

The LOFAR Clusters

This page describes the LOFAR clusters for (potential) users. The CEP2 cluster is used to store the data from the BlueGene and to run the pipelines that will do the “standard” reduction and calibration of this data. After this pipeline processing, the results are stored in the Lofar Export (staging) Archive. For further processing the CEP1 cluster can be used.

We welcome authorized users on the LOFAR clusters.

User Access

Access through portal to Cluster frontend

You can access the Lofar cluster through the portal. On your own system, try:

```
> ssh -X portal.lofar.eu
```

to connect. We maintain a ssh whitelist, so only registered institutes are able to login. Please send an email to *grit at astron.nl* or *h.paas at rug.nl* to add your institute or personal IP number to this white list. When you are logged in for the first time, you'll have an empty home directory.

To get onto the CEP1 cluster, you first have to login at one of the two frontend nodes `lfe001` or `lfe002` using `ssh -X` (as above). When you don't know which one, please use `lfe001` by default.

To get onto the CEP2 cluster, you first have to login at one of the two frontend nodes `lhn001` or `lhn002`



For more sophisticated useage of ssh read [this page](#).

LOGIN environment

See [The Lofar Login Environment](#) page for information on your default login environment. Only the default environment is supported, so if you defer from this, you're on your own..!

Do and Don'ts

DON'T:

-  Store data in your \$HOME on `portal.lofar.eu` ; this system has a ridiculously small disk and filling it will prevent other users to access properly 
- Pump around large amounts of data between nodes.
- Leave unused data on the systems. Clean it up, so others can make use of the available disk

space.

DO:

- Help improve the system by sending suggestions for improvements or new packages to the administrators

Contributed Software

Some groups or people would like their favourite tools or pipelines added to the cluster software. That can usually be accommodated, but there are some guidelines. Click [here](#) to learn more!

Short CEP1 LOFAR Cluster layout

Frontend

A frontend has 2 Intel Xeon L5420 quad core processors, 16GB internal memory, 2 GbE interfaces and 2TB disks in RAID5 configuration. There are actually two identical frontends: `lfe001` and `lfe002`. Both of them serve a specific group of subclusters. The frontends are used to build the software and regulate the workload on the subclusters.

Processing units

The compute elements have 2 Intel Xeon L5420 quad core processors, 16GB internal memory, 2 GbE interfaces and 1TB disks in RAID0 configuration. They can be accessed by secure shell.

Storage Units

The storage nodes are HP DL180G5 boxes, having 2 Intel Xeon L5420 quad core processors, 16GB internal memory, 6 GbE network interfaces and a total of 24TB disk space. The nodes have been named `lsexxx` with `xxx` a number from `001` to `024`.

The disk space is divided into 4 partitions of 6 disks each, set up in RAID5 configuration. The partitions are called `"/data1"` - `"/data4"`. The installed filesystem is the [XFS](#) filesystem.

Subclusters

The concept of subclusters is gone. All "Ice" compute nodes can reach all "Ise" storage nodes. There are 24 Ise nodes with 4 data partitions each. They can be accessed on Ice nodes using automounts like `/cep1/lse023_data2`. Automounts are setup when access is needed and will drop off after 60 minutes being unused.

To achieve a proper distribution, you can select a few Ice nodes from the same row as their

corresponding lse nodes.

lce-nodes	lse-nodes
=====	=====
lce001-lce009	lse001-lse003
lce010-lce018	lse004-lse006
lce019-lce027	lse007-lse009
lce028-lce036	lse010-lse012
lce037-lce045	lse013-lse015
lce046-lce054	lse016-lse018
lce055-lce063	lse019-lse021
lce064-lce072	lse022-lse024

Short CEP2 cluster layout

Frontend

The CEP2 frontend are called “lhn001” and “lhn002”. Each of them has the following components:

- 64GB memory
- 16 cpu cores
- 512GB Solid State Disk
- 5.4TB disk RAID5
- 40 Gbps Mellanox InfiniBand interconnect

Processing/storage nodes

The CEP2 cluster consists of 100 processing/storages nodes, called “locus...”. Each node has:

- 64GB memory
- 24 cpu cores
- 20TB disk RAID5
- 40 Gbps Mellanox InfiniBand interconnect

Hoover nodes

Two additional nodes locus010 and locus102 act as “hoover” nodes. They can access the storage partition of each locus node and provide a means to “gather” all relevant data. These nodes use the automount mechanism. A “hoover” node has:

- 64GB memory
- 24 cpu cores
- 2TB disk (RAID0)
- 40 Gbps Mellanox InfiniBand interconnect

You can get access to a locus data partition just by addressing it like “cp -a /cep2/locus_data <dest>”. The result is a NFS mount on these data and the mount will drop again when you are ready.

Bootleg mechanism

The new cluster is subject to the “bootleg” deployment regime. This service and installation facility was developed by Harm Paas of the CIT department of the RUG University in Groningen. He already had many years of experience with this system at the Computer Science Faculty. One server in the CEP domain is capable of installing many cluster nodes at once in less than 10 minutes. The administration of all nodes is done on this single Bootleg server. More details can be found on the [Bootleg page](#).

Central administration

The way the CEP cluster has been set up is as follows: We are using Ubuntu Linux LTS-8.04 as Operating System within the whole cluster. Bootleg takes care of all of the server management in the cluster by creating new images, administering servers, keeping track of updates and correct start up of servers by sending them new images when it finds outdated versions at boot time.

Profiles

Bootleg also administers machine profiles: Depending on the function of the server in the cluster it needs other settings and profiles. This management strategy implies that we have to keep track of all programs installed in the Linux system itself, so we will be able to add them to the image. We plan to build fresh images every month (at the maintenance days).

Instantaneous changes

We are also able to respond quickly on extra-program demands for system programs. Within bootleg there is a mechanism to roll-out extra programs through the cluster from a central administration point (CAP). From the CAP we order to deploy a program or a settings-change and after 1 minute the whole cluster will be updated with the change. So if you need extra system programs, simply ask, and you get it almost instantly on every server in the cluster!

Application programs

For an application program change or addition there is another method. Of course you develop your programs in your own user space and/or svn repositories. So you have completely control over all of this. Program releases for the cluster itself reside on a central disk. Normally we (as system administrators) are not involved in the application program deployment. Only if a completely new package is added under a new directory name we need to make this available in the cluster by connecting it in the O.S. via a link to the newly created package name. So there is a clear interface between system programs and applications: The connection between them is at a well defined point, the /opt directory is intended for this. The same goes for database content (mysql, postgresql, ...). Although it might involve OS program installation and services, the database content itself will never be on the Linux system image but stored on a separate data server.

Again: Simply ask, we can connect your program and make it available cluster-wide.

From:

<https://www.astron.nl/lofarwiki/> - **LOFAR Wiki**

Permanent link:

https://www.astron.nl/lofarwiki/doku.php?id=public:lofar_cluster&rev=1320756947

Last update: **2011-11-08 12:55**

