

# Reducing the TCO for Grand Scale Applications

Arndt Bode, Herbert Huber, Leibniz-Rechenzentrum

## **Job profile on current LRZ supercomputer SuperMUC**

- Execution of more than 170 different parallel applications
- Most applications need more than 1.0 GByte of main memory per CPU core
- Steadily growing number of data centric applications with new demands with respect to:
  - Data bases and long term online storage resources
  - Parallel I/O (e.g., small file I/O and local non volatile RAM as I/O buffer)

## **LRZ demands for future supercomputer architectures and system software**

- Highly scalable and energy efficient general purpose system architecture and system software
- Highly optimized programming environment (parallel environment, compilers, numerical libraries and performance tools)

## Holistic Optimization Strategy

- Reduce the power losses in the power supply chain
- Exploit your possibilities for using compressor-less cooling and use energy-efficient cooling technologies (e.g. direct liquid cooling)
- Re-use waste heat of IT systems

**Energy efficient infrastructure**

- Use newest semiconductor technology
- Use of energy saving processor and memory technologies
- Consider using special hardware or accelerators tailored for solving specific scientific problems or numerical algorithms

**Energy efficient hardware**

- Monitor the energy consumption of the compute systems and the cooling infrastructure
- Use energy aware system software to exploit the energy saving features of your target platform
- Monitor and optimize the performance of your scientific applications

**Energy aware software environment**

- Use most efficient algorithms
- Use best libraries
- Use most efficient programming paradigm

**Energy efficient applications**

LRZ has long term experience in all of the 4 pillars and their coupling (Infrastructure, Monitoring, Tools, Projects, ...)

- Enable chiller-free system cooling and heat waste re-use for central and south European climate zones
  - Use of direct liquid cooling technologies for the following HPC system components
    - ✓ Compute nodes
    - ✓ Communication network
    - ✓ Power supplies
  - Hardware support for ASHRAE W5 inlet temperatures spectrum (up to 60°C)
- Standardization of system energy sensors and corresponding APIs
  - Internal sampling rate
  - Sensor accuracy
- Energy-aware system software and performance tools (e.g, energy- and topology-aware batch schedulers, energy-aware performance monitoring tools)

## **Enhancement of application scalability and energy-efficiency**

- a) Promotion of close collaboration between numerical mathematics community and scientific domain experts as well as code developers through the foundation of European Centres of Excellence for various scientific domains
- b) R&D in new numerical algorithms and libraries that can take sufficient advantage of the available computing and communication resources

## **System software fault tolerance and resiliency**

- a) Fault tolerant parallel run time systems, in particular MPI run time system
- b) Fault tolerant network routing algorithms
- c) Parallel file systems with
  - a) substantially enhanced metadata performance
  - b) substantially enhanced data protection and data recovery mechanisms
  - c) enhanced scalability and reliability