The Cray XT3 at CSCS
Environment and Configuration

Swiss National Supercomputing Centre

CSCS
Manno, September 27th, 2006
XT3 Support Teams

USER AND APPLICATION
Maria Grazia Giuffreda
Claudio Bareato
Francesco Benvenuto
Angelo Mangili
Claudio Redaelli

+ Cray Support

SOFTWARE DEVELOPMENT
Mauro Ballabio
Mark Cheeseman
Nello Nellari
Neil Stringfellow

SYSTEM ADMINISTRATION
Richard Alexander
Hussein El-Harake
Davide Tacchella
On line documentation

- https://www-users.cscs.ch/
- Log in with your CSCS username/password combination

The CRAY XT3 at CSCS: General Information

1. XT3 Hardware and System Software
2. Login, remote access
3. Home, Quota and Backup
4. Documentation/On-line manuals
5. XT3 News and Current System Status

XT3 Hardware and System Software

The CSCS XT3 system consists of 18 cabinets, containing 26 service processing elements (PEs), subdivided into 13 service blades, and 1664 compute PEs, subdivided into 416 compute blades. The system runs the UNICOS/α operating system, implementing the Linux kernel and the Catamount lightweight kernel.

Both kinds of PE are based on an AMD Opteron CPU running at a frequency of 2.6 GHz (5.2 GigaFlops of peak performance), service and compute nodes may access 4 and 2 GB of RAM respectively.

The whole XT3 compute system has therefore a peak performance of 8.65 TeraFlops, and totals of 3.3 TeraBytes of RAM.

Each Opteron processor is directly connected to a dedicated Cray SeaStar chip (based on the IBM Power architecture), providing a reliable interconnect; each SeaStar contains a 6-Port router and communications engine and also provides a serial connection to the Cray RAS and Management system.

The CSCS system is configured as a so-called Class 2 Topology, a full 3D torus of size 9×12×16.

The service PEs specialize by function, each one may be configured as a Login PE, an I/O PE, a System PE or a Network PE. They run a customized version of Linux as their operating system.
Current system news


- Current News
- Previous Entries

Current News

2006-09-27

Next planned weekly maintenance

Previous Entries

2006-09-15

Service nodes lost connectivity with the HPN (high-performance network) yesterday around 18:30, the system was rebooted at 20:30, the problem reoccurred again this morning, and was solved by warm rebooting the login nodes only. Compute jobs should not have been affected. CSCS and Cray are looking into the cause of the problem.

2006-09-11 17:00 / 20:00

Unplanned shutdown due to a PBS MOM node failure and related problems with the communication network. Three hours of planned downtime.

2006-09-04 17:00 / 19:00

Unplanned shutdown due to cooling problems in the machine room. Two hours of planned downtime, in the worst case scenario outages might occur in the next two days.

2006-09-02 15:00 / 17:00

An hardware problem on OSS ndi00027 (Object Storage Server, an I/O service node managing Object Storage Targets, or OSTs) is causing severe trouble with the Lustre filesystem on palu.cscs.ch. I/O operations may have experienced failures due to time outs, the relative error messages should look like:

```plaintext
1159841364.469617:0-1820:client.dh9@rhpc.expire_one_request(): 808:
```

In order to fix the issue the problematic a reboot was required.

2006-08-30

Programming environment on palu.cscs.ch upgraded to version 1.4.26, service nodes have been upgraded to dual-core processors.

2006-08-23/2006-08-25

The number of available batch nodes on getc.cscs.ch will be reduced to accommodate a parallel programming course being offered at CSCS. We ask regular CSCS users to not use the additional interactive nodes that have been added for the use of the course participants. Currently, users will be able to only submit jobs using up to 16 nodes for a maximum of 1 hour. We thank you for your understanding and apologize for any inconvenience that this may cause.
Getting support

- Send mail to: help@cscs.ch

- Many new/open requests on 26/9 (16/2):
  - hpc-rt: 15 (54) [system issues]
  - xt3-rt: 81 (36) [XT3 specific, many of the requests are internal]
  - pe-rt: 9 (10) [programming environment]
  - chem-rt: 3 (8) [chemistry applications]
  - fem-rt: 1 (2) [CFD/engineering applications]

- Somewhat long response times
Two stages

- **gele.cscs.ch (Stage 0)**
  - new users, porting, development, testing, debugging
  - 76 dual-core compute nodes (68 batch, 8 interactive)
  - 10 dual-core service nodes (2 login, 4 I/O, 1 PBS MOM)

- **palu.cscs.ch (Stage 1)**
  - production platform
  - 1664 single-core compute nodes (all batch)
  - 32 dual-core service nodes (4 login, 14 I/O, 3 PBS MOM)

- **fred (Stage 2)**
  - small platform for system administrators to play experiment: *off limits!*
  - 48 single-core nodes (ever changing configuration)
Strong Points and Limitations

• Plus
  • Fast interconnect (SeaStar)
  • Very scalable balanced architecture (»10,000 CPUs, parallel FS)
  • Well-supported performance measuring tools

• Minus
  • Relatively limited memory per CPU
  • Cross-OS environment
  • limited libc on compute nodes (not a complete Linux/UNIX implementation)
  • Recent architecture (growth problems)
Stage 1 configuration

- 18 cabinets
  - approximately 96 nodes per cabinet
  - AMD Opteron 2.6 GhZ
- 4 login nodes with DNS rotary name
- Software maintenance workstation
  - Programming Environment 1.4
  - Suse Linux with a 2.6 kernel
  - PBS/Pro 5.3
  - TotalView 7.2
  - PGI Compilers 6.1.4 / ACML 3.0
File systems

- /users  \$\{HOME\}
  - home directories “/users/\$\{USER\}”, shared on palu and gele
  - NFS mounted (slow, reliable)
  - permanent storage, backed up, limited disk quota (≥ 1940 Mbytes)

- /scratch  \$\{SCRATCH\}
  - work directories “/scratch/\$\{USER\}”, system specific
  - Lustre parallel filesystem (fast, unreliable)
  - temporary storage, not backed up, no disk quota, cleanup policy

- /archive  \$\{ARCHIVE\}
  - archive directory “/users/\$\{USER\}/archive”, CSCS-wide
  - NFS mounted (slow, tape archive featuring disk caching), FTP, scp
  - long-term storage
**Lustre file system**

- Genuine parallel file system
- Great performance (at least in theory)
- Under development: immature feature set
- Cray is continuously improving Lustre
- When Lustre is unhealthy the system must be rebooted
- It is difficult to differentiate actual Lustre errors from problems with disks and/or controllers
Job Scheduling

- PBS/Pro batch system version 5.3.4
  - There are several limitations in the distributed version
- Custom scheduler maintained at CSCS
  - Written jointly in TCL by Cray, Altair and CSCS
  - Two resource pools: `walltime` and `size` (i.e. CPUs)
- Implemented policies:
  - Priority depends on used budget vs. allocated budget
  - Priority to large jobs (may cause resource draining)
  - Backfilling (small jobs are fitted in the “holes” left by large jobs)
  - Reservations (some jobs must run at a given time no matter what)
- To do:
  - Merge accounting on Palu and Gele
  - Limit the maximum number of queued jobs per user
### PBS configuration

- Jobs are never rejected
  - If the allocated budget is exhausted, jobs are queued at low priority and get to run when the system would otherwise be idle

- Large jobs are favored

- Small jobs are backfilled around large jobs
  - A low “walltime” value allows a better backfill and shortens the waiting time in the queue

- Some jobs can be set to start at a given time, in the meantime they drain resources and jobs are backfilled around them

- The number of jobs per user eligible for execution will be limited, excess jobs will be held, to be automatically submitted when another one completes
Example PBS job

#!/usr/bin/ksh
#PBS -l size=128,walltime=08:00:00
#PBS -N test_128
#PBS -e test_128.stderr
#PBS -o test_128.stdout
#PBS -q production
#PBS -V
#PBS -m abe
#PBS -M <your@email_address>
# WORKING directory
export WRK=$PBS_O_WORKDIR
cd $WRK
# set the SCRATCH directory
export SCR=/scratch/${USER}/${PBS_JOBNAME}
mkdir -p $SCR
(...)
yod -np $PBS_NNODES <application>
Budget Quotas

• Budgets are not automatically enforced, fair share is ensured by queueing policies.

• A command line tool for checking used vs. allocated monthly budget ("sbuchek") is planned, but development has not yet begun.

• In the meantime budgets can be checked via the www-users web site.
Supported Software

- The fast system software update cycle often makes it difficult to fully support and test application software at each PE release

- ACML is included with the Programming Environment, providing:
  - BLAS, LAPACK, FFTs

- More numerical libraries are supported by CSCS:
  - FFTW, METIS, parMETIS, LibSci, netCDF, HDF5...

- Applications
  - AMBER, CPMD, NAMD, ECHAM5...

- More software is available (i.e. Trilinos) even if not centrally installed
<table>
<thead>
<tr>
<th>Software provided at CSCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>• PGI V.6.1.4 (and 6.0.8)</td>
</tr>
<tr>
<td>• GCC V.3.3</td>
</tr>
<tr>
<td>• MPICH2 and SHMEM</td>
</tr>
<tr>
<td>• TotalView V.7.3.0 (and 7.2.0)</td>
</tr>
<tr>
<td>• ACML V.3.0</td>
</tr>
<tr>
<td>• FFTW V.2.1.5 and V.3.0.1</td>
</tr>
<tr>
<td>• Metis V.4.0.1</td>
</tr>
<tr>
<td>• ParMetis V.3.1.0</td>
</tr>
<tr>
<td>• NetCDF V.3.6.1</td>
</tr>
<tr>
<td>• LibSci V.1.4</td>
</tr>
<tr>
<td>• HDF5 V.1.6.5</td>
</tr>
<tr>
<td>• PAPI 3.2.1</td>
</tr>
<tr>
<td>• PBS-Pro V5.3.4</td>
</tr>
<tr>
<td>• Amber 9.0</td>
</tr>
<tr>
<td>• CPMD 3.11.1</td>
</tr>
<tr>
<td>• v-Espresso 3.1.1</td>
</tr>
<tr>
<td>• Gamess/US V.R2 22.02.06</td>
</tr>
<tr>
<td>• NAMD V.2.6b2</td>
</tr>
<tr>
<td>• aLMo</td>
</tr>
<tr>
<td>• ECHAM V5.3.02</td>
</tr>
<tr>
<td>• ECHAM-HAM V5.3.02/1.52a</td>
</tr>
<tr>
<td>• CCSM 3</td>
</tr>
</tbody>
</table>
## Modules provided by CSCS

- amber/9.0
- cpmd/3.11.1
- echem/5.3.02
- espresso/3.1.1
- fftw/2.1.5
- fftw/3.1.1
- gamess/22.11.06
- hdf5/1.6.5
- memcom/7.2.5
- metis/4.0
- namd/2.6b2
- netcdf/3.6.1
- parmetis/3.1
- petsc/2.2.1
- szip/2.0
- vasp/4.6
- wien2k/06.4
- zlib/1.2.3
modules and the environment

- **module list** (shows loaded modules)
- **module avail** (shows available modules)
- **module load / add** (loads a new module)
  - module load totalview
  - module load craypat
- **module swap** (exchanges two modules)
  - module swap PrgEnv-pgi PrgEnv-gnu
- **module unload / rm** (removes a module)
  - module unload totalview
- **module update** (refreshes the environment)
Excellent scalability properties, the largest tested case (T106) reaches 224 GFlops (6.7% of peak) on 640 CPUs with a parallel efficiency almost reaching 70%.
Preventive Maintenance

- On a weekly basis (on Wednesdays)
- Includes a planned reboot to put the system back into shape
- May include a system software/hardware upgrade
- Planned downtimes will be announced, as far as possible, in `/etc/motd` and on the news/status page on `www-users`
Issues

• The upgrade from PE 1.3 to PE 1.4 negatively affected the overall stability
  • support for dual-core processors

• The High-Speed Network (HSN) is not yet 100% stable
  • temporary problems may lead to permanent outages

• The Lustre parallel file system is still immature
  • write performance is poorer than expected

• PBS has limited functionality
  • CSCS installed a custom enhanced scheduler

• Bugs lead to intermittent node failures
  • under control in 1.3, problems increased again with 1.4
Future

- Software upgrade (PE 1.4.32 or 1.4.33)
- Single node reboot
  - the current experimental implementation may occasionally hang the system
- Enable dual-core on service nodes
  - dual-core service nodes are running in single-core mode
- Test Linux on compute nodes (maybe)
Summary

- System is available since January for production and development work
- System is slowly getting more mature, CSCS and CRAY hold weekly conference calls where open issues are discussed
- Several issues are still open
  - Parallel file system maturity
  - Scheduling system (updates, time outs, cleanup)
  - Node failures

*There are still more interruptions than we would like*