

Gaea System Details

From GFDL

Gaea System Details

To view the current state of gaea: please reference <http://www.ncrc.gov/>. This site will display the current status of compute nodes, login nodes, ldn's, rdtn's and the network file system. It also provides information regarding the last time a particular section was down.

Contents

- 1 Terminology
- 2 Current Hardware
- 3 System Architecture
- 4 Partitions
- 5 Filesystems
 - 5.1 Home
 - 5.2 Fast Scratch (FS)
 - 5.3 Long Term Scratch (LTFS)
- 6 Job Submission
- 7 Queues and Job States
- 8 Queue polices are as follows:

Terminology

Moab - The scheduler

Partition - A section of gaea that has its own scheduler. It is a logical unit in Moab.

Class - A Moab term for queue

DTN - data transfer node

Torque - The resource manager

CMRS - Another name for Gaea. You may see/hear this some place.

Current Hardware

- 30912 Cores
- 4 Login Nodes (Gaea1-4)
- RDTNS (remote data transfer nodes)
- LDTNS (local data transfer nodes)
- 3 Files Systems

System Architecture

Node types:

Compute Nodes

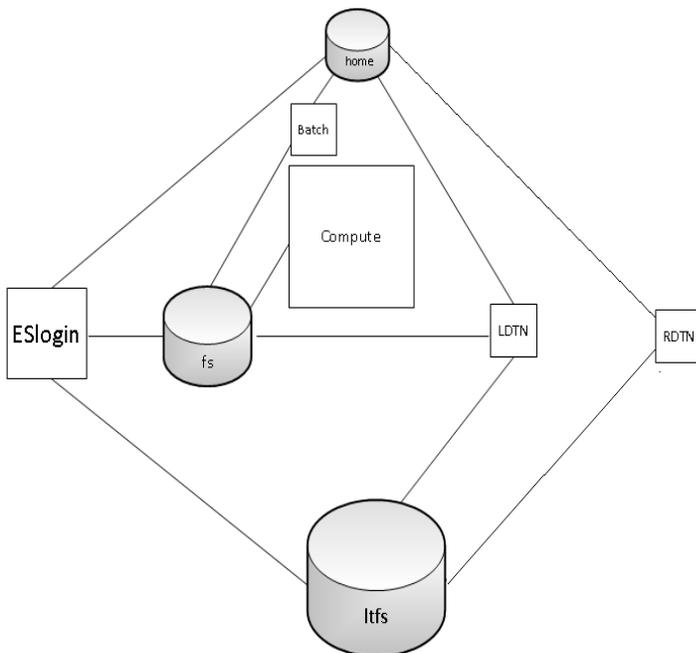
24 cores, 48GB memory, run model executable, filesystem mounts - FS

Batch Nodes

2 cores, 8GB memory, run scripts only (cores are not charged)

ESLogin Nodes

16 cores, 128GB memory, run interactive sessions, data jobs, Matlab



Partitions

t1ms - Testing and Development System (TDS) Partition is a partition that users usually will not use. As its name implies, it is used to test and develop new hard- and software. A few users may occasionally be asked to test models/codes/scripts on these test partitions. The TDS is currently split into two partitions. Currently one section is using the Gemini fabric for testing while the other is still using the C-Star fabric.

c1ms - Gaea's large compute partition

es - login nodes, local data transfer nodes (ldtn) and remote data transfer nodes (rdtn)

Examples-

```
msub -l partition=t1 ms scriptname
```

1. PBS -l partition=t1 ms

Filesystems

Gaea has 3 filesystems.

Home

The home filesystem is split into two sections which are backed up. There is a home1 and home2 for load balance purposes. Each user has a 5GB limit.

Home is mounted:

- Login nodes
- Batch nodes
- RDTN

A nightly snapshot can be accessed at `/ncrc/home1|2/.snapshot/nightly.0/$USER`

You can use this path to restore any files or sub-directories that are contained within that directory from last night. Use `nightly.1` for files from 2 nights ago. All files and sub-directories contained there will carry the same permissions as the originals. Users can simply copy from that location to any destination.

Fast Scratch (FS)

The FS is a 1PB lustre filesystem. User allocations are available at `/lustre/fs/scratch/$USER/`. All files over 2 weeks old will be scrubbed within the `/lustre/fs/scratch/$USER/` directories. This means files that have not been accessed or used in at least 2 weeks will be scrubbed. FS is NOT backed up. Users are responsible for monitoring their files and transferring what they do not want to lose to a location without a scrubbing policy.

FS is mounted:

- Login nodes
- c1ms (batch and compute nodes)
- LDTN

Long Term Scratch (LTFS)

The LTFS is a 3PB lustre filesystem. User allocations are available at `/lustre/ltfs/scratch/$USER/`. `/lustre/ltfs/stage` used for data transfers. The stage directory on the LTFS follows the same scrubbing policy as the FS scratch, files over 2 weeks old will be scrubbed. The LTFS scratch directories are not scrubbed. LTFS is NOT backed up.

LTFS is mounted:

- Login nodes
- LDTN
- RDTN

Job Submission

Please see the following Moab details.

There are two job types.

- Batch
 - Regular jobs - use msub
- Interactive/Debug
 - `msub -I -X -l partition=c1ms,size=48`

Queues and Job States

There are currently 3 different queues.

- Batch - no specification needed
- LDTN - data movement queue
- eslogin - compiles and data processing jobs

Example:

```
msub -q eslogin scriptname
```

Queue polices are as follows:

- Persistent - jobs that run continuously
- Urgent - heightened priority. Please see your group head for access to this queue.
- Novel - Jobs that require more than 25% of the system are held until a post preventive maintenance.

The following are job monitoring commands:

- showq
- checkjob - now applicable on jobs of other users
- showstart
- mdiag
- mjobctl

Retrieved from "http://wiki.gfdl.noaa.gov/index.php/Gaea_System_Details"

- This page was last modified on 13 July 2011, at 14:32.
- Privacy policy
- About GFDL
- Disclaimer

