



CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Systems-Level Configuration and Customization of Hybrid XC30

CUG 2014

Nicola Bianchi

Colin McMurtrie

Sadaf Alam

Swiss National Supercomputing Centre (CSCS)

Agenda

- **CSCS Hybrid XC 30 Platform**
- **Configuration details**
- **Customization details**
- **Conclusions**



What system did we get ?

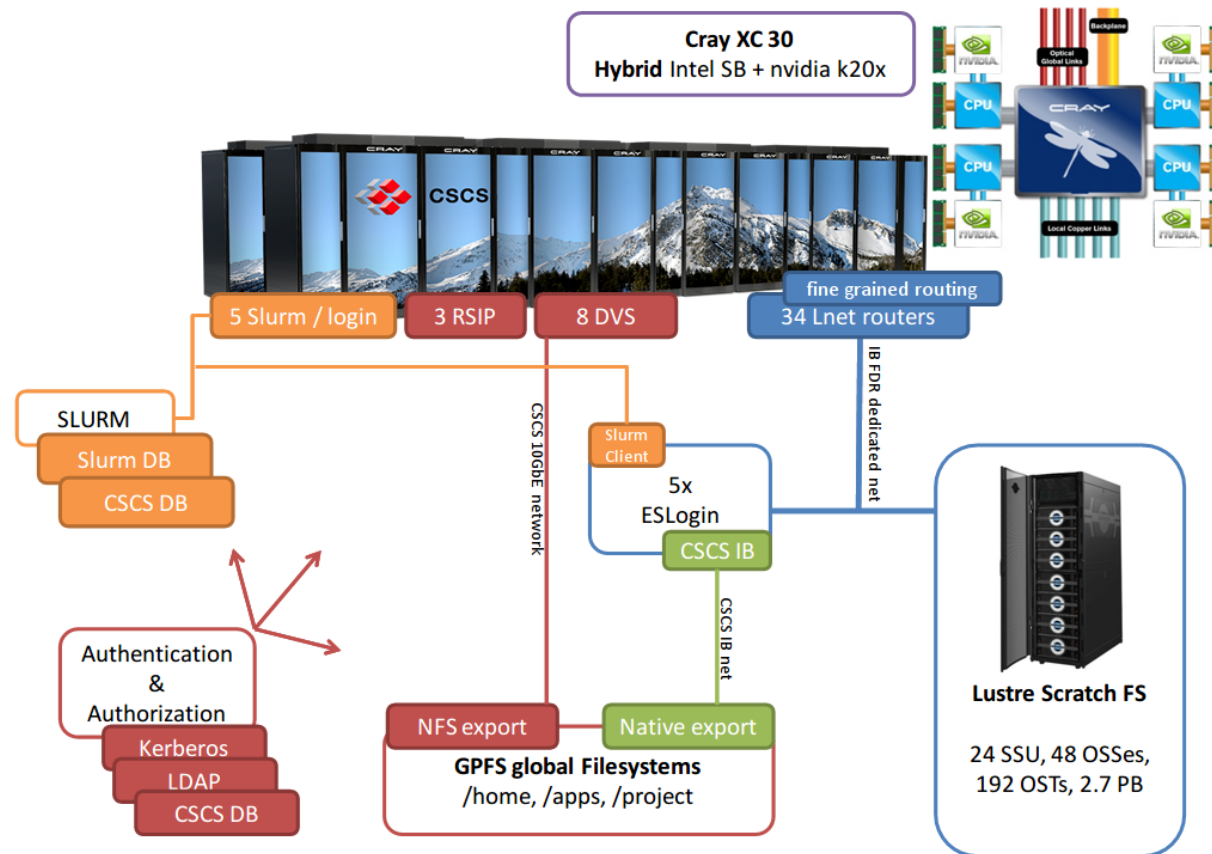
XC30 Platform at CSCS

- **Currently largest Hybrid XC30 worldwide**
- **28 cabinets**
 - 5272 compute nodes
 - 52 service nodes
- **5 EsLogin servers**
- **24 SSUs Sonexion1600 Lustre Appliance**
- **SLURM workload manager**



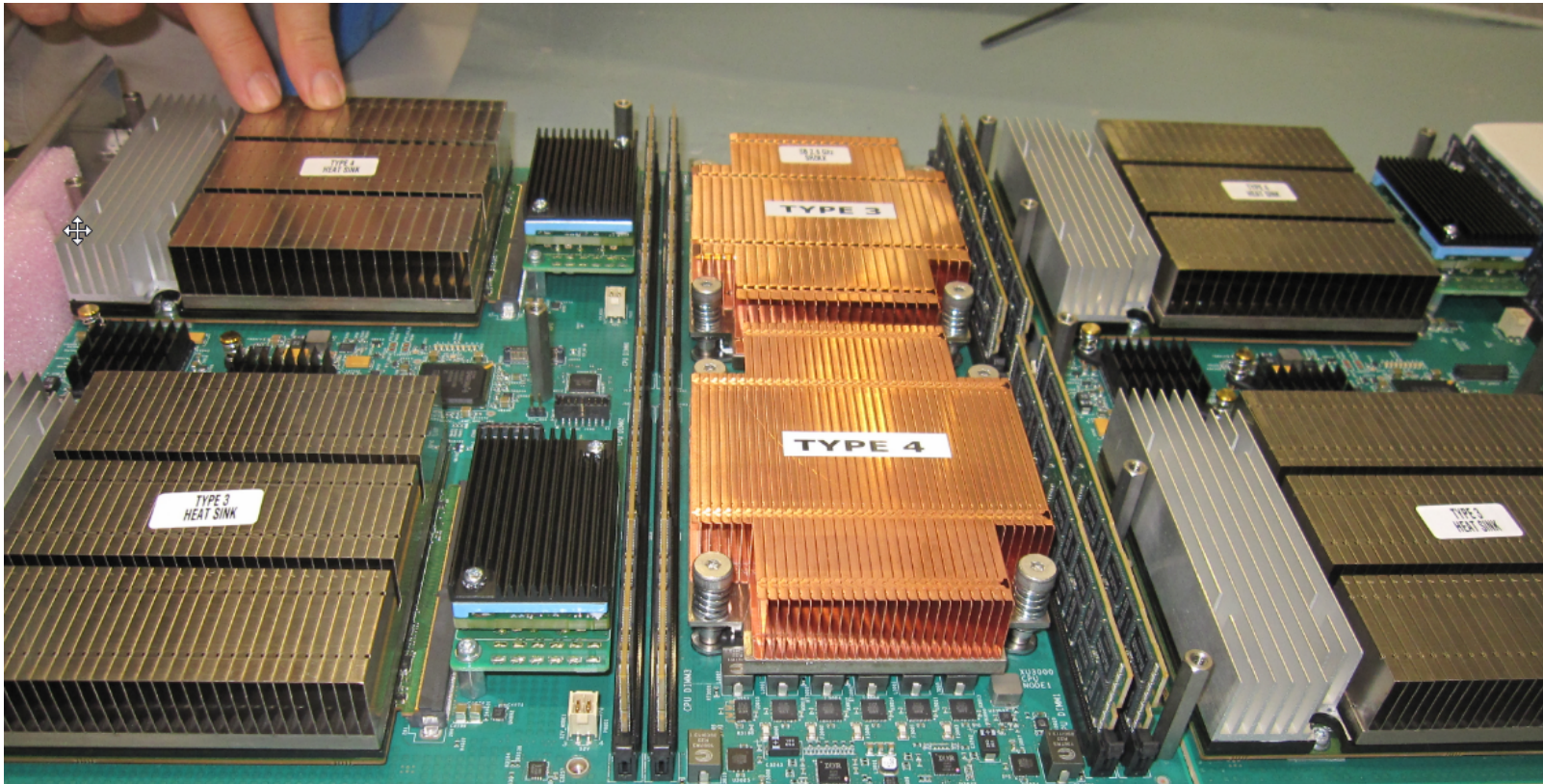
Integration overview ...

XC30 Platform integration



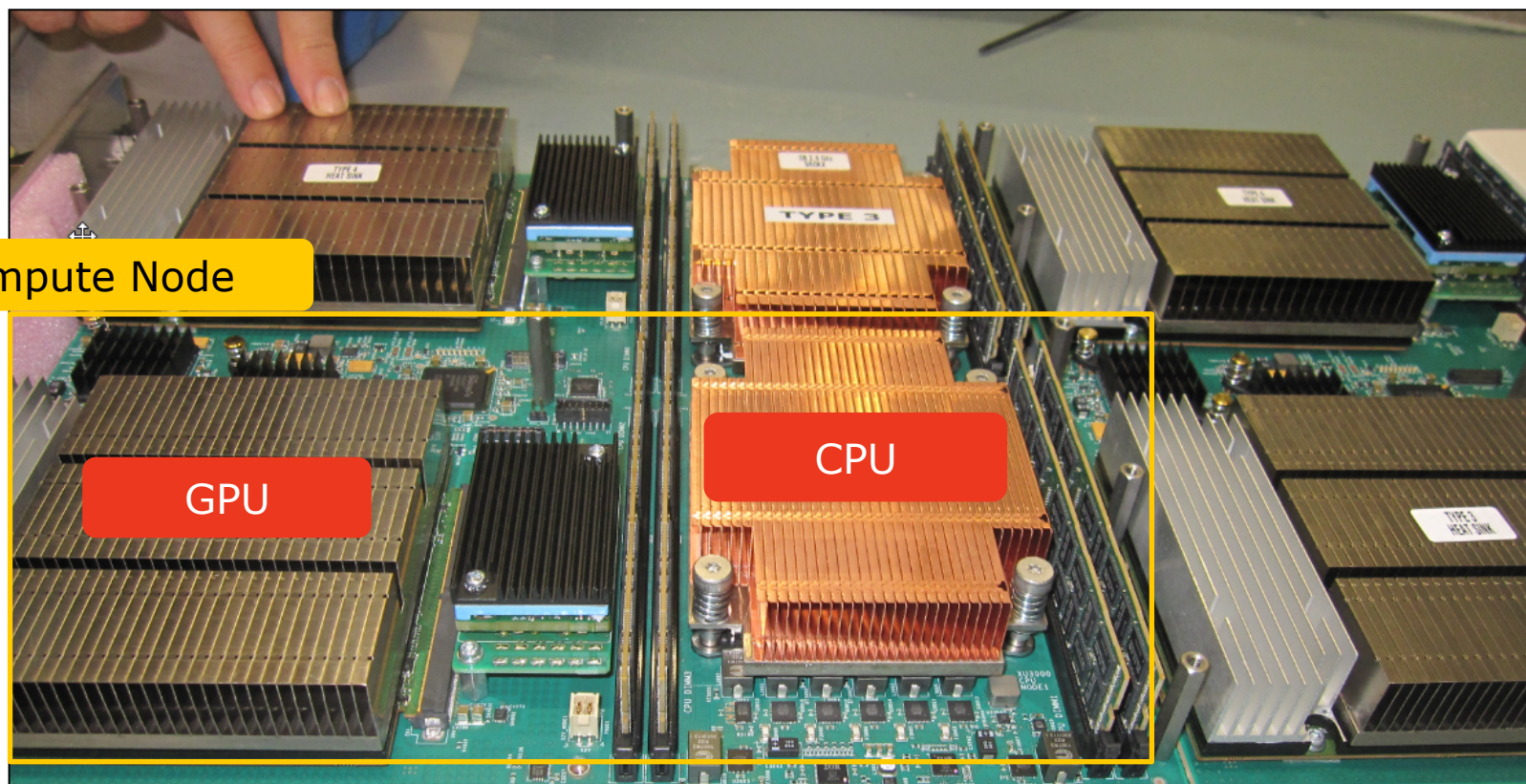
What does a hybrid node look like?

Hybrid compute blade



What's inside?

Hybrid compute blade



How to manage and monitor?

System tools & monitoring: nvidia-healthmon

- **A tool to perform health diagnostics for a system. This program returns a non-zero exit code if a problem is detected that MAY prevent any targeted GPU from running a GPU job**

Loading Config: SUCCESS

Global Tests

Black-Listed Drivers: SUCCESS

Load NVML: SUCCESS

NVML Sanity: SUCCESS

Tesla Devices Count: SKIPPED

Global Test Results: 3 success, 0 errors, 0 warnings, 1 did not run

0000:01:00.0

NVML Sanity: SUCCESS

InfoROM: SKIPPED

Multi-GPU InfoROM: SKIPPED

ECC: SUCCESS

PCIe Maximum Link Generation: SUCCESS

PCIe Maximum Link Width: SUCCESS

- **Implemented in a simple Node Health Check (NHC) script**

GPU configuration tricks...

System tools & monitoring: GPU operating mode

- **Allow users to use privileged command through the resource manager SLURM**
- **Set up the SLURM prolog/epilog scripts to configure the GPU with specific features:**
 - “gpu mode default”
 - “clock frequency”
- **Usually the prolog/epilog scripts relay on the nvidia-smi tool capabilities**
- **Simple usage for users**
 - `salloc -N 4 -C gpumoddefault`

Resources ?

System tools & monitoring: RUR

- **RUR: Resource Usage and Reporting**

- Cray tool for gathering statistics on how system resources are utilized by applications
- Modular approach, plugin for input and output
- Writing plugins is relatively simple (python)

- **We report back directly to the users the information we collect**

uid: 21542, apid: **2293684**, jobid: 264355, cmdname: bibw_621_daint **taskstats** ['utime', 5624000, 'stime', 524000, 'max_rss', 88028, 'rchar', 725336, 'wchar', 2624, 'exitcode:signal', ['0:0'], 'core', 0]

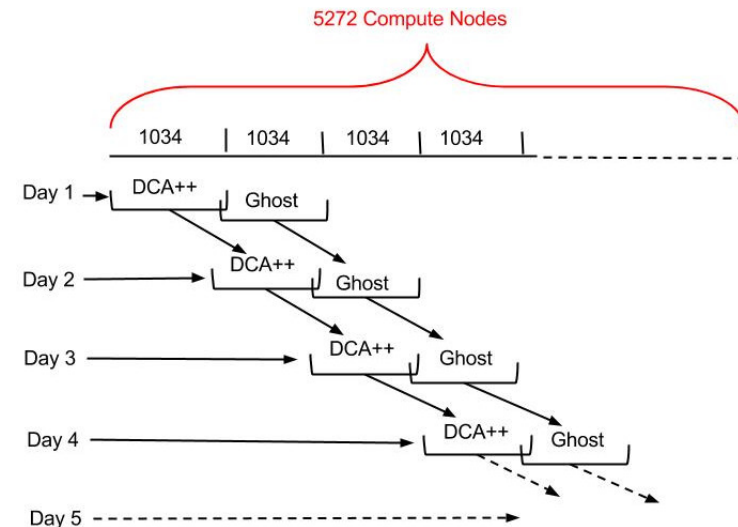
uid: 21542, apid: **2293684**, jobid: 264355, cmdname: bibw_621_daint **gpustat** ['maxmem', 79167488, 'summem', 158334976, 'gpusecs', 5]

uid: 21542, apid: **2293684**, jobid: 264355, cmdname: bibw_621_daint **energy** ['energy_used', 931]

Check the machine status

System screening mechanism

- **We needed a method to screen the health of each node of the system during production**
 - Exercise each node (x hrs)
 - Transient problems
 - Diag tools not applicable
- **We wanted an easy to use mechanism**
 - Transparent to users
- **SLURM reservations made the job**
 - Submission of hundred of jobs inside the reservation and the system would take care of run them around the system



Regression Testing

Holistic Regression Suite

- **CSCS has developed a regression test suite for Piz Daint that provides an overview of system health over a broad range of metrics including both the HW and the SW configurations**
- **Used as reference after PE installation, maintenance intervention and software upgrades**

```
robinson@daint102:~/daint-regression> ./run_regression
Regression test started by robinson
==> start date Tue May 6 09:35:09 CEST 2014
Running on daint external login ( daint102 )
Regression testing output will appear in:
/users/robinson/daint-regression/output/06-05-2014_09-35-09/edaint

Test 1000: Check system health and sanity
=====
| Checking checkSystem.out was last modified within last 2 hours [ OK ]
| Checking the load average [ OK ]
| Checking for zombie processes [ OK ]
| Checking ntpd process [ OK ]
| Checking free space in / [ OK ]
| Checking free space in /snv [ OK ]
| Checking free space in /rr [ OK ]
| Checking if /apps is mounted on all nodes [ OK ]
| Checking if /users is mounted on all nodes [ OK ]
| Checking if $SCRATCH is mounted on all nodes [ OK ]
| Checking the service nodes [ OK ]
| Checking the DVS nodes [ OK ]
| Checking lustre [ OK ]
| Checking lnet router nodes [ OK ]
| Checking login nodes [ OK ]
| Checking down nodes [ OK ]
| Checking scratch usage [ OK ]
| Checking slurm frontend nodes [ OK ]
| Checking slurm partition [ OK ]
| Check system health and sanity [ PASSED ]

Test 5040: Modules loaded at login match for ext and int login nodes
=====
| Modules loaded at login match for ext and int login nodes [ PASSED ]

Test 5041: Modules in PrgEnv-intel match for ext and int login nodes
=====
| Modules in PrgEnv-intel match for ext and int login nodes [ PASSED ]

Test 5042: Modules in PrgEnv-gnu match for ext and int login nodes
=====
| Modules in PrgEnv-gnu match for ext and int login nodes [ PASSED ]
```

... and more ...

Holistic Regression Suite

```
robinson@daint102:~/daint-regression> ./run_regression
Regression test started by robinson
==> start date  Tue May 6 09:29:36 CEST 2014
Running on daint external login ( daint102 )
Regression testing output will appear in:
/users/robinson/daint-regression/output/06-05-2014_09-29-36/edaint

Test 5004: Compile hello world in PrgEnv-cray
=====
| Compile hello world in PrgEnv-cray                [ PASSED ]

Test 5009: craype-accel-nvidia35 resolves to libsci_acc in PrgEnv-gnu
=====
| craype-accel-nvidia35 resolves to libsci_acc in PrgEnv-gnu  [ PASSED ]

Test 5014: Compiler generates AVX instructions PrgEnv-cray
=====
| craype-sandybridge module loaded                  [ OK ]
| Compiler flags for AVX found                      [ OK ]
| Compilation                                       [ OK ]
| AVX instructions generated in assembly code      [ OK ]
| Compiler generates AVX instructions PrgEnv-cray  [ PASSED ]

Test 5018: libsci_acc library symlinks in /opt/cray/lib64 for PrgEnv-cray
=====
| libsci_acc library symlinks in /opt/cray/lib64 for PrgEnv-cray [ PASSED ]

Test 5100: Compile and run HDF5 test case with cray-hdf5 in PrgEnv-cray
=====
| Compile and run HDF5 test case with cray-hdf5 in PrgEnv-cray [ PASSED ]

Test 5203: Run existing NetCDF executable r/w in PrgEnv-cray
=====
| Run existing NetCDF executable r/w in PrgEnv-cray [ PASSED ]

Test 5200: Compile and run NetCDF r/w with cray-netcdf in PrgEnv-cray
=====
| Compile and run NetCDF r/w with cray-netcdf in PrgEnv-cray [ PASSED ]
```

Conclusions ...

Conclusions

- **The first hybrid Cray XC30 system incorporates unique features in its system administration, management, monitoring and accounting environments.**
- **In future, we plan on continuing our collaboration with Cray and Nvidia to improve coverage of their system management and diagnostics tools**
- **CSCS will continue to invest in the regression suite so as to improve on early detection and diagnosis.**
- **We are currently investigating additional GPU operating modes and features, specifically for OpenGL, to facilitate interactive visualization using Piz Daint compute nodes.**



Q&A
