CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Early experiences with the Cray XK6 hybrid CPU and GPU MPP platform

Sadaf Alam, Jeffrey Poznanovic, Ugo Varetto and Nicola Bianchi (Swiss National Supercomputing Centre), Antonio Penya (UJI) and Nina Suvanphim (Cray Inc.)

Cray User Group Meeting
April 29 – May 3, 2012

Photo Gallery (more at: http://www.cscs.ch/)

(Cray XK6, before March '12)

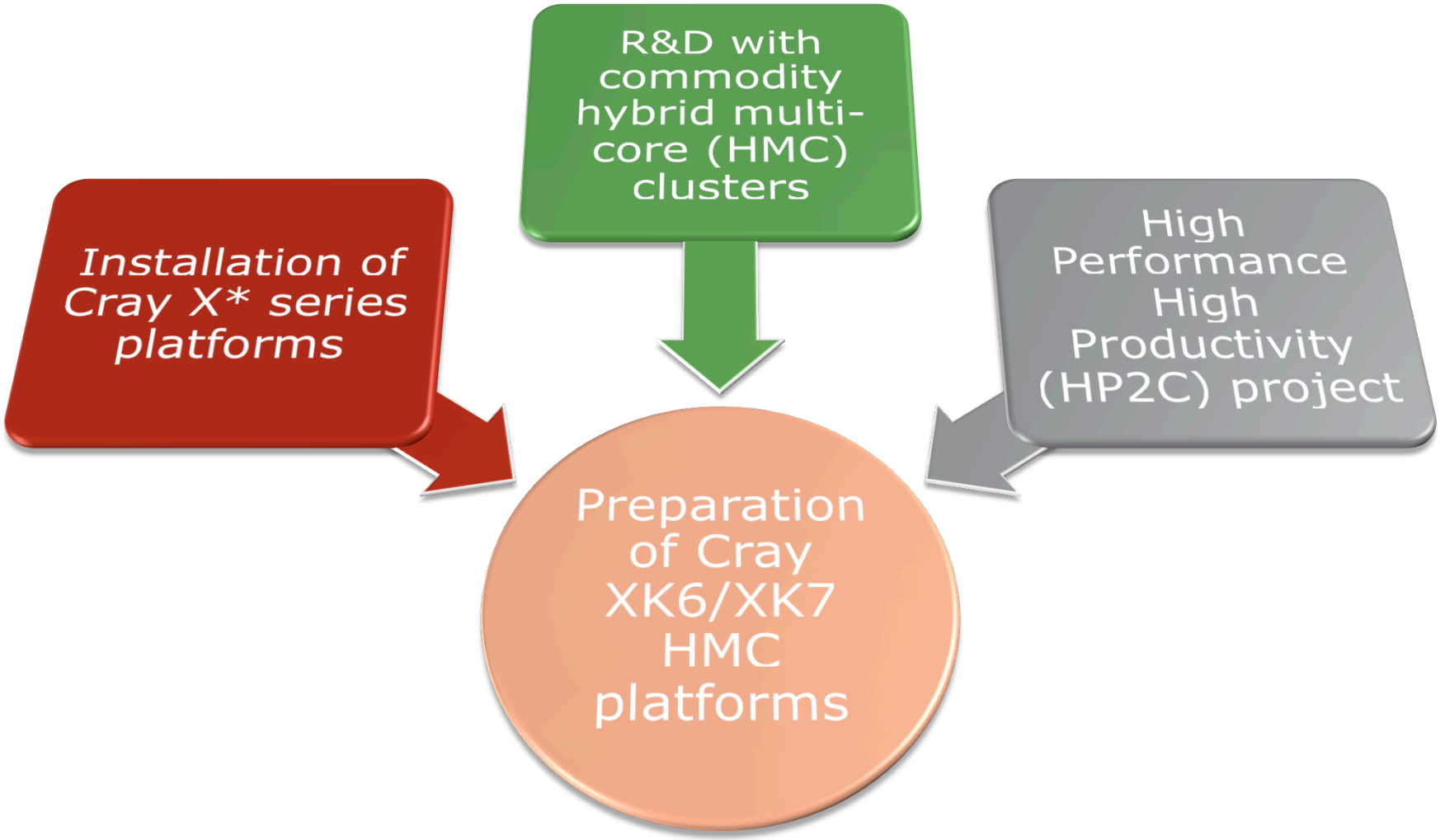
(Cray XE6, before March '12)

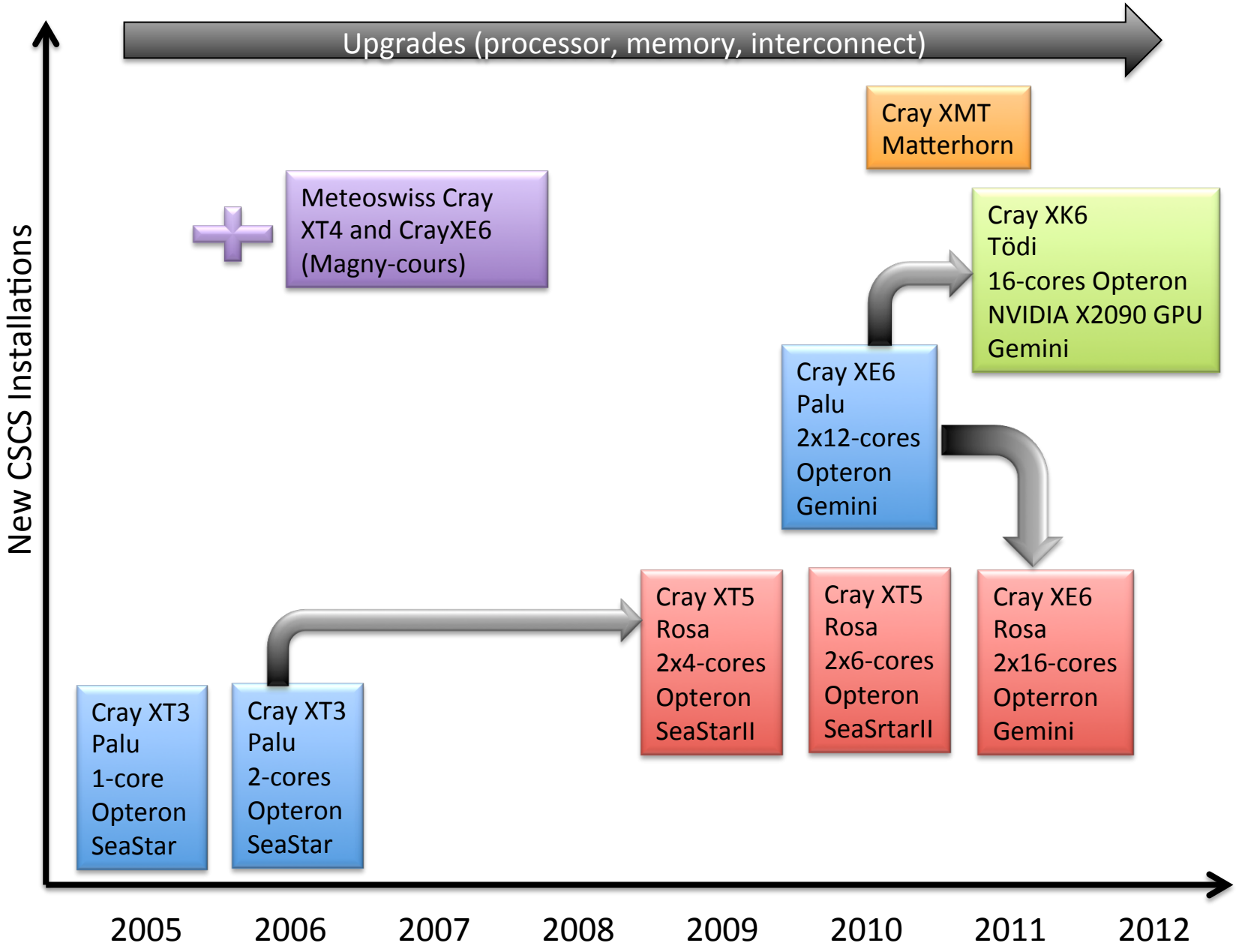
(New building, 2012)

(New machine room)

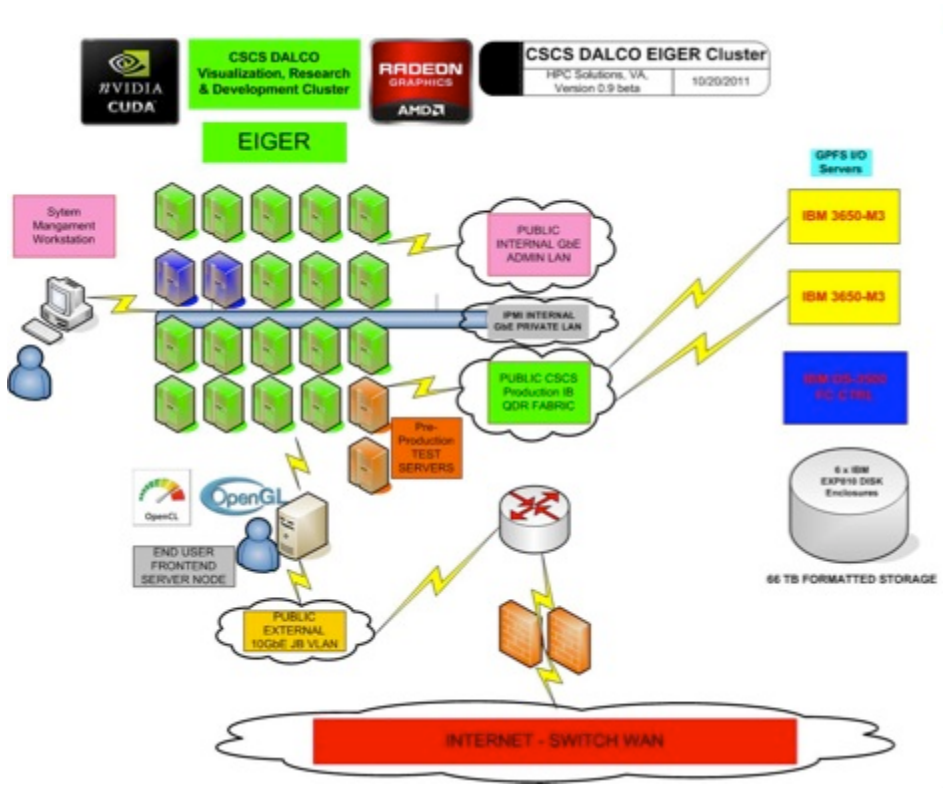
(Lake water cooling)

PROJECT PLANNING AND EXPERIENCES

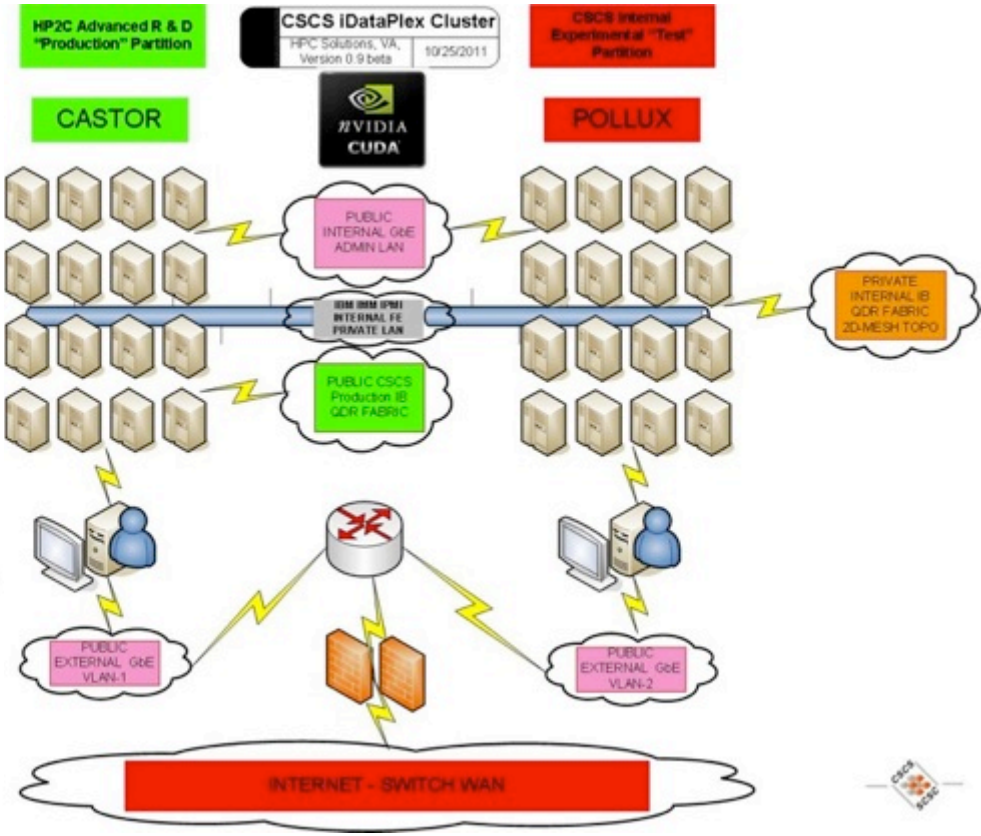



Upgrades (processor, memory, interconnect)
New CSCS Installations
Cray XMT
Matterhorn
Meteoswiss Cray
XT4 and CrayXE6
(Magny-cours)
Cray XK6
Tödi
16-cores Opteron
NVIDIA X2090 GPU
Gemini
Cray XE6
Palu
2x12-cores
Opteron
Gemini
Cray XT5
Rosa
2x4-cores
Opteron
SeaStarII
Cray XT5
Rosa
2x6-cores
Opteron
SeaSrtarII
Cray XE6
Rosa
2x16-cores
Opterron
Gemini
Cray XT3
Palu
1-core
Opteron
SeaStar
Cray XT3
Palu
2-cores
Opteron
SeaStar
2005
2006
2007
2008
2009
2010
2011
2012

2010-present



2011-present



CPU nodes: dual socket 6-cores AMD Istanbul, 12-cores AMD Magny-cours

GPU devices: NVIDIA M2050, C2070, S1070, GTX 480, GTX 285

Interconnect: Infiniband QDR

CPU: dual socket Intel Westmere

GPU: NVIDI M2090

Interconnect: InfiniBand QDR (2 networks)

Platforms for programming, system software, management and operational R&D

CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

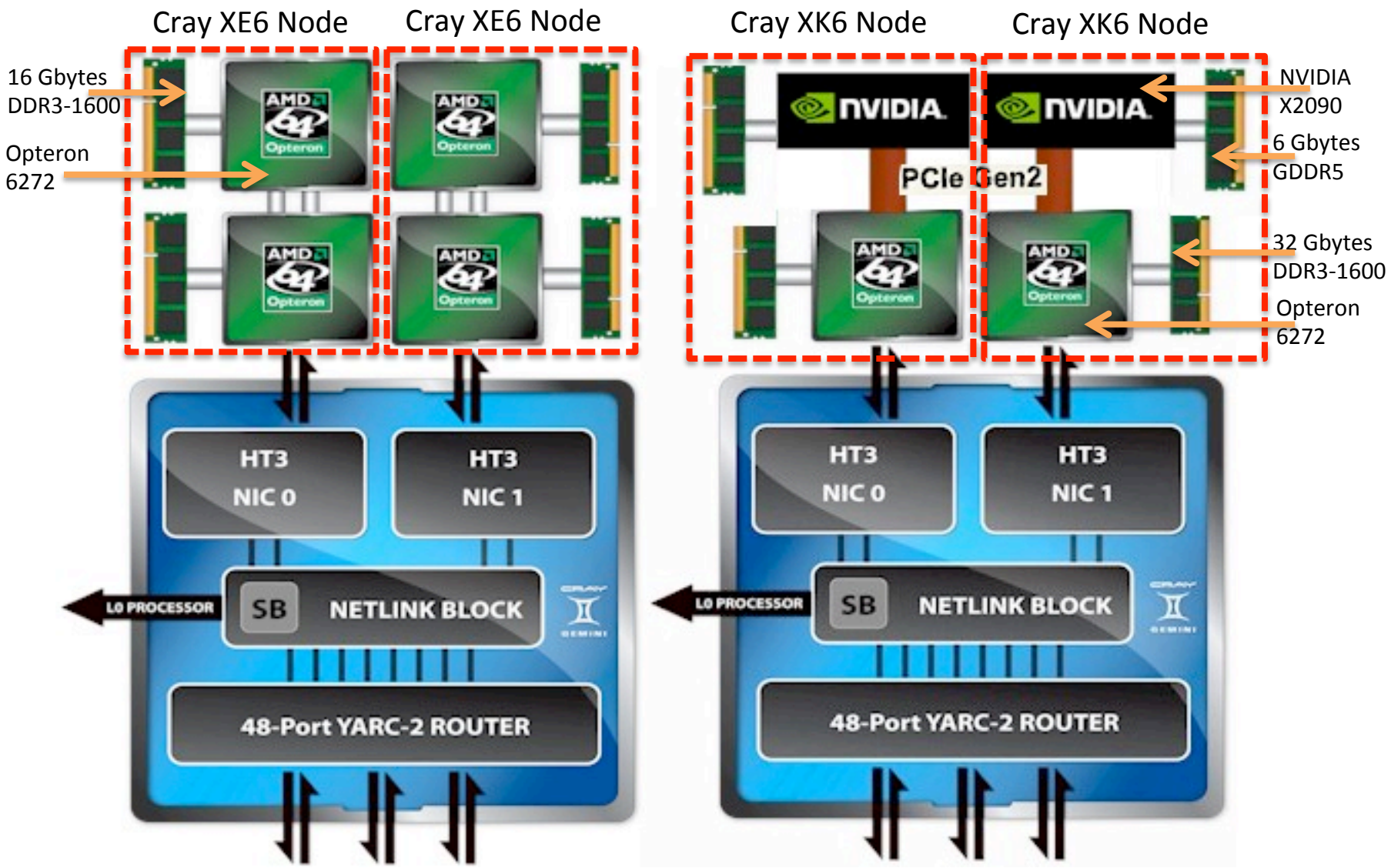
High Performance High Productivity (HP2C)

- **BigDFT - Large scale Density Functional Electronic Structure Calculations in a Systematic Wavelet Basis Set; Prof. Stefan Goedecker, University of Basel**
- **Cardiovascular - HPC for Cardiovascular System Simulations; Prof. Alfio Quarteroni, EPF Lausanne**
- **COSMO-CCLM - Regional Climate and Weather Modeling on the Next Generations High-Performance Computers: Towards Cloud-Resolving Simulations; Dr. Isabelle Bey, ETH Zurich**
- **Cosmology - Computational Cosmology on the Petascale; Prof. George Lake, Uni Zürich**
- **CP2K - New Frontiers in ab initio Molecular Dynamics; Prof. Juerg Hutter, Uni Zürich**
- **Ear Modeling - Numerical Modeling of the Ear: Towards the Building of new Hearing Devices; Prof. Bastien Chopard, University of Geneva**
- **Gyrokinetic - Advanced Gyrokinetic Numerical Simulations of Turbulence in Fusion Plasmas; Prof. Laurent Villard, EPF Lausanne**
- **MAQUIS - Modern Algorithms for Quantum Interacting Systems; Prof. Thierry Giamarchi, University of Geneva**
- **Neanderthal Extinction - Individual-based Modeling of Humans under climate Stress; Prof. C. P. E. Zollikofer, University of Zurich**
- **Petaquake - Large-Scale Parallel Nonlinear Optimization for High Resolution 3D-Seismic Imaging; Prof. Olaf Schenk, Uni Basel**
- **Selectome - Selectome, looking for Darwinian Evolution in the Tree of Life; Prof. Marc Robinson-Rechavi, University of Lausanne**
- **Supernova - Productive 3D Models of Stellar Explosions; Prof. Matthias Liebendörfer, University of Basel**

Outline

- **Comparison of Cray XK6 vs. Cray XE6**
  - Node and system architecture
  - Programming and operating environment
  - System management and control
- **Issues unique to Cray XK6**
- **Benchmarking results**
- **Applications status**
- **Summary and future outlook**

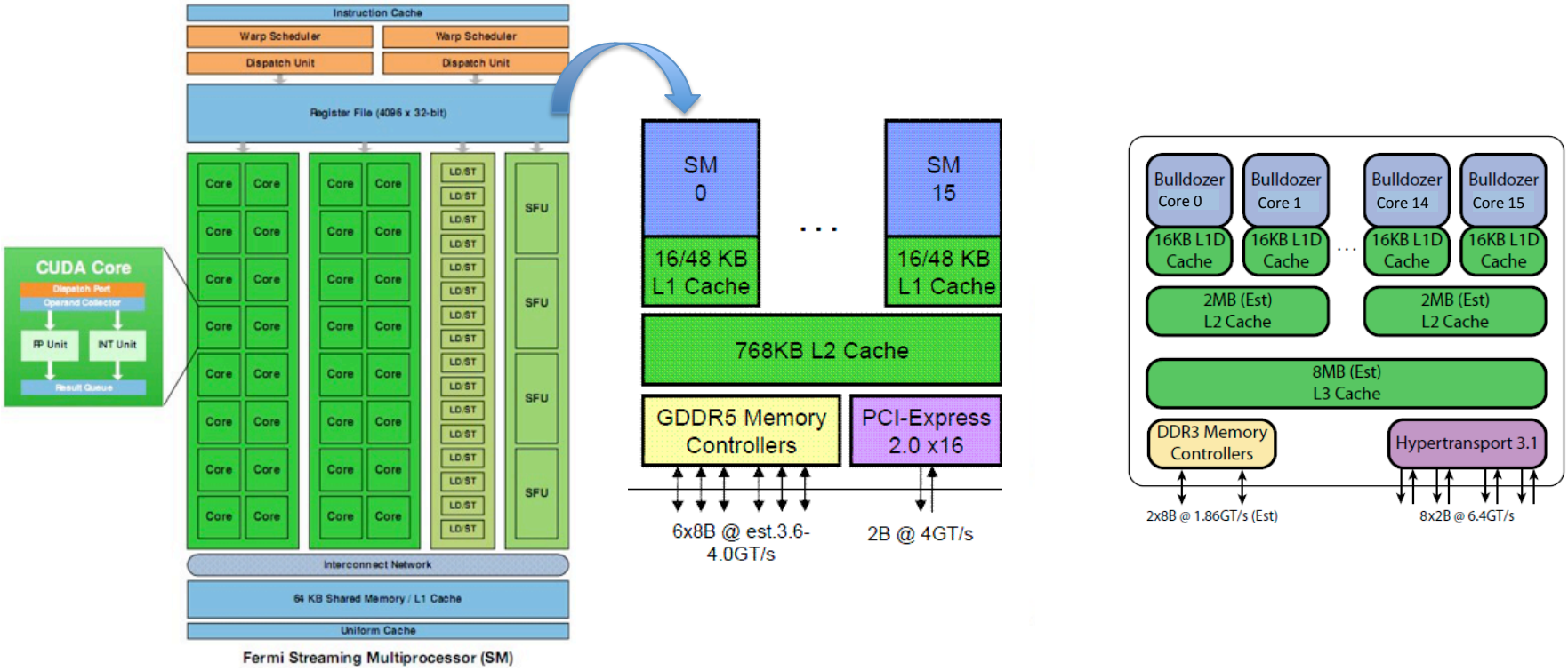
Node Architecture



CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

Cray XK6 vs. XE6 (Compute Node Characteristics)

	Cray XK6 (Tödi)	Cray XE6 (Monte Rosa)
Cores	512	16
Clock frequency	1.15 GHz	2.1MHz
FP performance	665 GFlops (double-precision)	134.4 GFlops (double-precision)
Memory interface	GDDR5	DDR3 (1600)
Power envelope	225-250 W	90-115 Watts

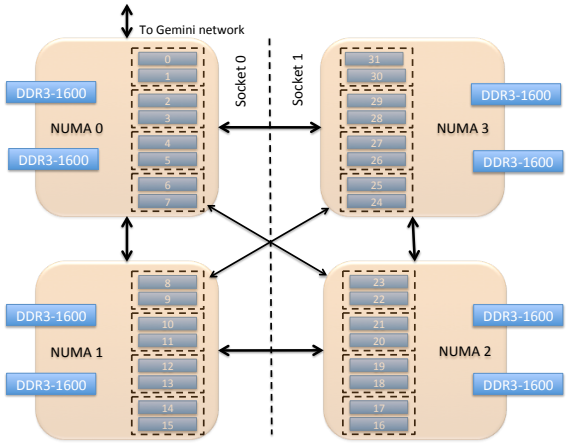


Cray XK6 vs. XE6 (Memory Characteristics)

	Cray XK6 (Tödi)	Cray XE6 (Monte Rosa)
L1 cache (size)	16-48 KB	16 KB
L1 (sharing)	SM (32 cores)	Core
L2 cache (size)	768 KB	2028 KB
L2 (sharing)	All SMs	Module (2 cores)
L3 cache	--	8 MB
L3 (sharing)	--	Socket
Shared memory	16-48 KB per SM	--
Global memory	6 GB	32 GB

Stream Copy BW (GPU) = ~133 GB/s

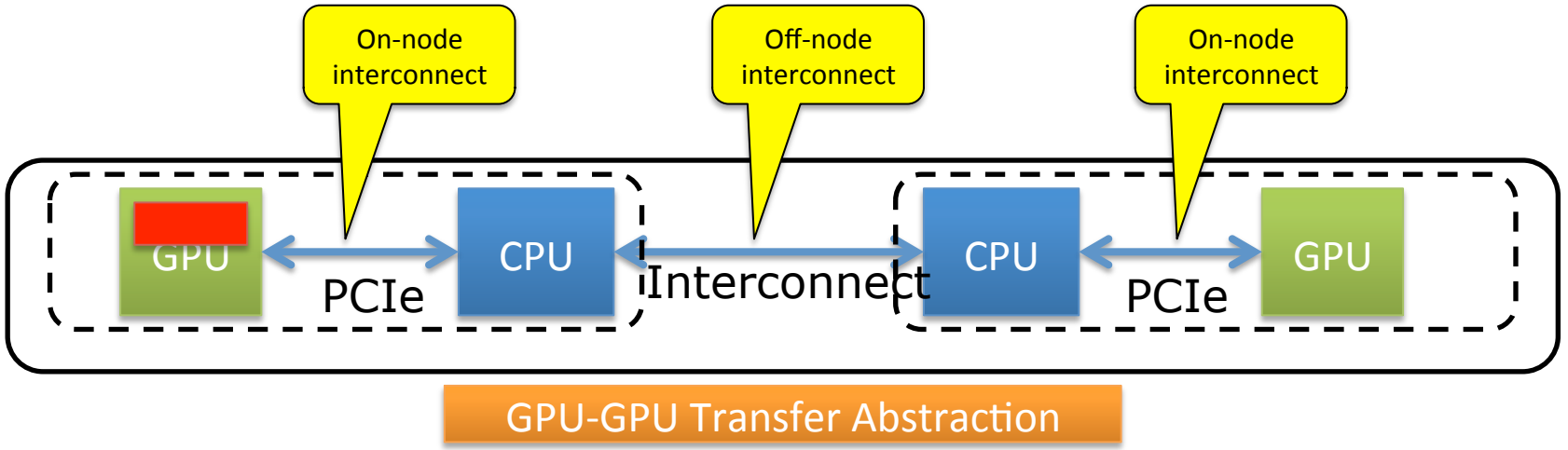
Stream Copy BW (CPU) = ~ 33 GB/s



Ecosystem (Cray XK6 and Cray XE6)

	Cray XK6	Cray XE6
Interconnect	Gemini	Gemini
Parallel file system	Lustre	Lustre
Operating system	CLE	CLE
Host compilers	Cray, GNU, Intel, PGI, Pathscale	Cray, GNU, Intel, PGI, Pathscale
Accelerator compilers	NVIDIA, Cray, PGI	--
Job scheduler	SLURM	SLURM
MPI library	Cray MPT	Cray MPT
Numerical libraries	Cray libsci including libsci_acc, CUBLAS	Cray libsci

Inter-node MPI on Cray XK6

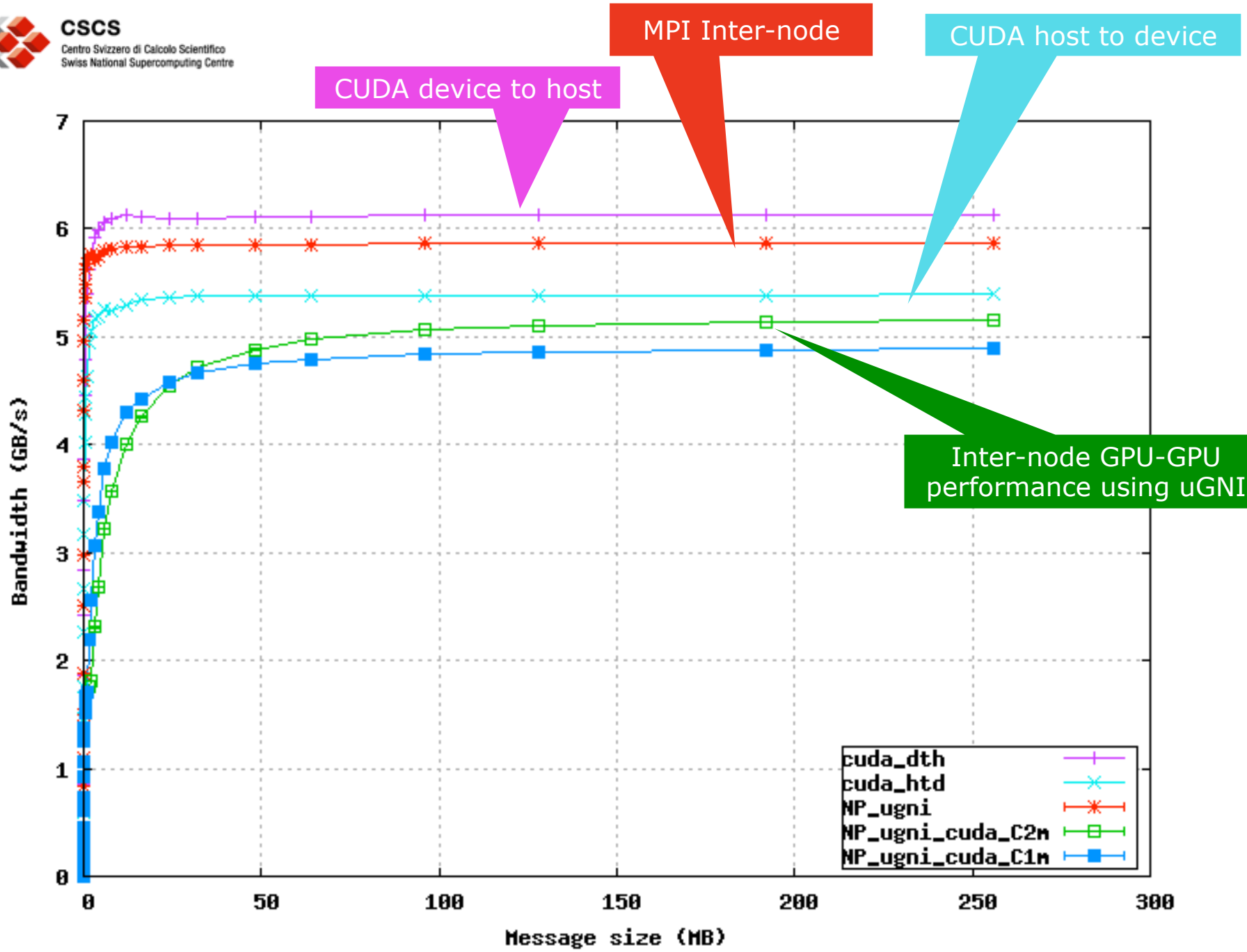



```
Copy data from GPU to CPU (device to host)

MPI message transfer between two CPUs (host to host MPI)

Copy data from CPU to GPU (host to device)
```

CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre



CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

HP2C and PRACE 2IP WP8 Initiatives

Application Names	Domains	Project & Development Teams Details	Collaborators (GPU Development)
BigDFT	Materials science / Nanoscience	http://inac.cea.fr/L_Sim/BigDFT	S. Goedecker (University of Basel), L. Genovese (CEA, France)
COSMO	Regional climate / meteorology	http://www.clm-community.eu	O. Fuhrer & X. Lapillonne (MeteoSwiss), T. Gysi (SCS, Switzerland)
CP2K	Chemical science / Nanoscience	http://www.cp2k.org	J. VandeVondele & U. Borstnik (ETHZ), C. Ribeiro (University of Zurich)
Gromacs	Life science	http://www.gromacs.org	E. Lindhal, S. Páll (Stockholm University)
PKDGRAV2	Cosmology	https://hpcforge.org/projects/pkdgrav2	J. Stadel, D. Potter (University of Zurich)
RAMSES	Cosmology	http://irfu.cea.fr/Projets/COAST/software.htm	R. Teyssier (University of Zurich), P. Kestener (CEA, France), A. Hart (Cray)
SPECFEM3D	Seismology	http://www.geodynamics.org/cig/software/specfem3d	O. Schenk, M. Riethmann (University of Lugano)

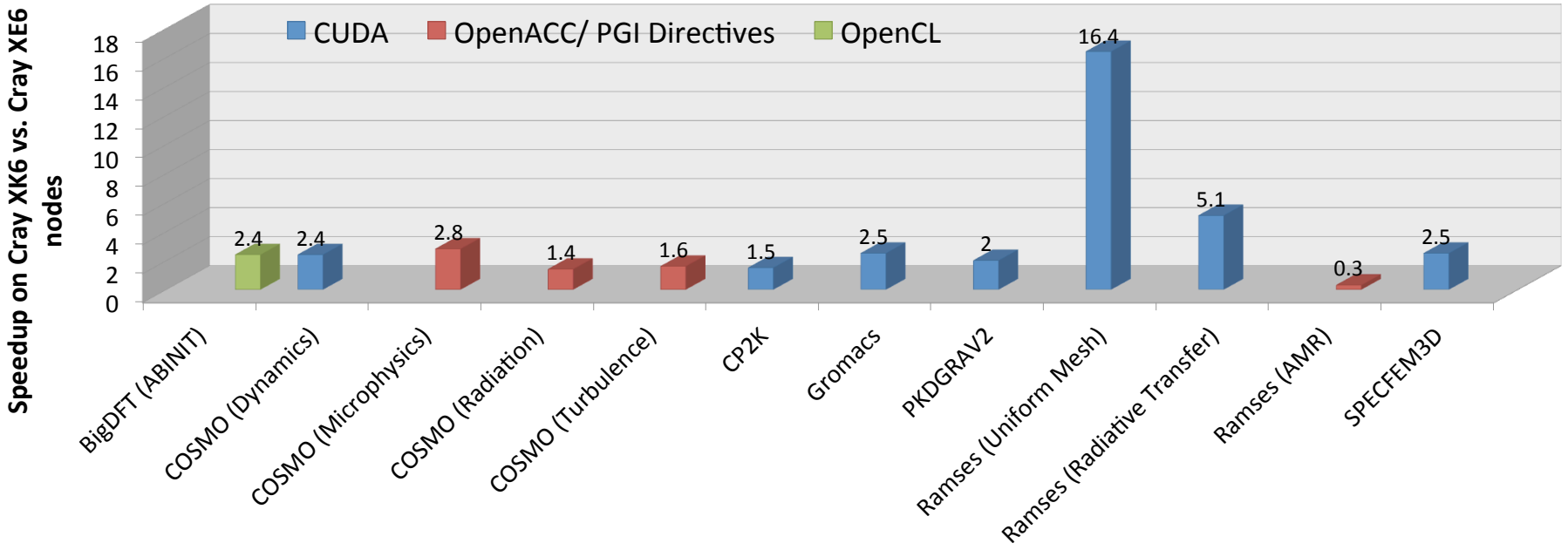
HP2C High Performance and High Productivity Computing

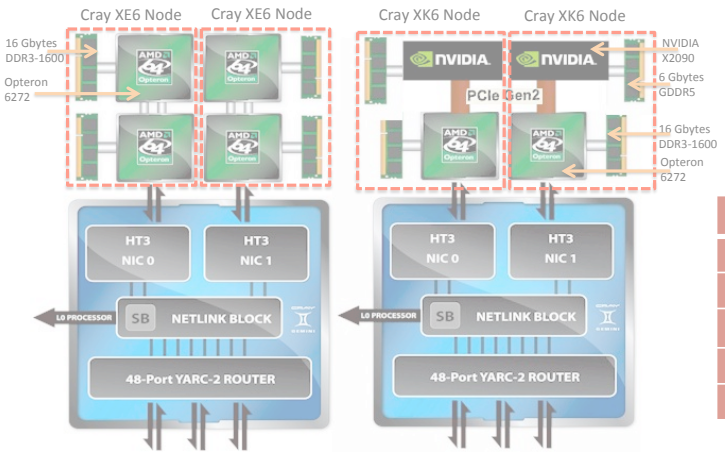
http://www.hp2c.ch

PRACE (Partnership for Advanced Computing in Europe), 2nd Implementation Phase
WP8 (Community Code Development)
http://www.prace-ri.eu/

CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

Application Status (March, 2012)





	Cray XK6 (Tödi)	Cray XE6 (Monte Rosa)
Cores	512	16
Clock frequency	1.15 GHz	2.1MHz
FP performance	665 GFlops (double-precision)	134.4 GFlops (double-precision)
Memory interface	GDDR5	DDR3 (1600)
Power envelope	225-250 W	90-115 Watts

Summary and future directions

- **A familiar environment for migration:**
  - From existing Cray platforms
  - For clusters based on accelerators
- **Critical issues:**
  - Up-to-date software stack (not nice-to-have but a must-have requirement for code developers)
  - Performance of MPI & I/O comparable to XE6
  - Interlagos compiler and optimization issues
  - Parallel tools (compilers, debuggers & performance measurements)
- **Future R&D investment priorities (XK6->XK7):**
  - Kepler readiness (for CUDA, OpenCL and others)
  - Kepler: MPI & I/O stacks
  - Interlagos issues (compute, memory & I/O performance)
  - OpenACC development and integration with other programming models

CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

Thank you.