

HP XC How To

Installing standard LSF on HP XC



© 2005 Hewlett-Packard Development Company, L.P.

The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Linux is a U.S. registered trademark of Linus Torvalds.

LSF, Platform Computing, and the LSF and Platform Computing logos are trademarks or registered trademarks of Platform.

Contents

Preparing to install standard LSF

Prerequisites	5
Removing or disabling SLURM and LSF-HPC for SLURM	5

Installing standard LSF

Overview of the example configuration.....	7
Configuring LSF	7
Reconfiguring the XC cluster	10

Revision history

Revision tables

Table 1 Revisions

Date	Edition	Revision
Oct 2005	V1.0	First

Preparing to install standard LSF

You can install standard LSF on XC in place of SLURM and LSF-HPC. This procedure removes SLURM and LSF-HPC and configures standard LSF daemons on every compute node in the XC cluster.

Prerequisites

The following software is needed to install standard LSF on XC:

- The appropriate standard LSF binary tar file. For XC v2.n this is as follows (where X.Y is the selected LSF version, such as 6.1):
 - `lsfX.Y_linux2.4-glibc2.3-ia64.tar.z` for an XC6000 Itanium-based system.
 - `lsfX.Y_linux2.4-glibc2.2-amd64.tar.z` for an XC4000 Opteron-based system.
 - `lsfX.Y_linux2.4-glibc2.3-ia32e.tar.z` for an XC3000 64-bit Intel-based system
- The corresponding standard LSF installation tar file, `lsfX.Y_lsfinstall.tar.z`.
- A standard LSF license file from Platform, appropriately licensed for the total number of CPUs on your XC cluster.

The procedure makes the following assumptions about your system:

- LSF is only to be installed on XC nodes. To extend the LSF configuration outside of XC to include non-XC nodes is beyond the scope of this document.
- All nodes are up and running, so that they can be updated with the changes described in this procedure. If there are nodes that are down, you will need to ensure that those nodes are updated as soon as they come online again.

Removing or disabling SLURM and LSF-HPC for SLURM

Use the following procedure to remove SLURM and LSF-HPC, and set up standard LSF to automatically start at boot time on every cluster compute node.

1. Remove the SLURM and LSF-HPC cluster configuration scripts.

Remove these scripts to prevent further cluster maintenance activities from attempting to reconfigure and re-enable SLURM and LSF-HPC. Use the following commands:

```
root@n16 root]# cd /opt/hptc/etc/gconfig.d/
[root@n16 root]# rm -f C55lsf C50slurm_controller
[root@n16 root]# cd /opt/hptc/etc/nconfig.d/
[root@n16 root]# rm -f C55lsf C51slurm_compute C51slurm_controller
```

Do not delete the `*slurm_launch` link in the `/opt/hptc/etc/nconfig.d/` directory. (See the next step.)

2. Configure standard LSF to automatically start on node boot.

To do this, you must convert a `slurm_launch` script from starting `munge` on every node to starting LSF on every node. The MUNGE service is required only by SLURM. The `slurm_launch` script insures that the configured service (in this case, LSF) is enabled for automatic startup during node boot up, via the `chkconfig` command. Make the following file changes:

1. Edit the file `/opt/hptc/etc/nconfig.d/C50slurm_launch`, changing the following line:

Original line: `my @services = ("munge");`

New line: `my @services = ("lsf");`

2. Set the value of `$ENABLE_PAM_SLURM` to 0, as follows:

```
my $ENABLE_PAM_SLURM = 0;
```

Save and exit the file.

3. If your XC system has just been installed you have not yet run the `cluster_config` utility, you can ignore the rest of these steps and proceed to the procedure in Installing Standard LSF.

4. Stop LSF-HPC by using the following command:

```
[root@n16 root]# controllsf stop
```

5. Stop SLURM by using the following command:

```
root@n16 root]# scontrol shutdown
```

6. Disable MUNGE and SLURM-controlled user access (if enabled) by using the following command:

```
[root@n16 root]# pdsh -a '/opt/hptc/slurm/etc/nconfig.d/munge_nconfig.pl nunconfigure'
```

7. Delete the LSF-HPC installation on XC by using the following command:

```
[root@n16 root]# /opt/hptc/lsf/etc/gconfig.d/lsf_gconfig.pl
gunconfigure
removing /opt/hptc/lsf/top/conf...
removing /opt/hptc/lsf/top/6.0...
removing /opt/hptc/lsf/top/work...
removing /opt/hptc/lsf/top/log...
removing /hptc_cluster/lsf/conf...
removing /hptc_cluster/lsf/work...
removing /var/lsf..
```

You must now run the `cluster_config` command to make the changes permanent. However, first install standard LSF so that when you run the `cluster_config` command, it will also configure the new installation.

Installing standard LSF

Overview of the example configuration

Perform the following steps to install standard LSF on an XC system. The example provided installs LSF v6.1 on a Cluster Platform 3000.

The most important decision when installing standard LSF on XC is where to put the installation files. The Platform documentation for installing standard LSF recommends establishing a location that is shared among all the nodes participating in the LSF cluster. On an HP XC system this shared file location (via NFS) is `/hptc_cluster`. For small XC systems it might be adequate. However, as the node count increases the scalability of the network traffic becomes a concern.

You can install standard LSF on XC in any location, providing you have carefully considered how files are managed on XC. The following example installs standard LSF in the shared location (`/hptc_cluster/lsf`), modified so that the binary directory (`/hptc_cluster/lsf/6.1`) and the daemon logging directory (`/hptc_cluster/lsf/log`) are local to each node.

Moving the binary directory to a node-local location (where the XC System Imager component will copy it to each node) reduces the network traffic each time a binary is invoked. Moving the daemon logging location off of the shared location also reduces the network traffic. This approach will leave the configuration directory (`/hptc_cluster/lsf/conf`) and the runtime working directory (`/hptc_cluster/lsf/work`) in the shared location, which is congruent with their purpose (one set of configuration files for the cluster, and a central accounting or runtime location for the cluster).

Configuring LSF

1. Prepare the LSF installation.

Obtain the 2 tar files and license file listed in the prerequisites, and copy them to a temporary directory on the XC head node. unpack the installation tar file and change into the resulting directory as follows:

```
[root@n16 root]# cd lsftmp
[root@n16 lsftmp]# ls
license.dat  lsf6.1_linux2.4-glibc2.3-ia32e.tar.Z  lsf6.1_lsfinstall.tar.Z
[root@n16 lsftmp]# tar xzf lsf6.1_lsfinstall.tar.Z
[root@n16 lsftmp]# cd lsf6.1_lsfinstall
[root@n16 lsftmp]# ls
COPYRIGHT      instlib          lsfinstall      rpm              upgrade.html
hostsetup      license_agreement.txt  README          scripts
install.config lsf6.1_unix_install.pdf  rhostsetup     slave.config
```

2. Edit the `install.config` file with appropriate values. Using the example described in the introduction:
 - a. `LSF_TOP` is required and is the location of the LSF installation

- b. `LSF_ADMINS` is required and can be set to the same account that was created for `LSF_HPC` (the default is `lsfadmin`). Whatever value is entered, it must be a user account that exists on all nodes in the XC cluster
- c. `LSF_CLUSTER_NAME` is required, and can be anything that complies with the naming rules.
- d. `LSF_ADD_SERVERS` is where you specify which nodes you want running LSF. HP recommends that you specify a filename for this value.

There are other variables that you can configure, but the four specified are the basic ones necessary to get LSF running on XC. The settings specified in the example are as follows:

```
LSF_TOP="/hptc_cluster/lsf"
LSF_ADMINS="lsfadmin"
LSF_CLUSTER_NAME="xc_cluster"
LSF_ADD_SERVERS=:xc_host_servers
```

3. The sample cluster contains 4 nodes, thus the `xc_host_servers` file should contain a line for each node, as follows:

```
n16
n15
n14
n13
```

4. Run the `lsfinstall` program specifying the `install.config` file as follows:

```
root@n16 lsf6.1_lsfinstall]# ./lsfinstall -f install.config
```

5. Optionally, move the binary and daemon log directories to a node-local location to reduce network traffic on the shared location. For the example, move the binary directory to `/opt/hptc/lsf/install` and the log directory to `/var/log/lsf` as follows:

```
[root@n16 lsf6.1_lsfinstall]# mkdir -p /opt/hptc/lsf/install
[root@n16 lsf6.1_lsfinstall]# mkdir -p /var/log/lsf
[root@n16 lsf6.1_lsfinstall]# cd /hptc_cluster/lsf
[root@n16 lsf]# ls
6.1  conf  home  log  work
```

Note the following concerning the directories:

3. The `home` directory in `/hptc_cluster/lsf` is for the `lsfadmin` account created during `cluster_config` when this HP XC system was originally installed. It was not created during the standard LSF installation.
 4. The `6.1` directory, which corresponds to the installed version of LSF, and contains the binaries and other files that remain static (unchanged) during the operation of LSF. In a later step, this directory is moved to a node-local location that is subsequently copied to all the other XC nodes.
 5. The `conf` directory contains the configuration files that may need to be edited during the normal operation of LSF, so this directory will remain on the shared location.
 6. The `log` directory is where each daemon on every node writes its error and debug messaging. Since there is the potential for heavy write traffic to this directory, it is moved to a node-local location. Consequently, you must log into each node to access the daemon logging that occurs on each node.
 7. The `work` directory is where runtime event and accounting files are written and stored. This directory remains on the shared location.
6. Move the following directories:

```
[root@n16 lsf]# mv 6.1/ /opt/hptc/lsf/install/
```

```
[root@n16 lsf]# ln -s /opt/hptc/lsf/install/6.1 6.1
```

7. Verify that the log directory is empty:

```
[root@n16 lsf]# ls log
[root@n16 lsf]# rmdir log
[root@n16 lsf]# pdsh -a mkdir -p /var/log/lsf
[root@n16 lsf]# pdsh -a chown lsfadmin /var/log/lsf
[root@n16 lsf]# ln -s /var/log/lsf log
```

8. The `/var/log/lsf` location is created cluster-wide. The `/var/log` directory is not maintained by System Imager because of its node-local nature (System Imager must not be configured to destroy the actual node-specific logs on each node during execution of an `updateclient` command).
9. The final layout of `LSF_TOP` appears as follows:

```
[root@n16 lsf]# ls -l
total 12
lrwxrwxrwx  1 root    root      25 May 19 13:34 6.1 -> /opt/hptc/lsf/install/6.1
drwxr-xr-x  3 lsfadmin root      4096 May 17 15:56 conf
drwx-----  4 lsfadmin lsfadmin 4096 May 16 08:22 home
lrwxrwxrwx  1 root    root      12 May 19 13:34 log -> /var/log/lsf
drwxr-xr-x  3 lsfadmin root      4096 May 17 15:33 work
```

10. To correct a problem with LSF V6.0 and v6.1, you must edit the environment setup files. The `profile.lsf` and `cshrc.lsf` LSF environment setup files look for the XC system file `/etc/hptc-release`, and if found they assume that LSF-HPC for SLURM was installed. To fix this problem, use the following commands to cause LSF to search for a non-existent file:

```
[root@n16 lsf6.1_lsfinstall]# cd /hptc_cluster/lsf/conf
[root@n16 conf]# mv profile.lsf profile.lsf.orig
[root@n16 conf]# mv cshrc.lsf cshrc.lsf.orig
[root@n16 conf]# sed -e "s?/etc/hptc-release?/var/lsf/lsfslurm?g" \
< profile.lsf.orig > profile.lsf
[root@n16 conf]# sed -e "s?/etc/hptc-release?/var/lsf/lsfslurm?g" \
< cshrc.lsf.orig > cshrc.lsf
```

In the preceding example, `/var/lsf/lsfslurm` is used as the non-existent file in anticipation of a correction planned for LSF V6.2 and beyond.

11. Once done, source the appropriate command as follows:

```
root@n16 conf]# . profile.lsf
```

12. Optionally, add LSF paths to the default environment upon login as follows:

- a. For `cs`h users, create `/etc/profile.d/lsf.csh` with the following contents (be sure to substitute `/hptc_cluster/lsf` with your `LSF_TOP`):

```
if ( "${path}" !~ *-slurm/etc* ) then
    if ( -f /hptc_cluster/lsf/conf/cshrc.lsf ) then
        source /hptc_cluster/lsf/conf/cshrc.lsf
    endif
endif
```

- b. For `bash` and other `sh`-based users, create `/etc/profile.d/lsf.sh` with the following contents (be sure to substitute `/hptc_cluster/lsf` with your `LSF_TOP`).

```
[root@xcl28 profile.d]# cat lsf.sh
```

```

case $PATH in
    *-slurm/etc:*) ;;
    */hptc_cluster/lsf/*) ;;
    *)
        if [ -f /hptc_cluster/lsf/conf/profile.lsf ]; then
            . /hptc_cluster/lsf/conf/profile.lsf
        fi
    esac

```

13. Create the following softlink to set up the control script:

```

[root@n16 conf]# which lsf_daemons
/hptc_cluster/lsf/6.1/linux2.4-glibc2.3-ia32e/etc/lsf_daemons
[root@n16 conf]# ln -s `which lsf_daemons` /etc/init.d/lsf

```

Reconfiguring the XC cluster

Use the `cluster_config` command to update the role changes in the XC database and update the golden image with the new configuration as follows:

1. Configure all nodes to reimage on bootup as follows:

```
[root@n16 conf]# setnode --resync -all
```

All nodes must reimage in order to update themselves with the latest changes.

2. Shutdown the cluster as follows:

```
[root@n16 conf]# stopsys
```

3. Run the `cluster_config` command

```

[root@n16 conf]# cd /opt/hptc/config/sbin
[root@n16 sbin]# ./cluster_config

```

specifying the following:

- a. Adjust any role assignments. The compute role will no longer be applicable to resource management on XC, because standard LSF is installed on every node. If you want to limit which nodes are to be used for actual user applications, you can create `hostgroups` within LSF. See the Platform documentation for more details.
 - b. The `resource_management` role is also no longer valid. You can remove this role from all nodes, or leave it with the understanding that nothing special will be occurring on the nodes configured with the `resource_management` role.
 - c. After adjusting the roles, select `p` to proceed. Provide the appropriate answers to the existing services. You should not be prompted with questions about SLURM or LSF-HPC.
4. When `cluster_config` completes, the golden image is updated. Start the cluster as follows, remembering that all nodes are reimaging so booting will take some time:

```
[root@n16 sbin]# startsys
```

5. When the cluster is fully up and running, standard LSF is running on all the compute nodes. The following example shows output from standard LSF commands on XC cluster configured in the preceding example:

```

[root@n16 root]# lsid
Platform LSF 6.1, Apr 12 2005

```

My cluster name is xc_cluster

My master name is n16

[root@n16 root]# lshosts

HOST_NAME	type	model	cpuf	ncpus	maxmem	maxswp	server	RESOURCES
n16	LINUX64	Intel_EM	60.0	2	3456M	6143M	Yes	()
n15	LINUX64	Intel_EM	60.0	2	3456M	6143M	Yes	()
n14	LINUX64	Intel_EM	60.0	2	3456M	6141M	Yes	()
n13	LINUX64	Intel_EM	60.0	2	3456M	6143M	Yes	()

[root@n16 root]# bhosts

HOST_NAME	STATUS	JL/U	MAX	NJOBS	RUN	SSUSP	USUSP	RSV
n13	ok	-	2	0	0	0	0	0
n14	ok	-	2	0	0	0	0	0
n15	ok	-	2	0	0	0	0	0
n16	ok	-	2	0	0	0	0	0

[root@n16 root]# su - lsfadmin

[lsfadmin@n16 home]\$ bhosts

HOST_NAME	STATUS	JL/U	MAX	NJOBS	RUN	SSUSP	USUSP	RSV
n13	ok	-	2	0	0	0	0	0
n14	ok	-	2	0	0	0	0	0
n15	ok	-	2	0	0	0	0	0
n16	ok	-	2	0	0	0	0	0

[lsfadmin@n16 home]\$ bsub -n8 -I printenv LSB_HOSTS

Job <101> is submitted to default queue <normal>.

<<Waiting for dispatch ...>>

<<Starting on n16>>

n16 n16 n15 n15 n14 n14 n13 n13