View for ClusterStor™ Installation and Configuration Guide

(1.2.0)

S-3025

Contents

1 About View for ClusterStor™ Installation and Configuration 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 (IB) fabric if connected to the ClusterStor high speed InfiniBand network.

View for ClusterStor can be integrated with the Cray System Management Workstation (SMW), and will collect job information such as start/stop, job id, ap id, user id, and duration for jobs launched on attached Cray computers. Administrators can view this information as it occurs or look at information collected at different points in the past through data dashboards and workflows.

The *View for ClusterStor Installation and Configuration Guide (S-3025)* covers all of the necessary procedures for preparing an on-site server for initial View for ClusterStor use. This guide assumes the reader is familiar with ClusterStor storage systems.

Release Information

The *View for ClusterStor Installation and Configuration Guide* supports View for ClusterStor version 1.2.0.

Product Requirements

Supported Versions

The following versions are supported by View for ClusterStor:

- CentOS 7.2
- Software Docker CE 17.06.2
- Docker Compose 1.14.0
- Slurm 17.11.7 or greater

The ClusterStor Storage System must be running the following software releases or greater:

- ClusterStor 2.0 SU26
- ClusterStor 3.0 SU10
  - For best results at scale, ClusterStor 3.0 SU11 or greater is suggested
- ClusterStor 3.1-010

All Lustre Clients must be running the following release level and patch level or greater:

- CLE 5.2UP04 patch number PS281
- CLE 6.0UP02 patch number PS44
- CLE 6.0UP03 patch number PS15

The following ports are used by View for ClusterStor:

- 80
- 441
- 9092
- 2128
- 8514

Hardware

- Standalone server with:
  - 128GB of memory or more
  - 8 CPU cores or more
  - SSD storage of 500GB or more (for View for ClusterStor data)

Time Synchronization

Time and time zone must be synchronized across any and all servers, especially:

- System Management Workstation (SMW)
- Storage server
- View for ClusterStor server
- Cray compute system

Other Considerations

To ensure accurate tracking of IB statistics, it is required that no more than one View for ClusterStor IB port be attached to the same IB fabric.

Record of Revision

Revision	Date	Content Information
View for ClusterStor Installation and Configuration Guide (S-3025) 1.2.0	12/7/2018	Release 1.2.0
View for ClusterStor Installation and Configuration Guide (S-3025) 1.1.0	08/13/2018	Release 1.1.0
View for ClusterStor Installation and Configuration Guide (S-3025) 1.0.1	05/16/2018	Release 1.0.1
View for ClusterStor Installation and Configuration Guide (S-3025) Rev A	03/20/2018	Revision A
View for ClusterStor Installation and Configuration Guide (S-3025)	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.

Proportional Bold 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.

`smaller font size` 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.

smw#	Run the command on the SMW as root.
sdb#	Run the command on the SDB node as root.
boot#	Run the command on the boot node as root.
login#	Run the command on any login node as root.
hostname#	Run the command on the View for ClusterStor system as root.
user@*hostname*>	Run the command on the specified system as any non-root user.
[*node*]$	Run the command on the specified ClusterStor node as the admin user.
[*node*]#	Run the command on the specified ClusterStor node as root.

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 sysctl Required Configuration for View for ClusterStor

View for ClusterStor™ requires some sysctl settings in the base host to operate correctly. These settings are in addition to the system requirements to run and operate Docker and to have both docker and docker-compose commands installed on the system.

Recommended Values

The following sysctl values need to be set in the base host:

- net.ipv4.tcp_keepalive_time
  - Recommended Value: 600
  - Maximum Value: 900
  - This is the time in seconds between keepalive packets on TCP IP version 4. The current system default is 7200 (2 hours), but Docker networks will disconnect with a period of inactivity of 15 minutes.
- vm.max_map_count
  - Recommended Value: 262144 (This value is recommended as minimum for Elasticsearch.)
  - System default: 65536

3.1 Change sysctl Settings

The sysctl settings can be changed in one of two ways. Using a sysctl command on the base host will have an immediate effect but will only be effective until the next reboot. For permanent changes, place the appropriate values in sysctl configuration files. Changing the configuration files requires a reboot to take effect.

Change sysctl Settings with sysctl Commands

Make these settings by using a sysctl command on the base host. Note that any settings made this way will only be effective until the next reboot of the base host operating system.

```
hostname# sysctl -w net.ipv4.tcp_keepalive_time=600
hostname# sysctl -w vm.max_map_count=262144
```

There are no spaces either before or after the equals signs (=).

Change sysctl Settings in sysctl Configuration Files

To make the sysctl settings permanent, place the values in either the /etc/sysctl.conf file or in separate files in the /etc/sysctl.d/ directory. Add the following lines:

```
net.ipv4.tcp_keepalive_time = 600
vm.max_map_count = 262144
```

Please note that spaces are required before and after the equals signs (=).

These settings are only processed during system start-up and will take effect when the base host reboots. Either reboot the system manually after updating the files or, to apply the new settings without rebooting use the sysctl -w command.

4 Install View for ClusterStor Software

Prerequisites

- All site-specific configuration values have been collected.
- A directory has been created on the SSD and, if applicable, on a separate partition from the root partition to store all View for ClusterStor™ data.
- CentOS has been installed.
- The InfiniBand (IB) interface has been configured and is up and running.
- Firewalld must be disabled before starting Docker, because it is not currently supported for View for ClusterStor. Alternatively, the `iptables` utility can be used to manage the firewall directly.
- Docker CE 17.06.2 has been installed from *https://docs.docker.com/install/linux/docker-ce/centos/*.
- Docker Compose 1.14.0 has been installed via `curl`:

  ```
  hostname# curl -L "https://github.com/docker/compose/releases/download/1.14.0/docker-compose-$(uname
  -s)-$(uname -m)" -o /usr/local/bin/docker-compose
   hostname# chmod +x /usr/local/bin/docker-compose
   hostname# ls -l /usr/local/bin/docker-compose
  -rwxr-xr-x. 1 root root 8278112 Nov 8 14:03 /usr/local/bin/docker-compose
  ```

- `sysctl` settings have been changed (see *sysctl Required Configuration for View for ClusterStor* on page 8).
- The OpenSSL version on the job-event daemon's host supports TLS v1.2.

  Run the following commands on the host to verify support for TLS v1.2:

  1. As `root`, log into the machine where the job daemon will run (e.g., SMW).
  2. Determine the OpenSSL version:

     ```
     smw# openssl version
     OpenSSL 1.0.2o 27 Mar 2018
     ```

  3. Determine the TLS version:

     ```
     smw# openssl ciphers -v | awk '{print $2}' | sort -u
     SSLv3
     TLSv1.2
     ```

  If support for TLS v1.2 is not indicated, upgrade the host to a more recent OpenSSL version. If that is not possible, contact Cray Customer Service for further instructions.

- The View for ClusterStor software package has been downloaded from *https://crayport.cray.com*.

Procedure

1. Log in to the View for ClusterStor server as `root`.

2. Copy the View for ClusterStor software archive to `/root`.

   The archive name will be a variation of *sma*`-1.2.0-`*<time-date-stamp>.tgz*.

3. Unpack the archive using the `tar` command:

   ```
   hostname# tar xvf sma-1.2.0-<time-date-stamp>.tgz
   ```

4. Change to the newly created `/root/sma-install` directory.

   ```
   hostname# cd /root/sma-install
   ```

5. Run the installation script:

   ```
   hostname# ./setup.sh
   ```

   If a default value is defined when prompted for site-specific information, press **Enter** to use the default value.

   Installation should take no more than forty minutes.

   a. Fill in the data directory path when prompted.

      This is the directory on the SSD that will store all View for ClusterStor data and configuration files.

      ```
      Enter path for data directory (default= /var/sma/data):/site/data/directory
      ```

   b. Enter the FQDN hostname of the View for ClusterStor system.

      ```
      Enter FQDN hostname (default= <detected hostname>):
      hostname.customer.site.com
      ```

   c. Enter the site email relay host.

      ```
      Enter site email relay host: mail-relay.customer.site.com
      ```

      The system will ask if you want to change any of these answers.

      ```
      Do you need to change any of the previous answers? (possible choices=[yes
      no]):
      ```

   d. Type `yes` and press **Enter**, if necessary, to go back and change any settings. Otherwise type `no` and press **Enter**.

      ```
      Do you need to change any of the previous answers? (possible choices=[yes
      no]):no
      ```

6. Verify that the View for ClusterStor service is running.

   ```
   hostname# systemctl status sma
   ```

   Sample output showing the service running:

   ```
   sma.service - Manage Docker containers
      Loaded: loaded (/etc/systemd/system/sma.service; enabled; vendor preset: disabled)
      Active: active (running) since Mon 2018-10-15 02:28:44 CDT; 4h 38min ago
     Process: 89987 ExecStop=/usr/bin/bash -c    HOSTNAME=$$(/usr/bin/hostname);    IMAGE_ID=$$(docker image ls
   cray_sma/elasticsearch --format {{.ID}});
      /usr/bin/docker run --rm --network sma_default --entrypoint /post.sh $IMAGE_ID -h $HOSTNAME -c "systemctl stop
   sma" -t "SMA service stop" -m "SMA service stop request issued";
      /usr/local/bin/docker-compose -f docker-compose.yml down   (code=exited, status=0/SUCCESS)
     Process: 92161 ExecStartPost=/usr/bin/chown :daemon /var/run/docker.sock (code=exited, status=0/SUCCESS)
    Main PID: 92160 (bash)
      Memory: 51.5M
   ```

```
   CGroup: /system.slice/sma.service
           ├─92160 /usr/bin/bash -c    MYDATE=$(/usr/bin/date -u +%Y-%m-%dT%H:%M:%S.%N);    IMAGE_ID=$(docker
image l...
           ├─92858 /usr/local/bin/docker-compose -f docker-compose.yml up
           └─92860 /usr/local/bin/docker-compose -f docker-compose.yml up

Oct 15 02:28:44 caribou12 systemd[1]: Starting Manage Docker containers...
Oct 15 02:28:44 caribou12 systemd[1]: Started Manage Docker containers.
```

View for ClusterStor has now been installed.

5 Configure ClusterStor System for Metrics

About this task

This procedure is performed on any ClusterStor™ system that is to be monitored by View for ClusterStor. The procedure can be performed on a live ClusterStor system, i.e., the Lustre file system does not need to be taken offline.

When complete, the ClusterStor system will be ready to stream metrics data to the View server.

> **IMPORTANT:** If a System Update (SU) is subsequently installed on the ClusterStor system, the configuration will be cleared and this procedure will need to be performed again.

Procedure

1. Log in to the target ClusterStor system's active management node (n000 or n001) as `admin`.

2. Create a new user with read-only access.

   There will be a prompt to create and confirm a password for the new user.

```
[MGMT0]$ cscli admins add --username=smauser --role=readonly --disable-ssh --
enable-web

Enter the password : password
Confirm the password : password

[MGMT0]$ cscli admins list
--------------------------------------------------------------
 Username  Role      Uid   SSH Enabled  Web Enabled  Policy
--------------------------------------------------------------
 smauser   readonly  1016     False         True      default
--------------------------------------------------------------
```

3. Enable the REST API:

```
[MGMT0]$ cscli service_console configure rest_api enable
```

4. Add the user created in step 2 on page 13 as an authorized REST API user.

```
[MGMT0]$ cscli service_console configure rest_api user_add --username smauser
User 'smauser' has been added to REST API authorized users list
```

5. Confirm the REST API configuration.

```
[MGMT0]$ cscli service_console configure rest_api show
REST API access: enabled

REST API authorized users:
  smauser
```

6. Check if the ClusterStor system identifier is set. In this example, `cls12345n000` is used, along with its corresponding serial number.

```
[MGMT0]$ cscli service_console configure system show
System settings:
        System serial number: CSSX0G4DE5
        System identifier name: [not-set]
```

   a. Assign the system identifier if it is not set.

   ```
   [MGMT0]$ cscli service_console configure system identifier -n cls12345n000
   ```

   b. Confirm the system identifier setting.

   ```
   [MGMT0]$ cscli service_console configure system show
   System settings:
           System serial number: CSSX0G4DE5
           System identifier name: cls12345n000
   ```

The ClusterStor system has been configured for streaming metrics.

6 Configure ClusterStor System for Log Forwarding

About this task

Use this procedure to configure a ClusterStor™ system to forward logs to View for ClusterStor.

This procedure is performed on any ClusterStor system that is to be monitored by View for ClusterStor. The procedure can be performed on a live ClusterStor system, i.e., the Lustre file system does not need to be taken offline.

When complete, the ClusterStor system will be ready to forward designated system logs to the View server.

This procedure requires specifying:

- The destination to which the ClusterStor system will send logs
- Which logs to send

Follow the procedure below that is required for the software version running on the ClusterStor system.

> **IMPORTANT:** If a System Update (SU) is subsequently installed on the ClusterStor system, verify that the log forwarding configuration has not been changed. If it has, perform this procedure again.

6.1 Log Forwarding Configuration for ClusterStor 3.1 and Above

About this task

This task describes how to configure ClusterStor™ systems running software release 3.1 and above for log forwarding.

This procedure requires specifying:

- The destination to which the ClusterStor system will send logs (i.e., View for ClusterStor server)
- Which logs to send

The example that follows uses the following values:

- Destination (View server hostname) - *`SMA1234`*. URL form: *`SMA1234.sitename.gov`*
- View for ClusterStor system's IP address: *`172.30.76.13`*

Procedure

1. Log in to the target ClusterStor system's active management node (n000 or n001) as `admin`. Examples assume the primary MGMT node is active.

2. Register a new Syslog consumer.

```
[MGMT0]$ cscli syslog_consumer add --host SMA1234.sitename.gov --port 8514 --
proto udp --format bsd
syslog_consumer: Registering new consumer of Syslog.
syslog_consumer: consumer udp://SMA1234.sitename.gov:8514/bsd is now registered.
```

3. List the Syslog consumers to confirm.

```
[MGMT0]$ cscli syslog_consumer show
--------------------------------------- Consumer
---------------------------------------

   udp://172.30.76.13:8514/bsd

   udp://172.30.76.77:8514/bsd

   udp://SMA1234.sitename.gov:8514/bsd

   ---------------------------------------
```

The ClusterStor system has been configured to forward logs to View for ClusterStor.

6.2 Log Forwarding Configuration for ClusterStor 3.0 and Below

About this task

This task describes how to configure ClusterStor™ systems running software release 3.0 or below for log forwarding.

This procedure requires specifying:

- The destination to which the ClusterStor system will send logs (i.e., View for ClusterStor server)
- Which logs to send

The example that follows uses the following values:

- Destination (View server hostname) - *SMA1234*. URL form: *SMA1234.sitename.gov*
- View for ClusterStor system's IP address: *172.30.76.13*

Procedure

1. Log in to the target ClusterStor system's active management node (n000 and n001) as `admin`. Examples assume the primary MGMT node is active.

Repeat this step on MGMT1.

2. Change to the root user.

```
admin@MGMT0> sudo su -
password for admin: password
Last login: Thu Sep 29 12:47:14 CDT 2016 from 172.16.2.3 on ssh
```

Repeat this step on MGMT1.

3. Save a backup copy of the `syslog-ng_receiver.erb` config file to a new file name on both MGMT nodes.

```
MGMT0# cp /etc/puppet/modules/syslog_ng/templates/syslog-ng_receiver.erb \
/etc/puppet/modules/syslog_ng/templates/syslog-ng_receiver.erb.save
MGMT1# cp /etc/puppet/modules/syslog_ng/templates/syslog-ng_receiver.erb \
/etc/puppet/modules/syslog_ng/templates/syslog-ng_receiver.erb.save
```

4. Add lines to the `/etc/puppet/modules/syslog_ng/templates/syslog-ng_receiver.erb` file on both nodes.

```
MGMT0# vi /etc/puppet/modules/syslog_ng/templates/syslog-ng_receiver.erb
```

   a. Add the following line ***BEFORE*** any existing `destination` lines:

```
destination SMA1234 { udp("172.30.76.13" port(8514) time-zone("+00:00")); };
```

   b. Add the following line ***BEFORE*** any existing `log` lines:

```
log { source(s_sys); source(s_udp); destination(SMA1234); };
```

Repeat this step on MGMT1.

5. Run puppet to update the official files and integrate changes into the ClusterStor cluster on the active management node. The changes will appear in `/etc/syslog-ng/syslog-ng.conf`.

```
MGMT0# puppet agent -tv
```

6. Verify that the log forwarding changes have been successfully propagated on the active management node.

```
MGMT0# grep SMA1234 /etc/syslog-ng/syslog-ng.conf
```

The passive MGMT node will not update the `syslog-ng.conf` file until it is made the active node.

The ClusterStor system has been configured to forward logs to View for ClusterStor.

7 Configure Lustre Job Statistics

Prerequisites

The ClusterStor™ system has been configured for Metrics.

About this task

The default Lustre jobstats code on the client extracts the unique JobID from an environment variable within the user process, and sends this JobID to the server with the I/O operation. This environment variable lookup on the client causes Lustre I/O performance degradations when jobstats are enabled. A Cray® Lustre client patch is required to work around this performance issue and is available for CLE 5.2 versions UP04 and above, as well as CLE 6.0 versions UP02-UP04. With this patch, enabling Lustre jobstats will be done with a WLM prologue script when the job is launched.

The following procedure assumes that a prolog script and epilog script do not currently exist. Since ALPS (Application Level Placement Scheduler) only supports one prolog/epilog script, additional scripts, such as RUR scripts, need to be combined with the Lustre jobstats prolog/epilog scripts. For example, run the RUR prolog and epilog scripts inside the Lustre jobstats scripts.

The procedure to configure a ClusterStor system for Lustre job statistics depends on the CLE version being run. Refer to the table below to find the required procedure.

CLE Version	Required Procedure
CLE 5.2 UP04 and above	Perform the procedure in *Enable Lustre Job Statistics for CLE 5.2* on page 18 before proceeding to *Enable Jobstats on the ClusterStor System* on page 21.
CLE 6.0 UP02-UP04	Perform the procedure in *Enable Lustre Job Statistics for CLE 6.0* on page 20 before proceeding to *Enable Jobstats on the ClusterStor System* on page 21.
CLE 6.0 UP05 and above	Proceed directly to *Enable Jobstats on the ClusterStor System* on page 21.

7.1 Enable Lustre Job Statistics for CLE 5.2

Procedure

1. Edit the `alps.conf` file to set the prologue/epilogue paths. The file is located on the shared-root file system on the boot node.

```
boot# xtopview -m "add prolog/epilog scripts to alps.conf for Lustre nodelocal jobstats"
boot# vi /etc/opt/cray/alps/alps.conf

prologPath      /ufs/alps_shared/lustre_jobstats_prolog.sh
epilogPath      /ufs/alps_shared/lustre_jobstats_epilog.sh
```

2. Create and install the `lustre_jobstats_prolog.sh` script on one of the login nodes to set "nodelocal" mode before the application starts.

   ```
   login# vi /ufs/alps_shared/lustre_jobstats_prolog.sh

   #!/bin/bash
   export CRAY_ROOTFS=INITRAMFS
   /opt/cray/nodehealth/default/bin/pcmd -f $ALPS_PREP_NIDFILE "/sbin/lctl
   set_param jobid_var=nodelocal"
   /opt/cray/nodehealth/default/bin/pcmd -f $ALPS_PREP_NIDFILE "/sbin/lctl
   set_param jobid_name=$ALPS_PREP_APID"
   ```

3. Create and install the `lustre_jobstats_epilog.sh` script on one of the login nodes to disable Lustre job statistics when the application exits.

   ```
   login# vi /ufs/alps_shared/lustre_jobstats_epilog.sh

   #!/bin/bash
   export CRAY_ROOTFS=INITRAMFS
   /opt/cray/nodehealth/default/bin/pcmd -f $ALPS_PREP_NIDFILE "/sbin/lctl
   set_param jobid_var=disable"
   ```

4. Set the correct permissions on the prolog and epilog scripts on one of the login nodes.

   ```
   login# chmod 700 /ufs/alps_shared/lustre_jobstats_prolog.sh \
   /ufs/alps_shared/lustre_jobstats_epilog.sh
   ```

5. Send a SIGHUP signal to apsys to reread the updated apsys configuration file.

   This must be performed on each login node.

   ```
   login# pkill -HUP apsys
   ```

To stop collecting jobstats at a later date, services must be disabled, as described in *Disable Lustre Jobstats for CLE 5.2* on page 19

7.1.1 Disable Lustre Jobstats for CLE 5.2

About this task

To stop collecting jobstats, services must be disabled using this procedure.

Procedure

1. Remove the prologue/epilogue paths from the ALPS configuration or disable the lctl `set_param` commands in the prolog/epilog scripts on the login node.

2. Send a SIGHUP signal to apsys to reread the update apsys configuration file.

This must be performed on all of the login nodes.

```
login# pkill -HUP apsys
```

7.2 Enable Lustre Job Statistics for CLE 6.0

Procedure

1. Invoke the configurator in interactive mode on the SMW for the `cray_alps` service to set the prologue/epilogue paths.

   Replace ***p0.staging*** with the site-specific config set name.

   ```
   smw# cfgset update -m interactive -s cray_alps p0.staging

   cray_alps.settings.apsys.data.prologPath: /home/crayadm/bin/lustre_jobstats_prolog.sh
   cray_alps.settings.apsys.data.epilogPath: /home/crayadm/bin/lustre_jobstats_epilog.sh
   ```

2. Run the configurator to confirm the settings.

   ```
   smw# cfgset search -s cray_alps --level basic p0.staging

   # 4 matches for '.' from cray_alps_config.yaml
   #-------------------------------------------------------------------------------
   -
   cray_alps.settings.common.data.xthostname: crayxc
   cray_alps.settings.common.data.alps_node_groups: [ ] # (empty)
   cray_alps.settings.apsys.data.prologPath: /home/crayadm/bin/
   lustre_jobstats_prolog.sh
   cray_alps.settings.apsys.data.epilogPath: /home/crayadm/bin/
   lustre_jobstats_epilog.sh
   ```

3. Update the alps configuration set on the boot, sdb, and all login nodes.

   ```
   boot# /etc/init.d/cray-ansible start
   sdb# /etc/init.d/cray-ansible start
   login# /etc/init.d/cray-ansible start
   ```

4. Create and install the `lustre_jobstats_prolog.sh` script on one of the login nodes to set "nodelocal" mode before the application starts.

   ```
   login# vi /home/crayadm/bin/lustre_jobstats_prolog.sh

   #!/bin/bash
   export CRAY_ROOTFS=INITRAMFS
   /opt/cray/nodehealth/default/bin/pcmd -f $ALPS_PREP_NIDFILE "/sbin/lctl
   set_param jobid_var=nodelocal"
   /opt/cray/nodehealth/default/bin/pcmd -f $ALPS_PREP_NIDFILE "/sbin/lctl
   set_param jobid_name=$ALPS_PREP_APID"
   ```

5. Create and install the `lustre_jobstats_epilog.sh` script on one of the login nodes to disable Lustre job statistics when the application exits.

```
login# vi /home/crayadm/bin/lustre_jobstats_epilog.sh

#!/bin/bash
export CRAY_ROOTFS=INITRAMFS
/opt/cray/nodehealth/default/bin/pcmd -f $ALPS_PREP_NIDFILE "/sbin/lctl
set_param jobid_var=disable"
```

6. Set correct permissions on the prolog and epilog scripts on one of the login nodes.

   ```
   login# chmod 700 /home/crayadm/bin/lustre_jobstats_prolog.sh \
   /home/crayadm/bin/lustre_jobstats_epilog.sh
   ```

7. Restart the apsys daemon to reread the updated apsys configuration file.

   This must be performed on all login nodes.

   ```
   login# systemctl restart apsys
   ```

To stop collecting jobstats at a later date, services must be disabled, as described in *Disable Lustre Jobstats for CLE 6.0* on page 21

7.2.1 Disable Lustre Jobstats for CLE 6.0

About this task

To stop collecting jobstats, services must be disabled using this procedure.

Procedure

1. Remove the prologue/epilogue paths from the ALPS configuration or disable the `lctl set_param` commands in the prolog/epilog scripts on the login node.

2. Restart the apsys daemon to reread the updated apsys configuration file.

   This must be performed on all login nodes.

   ```
   login# systemctl restart apsys
   ```

7.3 Enable Jobstats on the ClusterStor System

Prerequisites

- Users running CLE 5.2 versions UP04 and above have completed the procedure in *Enable Lustre Job Statistics for CLE 5.2* on page 18.
- Users running CLE 6.0 versions UP02-UP04 have completed the procedure in *Enable Lustre Job Statistics for CLE 6.0* on page 20.

- Determine what workload manager is being used for job statistics on the host compute system that uses the ClusterStor™ system for storage. For more information, see *Configure System Workload Managers for Job Events* on page 27.

  ATTENTION: Enabling job statistics collection on a ClusterStor system may cause Lustre I/O performance degradation because of the way in which the default jobstats code on Lustre clients works. See *Configure Lustre Job Statistics* on page 18 for more information about this issue and a workaround patch for CLE 5.2 versions UP04 and above, as well as CLE 6.0 versions UP02-UP04. Note that no guidance is available for non-Cray Lustre clients and configurations.

About this task

The ClusterStor system must be configured with job statistics collection enabled before it will collect job and application start/stop events that are forwarded from the workload manager that is running on the host compute system. In addition, the ClusterStor system must be aware of the type of workload manager being used to send the job statistics. This is done by specifying the appropriate scheduler ID with the `cscli lustre jobstats modify` command.

Table 1. Scheduler ID Arguments for Supported Workload Managers

Workload Manager	Host Compute System	Scheduler ID Argument
ALPS/SMW	Cray® XC only	`ALPS_APP_ID`
Slurm	Any except Cray XC	`SLURM_JOB_ID`
PBS Pro	Any except Cray XC	`PBS_JOBID`

Use the steps in the following procedure to configure the ClusterStor system.

IMPORTANT: If a System Update (SU) is subsequently installed on the ClusterStor system, verify that the jobstats configuration has not been changed. If it has, perform this procedure again.

Procedure

1. Enable Lustre job statistics collection by running these commands on the active ClusterStor management node (n000 or n001). Commands shown assume MGMT0 is the active management node.

   Note that the supported value for the collection frequency is 30 seconds. Use the appropriate scheduler ID argument from *Scheduler ID Arguments for Supported Workload Managers* on page 22 in place of the *`scheduler_id`* variable in the commands below.

   ```
   MGMT0$ cscli lustre jobstats collection --enable
   lustre: Enabling Lustre Job Statistics for cls12345
   lustre: Updating puppet configuration. This can take a while...
   lustre: Successfully enabled Lustre Job Statistics for cls12345.

   MGMT0$ cscli lustre jobstats modify --frequency 30 --scheduler scheduler_id
   lustre: Configuring Lustre Job Statistics for cls12345
   lustre: Updating configuration. This can take a while...
   lustre: Successfully configured Lustre Job Statistics for cls12345
   ```

2. Check the Lustre job statistics configuration to ensure it is enabled correctly. The following example shows the output when the scheduler ID was set for an XC host compute system running ALPS/SMW as the workload manager.

   ```
   MGMT0$ cscli lustre jobstats list
   -------------------------------------------------------
    FSName     Collection  Frequency(in sec)  Scheduler
   -------------------------------------------------------
    cls12345    Enabled    30                 ALPS_APP_ID
   -------------------------------------------------------
   ```

For information on disabling Lustre job statistics collection, see *Disable Jobstats on the ClusterStor System* on page 23

7.3.1 Disable Jobstats on the ClusterStor System

Prerequisites

- Users running CLE 5.2 versions UP04 and above have completed the procedure in *Disable Lustre Jobstats for CLE 5.2* on page 19.
- Users running CLE 6.0 versions UP02-UP04 have completed the procedure in *Disable Lustre Jobstats for CLE 6.0* on page 21.

Procedure

1. Disable Lustre job statistics collection by running these commands on the active ClusterStor™ management node (n000 or n001). Commands shown assume MGMT0 is the active management node.

   ```
   MGMT0$ cscli lustre jobstats collection --disable
   ```

2. Check the Lustre jobstats configuration to confirm that collection has been disabled. The following example shows the output for an XC host compute system running ALPS/SMW as the workload manager.

   ```
   MGMT0$ cscli lustre jobstats list
   -------------------------------------------------------
    FSName    Collection  Frequency(in sec)  Scheduler
   -------------------------------------------------------
    cls12345  Disabled    30                 ALPS_APP_ID
   -------------------------------------------------------
   ```

8 Configure ClusterStor System for SNMP

About this task

Enabling Simple Network Management Protocol (SNMP) on a ClusterStor system permits View for ClusterStor to obtain information about service alerts and determine what alarm state a system may be in at any given time. The View for ClusterStor graphical user interface will display these service alerts, and the SNMPwalk interface will create and send an email with more detailed information.

> **IMPORTANT:** If a System Update (SU) is subsequently installed on the ClusterStor system, verify that the SNMP configuration has not been changed. If it has, perform this procedure again.

Procedure

1. Log in to the target ClusterStor system as `admin`.

2. Enable SNMP by running the following command on the active management node (n000 or n001).

   The example assumes MGMT0 is the active management node.

   ```
   admin@MGMT0> cscli service_console configure snmp enable
   Attempting to enable SNMP, please wait
   .....
   SNMP has successfully been enabled.
   ```

SNMP is now enabled.

9 Configure Security Keys for Job Event Daemons

Job event daemons capture WLM job and application start/stop events and forward them to the View server. There are two implementations:

- Cray® XC systems: the job event daemon is installed on the System Management Workstation (SMW) of a Cray XC cluster.
- Non-XC systems: the job event daemon is used to monitor Slurm job events on systems other than Cray XC.

Both implementations use a Kafka REST API to forward job events to View for ClusterStor™. The Kafka REST API interface is secured. Only daemons that present valid secret keys are allowed to post job events to View for ClusterStor. The administrator must configure these daemons and their secret keys in View for ClusterStor.

In View's security parlance, job event daemons are called *consumers*. For a specific View for ClusterStor installation, the administrator must configure one consumer for each job event daemon that will forward job events to View for ClusterStor. Each consumer must have at least one secret key defined. Once the secret key is created, the administrator must install it into the appropriate location in the job event daemon's configuration file.

The remainder of this topic provides procedures to configure and work with consumers and secret keys, as well as remove job event daemons.

Configure Consumers and Their Secret Keys

Use `sma-kafka-cli` to configure consumers and their secrets. The following example illustrates the creation of a consumer named `pollux-smw-job-events` with the auto-generated secret key `JpAKDSc9bFTc3P7VZ2JdihmX1VZenRG5`. The consumer name can be any string.

```
hostname# cd /root/sma-install
hostname# ./sma-kafka-cli consumer list
No consumers exist
hostname# ./sma-kafka-cli consumer add pollux-smw-job-events
JpAKDSc9bFTc3P7VZ2JdihmX1VZenRG5
```

List Existing Consumers

```
hostname# ./sma-kafka-cli consumer list
pollux-smw-job-events
```

Add Additional Consumers

This example shows the addition of a consumer named `slurm-job-events` with an administrator supplied secret key of `OpenSesame`, rather than an auto-generated secret key.

```
hostname# ./sma-kafka-cli consumer add slurm-job-events OpenSesame
OpenSesame
hostname# ./sma-kafka-cli consumer list
pollux-smw-job-events
slurm-job-events
```

List a Consumer's Existing Secret Keys

```
hostname# ./sma-kafka-cli secret list pollux-smw-job-events
JpAKDSc9bFTc3P7VZ2JdihmX1VZenRG5    2018-06-19 15:46:29
```

The timestamp associated with each secret key shows the date/time it was created.

Add Additional Secret Keys for a Consumer

This example shows adding 2 new secret keys for the consumer named `pollux-smw-job-events`.

```
hostname# ./sma-kafka-cli secret add pollux-smw-job-events
y1I5IhdT6U9IUX1rMcXlzGPfzfo6oq6H
hostname# ./sma-kafka-cli secret add pollux-smw-job-events OpenSesame2
OpenSesame2
hostname# ./sma-kafka-cli secret list pollux-smw-job-events
JpAKDSc9bFTc3P7VZ2JdihmX1VZenRG5    2018-06-19 15:46:29
y1I5IhdT6U9IUX1rMcXlzGPfzfo6oq6H    2018-06-19 15:46:35
OpenSesame2                         2018-06-19 15:46:36
```

Remove Unused Secrets

```
hostname# ./sma-kafka-cli secret remove pollux-smw-job-events OpenSesame2
hostname# ./sma-kafka-cli secret list pollux-smw-job-events
JpAKDSc9bFTc3P7VZ2JdihmX1VZenRG5    2018-06-19 15:46:29
y1I5IhdT6U9IUX1rMcXlzGPfzfo6oq6H    2018-06-19 15:46:35
```

Remove an Existing Consumer (Job Event Daemon)

This example shows removing the consumers named `pollux-smw-job-events` and `slurm-job-events`.

```
hostname# ./sma-kafka-cli consumer remove pollux-smw-job-events
hostname# ./sma-kafka-cli consumer remove slurm-job-events
hostname# ./sma-kafka-cli consumer list
No consumers exist
```

10 Configure System Workload Managers for Job Events

View for ClusterStor™ collects job and application start/stop events that are forwarded from workload managers that are running on host compute systems that use ClusterStor systems for storage.

View for ClusterStor supports the following workload managers:

Table 2. Supported Hosts and Workload Managers

Workload Manager	Host Compute System	Minimum View for ClusterStor Release
ALPS/SMW	Cray® XC only	1.0.0
Slurm	Any except Cray XC	1.1.0
PBS Pro	Any except Cray XC	1.2.0

Workload managers must be configured first before they will forward job events to View for ClusterStor. This configuration involves:

- Installing the appropriate job event RPM for the workload manager, as noted in *Required Configuration Procedure for Specific Workload Managers* on page 27.
- Configuring job event options
- Other configuration and installation steps, such as installing startup scripts or enabling WLM scripts or hooks

Before proceeding:

1. Determine what host system is being used at the site.
2. Determine which workload manager is installed on the site's host system.
3. Confirm that the host system and workload manager combination is supported. See *Supported Hosts and Workload Managers* on page 27.
4. Ensure that the minimum version of View for ClusterStor is installed.
5. Determine the required configuration procedure from *Required Configuration Procedure for Specific Workload Managers* on page 27
6. Proceed to the required configuration procedure.

Table 3. Required Configuration Procedure for Specific Workload Managers

Workload Manager	Path and RPM File Name	Required Configuration Procedure

ALPS/SMW
(Cray XC systems only)
	`/root/sma-install/rpm/jobevent-rpm`
	`cray-jobevent_generic-<version>.rpm`
		Configure SMW for Job Events on page 28

Workload Manager	Path and RPM File Name	Required Configuration Procedure

Slurm
(non-XC systems)
	`/root/sma-install/rpm/wlm-jobevent-rpm`
	`sma-jobevents-<version>.rpm`
		Configure Slurm for Job Events on page 30

PBS Pro
(non-XC systems)
	`/root/sma-install/rpm/wlm-jobevent-rpm`
	`sma-jobevents-[`*`version`*`].rpm`
		Configure PBS Pro for Job Events on page 32

10.1 Configure SMW for Job Events

Prerequisites

Important: This procedure is to be performed only on Cray® XC systems. **DO NOT** use this procedure on non-XC systems.

About this task

This procedure describes the steps needed to configure a Cray XC system's System Management Workstation (SMW) to forward job events to View for ClusterStor. Once the job event RPM has been installed on the Cray XC system, the jobevent daemon will capture WLM job and application start/stop events on the SMW and forward them to the View server.

The following required steps can be used to install or upgrade the daemon. Configuration steps are dependent on the specific version of software installed on the SMW.

Table 4. Job Event RPM

Cray Job Event RPM
cray-jobevent_generic-*<version>*.rpm

The RPM name varies depending on the software version. For example:

`cray-jobevent_generic-2.x-1.0502.3a5e398.3.1.ari.x86_64.rpm`

Procedure

1. Copy the job event RPM located in the installation directory in `/root/sma-install/rpm/jobevent-rpm` to the crayadm user's home directory on the SMW.

   For upgrades, the RPM directory will be based of the build date (e.g., `/root/sma-install/rpm-`*`time-date-stamp`*).

   ```
   hostname# cd /root/sma-install/rpm/jobevent-rpm/
   hostname# ls cray-jobevent*
   cray-jobevent_generic<version>.rpm
   hostname# scp cray-jobevent_generic-<version>.rpm crayadm@smw-hostname:
   ```

2. Install the appropriate jobevent RPM on the SMW. The jobevent daemon will be installed in `/usr/bin/jobevent`.

a. Remove the old RPM.

```
smw# rpm -qa cray-jobevent_generic
cray-jobevent_generic-<version>
smw# rpm -e cray-jobevent_generic-<version>
```

b. Install new jobevent RPM.

```
smw# rpm -Uhv cray-jobevent_generic-<version>.rpm
```

3. Record the following for configuration:

- The name (`broker_name`) of the View server. This name must be able to be resolved by the SMW.
- The View API key (`api_key`). The connection to the View server is secure and requires a secret key. The API key can be created and listed from the `/root/sma-install/sma-kafka-cli` command installed on the View server.

4. Define the jobevent parameters in the `/opt/cray/sma/jobevent/config/jobevent.cfg` file.

Items in the `[global]` section can be overridden by subsequent sections.

a. Define a new section in brackets ([*`section_name`*]) for each Cray View server.

The section name is arbitrary.

b. Define `broker_name` and `api_key` under each section.

```
[global]
topics = metrics,jobevents

[production]
broker_name = prod.myproject.com
api_key = MlkrmF9TY3tZFnfmZWV30iFNT8UOvjgp

[test]
broker_name = test1.myproject.com
api_key = MlkrmF9TY3tZFnfmZWV30iFNT8UOv987
```

5. Install the startup script.

here are two methods for installing the jobevent daemon on an SMW shown below, one each for CLE 5.2 and CLE 6.0.

- For CLE 5.2, go to step *6* on page 29.
- For CLE 6.0, got to step *7* on page 30.

6. Install the startup script for CLE 5.2.

a. Install the rc startup scripts.

```
cle5-smw# cp /opt/cray/sma/jobevent/config/rc.jobevent /etc/init.d/jobevent
cle5-smw# ln -s /etc/init.d/jobevent /etc/init.d/rc5.d/S15jobevent
cle5-smw# ln -s /etc/init.d/jobevent /etc/init.d/rc5.d/K15jobevent
```

b. Reload systemd Enable/Restart.

```
cle5-smw# /etc/init.d/jobevent restart
```

The procedure is now complete for CLE 5.2.

7. Install the startup scripts for CLE 6.0

   a. Install the systemd unit file for SMWs.

   ```
   cle6-smw# cp /opt/cray/sma/jobevent/config/jobevent.service /usr/lib/systemd/system
   ```

   b. Reload systemd Enable/Restart.

   ```
   cle6-smw# systemctl daemon-reload
   ```

   c. Enable and start the jobevent daemon.

   ```
   cle6-smw# systemctl enable jobevent
   cle6-smw# systemctl start jobevent
   cle6-smw# systemctl status jobevent
   ```

   The procedure is now complete for CLE 6.0.

10.2 Configure Slurm for Job Events

Prerequisites

- **Important: DO NOT** use this procedure on Cray® XC systems. View for ClusterStor™ gets job event data from the System Management Workstation (SMW) on Cray XC systems. See: *Configure SMW for Job Events* on page 28 for more information.
- View for ClusterStor server is installed and configured
- Network connectivity exists between the slurmctld server node and port 443 of the View server.
- The standard Python modules, `requests` and `argparse`, have been installed on the slurmctld node. If they are not installed, install them using pip, rpm, or the preferred installer of the O/S running on the slurmctld node.

About this task

This procedure describes the steps needed to configure Slurm on host systems (**NOT** Cray XC systems) to forward job events to View for ClusterStor. Once the job event RPM has been installed on the host system, the jobevent daemon will capture Slurm job start/stop events on the slurmctld server and forward them to the View server.

Use the following required steps to install the Slurm jobevent daemon. The configuration steps are dependent on the specific operating system installed on the slurmctld server.

Table 5. Slurm Job Event RPM

Slurm Job Event RPM
sma-jobevents-*<version>*.rpm

For example: `sma-jobevents-1.0.1-0.0.41.x86_64.rpm`

Procedure

1. Copy the Slurm jobevent RPM located in the installation directory in `/root/sma-install/rpm/wlm-jobevent-rpm` of the View server to a usable location on the slurmctld server, such as a home directory or `/tmp`.

   ```
   viewserver# cd /root/sma-install/rpm/wlm-jobevent-rpm/
   viewserver# ls sma-jobevent*
   sma-jobevents-<version>.rpm
   viewserver# scp sma-jobevents-<version>.rpm root@slurmctld:/tmp
   ```

2. Install the RPM on the slurmctld server.

   The jobevent daemon will be installed in `/opt/cray/sma/wlm-jobevents/`.

   ```
   slurmctld# rpm -ivh sma-jobevents-<version>.rpm
   ```

3. Record the following for configuration:

   - The name (`broker_name`) of the View server. This name must be able to be resolved by the slurmctld server.
   - The View API key (`api_key`). The connection to the View server is secure and requires a secret key. The API key can be created and listed from the `/root/sma-install/sma-kafka-cli` command installed on the View server.

4. Define the jobevent parameters in the `sma-jobeventsd.cfg` file. The RPM installs a file `/etc/opt/cray/sma/wlm-jobevents/sma-jobeventsd.cfg.example`. Edit this file and rename it to `/etc/opt/cray/sma/wlm-jobevents/sma-jobeventsd.cfg`

   Items in the `[global]` section can be overridden by subsequent sections.

   a. Define a new section in brackets ([*section_name*]) for each View server.

      The section name is arbitrary.

   b. Define `broker_name` and `api_key` under each section.

   ```
   [global]
   topics = metrics,jobevents

   [production]
   broker_name = prod.myproject.com
   api_key = MlkrmF9TY3tZFnfmZWV30iFNT8UOvjgp

   [test]
   broker_name = test1.myproject.com
   api_key = MlkrmF9TY3tZFnfmZWV30iFNT8UOv987
   ```

5. Install the startup script.

   There are two methods for installing the jobevent daemon on an SMW shown below, one for base operating systems running systemd, and one for base operating systems using system V init scripts.

   - For Init Scripts operating systems, (such as Red Hat 6, SuSe 11, Ubuntu 14) go to *6* on page 32.
   - For systemd operating systems, (such as Red Hat 7, SuSe 12, Ubuntu 15) go to *7* on page 32.

6. Install the startup script for Init Scripts operating systems.

   a. Install the rc startup scripts.

   ```
   slurmctld# cp /etc/opt/cray/sma/wlm-jobevents/rc.sma-jobeventsd /etc/init.d/
   sma-jobeventsd
   slurmctld# ln -s /etc/init.d/sma-jobeventsd /etc/init.d/rc5.d/S15jobeventsd
   slurmctld# ln -s /etc/init.d/sma-jobeventsd /etc/init.d/rc5.d/K15jobeventsd
   ```

   b. Reload systemd Enable/Restart.

   ```
   slurmctld# /etc/init.d/sma-jobeventsd restart
   ```

7. Install the startup script for systemd operating systems.

   a. Install the systemd unit file for the slurmctld server.

   ```
   slurmctld# cp /etc/opt/cray/sma/wlm-jobevents/sma-
   jobeventsd.service /usr/lib/systemd/system
   ```

   b. Reload systemd Enable/Restart.

   ```
   slurmctld# systemctl daemon-reload
   ```

   c. Enable and start the jobevent daemon.

   ```
   slurmctld# systemctl enable sma-jobeventsd
   slurmctld# systemctl start  sma-jobeventsd
   slurmctld# systemctl status sma-jobeventsd
   ```

8. Configure slurmctld to call the jobevents prolog and epilog scripts.

   a. Edit the slurm config file to add the following lines:

   ```
   PrologSlurmctld=/opt/cray/sma/wlm-jobevents/slurm_sma_prolog.py

   EpilogSlurmctld=/opt/cray/sma/wlm-jobevents/slurm_sma_epilog.py
   ```

   b. Restart slurmctld.

   ```
   slurmctld# systemctl restart slurmctld
   ```

Copy the updated config file to all slurm nodes and restart slurmd to prevent errors due to mismatching config files.

10.3 Configure PBS Pro for Job Events

Prerequisites

- **Important: DO NOT** use this procedure on Cray® XC systems. View for ClusterStor™ gets job event data from the System Management Workstation (SMW) on Cray XC systems. See: *Configure SMW for Job Events* on page 28 for more information.
- View for ClusterStor server is installed and configured

- PBS Professional (PBS Pro) cluster is installed and configured
- PBS Pro cluster is attached to all Lustre file systems as appropriate
- Network connectivity between the PBS Pro server node and port 443 of the View server.
- The standard Python modules, `requests` and `argparse`, have been installed on the PBS Pro server node. If they are not installed, install them using pip, rpm, or the preferred installer of the O/S running on the PBS Pro server node.

About this task

This procedure describes the steps needed to configure PBS Pro on host systems (**NOT** Cray XC systems) to forward job events to View for ClusterStor.

View for ClusterStor supports job statistics collection from ClusterStor systems with non-XC systems running PBS Pro. To enable this functionality, job start/stop events must be generated on the PBS Pro scheduler node and forwarded to the View server over the secure restful interface to the message bus. This will allow non-XC PBS Pro jobs to have the same reported statistics as XC jobs or Slurm jobs.

Use the following required steps to install the PBS Pro jobevent daemon.

Table 6. PBS Pro Job Event RPM

PBS Pro Job Event RPM
sma-jobevents-*<version>*.rpm

The RPM name varies depending on the software version. For example: `sma-jobevents-1.0.1-0.0.1.x86_64.rpm`

Procedure

1. Log into the PBS Pro server as root.

2. Copy the `sma-jobevents-<version>.rpm` file from the `sma-install/rpm/wlm-jobevent-rpm` directory on the View server:

   ```
   root@pbsserver:/tmp# scp root@viewserver:/root/sma-install/rpm/wlm-jobevent-rpm/
   sma-jobevents-<version>.rpm ./
   ```

3. Install the rpm.

   - If sma-jobevents is not already installed, run:

     ```
     root@pbsserver:/tmp# rpm -ivh sma-jobevents-<version>.rpm
     ```

   - If sma-jobevents is already installed, update it by running the `rpm` command with the `-U` option:

     ```
     root@pbsserver:/tmp# rpm -Uvh sma-jobevents-<version>.rpm
     ```

4. Record the following for configuration:

   - The name (`broker_name`) of the View server. This name must be able to be resolved by the PBS Pro server.

- The View API key (`api_key`). The connection to the View server is secure and requires a secret key. The API key can be created and listed from the `/root/sma-install/sma-kafka-cli` command installed on the View server.

5. Copy `sma-jobeventsd.cfg.example` to `sma-jobeventsd.cfg`:

```
root@pbsserver:/tmp# cp /etc/opt/cray/sma/wlm-jobevents/sma-
jobeventsd.cfg.example /etc/opt/cray/sma/wlm-jobevents/sma-jobeventsd.cfg
```

6. Configure the network daemon.

a. Open the `sma-jobeventsd.cfg` file for editing:

```
root@pbsserver:/tmp# vim /etc/opt/cray/sma/wlm-jobevents/sma-jobeventsd.cfg
```

b. Make the following changes in `sma-jobeventsd.cfg`:

Note that items in the `[global]` section can be overridden by subsequent sections.

1. Define a new section in brackets ([*`section_name`*]) for each Cray View server. Note that the section name is arbitrary.

2. Define `broker_name` and `api_key` under each View server section.

3. Set the `topics` option to `metrics,job_events` in the `[global]` section.

4. Add `pbs_epilog=True` to the `[global]` section.

5. Add the `pbs_epilog_timer` option in the `[global]` section, if desired, with a value between 1 and 3600 seconds.

The default value is 60 seconds, which is the time interval for the `pbs_epilog` function to scan for and find completed jobs. Shorter intervals will result in the View server updating more quickly, but will increase overhead on the PBS Pro server node.

6. Set the `pbs_bin_path` option if necessary.

By default, the network daemon looks for PBS binaries in `/opt/pbs/bin`. If those binaries reside in a different location, the path must be specified in this configuration option.

Following is an example of a modified `sma-jobeventsd.cfg` file, with `pbs_epilog_timer` set to the default value of 60 seconds:

```
[global]
topics = metrics,jobevents
pbs_epilog=True
pbs_epilog_timer=60

[production]
broker_name = prod.myproject.com
api_key = MlkrmF9TY3tZFnfmZWV30iFNT8UOvjgp

[test]
broker_name = test1.myproject.com
api_key = MlkrmF9TY3tZFnfmZWV30iFNT8UOv987
```

c. Save the modified configuration file.

7. Perform the following steps for base operating systems running systemd:

a. Install the systemd file:

```
root@pbsserver:/tmp# cp /etc/opt/cray/sma/wlm-jobevents/sma-
jobeventsd.service /usr/lib/systemd/system
```

b. Restart systemd:

```
root@pbsserver:/tmp# systemctl daemon-reload
```

c. Enable sma-jobeventsd:

```
root@pbsserver:/tmp# systemctl enable sma-jobeventsd
root# systemctl start  sma-jobeventsd
```

8. Perform the following steps for base operating systems that use system V init scripts:

a. Install `rc.sma-jobeventsd` in `/etc/init.d`.

b. Link start and stop scripts to `rc.sma-jobeventsd` from the appropriate runlevel directories. In most cases, these would be runlevels 3 and higher.

9. Enable the PBS hooks.

```
root@pbsserver:/tmp# qmgr -c "create hook ViewProlog"
root@pbsserver:/tmp# qmgr -c "import hook ViewProlog application/x-python
default /opt/cray/sma/wlm-jobevents/pbs-sma-prolog.py"
root@pbsserver:/tmp# qmgr -c "set hook ViewProlog event=runjob"
root@pbsserver:/tmp# qmgr -c "set hook ViewProlog order=1"
root@pbsserver:/tmp# qmgr -c "set hook ViewProlog alarm=120"
root@pbsserver:/tmp# qmgr -c "set hook ViewProlog enabled=True"
```

10. Confirm that the network daemon is running:

- For base operating systems running systemd, run:

```
root@pbsserver:/tmp# systemctl status sma-jobeventsd
```

- For base operating systems that use system V init scripts, run:

```
root@pbsserver:/tmp# ps -ealf|grep jobeventsd
```

11. Run a simple job that makes use of the Lustre file system, to verify that job events are appearing in View for ClusterStor. For example:

```
echo -e "#\!/bin/sh\ncp -a /bin /$LUSTREDIR/\nuptime\n" > /tmp/sma-testjob.sh
chmod +x /tmp/sma-testjob.sh
qsub -lselect=4:ncpus=1,place=scatter /tmp/sma-testjob.sh
```

The View server should show an entry for the job in the **Jobs Table**. Remember that the job-end scanner only runs every 60 seconds, so there may be some lag between when the job completes and when it appears in the **Jobs Table**.

11 Add Initial ClusterStor System to View for ClusterStor

Prerequisites

View for ClusterStor™ has been installed.

Procedure

1. Go to the site specific URL for View for ClusterStor.

   If the browser displays a security exception, make the browser-specific selections needed to proceed to View for ClusterStor.

2. Log in to View for ClusterStor.

   User Name: `admin`

   Password: `admin`

3. Select the green **Add** icon at the lower right of the GUI's content pane.

   The **Add ClusterStor System** window will appear.

4. Complete the following requested information, and then select the **Add** button.

   - **Name**: File system name for the ClusterStor system. This must match the **Filesystem Name** value that displays when running the `csinfo` command on the ClusterStor system. For example:

     ```
     [root@cls12345n000 ~]# csinfo

     =================================================

     System Serial Number:            CSSG1E00HM
     Possible previous SSN:           N/A
     OEM System Serial Number:        oem
     System Identifier:               cls12345n000
     Cluster Name:                    cls12345n
     Filesystem Name:                 cls12345
     Filesystem Type:                 Lustre
     Hardware Platform:               SNX3000
     Data Network Type:               IB (FDR)
     Software Release:                3.0.0
     Full System Update:              011.79
     RAS-only System Update:          N/A
     Firmware-only System Update:     N/A
     FS-only System Update:           N/A
     ```

   - **Primary IP address**: IP Address of ClusterStor primary management node (n000)
   - **Backup IP address**: IP Address of ClusterStor secondary management node (n001)

- **User**: Username for the ClusterStor system read-only user that was created in *Configure ClusterStor System for Metrics* on page 13
- **Password**: Password for the ClusterStor system read-only user that was created in *Configure ClusterStor System for Metrics* on page 13

It should take no more than one minute for a new ClusterStor tile to appear in the View for ClusterStor **System Overview**. Metrics should appear within a few minutes after the ClusterStor system has been added.

5. Click the browser's refresh button to refresh the **System Overview**.

The displayed ClusterStor metrics will now be current.

12 Configure Notification Email

Prerequisites

View for ClusterStor™ has been installed and configured.

About this task

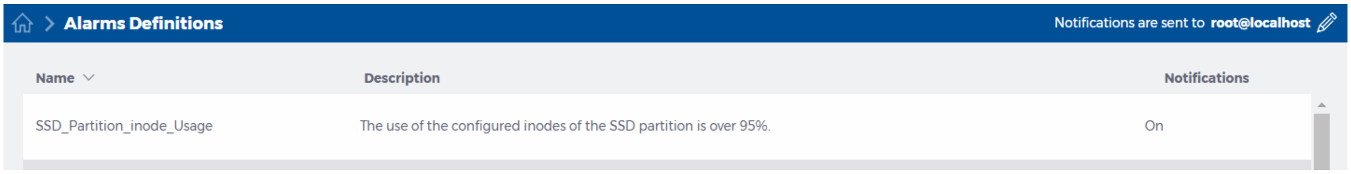
View for ClusterStor creates alarms that record the state of the system as indicated by the values in specified metric series. A notification email is a message sent automatically when the state of an alarm changes. The following procedure configures the email address the system uses when sending alarm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



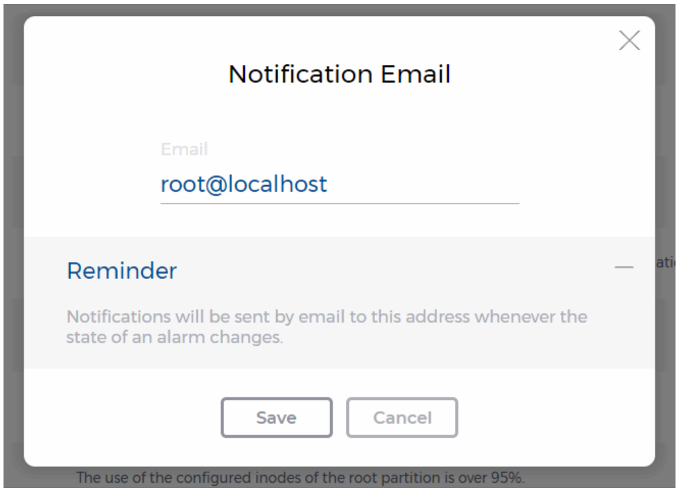
Alarms Definitions

Notifications are sent to **root@localhost**

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: root@localhost

Reminder

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

The use of the configured inodes of the root partition is over 95%.

4. Click **Save**.

13 Change the Default Password

Prerequisites

View for ClusterStor™ has been installed and a ClusterStor system has been added.

About this task

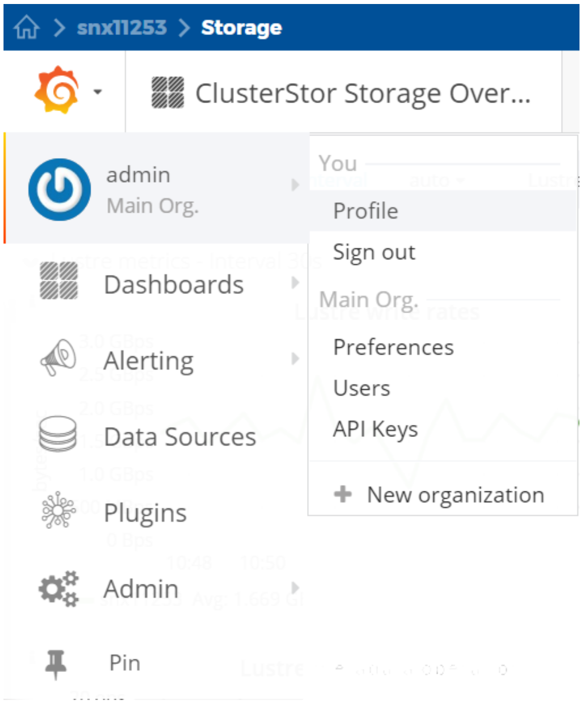
For security purposes, the password should be changed from its default configuration immediately after initial installation. Changing the password is a two step process:

- Change the password in Grafana, which is accessed from the **System Overview** in the GUI.
- Change the password on the View server.

Procedure

CHANGE THE PASSWORD IN GRAFANA

1. Navigate to Grafana by clicking any ClusterStor system name on the **System Overview**.
2. Select the main menu by clicking the Grafana icon in the top left part of the screen.
3. Select **Admin** and click **Profile**.



This will open the **User Profile** page in Grafana.

4. Select **Change Password**.

   Create a secure password.

CHANGE THE PASSWORD IN VIEW FOR CLUSTERSTOR

5. Log in to the View for ClusterStor server as `root`.

6. Navigate to the `/etc/sma-data/etc` directory.

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

7. Change the Grafana password in the configuration file, `site_config.yaml`.

   ```
   hostname# vi site_config.yaml
   grafana_params: {password: admin, user: admin}
   ```

8. Save the updated `site_config.yaml` file.

The password has been changed.

14 Update View for ClusterStor Software

Prerequisites

- View for ClusterStor has been previously installed.
- The procedure in *sysctl Required Configuration for View for ClusterStor* on page 8 has been completed.

About this task

This procedure is used to update the View for ClusterStor software to a newer release.

Procedure

1. Stop any running containers that are not View for ClusterStor.

   If the customer site uses Docker on their View server and has images loaded other than those for View for ClusterStor, it is recommended that those customer images be stopped gracefully before the View for ClusterStor software is updated.

   > **IMPORTANT:** During the software update, **all Docker images will be unloaded**, including those that are not View for ClusterStor.

2. Copy the software update package to `/root`.

3. Unpack the software update package with the `overwrite` option.

   ```
   hostname# tar xvf sma-1.2.0-<time-date-stamp>.tgz --overwrite

   ./
   ./sma-install/
   ./sma-install/sma-1.2.0-rpms-<time-date-stamp>.tgz
   ./sma-install/setup.sh
   ./sma-install/templates/
   ./sma-install/templates/.env
   ./sma-install/templates/site_config.yaml
   ./sma-install/templates/sma.conf
   ./sma-install/templates/sma.service
   ./sma-install/templates/docker-compose.yml
   ```

4. Change directories to `sma-install`.

   ```
   hostname# cd sma-install/
   ```

5. Run the installation script with the update option.

   The time-date stamp in the software package file name will match the time-date stamp in the RPM package file name. For example, `sma-1.2.0-201805151235.tgz` matches with `sma-1.2.0-rpms-201805151235.tgz`.

```
hostname# ./setup.sh -u sma-1.2.0-rpms-<time-date-stamp>.tgz
Updating SMA software
RPM package: sma-1.2.0-rpms-<time-date-stamp>.tgz

Thu Jan  4 14:10:19 CST 2018: Stopping SMA containers...

Thu Jan  4 14:10:19 CST 2018: Checking docker and docker-compose versions...

Thu Jan  4 14:10:20 CST 2018: Prompting user for site configuration
information...
```

6. Reload and restart any Docker images the site uses, other than View for ClusterStor, that were stopped before the software upgrade was performed.

Make sure the jobevent daemon is updated (see *Configure SMW for Job Events* on page 28).

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 `/etc/sma-data/seastream`
`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.