The NDGF Oslo View

Miguel Afonso Oliveira, Thomas Röblitz, Dmytro Karpenko

February 11, 2013

Current Hardware
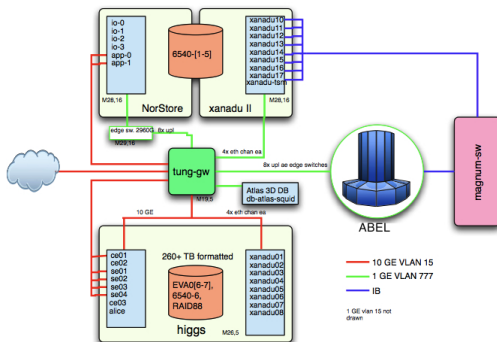
Software Updates

- Operating System updates:
  - Compute Elements: RHEL 5.9, 2.6.18-348.1.1.el5
  - Compute Nodes: CentOS 6.2, 2.6.32-279.5.1.el6.centos.plus
  - Storage Elements: RHEL 5.9, 2.6.18-238.1.1.el5
- ARC Middleware updates:
  - Compute Nodes: ARC 2.0.1
- dCache updates:
  - dcache-2.2.3-1

Monitoring

- Internal Nagios.
- Internal Ganglia.
- Central Nagios.
    - https://chaperon.ndgf.org/nagios/
    - https://mon-ngi.ndgf.org/nagios/
    - Our aggregate Nagios page.
- dCache - https://chaperon.ndgf.org:2288/

On site customisation

- **configure-SLURM-env.sh**
  - **Path to sacct:** We let the back-end know the path to sacct command, so then it can be issued by other scripts.
- **scan-SLURM-job**
  - **Memory report:** After the job is finished, we use sacct to populate the infosys parameter showing how much memory the job consumed during its execution.
- **Sysinfo.pm**
  - **gpfs support:** Available diskspace is determined by ndf command, that connects to the control server of disk pool and extracts information from it. Without modification, this script will execute ordinary df and thus get a wrong space report. Note, that ndf doesn't return mountpoint at the moment, so Sysinfo.pm has mountpoints hardcoded inside (/grid/nordugrid/$path).
- **cache-clean**
  - **gpfs support:** Available diskspace is determined by ndf command, that connects to the control server of disk pool and extracts information from it. Without modification, this script will execute ordinary df and thus get a wrong space report. Note, that ndf doesn't return mountpoint at the moment, so Sysinfo.pm has mountpoints hardcoded inside (/grid/nordugrid/$path).

On site customisation

- submit-SLURM-job
    - **Steering jobs to different projects:** It is implemented by determining the user account the job is composed for and requesting a corresponding SLURM project for the job.
    - **Reservation of the whole node:** This is an old NDGF requirement to provide the RTE that allows a job to request all the cores on one node. Our implementation is that RTE sets the corresponding environment variable, RESERVE_THE_FULL_NODE, and then it is checked in submit-SLURM-job if the variable is set. If it is, the job will request the maximum number of cores per node.
    - **Memory scaling:** For non-simulation and non-evgen jobs we scale the requested memory by 1.3 to prevent jobs from being killed by SLURM due to exceeding the requested default memory limit. Precautions are also taken not to leave joboption_memory empty, because SLURM refuses to submit jobs without memory specifications.
    - **Cleaning of job scratch:** Grid jobs use a scratch outside /work (namely /tmp) to run. The job cleans and removes the scratch after its execution, but only if it finishes successfully. If it fails A-REX cleans the sessiondir on the front-end, the scratch on the nodes is never cleaned. tmpwatch should take care of it, but not before the dir is 10 days old. So if there is massive failure of jobs on the node – /tmp partition can become completely full. To prevent this, we now manually insert removal of the scratch into .cleanup file that is always executed by SLURM when the job is terminated, regardless of the termination cause.

Problems Solved

- ▶ FHGFS could not be used - no support for hard links.

  We had to revert to the old Titan storage.

- ▶ OOM for ATLAS jobs.
  - ▶ We contained the ATLAS jobs within a rack on which cgroups are not active.
  - ▶ Many kernels were tried. Latest for RHEL 6.4 beta solves the OOM problem but introduces "BUG: soft lockup" problems....

Objectives

- Quickly alert to problems.
- Provide assistance.
- Minimize downtimes.
- Minimize tickets.

Operations

- ▶ NDGF Chat.
- ▶ Operations portal dashboard.
- ▶ Security portal.
- ▶ NDGF wiki.
- ▶ GGUS.
    - Our aggregate GGUS tickets page.
- ▶ GOCDB.
- ▶ Nagios
    - https://chaperon.ndgf.org/nagios/
    - https://mon-ngi.ndgf.org/nagios/
- ▶ NDGF Ganglia.
- ▶ NUNOC.
- ▶ dCache Service.
- ▶ FTS transfers.

Documentation and Information Sources

- Lack of documentation on procedures and priorities.

  Maybe we can help...

  https://wiki.uio.no/usit/suf/vd/hpc/index.php/NDGF_OOD

- Too many sources of info.

  We are trying to consolidate all source of info on one page...

  http://exp0.abel.uio.no/grid/

Upgrade Plans

- UiO will have a complete overhaul of its grid resources soon...
- New service hardware:
    - Computing Elements - in house....
    - Compute Nodes - ETA week 9...
    - Storage - being specked....