View for ClusterStor™ Installation and Configuration Guide

(1.1.0)

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, jobs metrics, and system events specific to the storage system. Additionally, system logs, system metrics, and system events from each ClusterStor storage system can be configured to be monitored and will collect and persist ibstats metrics from the InfiniBand 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 the Cray ClusterStor storage system.

Release Information

The *View for ClusterStor Installation and Configuration Guide* supports View for ClusterStor version 1.1.0.

Product Requirements

Supported Versions

The following versions are supported by View for ClusterStor:

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

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:
  - 128 GBytes of Memory or more
  - 8 CPU cores or more
  - SSD storage of 500GB or more

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

In order for the View for ClusterStor server to accurately track IB statistics, attaching more than one View IB port to the same fabric is not supported.

Record of Revision

Revision	Date	Content Information
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.

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.

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
```

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 sysctl Required Configuration for View for ClusterStor

Requires Settings

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

Recommended Values

The following sysctl values need to be set for View for Clusterstor to operate correctly.

- net.ipv4.tcp_keepalive_time
  - Recommended Value: 600
  - Maxiumum 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
  - Recommened Value: 262144 (This value is recommended as minimum for ElasticSearch.)
  - System default: 65536

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 change, place the appropriate values in either the /etc/sysctl.conf file or in separate files in the /etc/sysctl.d/ directory. Changing the /etc/sysctl files requires a reboot to take effect.

Change sysctl Settings with sysctl Commands

These settings may be effected by using a sysctl command on the base host. Note that any sysctl settings set by a sysctl command 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 sign (=).

Change sysctl Settings in /etc/sysctl.conf

To permanently effect the sysctl settings, place the values in either the /etc/sysctl.conf file or in separate files in the /etc/sysctl.d/ directory. Add the following lines with spaces before and after the equal sign:

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

These settings only get processed during system start-up. They will take effect on the next reboot of the base host. Either use the sysctl -w command for immediate effect, or reboot the system after updating the files.

3 Install View for ClusterStor Software

Prerequisites

- All site specific configuration values have been collected and the directory on the SSD for storing all SMA data has been created.
- CentOS has been installed and the IB interface has been configured and is up and running.
- Firewalld is not currently supported by View for ClusterStor and must be disabled before starting Docker. IPTables can be used to manage the firewall directly.
- Docker 17.06.2-ce or greater has been installed from *https://docs.docker.com/install/linux/docker-ce/centos/*. Docker Compose 1.14.0 or greater has been installed from *https://docs.docker.com/compose/install/*.
- `sysctl` settings have been changed (see *sysctl Required Configuration for View for ClusterStor* on page 7).
- Openssl version on the job-event daemon's host supports TLSv1.2.

  Run the following commands on the host to verify support for TLSv1.2:

  1. As `root`, log into the machine where the job daemon will run (e.g. SMW).
  2. Determine the SSL 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 TLSv1.2 is not present, upgrade your host to a more recent openssl version. If that is not possible, contact Cray Support for further instructions.
- View for ClusterStor software package has been downloaded from CrayPort.

Procedure

1. Login to the View server as `root`.
2. Copy the View for ClusterStor software archive to `/root.`

   The archive name will be a variation of *`sma-1.1.0-<time-date-stamp>.tgz`*.
3. Unpack the archive using the tar command as shown below.

   ```
   hostname# tar xvf sma-1.1.0-<time-date-stamp>.tgz
   ```

4. Move to newly created directory, `/root/sma-install`.

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

5. Run the installation script.

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.

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

a. Fill in the data directory path when prompted.

This is the directory on the SSD that will store all SMA 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
```

d. Change any previous answers if corrections need to be made.

Answer no if all config settings are correct.

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

6. Verify that the SMA service is running.

```
hostname# systemctl status sma
```

View for ClusterStor has now been installed.

4 ClusterStor System Configuration for Metrics

About this task

This step enables the View for ClusterStor system to distribute metrics. This SDK provides the ability to stream metric data off of a ClusterStor system and onto the SMA server. The following steps can be done on a “live” system, i.e. with Lustre targets mounted on the Lustre servers.

Procedure

1. Log in to the target ClusterStor management node (n000) as `admin`.

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
```

3. 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
--------------------------------------------------------------
```

4. Enable the REST API.

```
MGMT0# cscli service_console configure rest_api enable
```

5. Add the guest user as an REST API authorized user.

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

6. Confirm the REST API.

```
MGMT0# cscli service_console configure rest_api show
REST API access: enabled
REST API authorized users:

  smauser
```

7. Check if the system identifier is set. In this example, `snx11103n000` 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 snx11103n000
```

b. Confirm the system identifier setting.

```
MGMT0# cscli service_console configure system show
System settings:
        System serial number: CSSX0G4DE5
        System identifier name: snx11103n000
```

The ClusterStor system has been configured for Metrics.

⚠ **CAUTION:** After the installation of a System Update (SU) on the ClusterStor system, the configuration will be cleared and will need to be performed again.

5 ClusterStor System Configuration for Log Forwarding

About this task

Designate the destination to which the ClusterStor system will send logs as well as which logs to send. In this example, *SMA1234* is the given destination, *SMA1234.sitename.gov*, and the View for ClusterStor's IP address is *172.30.76.13*.

Use the appropriate procedure below depending on the software version.

5.1 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.

Procedure

1. Log in to the active target ClusterStor management nodes (n000 or n001) as `admin`.

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
```

3. Register a new consumer of Syslog.

```
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.
```

4. List the Syslog consumers. (via SMA-2953)

```
MGMT# 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

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

5.2 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.

Procedure

1. Log in to both the active and standby target ClusterStor management nodes (n000 and n001) as `admin`.

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
   ```

3. Save a copy of the `syslog-ng_receiver.erb` config file on both nodes.

   ```
   MGMT0# 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 before existing `destination` lines:

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

   b. Add before exisiting `log` lines:

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

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

The logs have now been configured for forwarding on the ClusterStor system.

6 Configuration for 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 workaround 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 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.

Users running CLE 5.2 versions UP04 and above are required to perform the procedure in *Enable Lustre Job Statistics for CLE 5.2* on page 15 before proceeding. Users running CLE 6.0 versions UP02-UP04 are required to perform the procedure in *Enable Lustre Job Statistics for CLE 6.0* on page 16 before proceeding. Users running CLE 6.0 UP05 and above can proceed directly to *Enable Jobstats on the ClusterStor System* on page 18.

6.1 Enable Lustre Job Statistics for CLE 5.2

Procedure

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

```
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 jobstats 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
```

6.2 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 jobstats 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
```

6.3 Disable Lustre Jobstats for CLE 5.2

About this task

In order to stop collecting jobstats, services must be disabled.

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
   ```

6.4 Disable Lustre Jobstats for CLE 6.0

About this task

In order to stop collecting jobstats, services must be disabled.

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
   ```

6.5 Enable Jobstats on the ClusterStor System

Prerequisites

Either *Enable Lustre Job Statistics for CLE 5.2* on page 15 or *Enable Lustre Job Statistics for CLE 6.0* on page 16 has been completed.

Procedure

1. Enable Lustre job statistics collection from the ClusterStor management node.

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

   MGMT0$ cscli lustre jobstats modify --frequency 30 --scheduler ALPS_APP_ID
   lustre: Configuring Lustre Job Statistics for snx
   ```

```
lustre: Updating configuration. This can take a while...
lustre: Successfully configured Lustre Job Statistics for snx
```

2. Check the Lustre job statistics configuration to ensure it is enabled.

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

6.6 Disable Jobstats on the ClusterStor System

Prerequisites

The procedure in *Disable Lustre Jobstats for CLE 5.2* on page 17 or *Disable Lustre Jobstats for CLE 6.0* on page 18 has been completed.

Procedure

1. Disable Lustre job statistics collection from the ClusterStor management node.

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

2. Check the Lustre jobstats configuration to confirm that collection has been disabled.

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

7 ClusterStor System Configuration for SNMP

About this task

Enabling Simple Network Management Protocol (SNMP) on ClusterStor 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 interface will display these service alerts, and the SNMPwalk interface will create and send an email with more detailed information.

Procedure

1. Log in to the target ClusterStor management node (n000) as admin.

2. Enable SNMP.

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

SNMP is now enabled.

8 Security Key Configuration 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; one which is installed on the SMW of an XC cluster, and the other is used to monitor SLURM job events on white-box clusters. Both implementations use a Kafka REST API to forward job events to View for ClusterStor. The Kafka REST API interface is secured. Only daemons which present valid secret keys are allowed to post job events to View. The administrator must configure the daemons and their secret keys in View in order to enable the daemons to post their job events.

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

Configure Consumers and Their Secrets

Use `sma-kafka-cli` to configure consumers and their secrets.

```
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
```

This example shows the creation of a consumer named `pollux-smw-job-events` with the auto-generated secret key `JpAKDSc9bFTc3P7VZ2JdihmX1VZenRG5`. The consumer name can be any string.

List Existing Consumers

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

Add Additional Consumers

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

This example shows the addition of a consumer with an admin supplied secret value, rather than an auto-generated secret.

List a Consumer's Existing Secrets

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

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

Add Additional Secrets for a Consumer

```
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 Job Event Daemons

```
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
```

9 SMW Installation and Upgrade for Job Events

About this task

The jobevent daemon captures the WLM job and application start/stop events on the SMW and forwards them to the View server. Following are the required steps to install or upgrade the daemon. Configuration steps are dependent on the specific version of software installed on the SMW.

Table 1. Job Event RPM

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

RPM name varies depending on the software version. For example: cray-jobevent_generic-2.x-1.0502.3a5e398.3.1.ari.x86_64.rpm

This jobevent daemon is meant to run on Cray XC systems' System Management Workstation (SMW) only.

Procedure

1. Copy the job event RPM located in the installation directory in `/root/sma-install/rpm/jobevent-rpm` to the crayadm users 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
topics = metrics
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 *6* on page 24.
- For CLE 6.0, got to *7* on page 24.

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
```

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
```

10 Install the jobevent Daemon on the slurmctld Server

About this task

The jobevent daemon captures Slurm job start/stop events on the slurmctld server and forwards them to the View server. The following are the required steps to install the slurm jobevent daemon. The configuration steps are dependent on the specific operating system installed on the slurmctld server.

This jobevent daemon is intended for clusters running Slurm that are not Cray XC systems.

Table 2. Slurm Job Event RPM

Slurm Job Event RPM
sma-jobevents-<*version*>.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 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
topics = metrics
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 Scipts operating systems, (such as Redhat 6, SuSe 11, Ubuntu 14) got to *6* on page 27.
- For systemd operating systems, (such as Redhat 7, SuSe 12, Ubuntu 15) got to *7* on page 27.

6. Install the startup script for Init Scipts 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.

11 Add Initial ClusterStor to View for ClusterStor

Prerequisites

View for ClusterStor has been installed.

Procedure

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

   A security exception may appear. Click **Advanced**, and select **Proceed**.

2. Log into View for ClusterStor.

   User Name: `admin`

   Password: `admin`

3. Select the gray tile with a plus sign.

   A pop-up form will appear.

4. Fill out the form and select **Add a ClusterStor System**.

   - **Name**: System identifier name for the ClusterStor System
   - **IP address**: IP Address of active ClusterStor management node (n000)
   - **IP address**: IP Address of backup ClusterStor management node (n001)
   - **User**: Username for ClusterStor System read-only user, created in *ClusterStor System Configuration for Metrics* on page 11
   - **Password**: Password for ClusterStor System read-only user, created in *ClusterStor System Configuration for Metrics* on page 11

   It should take no more than one minute for a new ClusterStor tile to appear. Metrics should appear within a few minutes once the ClusterStor system has been added.

5. Click **View for ClusterStor** to refresh the system.

   The browser's refresh button can also be used.

The ClusterStor tile will now display current information.

12 Configure Notification Email

Prerequisites

View for ClusterStor has been installed and configured.

About this task

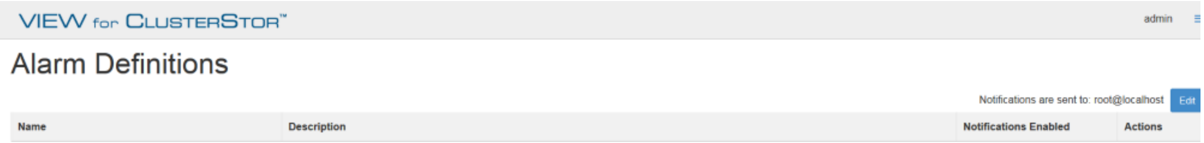
View for ClusterStor creates alarms that records the state of the system as indicated by the values in specified metric series. A notification email is a message sent automatically when the state of an alarm changes. The following procedure configures

Procedure

1. Navigate to the drop-down menu in the upper right corner of the Landing Page.

2. Select **Alarm Definitions**.

3. Select the **Edit** button located at the top right of the Alarm Definitions table.

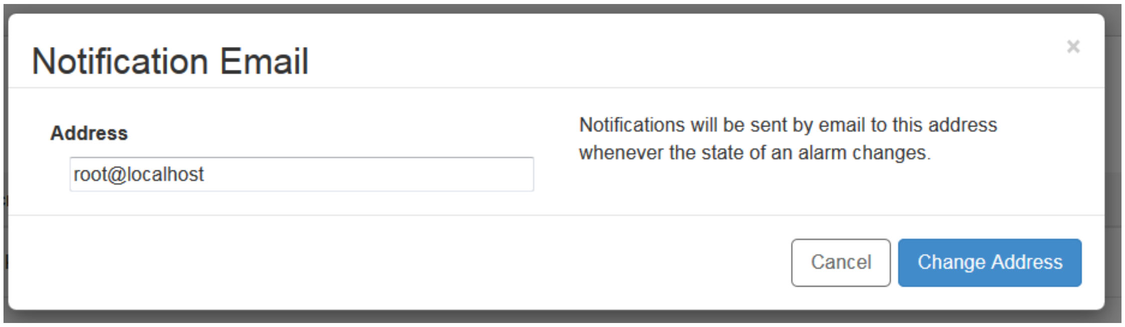


Alarm Definitions

Notifications are sent to: root@localhost

Name	Description	Notifications Enabled	Actions

4. Enter the email address of the intended recepient of notifications.



Notification Email

Address: root@localhost

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

5. Click **Change Address**.

13 Change Default Password

Prerequisites

View for ClusterStor has been installed and a ClusterStor 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.

Procedure

CHANGE THE PASSWORD IN GRAFANA

1. Navigate to Grafana by clicking the ClusterStor title in tile.
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 take you the 'User Profile' page in Grafana.
4. Select **Change Password**.

   Create a secure password.

CHANGE THE PASSWORD IN VIEW FOR CLUSTERSTOR™

5. Login to View for ClusterStor as `root`.
6. Navigate to `/etc/sma-data/etc`.

   ```
   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}
   ```

The password has been changed.

14 Update View for ClusterStor Software

Prerequisites

View for ClusterStor has been previously installed.

sysctl Required Configuration for View for ClusterStor on page 7

Procedure

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

   During the update, all Docker images will be unloaded, including any images that are not View for ClusterStor.

2. Copy software package to `/root`.

3. Unpack software package with `overwrite` option.

```
hostname# tar xvf sma-1.1.0-<time-date-stamp>.tgz --overwrite

./
./sma-install/
./sma-install/sma-1.1.0-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 installation script with 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-1.1.0-201805151235.tgz` matches with `sma-1.1.0-rpms-201805151235.tgz`.

```
hostname# ./setup.sh -u sma-1.1.0-rpms-<time-date-stamp>.tgz
Updating SMA software
RPM package: sma-1.1.0-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...
```

Make sure the jobevent daemon is updated (see *SMW Installation and Upgrade for Job Events* on page 23).

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 the workload is distributed over. Each site will need to set the appropriate length of the View for ClusterStor retention policy. The default retention policy 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.

The retention policy affects metric data in InfluxDB, log and event data in Elastic Search, and job event data in the MariaDB SQL database.

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

Procedure

1. Change to the SMA configuration directory.

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

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

   ```
   hostname# vi .env
   ```

3. Change the `RETENTION_POLICY_DURATION` variable to the desired value in days.

   Note that days are indicated with a suffix of “d”.

4. Restart SMA.

   ```
   hostname# systemctl restart sma
   ```