TECHNISCHE UNIVERSITÄT DRESDEN

Center for Information Services and High Performance Computing (ZIH)

Performance Tools for System Monitoring

1st CHANGES Workshop, Jülich

01069 Dresden
Zellescher Weg 12
Tel. +49 351 - 463 – 35450

September 5th, 2012

Wolfgang E. Nagel

ZIH
Center for Information Services & High Performance Computing

Dresden University of Technology

- Founded in 1828
- one of the oldest **technical** universities in Germany
- 14 faculties and a number of specialized institutes
- More than 36.500 Students, about 5.000 Employees, 500 professors
- International courses of studies, bachelor, masters
- One of the largest computer science faculties in Germany
- 200 million Euro annual third party funding
- One of the eleven "Excellence Universities" in Germany
- http://tu-dresden.de

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH
Center for Information Services & High Performance Computing

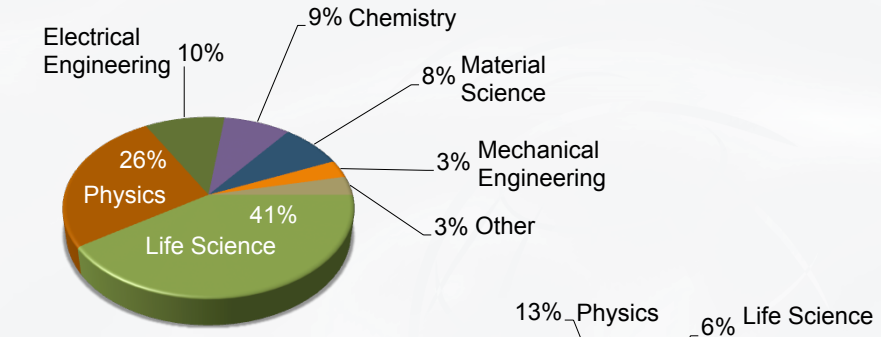
Current HPC Activities at TU Dresden

- Dresden is just preparing the next HPC procurement
  (15 Mio. EUR in 2013/14, additional machine room (about 5 MW),
  research focus on scalability, data intensive computing, energy awareness)
- Work on the performance analysis environment for the next Oak Ridge system (Titan, based on Nvidia Kepler)
  - Scalability (OS, information gathering, displays, …)
  - Heterogeneity (Cluster, Multicore, GPUs)
  - Integration of new functionalities/views
  - I/O support (mapping local and global I/O requests)
  - Energy measurement and mapping support
- MPI correctness tools (MARMOT/MUST), application scalability (COSMO/SPECS code, FD4 library)
- Running the Dresden CUDA Center of Excellence
- Member of the IESP/EESI groups, co-coordinator of the German SPPEXA (DFG priority program), to prepare the software research roadmap for Exascale

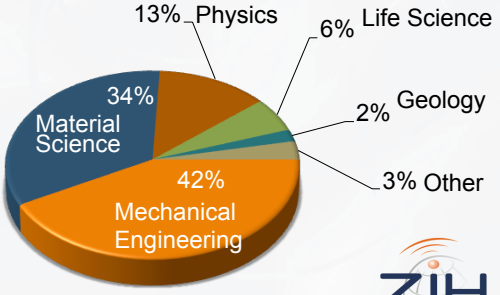
TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

Scientific Application Areas



PC Farm

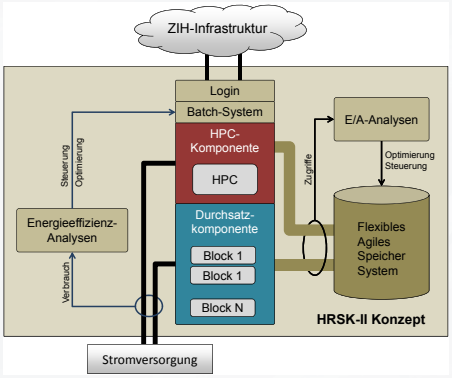


HPC Server

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

HPC at ZIH (Procurement Ongoing)



- Scalability
- Data intensive computing
- Energy efficiency

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

Challenges

- HPC systems evolve; Tools need to adapt:

MPI+ Accelerator+ Threads — **Heterogeneous/ Hybrid** — MPI+ PGAS ?

MPI+CUDA PGAS+CUDA MPI+Threads

Paradigms

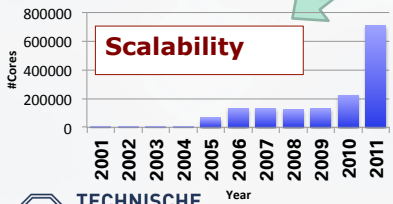
Co-Array Fortran, SHMEM, OpenCL

Challenges

Fault Tolerance

Towards Exascale …

TOP 500 – Largest Machine



Scalability

Usability

Make all this accessible!

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

TECHNISCHE UNIVERSITÄT DRESDEN

Scalability

01069 Dresden
Zellescher Weg 12
Tel. +49 351 - 463 – 35450

Wolfgang E. Nagel

ZIH
Center for Information Services & High Performance Computing

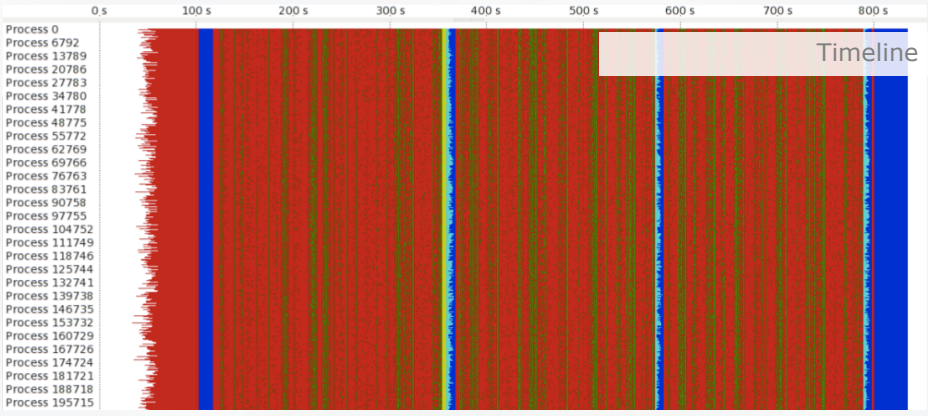
Meeting the Challenges – Collaboration with ORNL

- Goal:
  - Support hybrid system and software architecture at 10 Petascale
  - Show MPI and GPGPU programming
  - Do full system performance profiling **and** tracing
- Facts:
  - Jaguar / Titan
  - >220.000 cores
  - 200,448 monitored MPI processes
  - >20 Tera-bytes of performance data
  - 21,515 VampirServer processes

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH
Center for Information Services & High Performance Computing

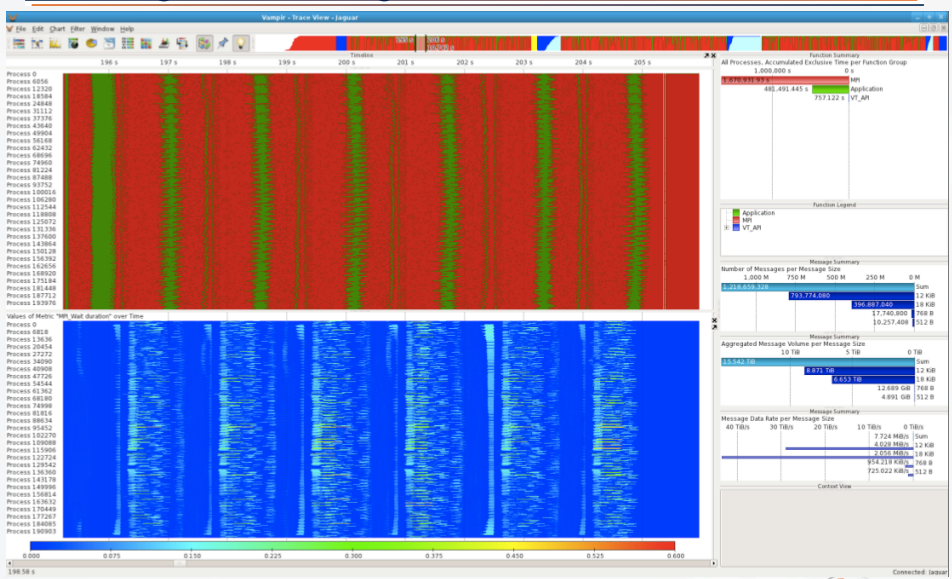
Meeting the Challenges – Collaboration with ORNL



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH
Center for Information Services & High Performance Computing

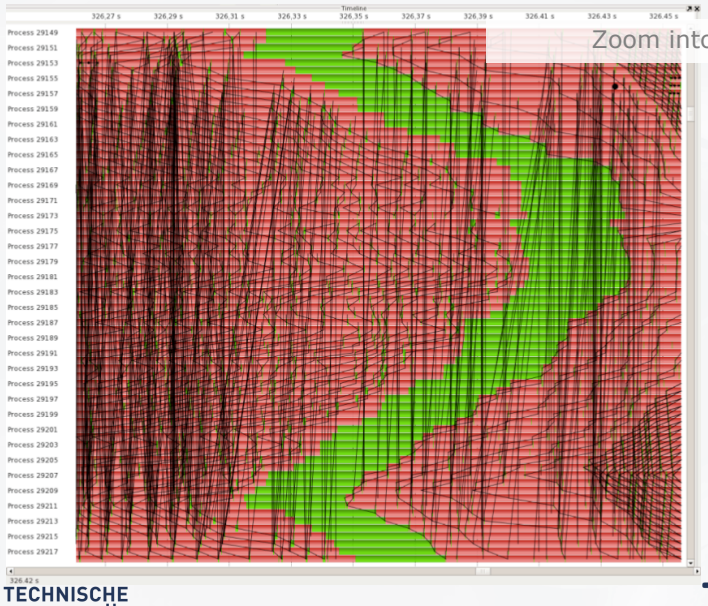
Meeting the Challenges – Collaboration with ORNL



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH
Center for Information Services & High Performance Computing

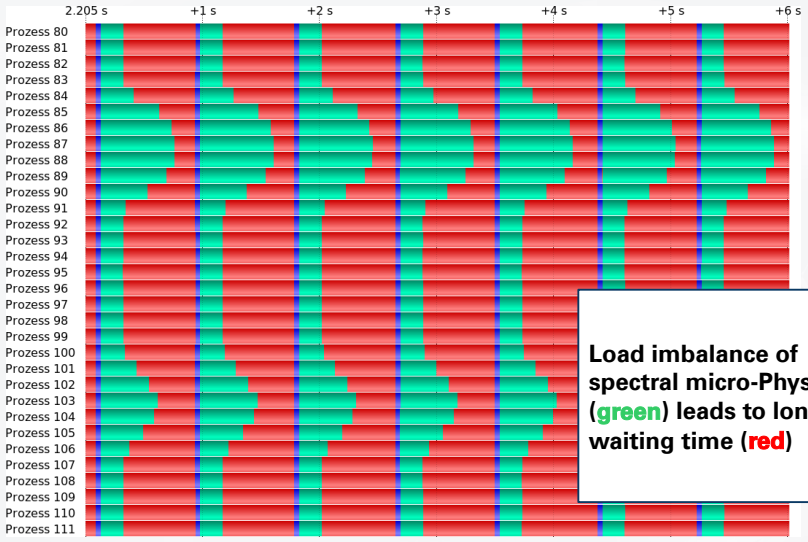
Meeting the Challenges – Collaboration with ORNL



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH
Center for Information Services & High Performance Computing

Load Imbalance in COSMO-SPECS

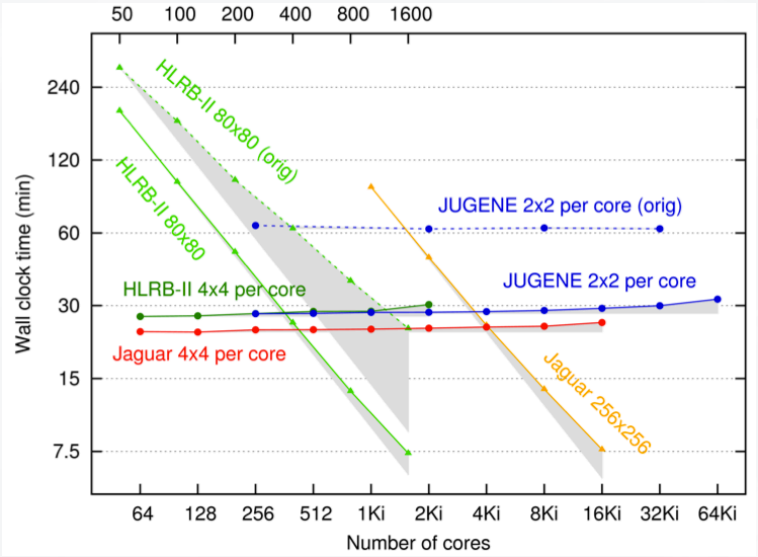


Load imbalance of spectral micro-Physics (green) leads to long waiting time (red)

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH
Center for Information Services & High Performance Computing

Scalability of COSMO-SPECS



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH
Center for Information Services & High Performance Computing

TECHNISCHE UNIVERSITÄT DRESDEN

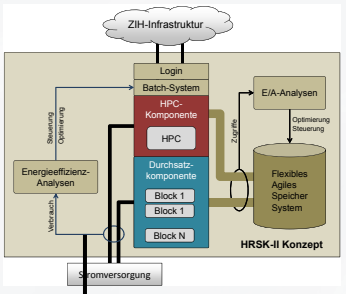
Energy

01069 Dresden
Zellescher Weg 12
Tel. +49 351 - 463 – 35450

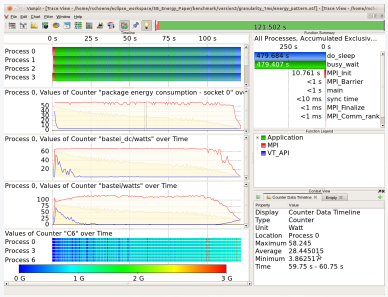
Wolfgang E. Nagel

ZIH
Center for Information Services & High Performance Computing

Power Consumption Monitoring



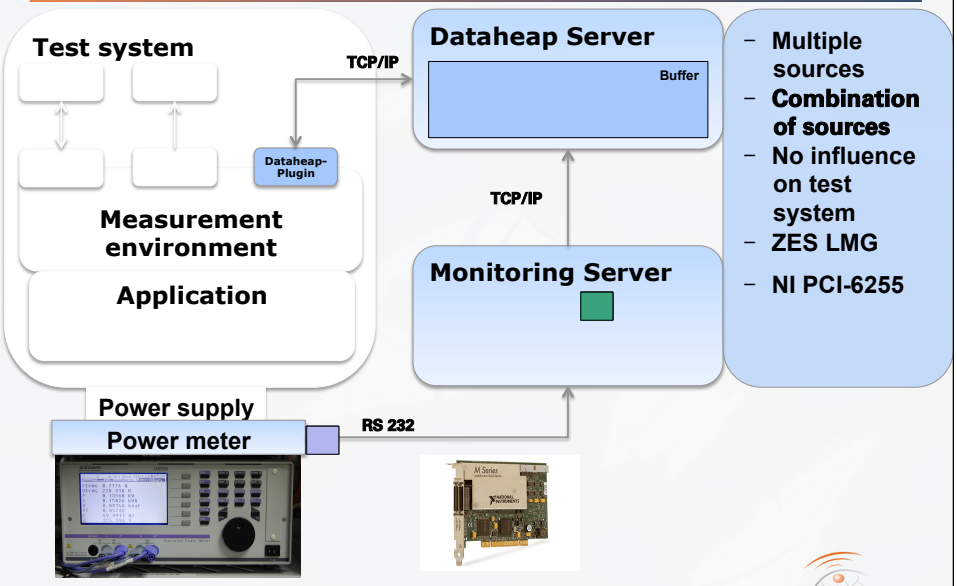
- High Precision
- High Frequency
- From complete system down to single nodes



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

Energy measurement



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

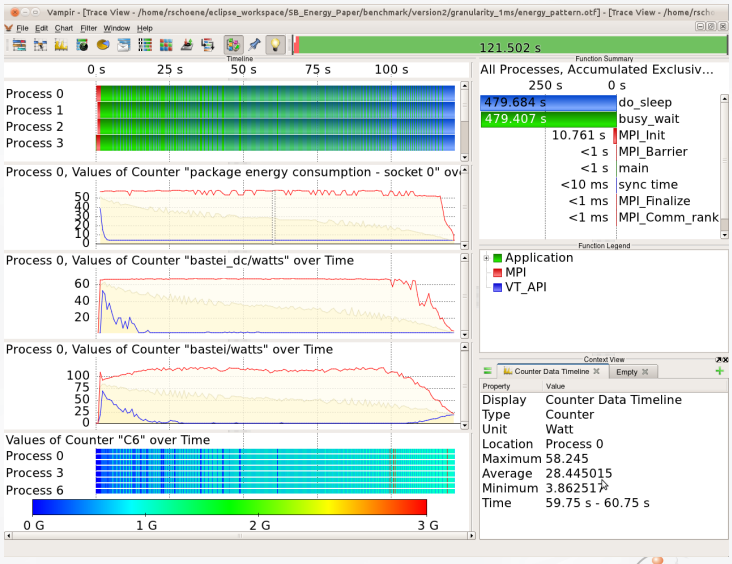
Combined Performance und Energy Analysis

Package energy consumption

12 V DC lane

AC power supply

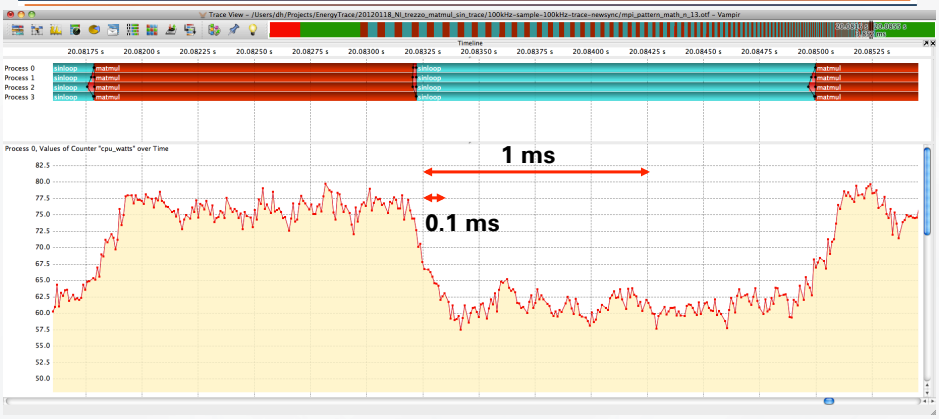
cycles in C-State 6



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

High Frequency DC-Sampling



- Measurements on Intel Sandy Bridge, 4 core, 12V
- National Instruments PCI-6255 measurement card (16 Bit, **750 kS/s sampling,** 80 analog inputs)

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

Energy Accounting

- Atlas System:
  - 90 nodes with 4x AMD Interlagos
  - QDR-IB Interconnect
- Megware ClustSaf
  - 1 Hz Sampling of all individual nodes
  - Integrated in ZIH software stack
  - Allows tracing and energy accounting
- Sample Output:

```
Run finished: Fri Apr 13 18:24:30 2012
total runtime:    6975.2 s
total energy :    5207.3 kJ
average power:     746.5 W
```

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

TECHNISCHE UNIVERSITÄT DRESDEN

I/O Analysis

01069 Dresden

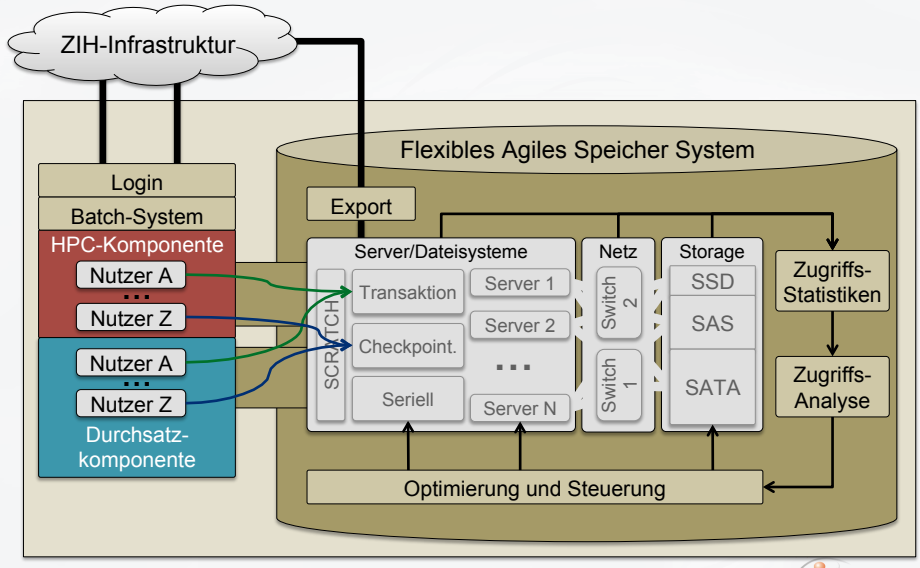
Zellescher Weg 12

Tel. +49 351 - 463 – 35450

Wolfgang E. Nagel

ZIH Center for Information Services & High Performance Computing

File System Concept: FASS



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

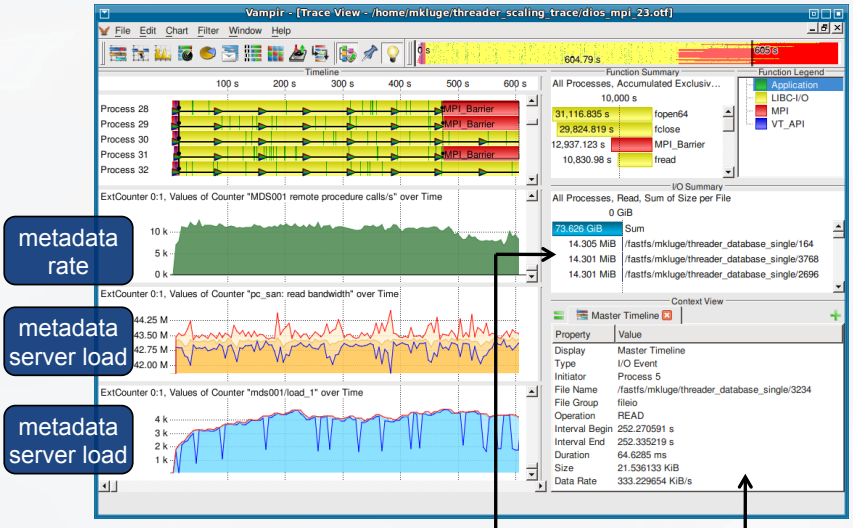
Vampir I/O Analysis

- recording of POSIX I/O and MPI I/O operations with VampirTrace
- embedding of node local data like InfiniBand statistics
- on demand inclusion of external performance data from:
  - I/O network
  - storage controller
  - file servers
- within Vampir:
  - counter timelines for the host based and the external data
  - specialized I/O display to show:
    - details for single I/O events
    - grouping of events for the current portion of the timeline based on the filename, the type of the I/O record (read, write, ...) and some more
    - I/O request size statistics

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

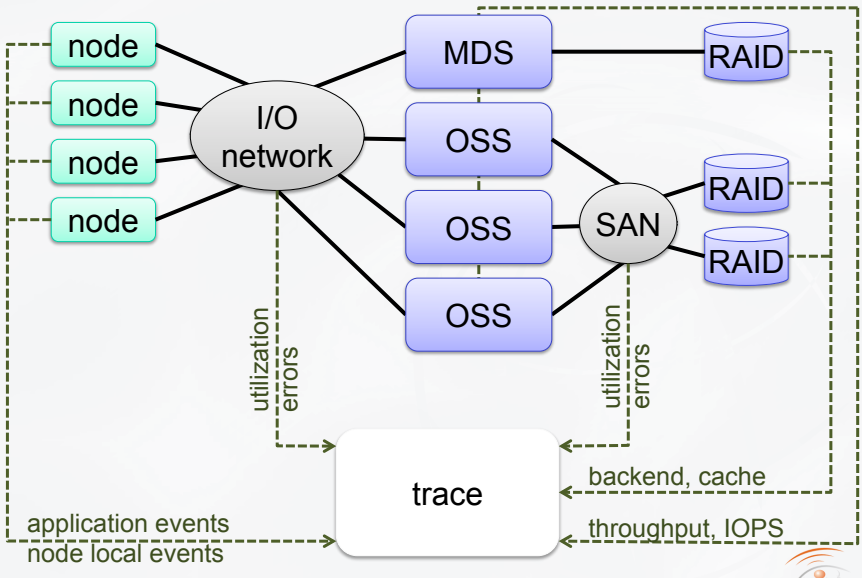
Vampir: Combined Application + Lustre Server Analysis



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

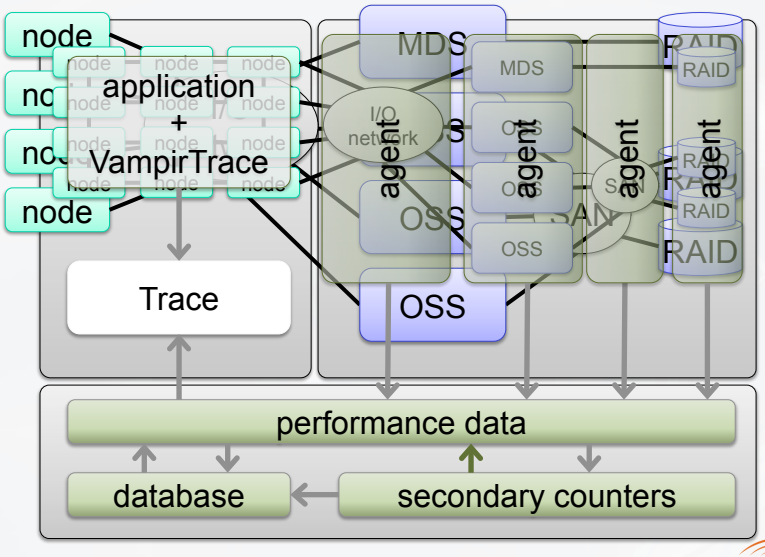
Target: Exploitation of all data sources relevant to I/O



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

DataHeap: Application + Infrastructure Events



TECHNISCHE UNIVERSITÄT DRESDEN

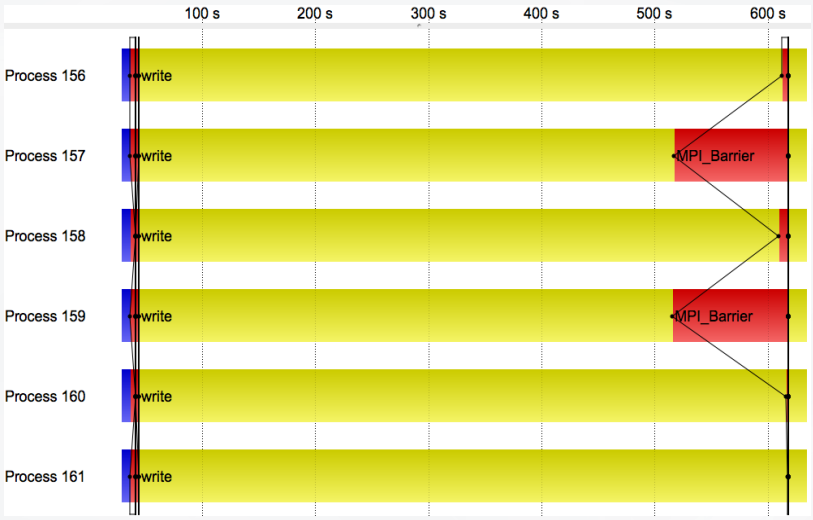
ZIH Center for Information Services & High Performance Computing

Starting Point: 6 from 576 IOR processes

TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

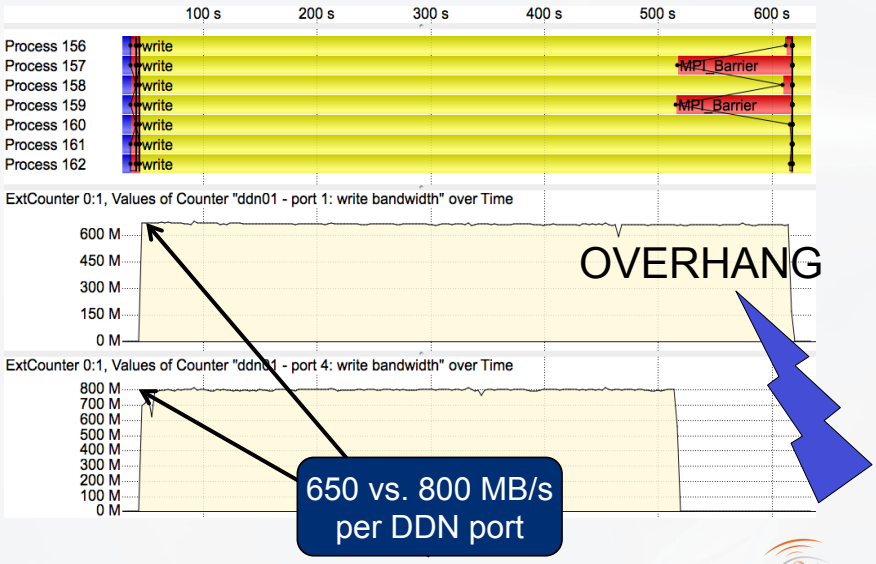
IOR: write phase



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

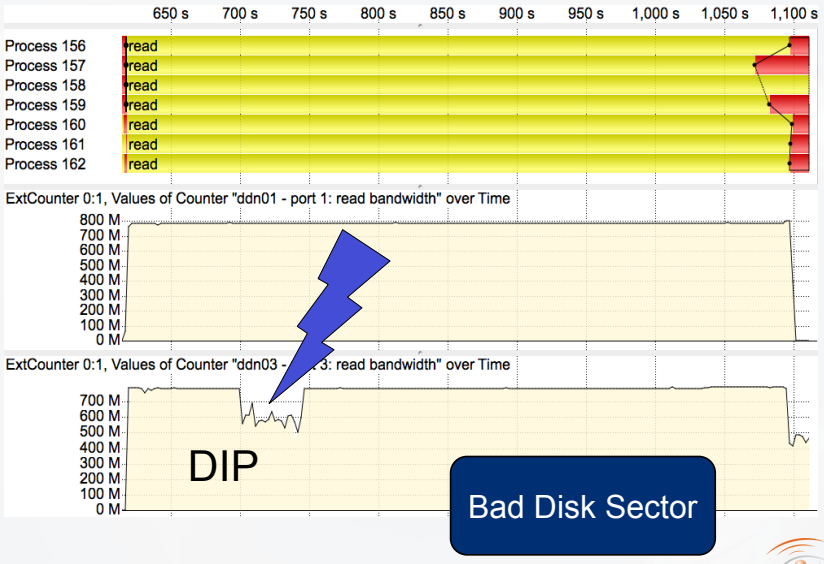
IOR: Write Phase



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

IOR: Read Phase



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH
Center for Information Services & High Performance Computing

TECHNISCHE UNIVERSITÄT DRESDEN

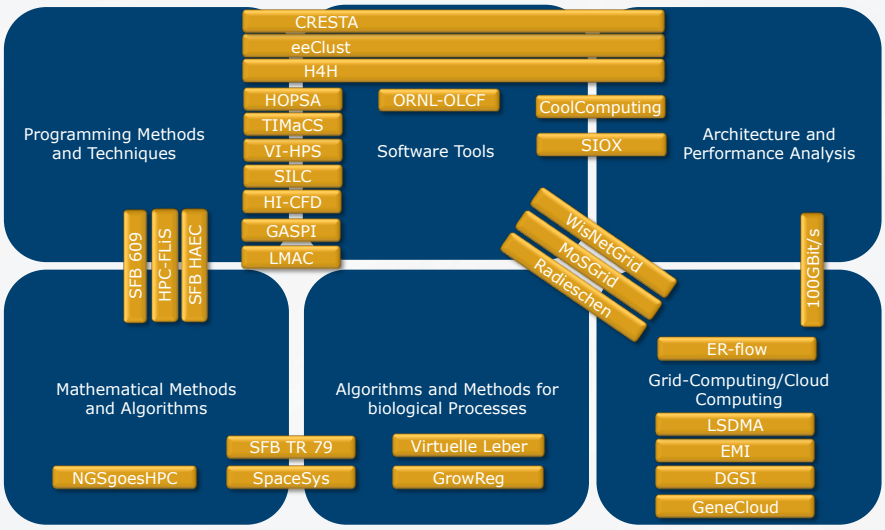
Outlook

01069 Dresden
Zellescher Weg 12
Tel. +49 351 - 463 – 35450

Wolfgang E. Nagel

ZIH
Center for Information Services & High Performance Computing

Projects



TECHNISCHE UNIVERSITÄT DRESDEN

ZIH Center for Information Services & High Performance Computing

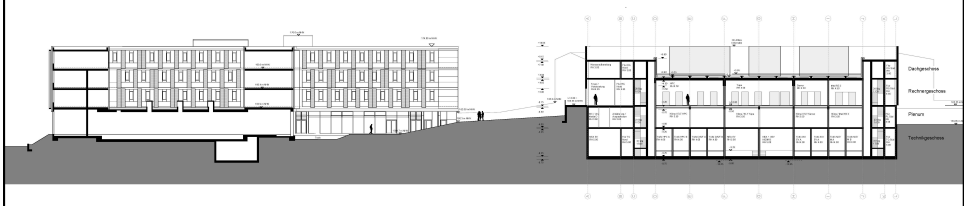
New Machine Room: Architectural View

TECHNISCHE UNIVERSITÄT DRESDEN

Dr. Peter Fischer

ZIH Center for Information Services & High Performance Computing

CS Department and New Machine Room



Planned Schedule

Start: January 2013

Availability for HRSK Phase II: October 1st, 2014

New Machine Room Infrastructure

- PUE < 1.3
- First building in university "research mile"
- HPC room:
  - 450 m^2 (12m x 36m, 140+ racks)
  - 2.5 MVA
  - Cold water: 20°C
  - Warm water: 30-50°C
  - 250 kW air cooling capacity


Thank you !