



# **Differences Between CLE 6.0/SMW 8.0 and CLE 5.2/SMW 7.2**

**(CLE 6.0.UP07)**

**S-2573**

# Contents

1 About This Publication.....	3
2 New Functionality in CLE 6.0 and SMW 8.0.....	4

# 1 About This Publication

---

The What's New for CLE 6.0 and SMW 8.0 (S-2573) publication describes the changes in the Cray Linux Environment operating system effective with CLE 6.0.UP00 through UP07 and SMW 8.0.UP00 through UP07. This publication is intended to outline core CLE functionality and behavior changes between the CLE 6.0/SMW 8.0 and CLE 5.2/7.2 release versions. For a list of changes and new features between CLE 6.0.UPxx/SMW 8.0.UPxx releases, refer to the release notes. This publication was previously titled *What's New for CLE 6.0 and SMW 8.0*.

The *Differences Between CLE 6.0/SMW 8.0 and CLE 5.2/SMW 7.2* (S-2573) document supports Cray software release CLE 6.0.UP07/SMW 8.0.UP07 for Cray XC™ Series systems, released 12 JUL 2018.

## **New in this release**

- There are no major changes to this document for the CLE 6.0.UP07 release.

## 2 New Functionality in CLE 6.0 and SMW 8.0

---

There are a number of significant changes to the functionality in CLE 6.0 and SMW 8.0, particularly with regard to installation and configuration of software. This document provides an overview of the changes effective as of the CLE 6.0.UP07 and SMW 8.0.UP07 releases.

### Operating System/SLES 12

There is no direct upgrade path from CLE 5.2 and SMW 7.2 to CLE 6.0 UP07 and SMW 8.0 UP07.

SMW 8.0 and CLE 6.0 are based on the Novell SLES 12 base operating system. Notable differences in SLES 12 are:

- The default filesystem type has changed from ext3 to btrfs for the SMW system and xfs for most other filesystems, such as /home, /alps\_shared, and data file systems. The SMW PostgreSQL database is on an ext4 filesystem.
- The bootloader has changed from grub1 to grub2. This affects the SMW, but not CLE nodes.
- The init system has changed from `sysvinit` to `systemd`. Use the `systemctl` command to inspect and manage services.
- Network configuration has changed from interface configuration files (`ifcfg`) to SUSE's Wicked network configuration framework.
- By default the `sudo` command requires the full path to the target. This was implemented for security reasons. On the SMW this can be changed by editing the `/etc/sudoers` file to comment out the following line:

```
#Defaults secure_path="/usr/sbin:/usr/bin:/sbin:/bin"
```

Because this is a Linux file in the image root, editing the `sudoers` file will not work for images that boot on CLE nodes or eLogin nodes. If necessary, this behavior can be modified with an Ansible play.

- The database system has changed from MySQL to MariaDB open source database. However, the daemon is still called `mysqld` and the names of the client commands to interact with the database are unchanged.
- Multi-queue SSD Access. This applies only to systems where an SSD is installed.
- `SUSEconfig.permissions` has been replaced by `chkstat`.
- Information that was kept in the `/etc/SuSE-release` file should now be in the `/etc/os-release` file. Use of the `/etc/SuSE-release` file is deprecated and the file itself will be removed in a future service pack or release.
- The Korn shell version from AT&T, `ksh-93`, is deprecated and will be removed in SLES 13. Going forward, the MirBSD Korn shell (`mksh` for interactive use with `lksk` for non-interactive script compatibility) is the standard `ksh` version. A legacy version, `ksh-93vu`, is available in SLES 12 to allow for a transition to `mksh`/`lksk` prior to SLES 13. In CLE 6.0, image recipes include both `ksh` versions, with `ksh-93` being the default. In SMW 8.0, only `mksh`/`lksk` is installed by default, but sites can use `zypper` to install `ksh-93`.

More details on the changes and features in SLES 12 can be found on this Novell website. [https://www.suse.com/releases/notes/x86\\_64/SUSE-SLES/12/](https://www.suse.com/releases/notes/x86_64/SUSE-SLES/12/)

Other noteworthy changes to CLE:

- Compute nodes now boot a full Linux image, instead of BusyBox embedded Linux.
- The 32-bit HSS controllers run openSUSE 13.2.

## Logical Volume Management (LVM)

Logical Volume Management (LVM) is used for all storage on Boot RAID. The SMW, boot node, and SDB node all have LVM volume groups on the Boot RAID.

## Image and Configuration Management

A key feature of CLE 6.0 and SMW 8.0 is the separation of the software from the configuration. This provides an easier path to upgrade software and configuration. To configure Cray XC systems and manage configuration content, system administrators use the Cray Management System (CMS).

The CMS comprises configuration data, the tools to manage and distribute that data, and software to apply the configuration data to the running image at boot time. Its major components include the Configuration Management Framework (CMF), Image Management and Provisioning System (IMPS), and Node Image Mapping Service (NIMS).

IMPS provides repository management to create and distribute repository content (rpms) and prescriptive image creation to create and update standard or custom images.

CMF provides centralized configuration on the SMW by means of configuration sets.

The configurator configures individual software services when the system image is being built. Cray provides configuration templates and other content for each software component or service. For an initial install the configurator creates worksheets, which sites fill out with site-specific settings. These worksheets are used to create a specific configuration set for each software service, which contains all of the information needed to configure (or start the configuration of) that particular service.

A site can choose to change configuration data using the configurator command, `cfgset` and to make changes in the data, either by direct editing of the config template YAML file or by editing of configuration worksheets. If the configuration template is changed, then the `cfgset` command must be used to update the configuration set with that changed data.

As part of the software image, Cray also provides an Ansible play for many of the services service. Sites can add their own content, including Ansible plays and related content to the configuration sets.

IMPS Distribution Services (IDS) provides the configuration set data to nodes. During the boot process for CLE nodes, the software in the image is customized with the configuration settings by Ansible plays.

Ansible automates configuration management and application deployment. Ansible activity is controlled by a script called `cray-ansible`. The `cray-ansible` script assembles Ansible plays from the image and the configuration set, gathers system level data, such as the OS, and host name, reads in the configuration set data, and executes plays to apply configuration data. Ansible runs first during Linux initialization, where it performs some initial configuration of Linux, shuts off default services that aren't needed, and sets up the high-speed network interfaces, allowing Linux/systemd to run. When Linux has finished its startup, Ansible runs again to configure software and start services. Ansible runs on every compute and service node to pull information, rather than the previous model where content was pushed from a centralized location.

NIMS determines which images get booted on specific nodes, any additional kernel parameters to pass to the nodes on boot, and which load file to use within a boot image.

For more information, see the *Configurator User's Guide*.

## Dynamic RDMA Credentials (DRC)

Dynamic RDMA Credentials (DRC) is a new XC system service that enables shared network access between different user applications. DRC enables user applications to request managed network credentials, which can be shared with other users, groups, or jobs. Access to a credential is governed by the application and DRC to provide authorized and protected sharing of network access between applications. DRC extends the existing protection domain functionality provided by ALPS without exposing application data to unauthorized applications. DRC can also be used with other batch systems, such as Slurm, without any loss of functionality.

## SimpleSync

The Cray SimpleSync service provides a straightforward mechanism for administrators to copy files onto their system without resorting to writing an Ansible play. When enabled, the service automatically copies files found in source directories in the configuration set on the SMW to one or more target nodes.

## Resource Utilization Reporting (RUR)

RUR replaces the legacy resource tracking tool, Comprehensive System Accounting (CSA). RUR provides a reliable, high-performance framework that supports plugins to collect data about the usage of a particular resource. RUR is highly configurable and extensible. Because RUR runs before and after an application, it introduces virtually no performance noise into the running application. Cray provides several plugins to support a variety of sets of collected data. Customers can also implement their own plugins to collect data on their systems. Plugins also provide multiple ways to output collected data.

The RUR performance switch, `compute_local_python: True` is no longer supported as that capability is now built into the system.

For more information, see the *XC System Management Guide*.

## External Login (eLogin)

External services formerly provided by esLogin are now supported by eLogin. eLogin uses Cray System Management Software (CSMS) installed on a separate Cray Management Controller (CMC) to deploy eLogin images to external Cray Development and Login (CDL) nodes. The CMC and the CDL nodes are each deployed to Dell 720s.

The CMC connects to the SMW, which provides shared image and configuration services. A recipe for eLogin nodes exists on the SMW so that an image root for eLogin can be created and packaged into the proper boot image format thus enabling the SMW to deliver it to the eLogin node.

For more information see the *eLogin Installation Guide* for CLE 6.0 and the *eLogin Administration Guide* for CLE 6.0.

## Programming Environment (PE)

The PE components (compilers, debuggers, optimizers, and run-time libraries) are installed and managed on the SMW. The same software can run on compute, login, or elogin nodes. The `craype-installer` targets the PE image root which can be pushed to the boot node or, if eLogin is enabled, to the CSMS node. Internal XC nodes use the image binding service to mount the various directories which contain PE software components from the

copy of the PE image root on the boot node. The PE image root is pulled to the eLogin node from the CMC node as the eLogin node boots and is stored on a local disk in the eLogin node.

## Dynamic Shared Libraries

Dynamic Shared Libraries are supported by default and are not dependent on Netroot. Netroot is still needed for large image-based libraries.

## Support for Rollbacks and Staged Upgrades

A new command, `snaputil`, creates and manages snapshots on the SMW to support rolling back to a prior state. This also enables staged upgrades by installing new software to a snapshot, switching context via `chroot` to the new snapshot to update the configuration, then shutting down CLE and booting the SMW to the updated snapshot.

For more information, see the *XC System Management Guide*.

## Support for Rolling Patches on Compute and Service Nodes

Software patches can now be rolled out to nodes via the workload manager as they become available when running applications complete.

## Third-party Software

Because there is no longer a shared root, third party software packages are installed into node-specific images.

## Miscellaneous Changes

There are no default passwords for root and crayadm on CLE nodes. You will be prompted to choose passwords when completing the configuration worksheets. When SLES12 is installed to the SMW, the root and crayadm accounts are provided with default passwords.

The Platform Management Database (PMDB) is now located on a dedicated drive on the SMW, rather than on the root drive as was the previous default location. There is also the option to move power and environment data from the SMW to a Postgres database on a remote device.

System Environment Data Collection (SEDC) information is now stored only on the PMDB, rather than in log files. A new script, `xtgetsedcvalues`, is available to more easily query the PMDB for SEDC information.

A new utility, `xtpgtune`, automatically tunes configuration of the PostgreSQL cluster for the Cray Platform Management Database (PMDB) for optimal performance.

Direct-Attached lustre (DAL) is supported on CentOS. DAL over InfiniBand is no longer supported.

Data Virtualization Service (DVS) is now an integral part of the operating system used, for example, to distribute the PE image to the compute and login nodes.

16Gbps Fibre Channel replaces 8Gbps Fibre Channel.

The runtime functionality previously provided by DSL nodes has been replaced by Tier 2 nodes. In addition, the Tier 2 nodes provide boot time assistance. Use repurposed compute nodes for Tier 2 provisioning at the rate of two per 400 nodes for 400 nodes or fewer, or one per 400 nodes in larger systems. If a system was configured with DSL nodes, these can be used for Tier 2 nodes.

The `xtcon` utility, which provides a two-way connection to the console of any running node, is now enabled by default.