



ETH Cluster Services

Olivier Byrde

Associate Director, CSCS

HPC-CH Meeting / ETH Zurich / 28.10.2010

Mission

- The IT Services' High Performance Computing Group was created on 1.4.2007
 - Its mission is to **evaluate**, **buy** and **manage** the central HPC facilities of ETH Zurich, and to provide **support** to their users
 - These facilities are intended primarily for the scientific community of ETH
 - People not affiliated with ETH may also use them for research in collaboration with an institute or laboratory of ETH
- The group has been **transferred to CSCS** on 1.9.2010
 - New CSCS division: **ETH Cluster Services**
 - The people and current systems stay in **Zurich**
 - Next cluster to be installed in new CSCS building in **Lugano**

Organization

CSCS Directorate

Thomas Schulthess, Director
Dominik Ulmer, General Manager

**National
Supercomp.
Service**

Dominik Ulmer (*a.i.*)

**HPC
Co-Location
Services**

Michele De Lorenzi

...

**ETH
Cluster
Services**

Olivier Byrde

People

- Teodoro Brasacchio
 - System design, testing and implementation, vendor relations
- Olivier Byrde (group leader)
 - Software development, user support and shareholder relations
- Eric Müller
 - Cluster administration and maintenance, user support
- Tonko Racic
 - System administration and user support for HP Superdome
- George Sigut
 - Applications support, benchmarking
- Adrian Ulrich
 - Storage, system management and software development

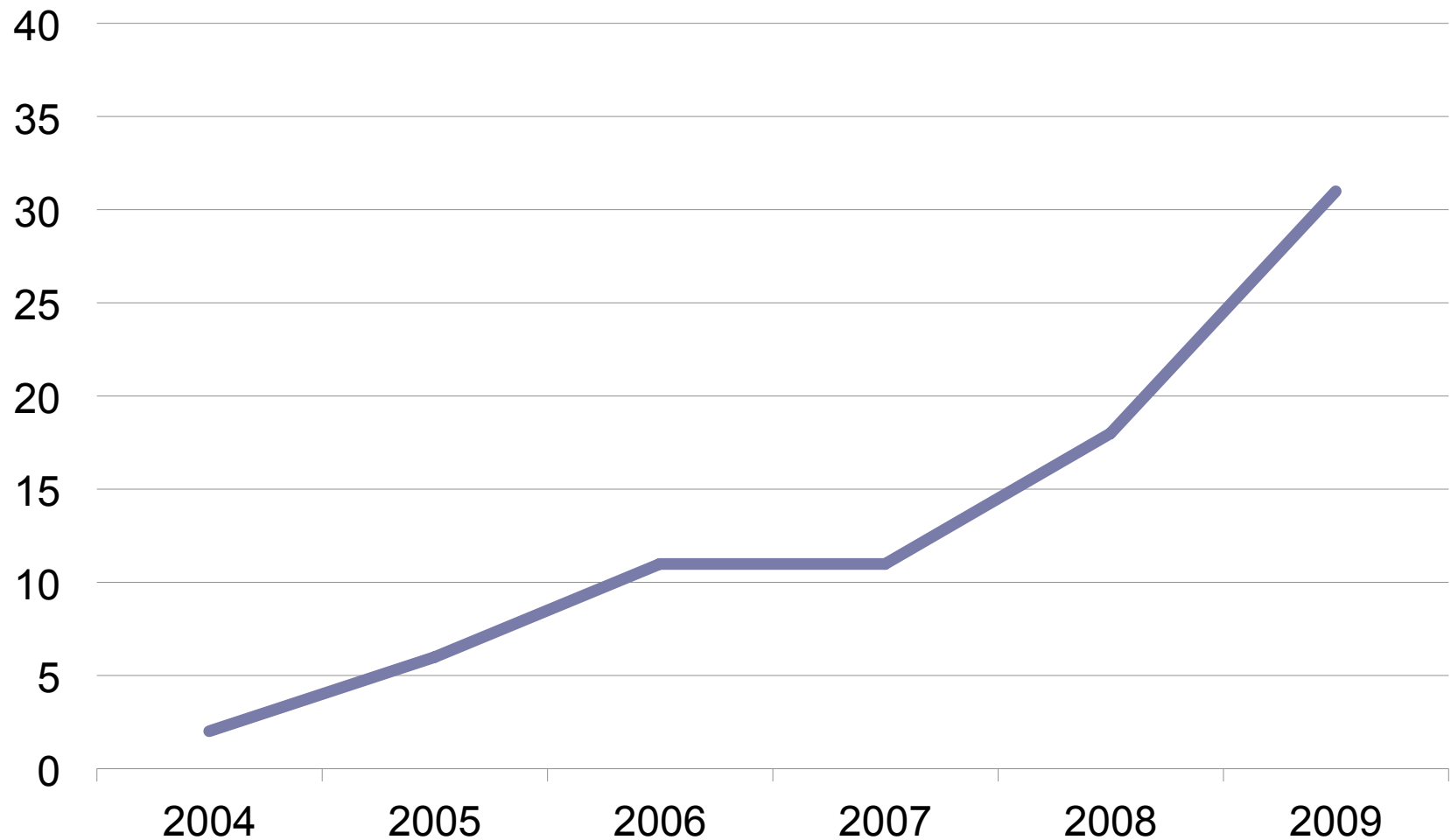
Facilities

- Linux clusters
 - Asgard (2000-2007; RIP)
 - Hreidar (2004-2008; partly integrated into Brutus)
 - Gonzales (2005-2008; integrated into Brutus)
 - Fortunato (2007-2008; integrated into Brutus)
 - **Brutus (2008-present)**
- Shared memory systems
 - HP Superdome **Stardust** & **Pegasus** (64+32 dual-core Itanium2)
 - Both systems approaching their end of life
 - Operation after 1.1.2010 not guaranteed
 - Migration to Brutus under way

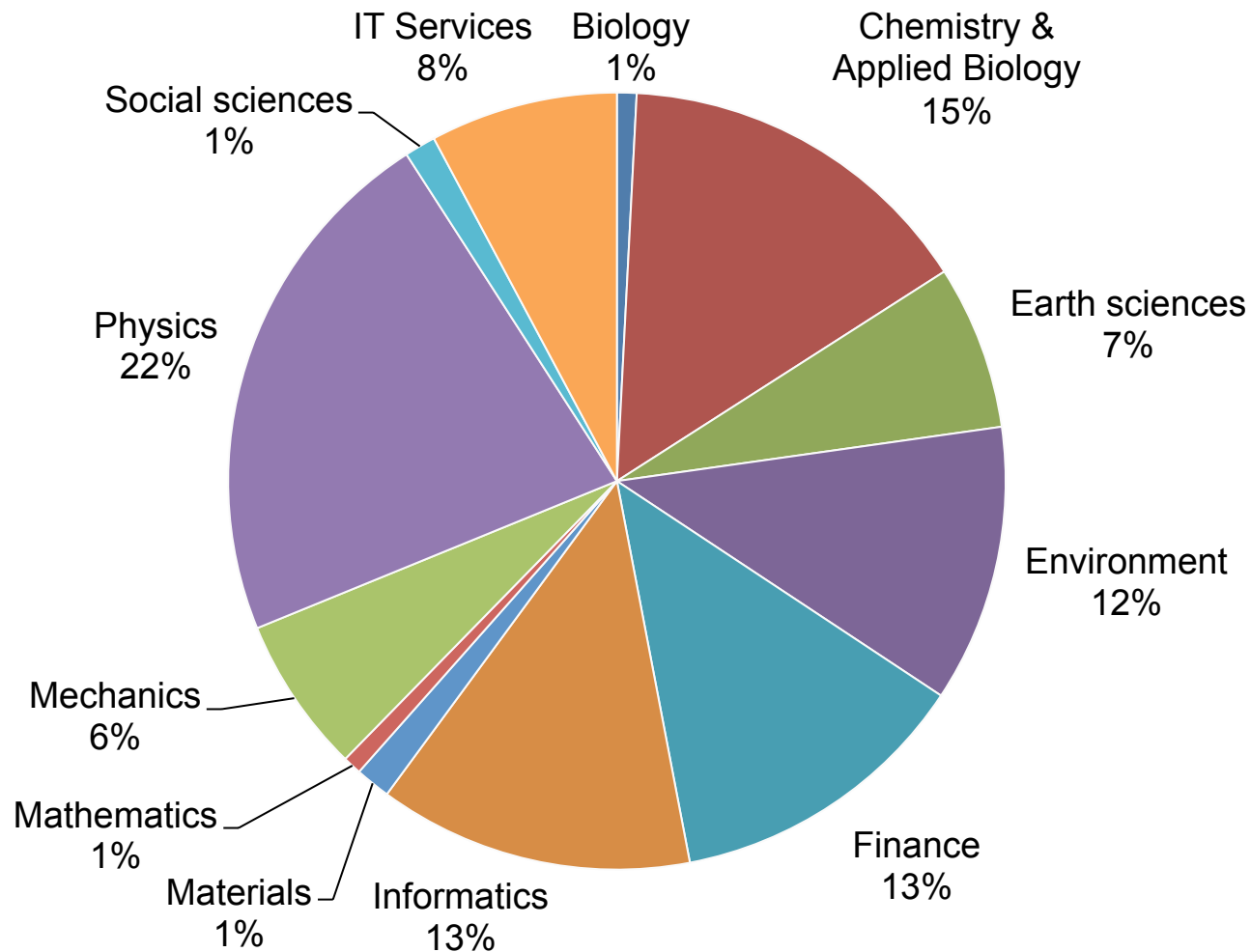
Shareholders

- All large clusters since Asgard have been **financed by their users**
- Brutus is owned jointly by about **50** professors in **12** departments and CSCS (formerly IT Services)
- These professors — called **shareholders** — get a share of CPU time proportional to their investment
- The share financed by the **IT Services** is made available to the scientific community of ETH **for free**
- CSCS (formerly IT Services) takes care of **system management** and **user support**

Number of shareholders



Shares by research department



Brutus platform

- Unified **platform** for all central clusters at ETH
- More **efficient** than small clusters
 - Lower cost thanks to economies of scale
 - Optimal resource utilization
- More **reliable** than typical Beowulf clusters
 - Redundant file servers, login nodes, management nodes
- Easier to **manage** and to **use**
 - Centralized management, single user environment
- Hardware **neutral**, **flexible** and **extensible**
 - 8 generations of compute nodes (2004-2010)
 - 4 networks (2x Ethernet, Quadrics, InfiniBand)

Brutus cluster

- Compute nodes
 - 4 x 2 hexa-core AMD 2435, each with 6 GPUs (SM/Nvidia)
 - 490 x 4 quad-core AMD 8380/84, each with 32 GB RAM (Sun)
 - 10 x 4 quad-core AMD 8380 with 128 GB RAM (Sun)
 - 8 x 8 dual-core AMD 8220 with 64-128 GB RAM (Sun)
 - 272 x 2 dual-core AMD 2220 with 16 GB RAM (Dalco)
 - 324 x 2 single-core AMD 250 with 8 GB RAM (Dalco)
- Networks
 - 648 x InfiniBand QDR (Sun)
 - 256 x QsNet^{II} (Quadrics)
 - Gigabit Ethernet (Force10, Enterasys)
 - Fast Ethernet (management network)

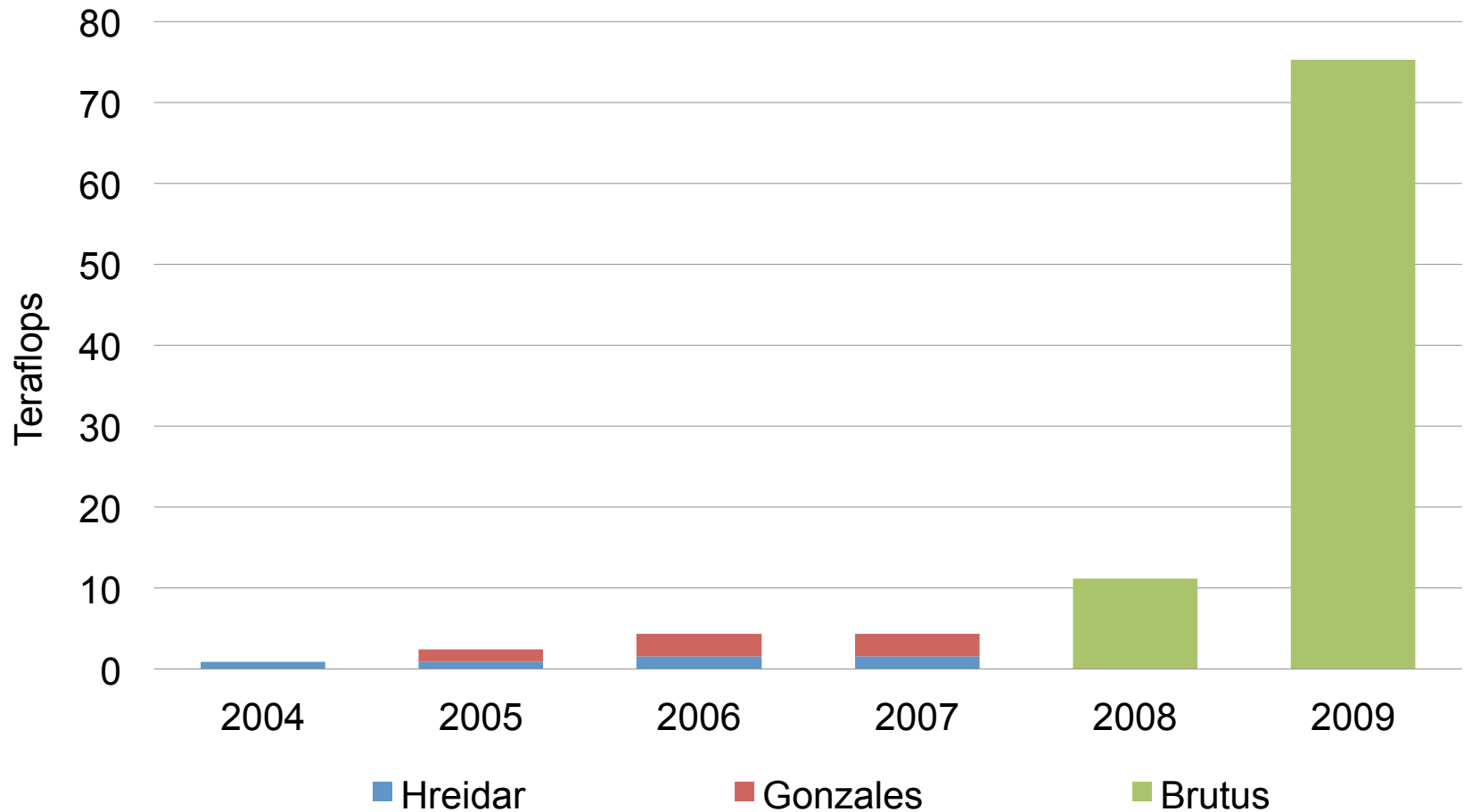
Brutus cluster (cont.)

- Storage
 - Home: [NFS / SAN](#) (6 TB)
 - Work: [Panasas](#) (45 TB)
 - Scratch: [Lustre](#) (230 TB)
 - Access to the ETH's [NAS](#) infrastructure (for specific groups)
- Software
 - OS: [CentOS](#) 5.5
 - Batch: Platform [LSF](#) 7.0 Update 6
 - Compilers: GCC, Intel, Open64, PGI and Sun Studio
 - MPI: [Open MPI](#), MVAPICH2, Quadrics and HP (comm. apps)
 - Commercial apps: Abaqus, Ansys CFX & FLUENT, Comsol, Gaussian, Marc, Materials Studio, MATLAB, StarCD, etc.

Key figures

- 1 single platform
- 34 shareholder groups (about 50 professors)
- 43 racks (20 water-cooled; 23 air-cooled)
- 90 teraflops peak
- 1000 active users
- 10,000 processor cores
- >1 million jobs per month!

Performance growth



Top 500

- Brutus was ranked 88th in the Nov '09 [Top 500](#) list
 - 2nd in Switzerland, behind the Cray XT5 at CSCS
- Fastest cluster in Switzerland
 - 10th fastest cluster in Europe
- Most energy efficient general purpose supercomputer in the world:
 - Der effizienteste Allzweckrechner ist gemäß den angegebenen Verbrauchswerten ein Opteron-Quadcore-Blade 6440 von Sun an der ETH Zürich mit 341 MFlops/Watt
(source: [Heise.de](#), <http://tinyurl.com/ye69v3c>)

Future goals

- Improve the **current** system
 - Stability, ease of use, applications
 - Performance (single-CPU optimization, scaling)
- **Extend** it based on the **needs** of shareholders
 - Compute nodes
 - GPU nodes
- Start work on **Brutus' successor**
 - Based on same shareholder model
 - System to be installed in new CSCS building in Lugano
 - Local HPC support group in Zurich
 - Call for tender in Q2 2011, delivery in Q1 2012

Thank you