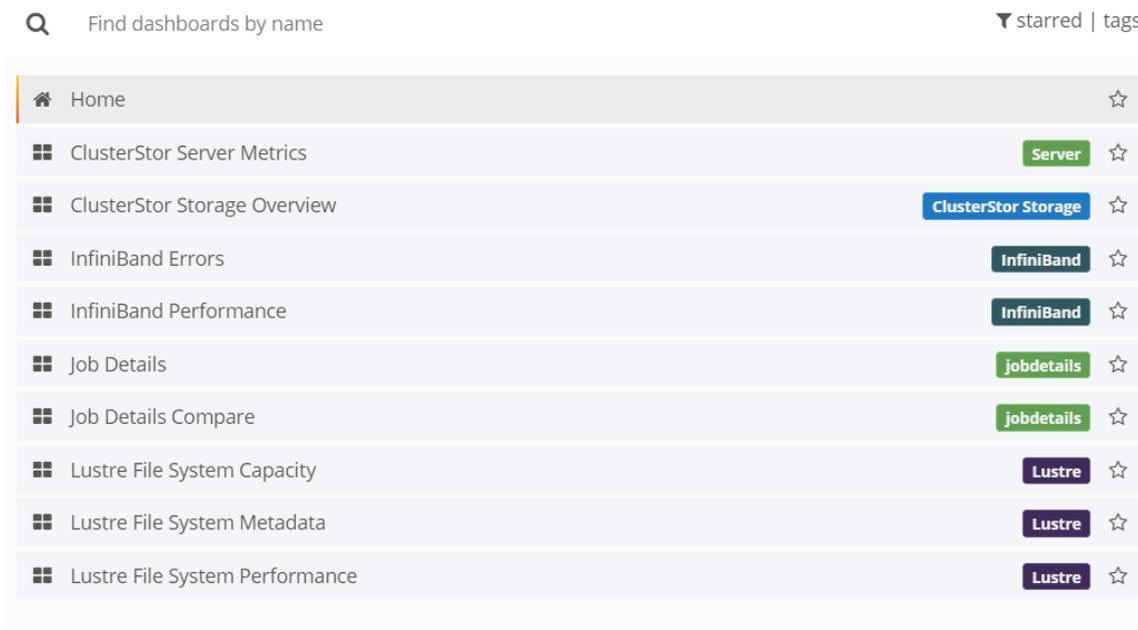
### InfiniBand Metrics

In addition to the InfiniBand error counters displayed in this dashboard, View for ClusterStor provides these additional InfiniBand related dashboards:

- **InfiniBand Errors Dashboard**

- **InfiniBand Performance Dashboard**

Navigate to these dashboards via the Grafana **Dashboard Selector**:

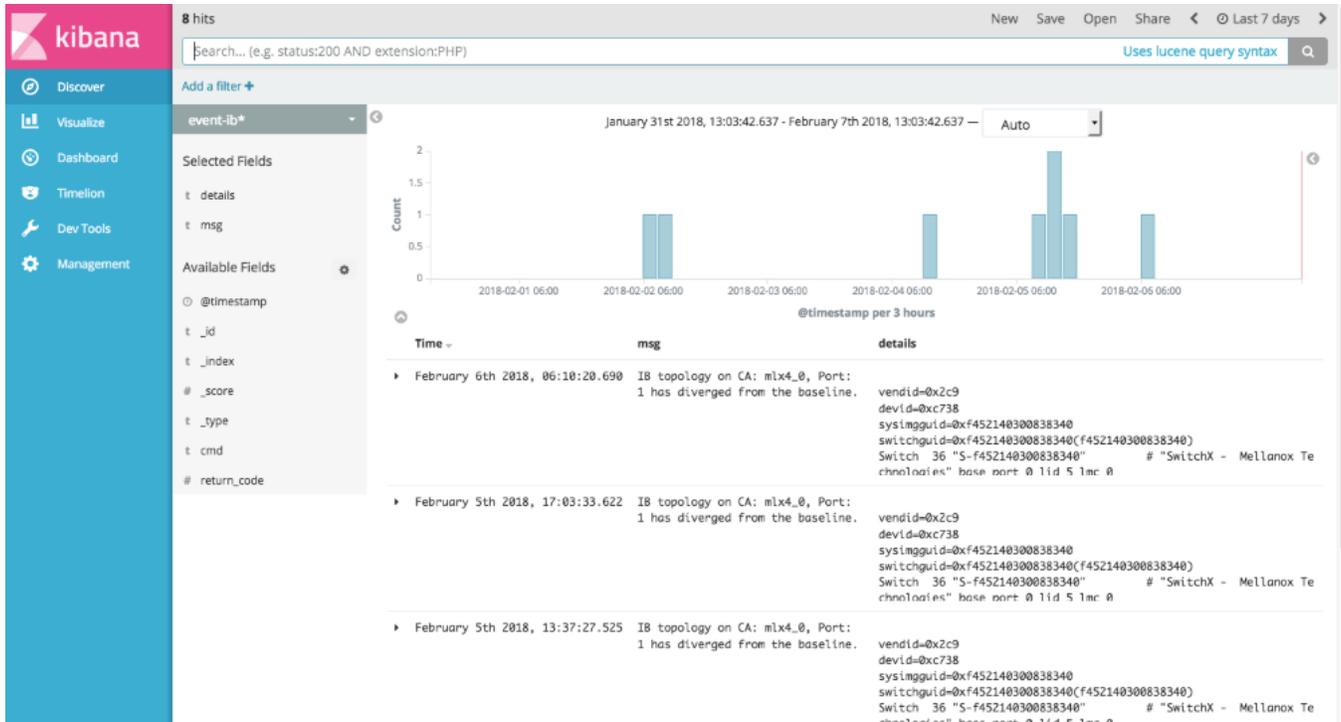


See [InfiniBand Errors Dashboard](#) on page 63 and [InfiniBand Performance Dashboard](#) on page 67 for additional information about these dashboards.

## 14.6.1 InfiniBand Logs

Logs are collected regarding InfiniBand topology changes. Individual topology changes are reflected in each Grafana dashboard. Topology changes are also recorded in a database that can be viewed in Kibana.

Select the **IB** icon in the header area of the **System Overview**, which opens the **IB Page** in Kibana. The index used for the **IB Page** is `event-ib*`.

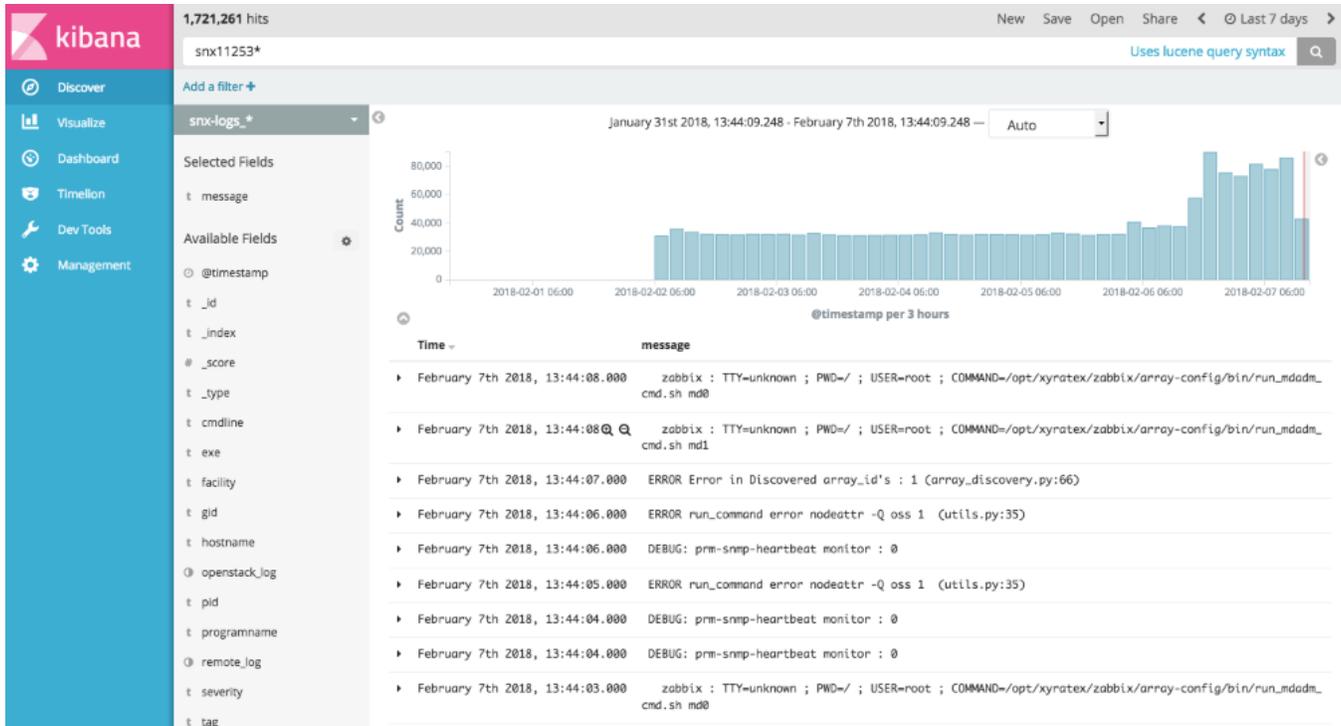


## 14.7 ClusterStor Logs

Logs are collected from the ClusterStor™ system and recorded in a database that can be viewed in Kibana.

Select the **Logs** information block from a tile on the **System Overview**, which opens the **Logs Page** in Kibana. The index used for the **Logs Page** is `snx-logs_*`.

Stored ClusterStor logs are displayed in the page. Logs can also be opened by selecting the **Details** icon located to the left of **Log Annotations** at the bottom of any dashboard.



## 14.8 Health Events

Health events are collected from the ClusterStor™ system and recorded in a database that can be viewed in Kibana.

Select the **Health** information block from a tile on the **System Overview**, which opens the **Health Page** in Kibana. The index used for the **Health Page** is `event-*`.

Browser address bar: [https://carlbou02/kibana/app/kibana#/discover?\\_g=\(refreshInterval:\(display:Off,pause:if,value:0\),time:\(from:now-7d,mode:quick,to:now\)\)&\\_a=\(columns:\(message,hostname,severity\),index:'event-\\*'\)](https://carlbou02/kibana/app/kibana#/discover?_g=(refreshInterval:(display:Off,pause:if,value:0),time:(from:now-7d,mode:quick,to:now))&_a=(columns:(message,hostname,severity),index:'event-*'))

Search:  [Uses lucene query syntax](#)

9 hits

**Discover**

Visualize

Dashboard

Timeline

Dev Tools

Management

event-\*

Selected Fields

- t hostname
- t message
- t severity

Available Fields

- @timestamp
- t \_id
- t \_index
- # \_score
- t \_type
- t component
- # metacount
- t msgformat
- t system\_name
- t tag
- # value
- t xref

Time	message	hostname	severity
July 31st 2018, 02:31:32.017	WARNING - There are currently 2 service alerts in progress. Please check your ClusterStor system. System: Serial Number: CSSF11103 Identifier: snx11103-00	snx11103-000	warning
July 30th 2018, 07:21:16.045	WARNING - There are currently 2 service alerts in progress. Please check your ClusterStor system. System: Serial Number: CSSF11103 Identifier: snx11103-00	snx11103-000	warning
July 30th 2018, 02:31:36.311	WARNING - There are currently 2 service alerts in progress. Please check your ClusterStor system. System: Serial Number: CSSF11103 Identifier: snx11103-00	snx11103-000	warning
July 29th 2018, 02:31:22.515	WARNING - There are currently 2 service alerts in progress. Please check your ClusterStor system. System: Serial Number: CSSF11103 Identifier: snx11103-00	snx11103-000	warning
July 28th 2018, 02:31:37.525	WARNING - There are currently 2 service alerts in progress. Please check your ClusterStor system. System: Serial Number: CSSF11103 Identifier: snx11103-00	snx11103-000	warning

## 15 Update Retention Policies for View for ClusterStor

### About this task

Persistent databases must be configured to the duration that metrics are preserved in the database. These retention policies are highly dependent on a site's job workload and the number of OSTs over which the workload is distributed.

Each site will need to set the appropriate length of the View for ClusterStor™ retention policies. The default retention policy period is 14 days. For the first 14 days, the data will be retained. On the fifteenth day, the first day's data will be pruned. On the sixteenth day, the second day's data will be pruned, and so on.

There are three retention policies, for which the retention policy period must be set:

Retention Policy Variable	Description
RETENTION_POLICY_DURATION	Handle InfluxDB and job events
LOG_RETENTION_DURATION	Defined to restrict CStream logs in <code>/etc/sma-data/seastream</code>
ELASTICSEARCH_DURATION	Defined to restrict Elasticsearch logs and events

Retention policies can be modified at any time. The retention policy duration is expressed in days.

### Procedure

1. Log in to the View for ClusterStor server as `root`.
2. Change to the View for ClusterStor configuration directory.

```
hostname# cd /etc/sma-data/etc
```

3. Open the `.env` configuration file in an editor.

```
hostname# vi .env
```

4. Change the retention policy duration variables, as desired, to a new value in number of days. Note that days are indicated with a suffix of `d`.

5. Save the modified `.env` configuration file.
6. Log out of any View for ClusterStor GUI sessions that are open in a browser.
7. Restart the View for ClusterStor service.

```
hostname# systemctl restart sma
```

- 8.** Restart the View for ClusterStor GUI, if desired, after the service has restarted:
  - a. Reload the browser tab that was previously used to work with the GUI, or open a new browser tab and navigate to the GUI URL.
  - b. Log in to the new GUI session.

## 16 Collect Current Infrastructure State

---

### About this task

If unexpected behavior is experienced with View for ClusterStor™, it is important to collect the current state of the View for ClusterStor infrastructure.

Beginning with release 1.2.0, System Snapshot Analyzer is now available for use with View for ClusterStor. Use this tool to securely collect, analyze, and upload (if upload is enabled) product health, configuration, or triage information about a View for ClusterStor system to Cray service.

For more information, please refer to *View for ClusterStor System Snapshot Analyzer (SSA) User Guide (1.7.1) S-3028*.