CoolEmAll - Tools for realising an energy efficient data centre

Wolfgang Christmann

christmann informationstechnik + medien GmbH & Co. KG

www.christmann.info

Outline

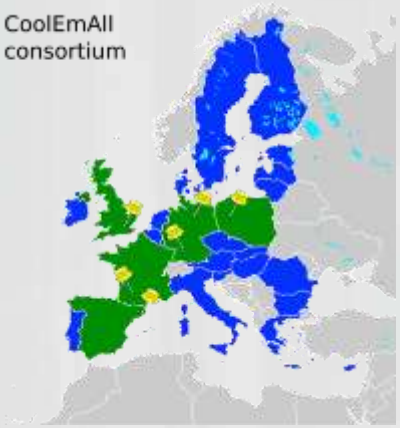
- CoolEmAll project
- RECS system – towards a data center in a rack
- Data center efficiency building blocks
- SVD Toolkit
- Conclusions and future plans

What is all about?

COOLEMALL PROJECT

CoolEmAll Project

- FP7 ICT Call 7
- Budget: 3614210€ (funded: 2645000€)
- Duration: 30 months
- Start date: 1st Oct 2011



- Consortium

POZNAN IRIT Atos the 451 group

HLRS IREC CHRISTMANN INFORMATIONSTECHNIK + MEDIEN

CoolEmAll Goal

- Goal: To enable **data center designers** and **operators** to reduce its **energy impact** combining **optimization** of **IT**, **cooling** and **workload** management

- Two main results
  - Open source simulation, visualization and decision support toolkit **(SVD Toolkit)**
  - Open designs of data center building blocks **(ComputeBox Blueprints)**

End Users

- For data center planners/designers
  - Planning equipment types and placement, cooling approach
- For data center operators
  - Optimization of equipment placement, cooling, workload and resource management
- For data center equipment vendors
  - Finding an optimal configuration for DC workloads and conditions
- For researchers
  - Studying impact of workload and resource management methods on DC energy efficiency

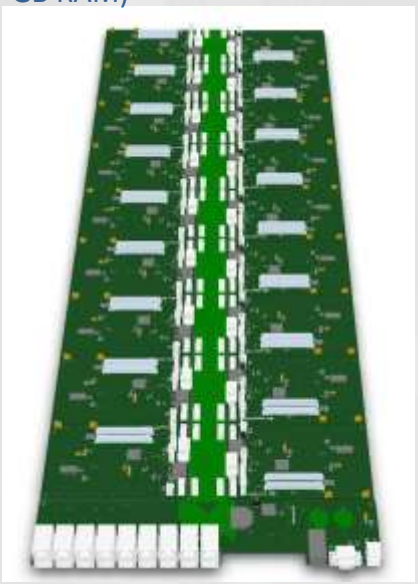
Towards data center in a rack

RECS SYSTEM

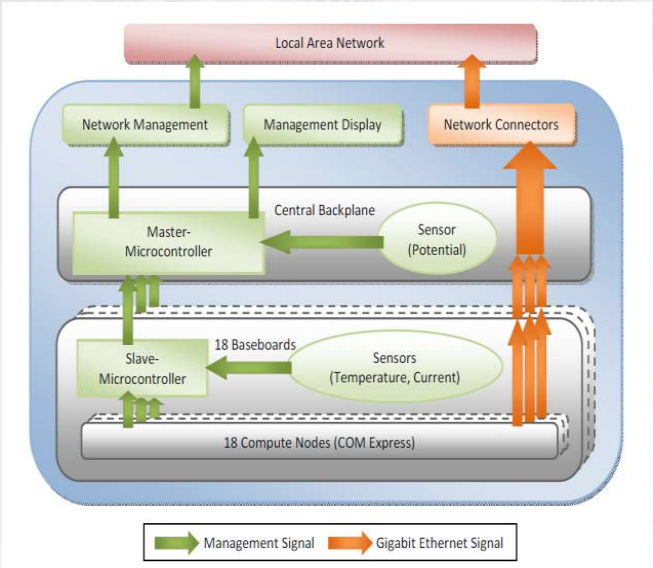
ComputeBox Prototype

- RECS – Resource Efficient Computing System
- Highly integrated monitoring and controlling infrastructure
- Airflow and liquid cooling, diverse density

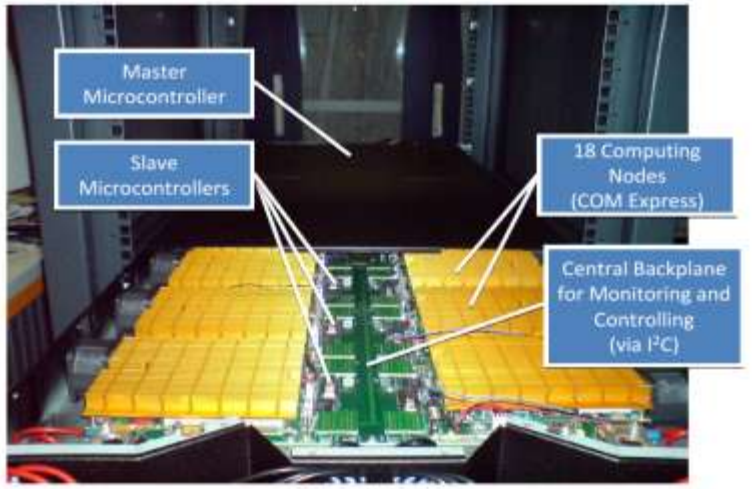
Skinless compute node for RECS (2 or 4 cores, max. 16 GB RAM)



RECS backbone and basebords



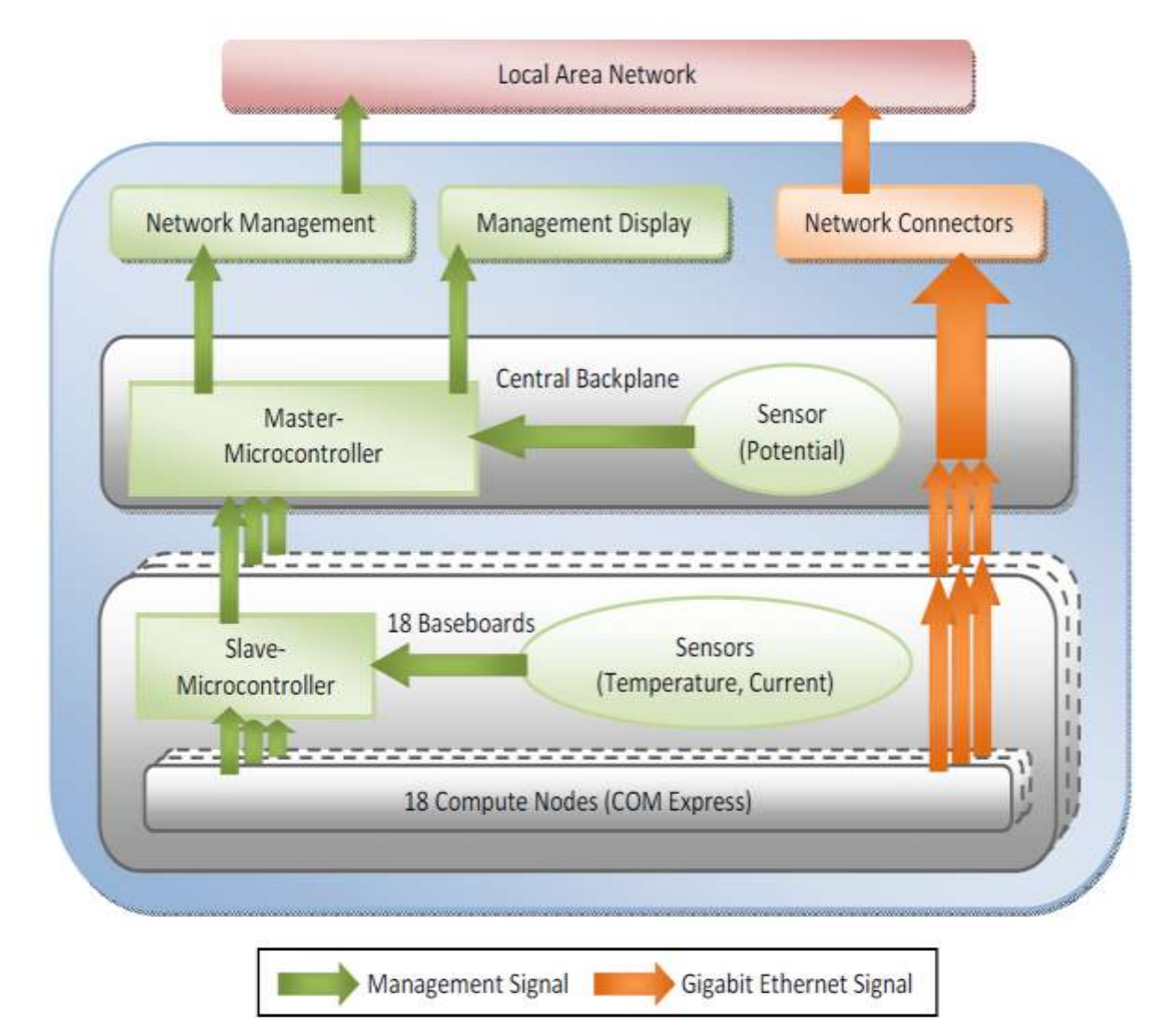
Architecture of the master-slave microcontroller system



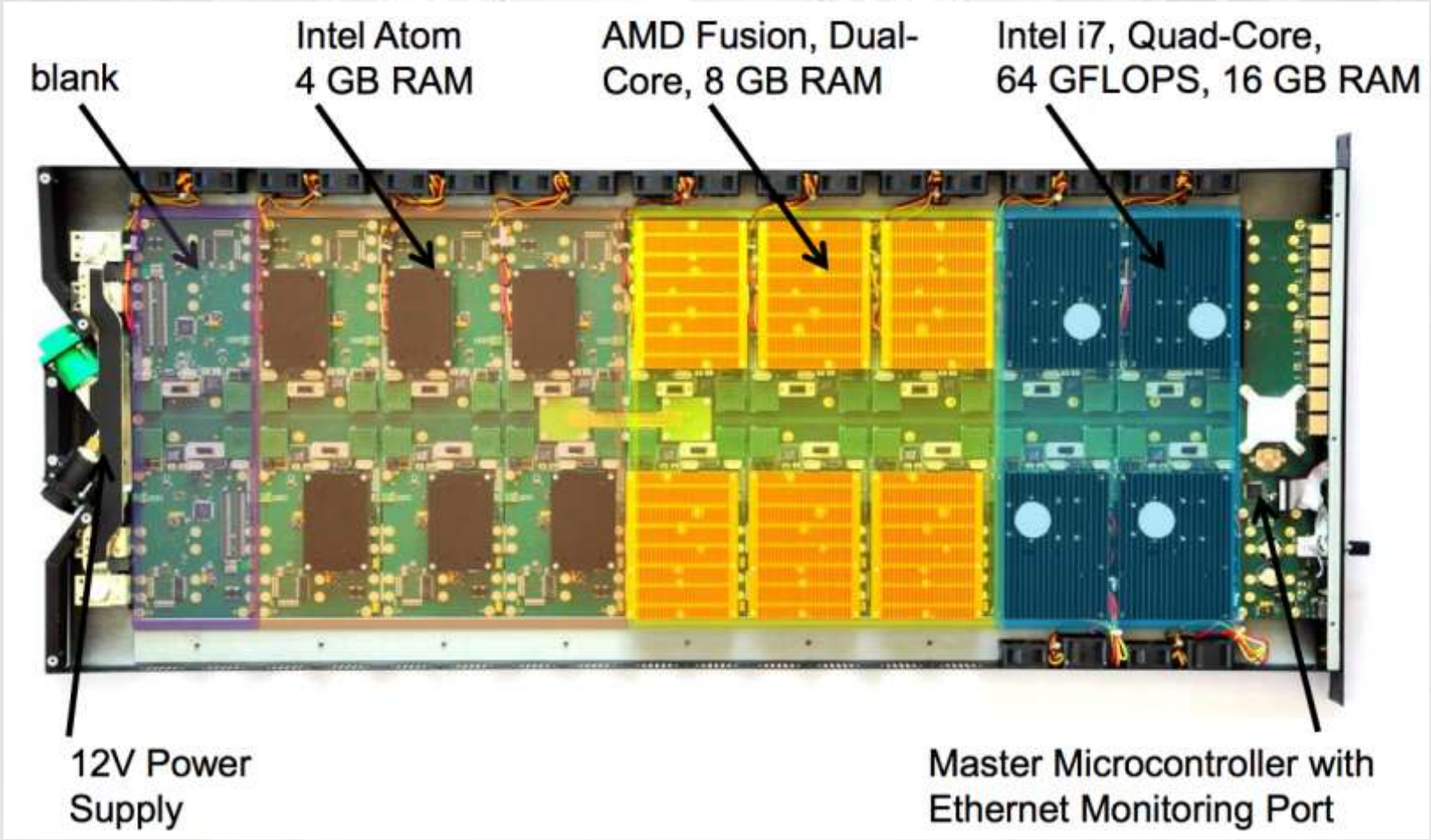
RECS microcontroller system

Integrated Monitoring & Control

- Fine grained architecture for computing, monitoring and controlling
- Negligible overhead
- Reconfigurable architecture using COM Express standard
  - Hot-swap CPU modules



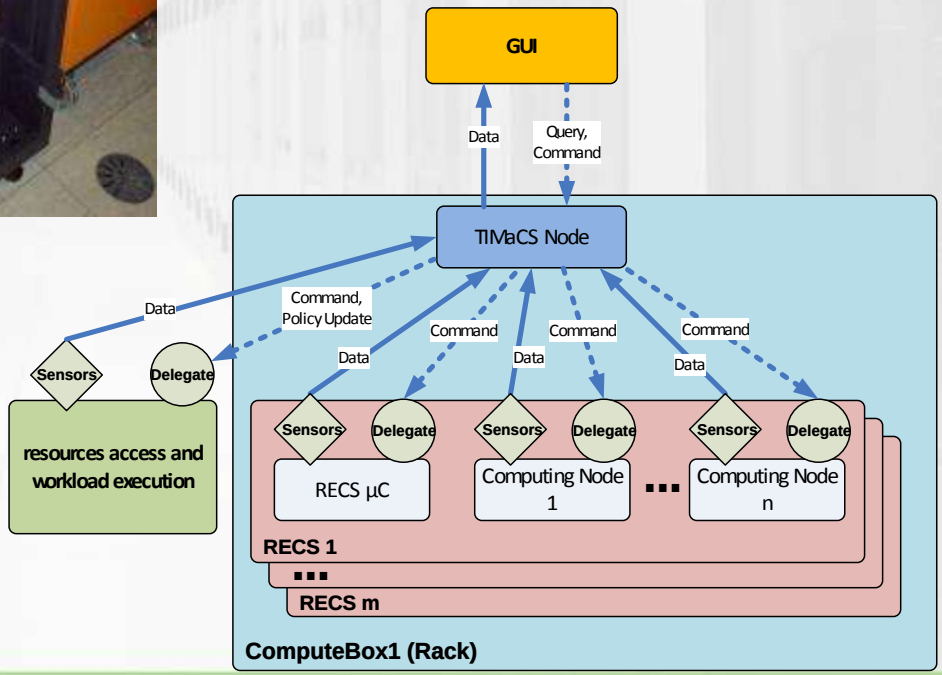
RECS Example



Validation Environment

- 4 1U units (PSNC)
  - Intel i7(Quad) x18
  - AMD Fusion x18
  - Intel Atom x18
  - ARM x72
- Additional equip.
  - Inteligent PDU
  - Temp. sensors
  - HX side cooling
- Monitoring and management layer
  - Timacs
  - Resource access and workload execution service
  - Master and slave Microcontroller

RECS testbed at HLRS



Towards Data Center in a Rack

CHRISTMANN
INFORMATIONSTECHNIK + MEDIEN

The midrange data center

RECS data centre in a rack

Servers	600*	600 (up to 32 TFlops)
Area (m^2)	100 - 600*	20
PUE	1.2 to 2.2*	1.05* to 1.2
Investment costs Including building and cooling	2 to 6 Million €	0.8 to 1.2 Million €
Power consumption	110 to 240 kW	15 to 40 kW
Energy costs / year Including cooling	120,000 to 250,000 €	24,000 to 65,000 €

*Average values, According to a 2010 study conducted by Borderstep Institute and Fraunhofer IZM (Materialbestand der Rechenzentren in Deutschland, Eine Bestandsaufnahme zur Ermittlung von Ressourcen- und Energieeinsatz, UBA)

* With geothermal cooling

CHRISTMANN
INFORMATIONSTECHNIK + MEDIEN

RECS data centre in a rack

Next Generation Data Centre

One single rack with up to 100 TFLOPS or more

- Up to 600 Microservers
- Up to 300 GPUs
- Up to 200 MICs (e.g. Xeon Phi)
- Up to 2.5 PB physical storage
- Up to 75 PB virtual storage
- Up to 70 TB of RAM
- Or a mix of this

integrated highly efficient cooling - very flexible configuration – modular interconnect – up to three physical networks - integrated monitoring and controlling - self management

ComputeBox & DEBB concepts

DATA CENTER EFFICIENCY BUILDING BLOCKS

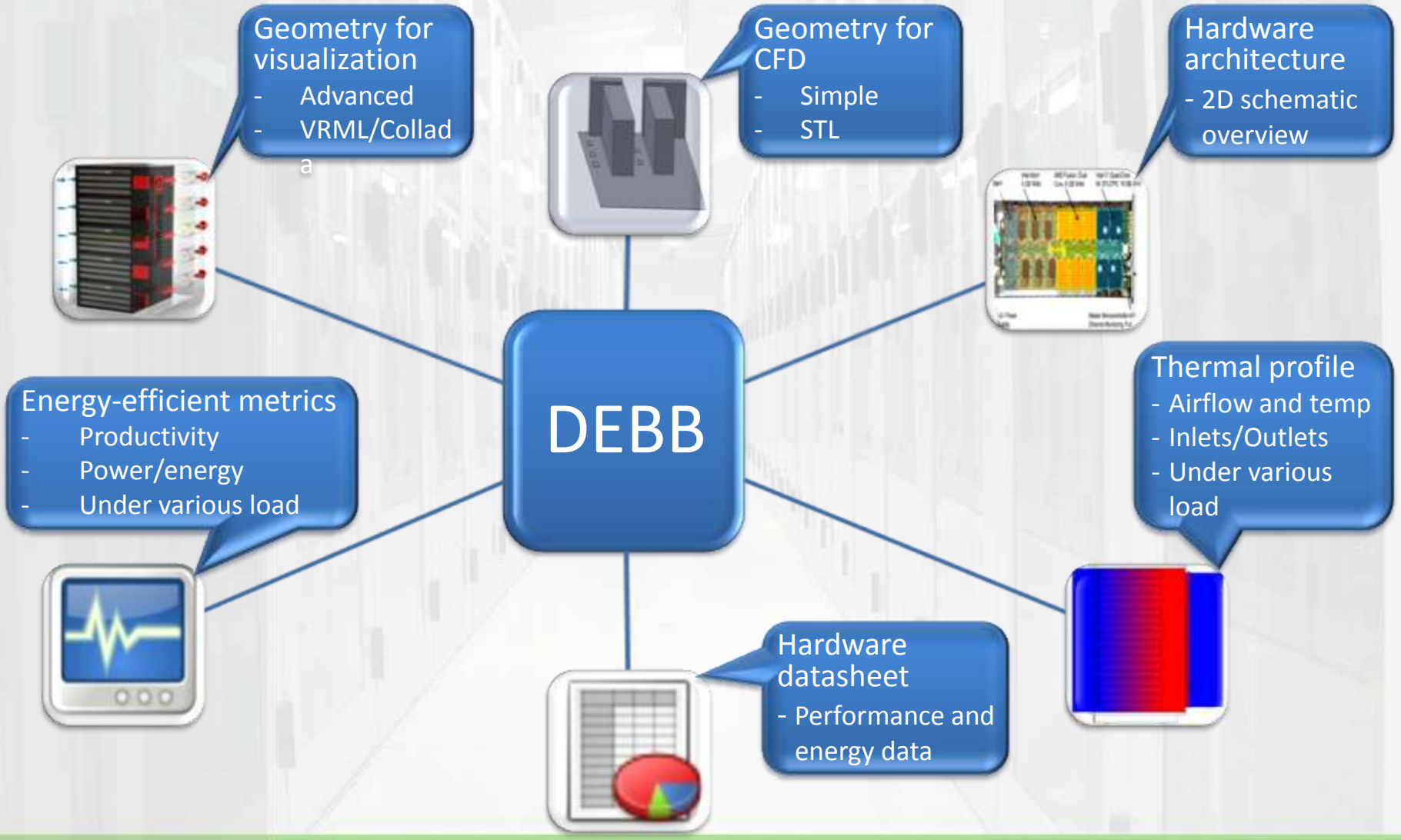
ComputeBox

- ComputeBox concept
  - Modular, self-contained pool of machines
  - Various levels of ComputeBox (e.g single or multiple racks)
  - Consist of Data Center Efficiency Building Blocks (DEBBs)
- Open ComputeBox Blueprints
  - Open designs of ComputeBoxes including hardware architecture, cooling solution, node density and performance
  - Similar idea to Open Compute Project initiated by Facebook (but more focused on DC building blocks)

DEBB Concept

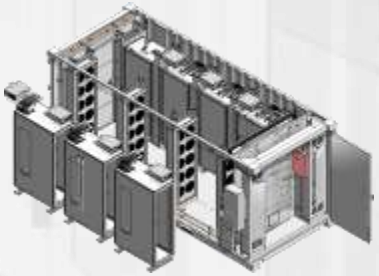
- What is a DEBB?
  - **D**ata Center **E**fficiency **B**uilding **B**lock
  - The DEBB is an abstraction for computing and storage hardware and describes energy efficiency of data-center building blocks on different granularity-levels.
- Purpose: To find the most energy efficient configuration while planning a data center
  - Used for thermodynamic modeling (SVD Toolkit)
  - Used for configuration and reconfiguration (for instance for capacity management)
- Availability
  - To be publicly available
  - Defined according to open specification

DEBB Specification



DEBB Granularity Levels

- Granularity-levels
  - **Node unit**
    single blade CPU unit
    (for instance a RECS CPU module)
  - **Node group**
    assembled unit of node units
    (for instance a complete RECS18)
  - **ComputeBox1**
    reflects a typical rack
  - **ComputeBox2**
    Reflects a container filled with racks and additional infrastructure



DEBBs Distribution Model

- Open specification to be delivered by CoolEmAll
  - Including reference model for RECS system
  - Can be applied to other building blocks (other servers or DC modules)!
- DEBB examples
  - Pizzabox-Design (1 Server in 1 U), Two Blades in 1 U, 4 Nodes in 1 U, 18 Nodes in 1 U, Bladecenters, Storage-modules
- Licensing and availability
  - The blueprints of the designed and evaluated ComputeBoxes published as a project outcome
  - DEBBS will be available as a free download on a website
    - Printed Circuit Board (PCB) layouts are not going to get published

Simulation, Visualization and Decision Support Toolkit

SVD TOOLKIT

SVD Toolkit Concept

Application types
- HPC
- Virtual machines
- Container(s)

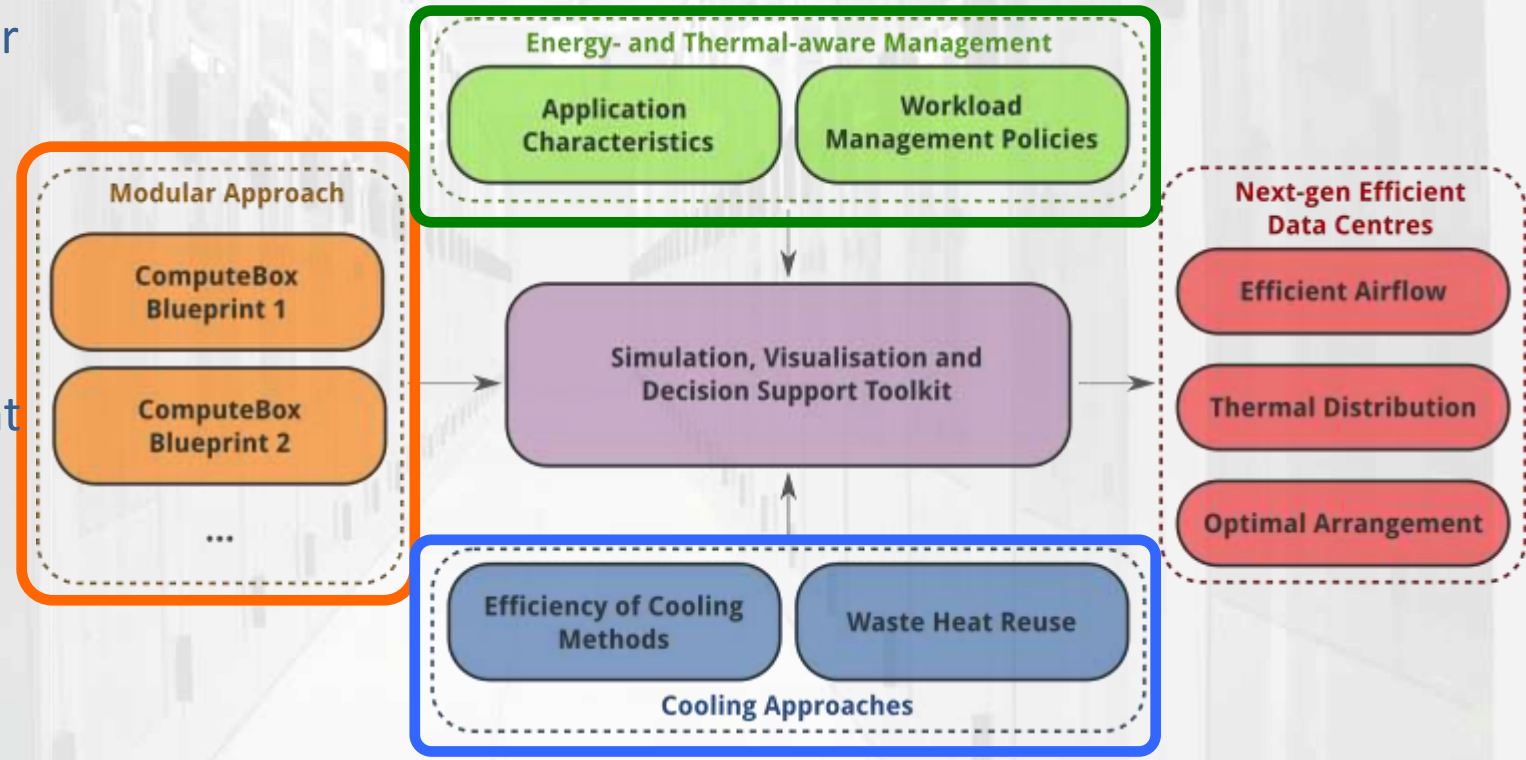
Application characteristics
- CPU-bound
- IO-bound
- Low density

Workload mngmt policies
- Workload consolidation
- Energy-aware policies
- Thermal-aware policies

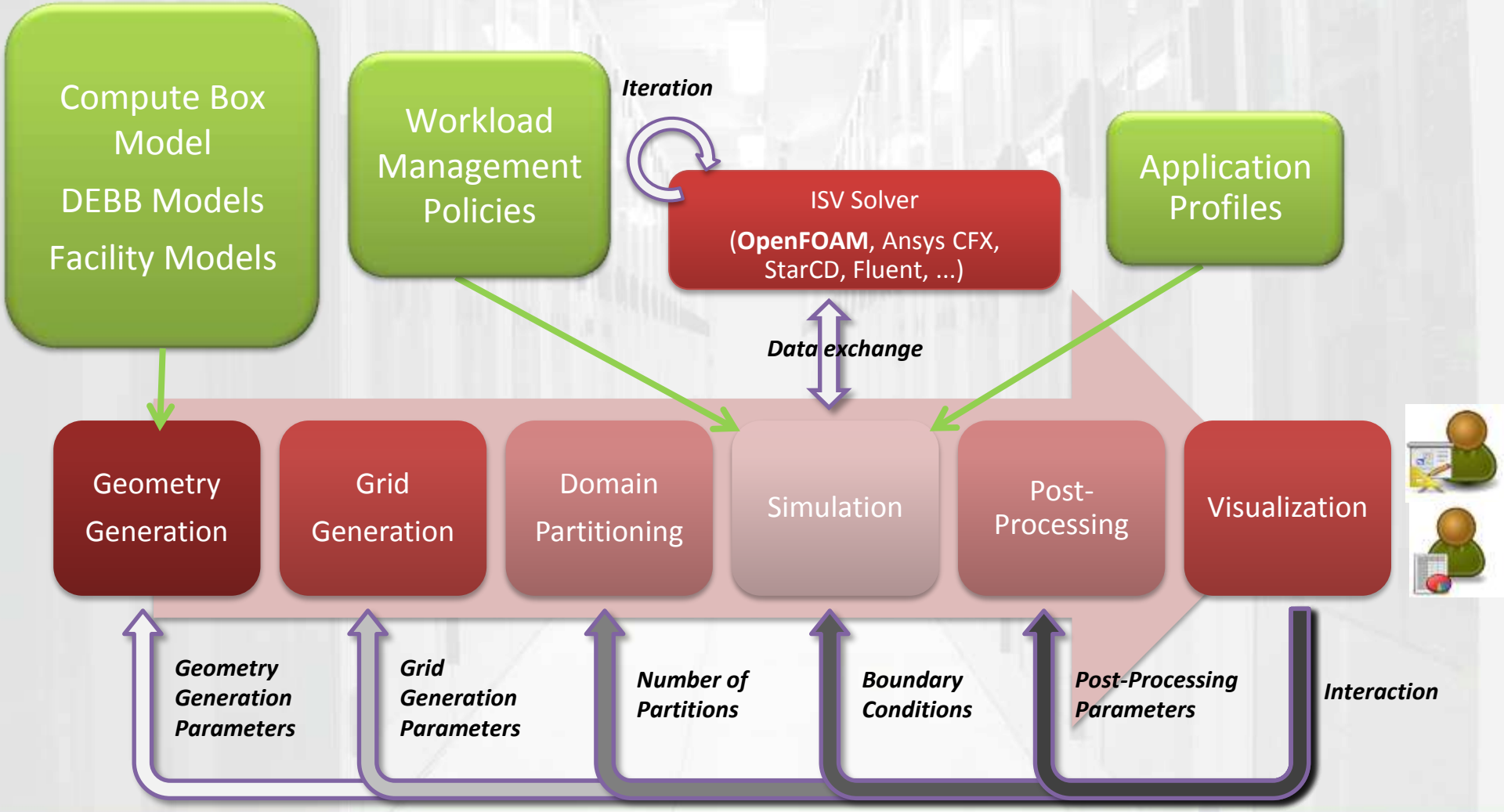
Scale
- DC
- Rack
- Server

Density
- High density (up to hundreds nodes in a rack)
- Low density

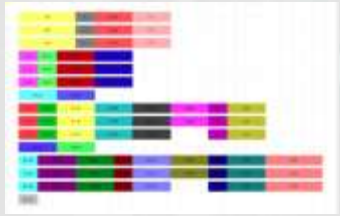
Cooling
- CRAC
- Indirect water cooling
- Direct water cooling
- Free air cooling (air economizer)
- Heat re-use integrated
- No integrated cooling

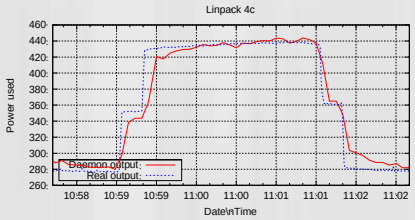


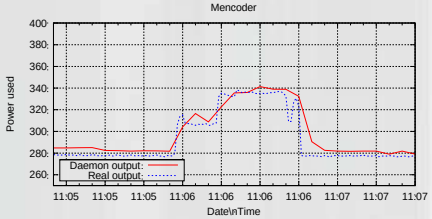
Simulation and Vizualization Flow



Modelling Workloads and Applications



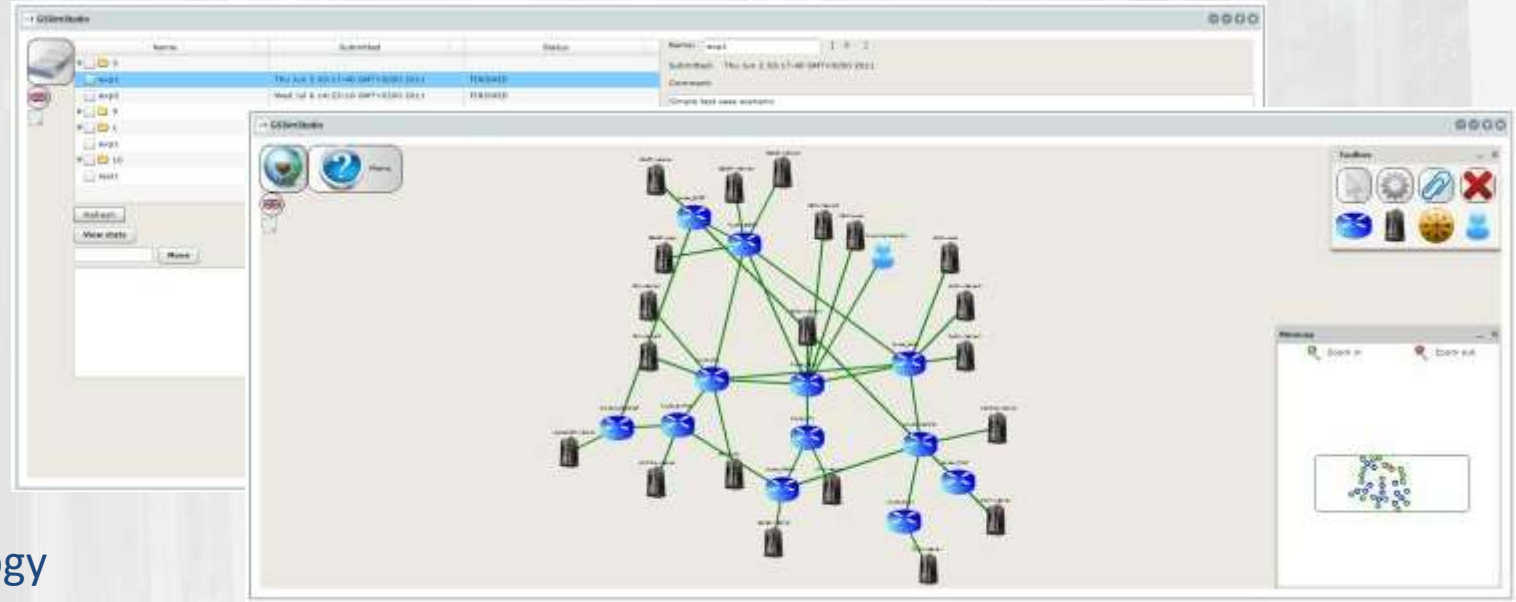




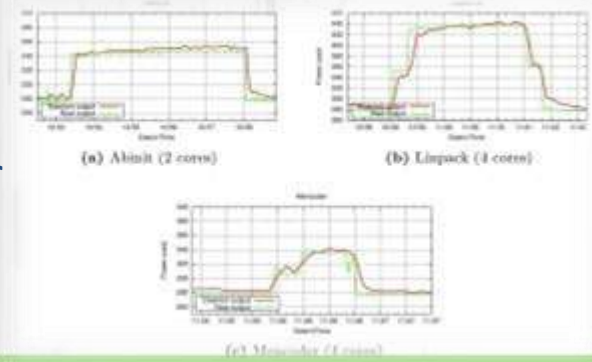
- Workload classes differ depending on centre type
  - HPC applications, high throughput jobs, virtualization, services
- Power usage and heat generation by various applications differs
  - Even more than 50W and 10-20C difference…
  - …and changes in time!
- Need for benchmarks and metrics that take into account space and time (PUE-like not sufficient)
  - Heat transfer processes depend on space and time
  - Benchmarks need to be classified according to performance, power usage, and thermal impact

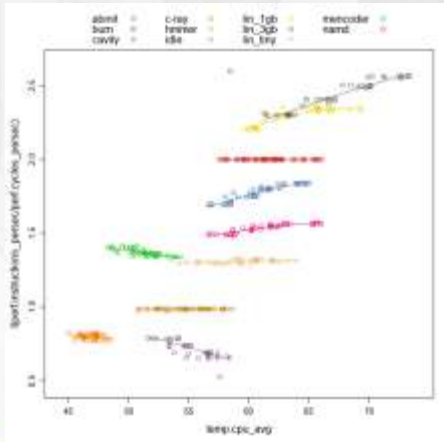
Simulation of Workloads

- Data center workload simulator
  - workload modelling
  - management policies
  - resource and network topology

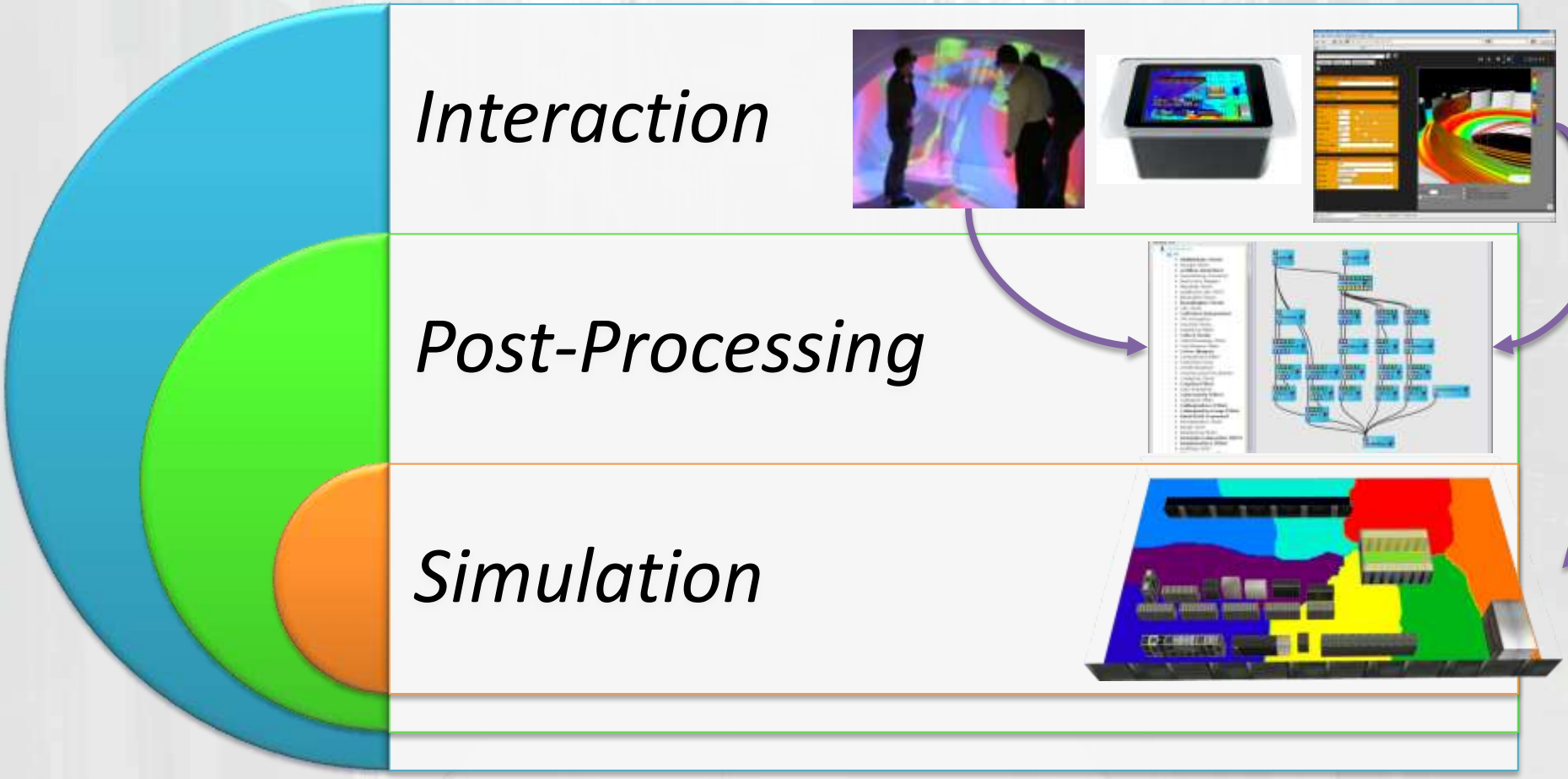


- Application models
  - estimated power usage
  - application characteristics

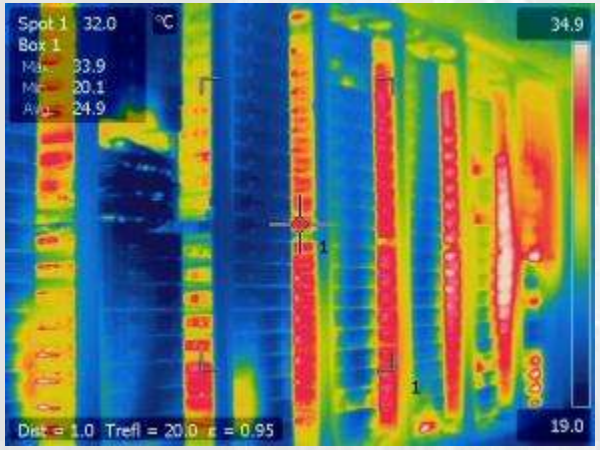


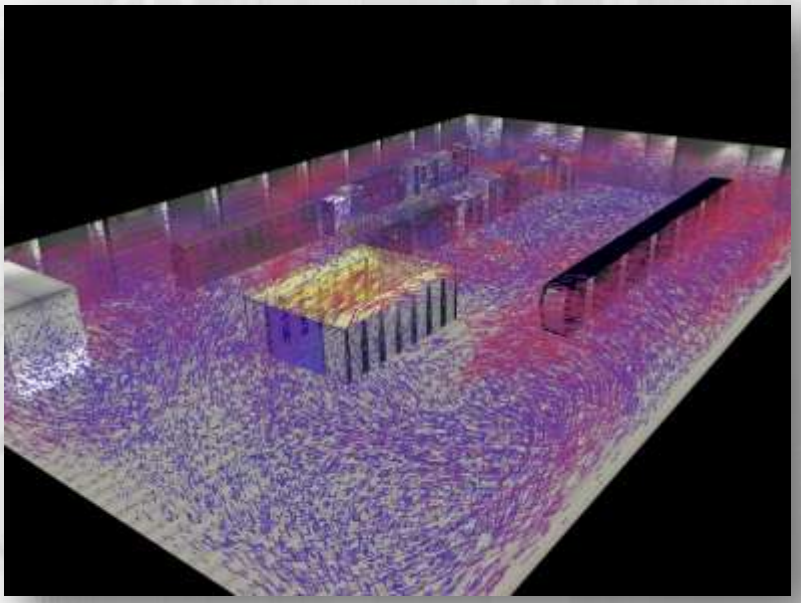


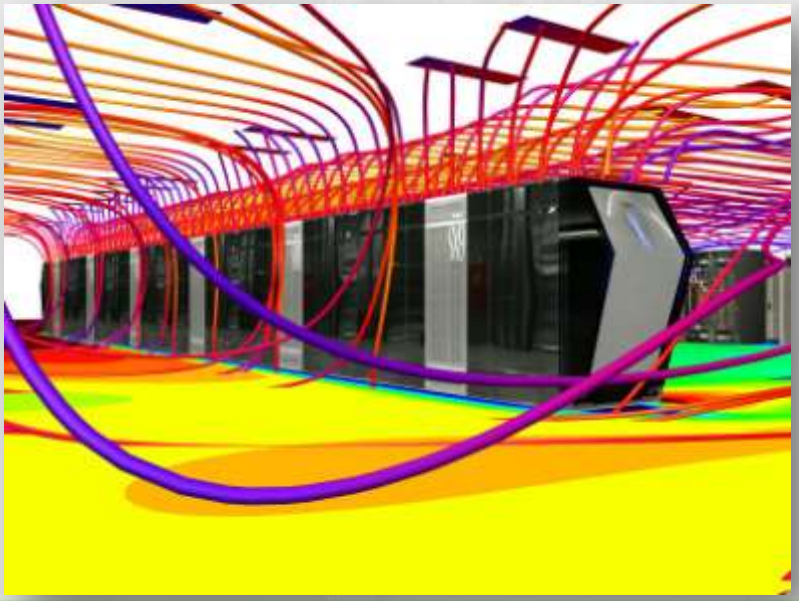
Advanced Visualization



SVD Toolkit Output







- Virtual thermal camera
- Heat transfer and air flow processes
- Remote, interactive visualization

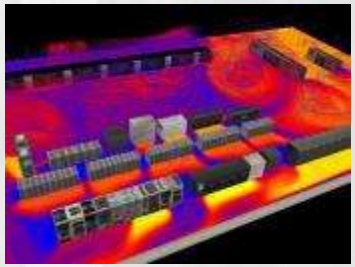
What's next?

CONCLUSIONS AND FUTURE PLANS

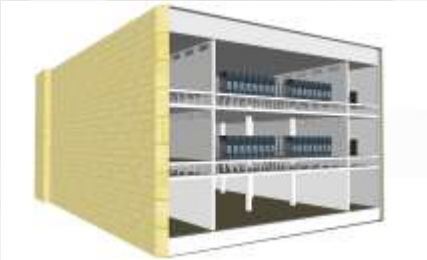
General Use Cases

- Christmann RECS designs
  - RECS with a higher level of integrated management
  - Towards "data centre in a rack"

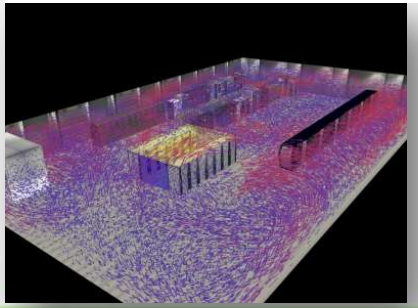
- HLRS existing centre
  - 5MW of power, ~2Petaflops
  - Dynamic, interactive visualization to optimize temperature distribution



- PSNC existing and new centre
  - Hot/cold air mixing problem in existing DC
  - Completely new centre in design phase
  - 2x800m^2, up to ~170 racks, 2MW-16MW



- ATOS Origin data centre
  - 45 data centres, total capacity ~40.000 m^2
  - Greener Data Centres global program
  - Requirements from industrial DCs



Conclusions and Plans

- CoolEmAll will deliver two main products:
  - Open source **SVD toolkit** to investigate data center energy-efficiency based on simulations
  - Best practises and open designs (**ComputeBox** and **DEBB** blueprints)

- These outcomes will allow minimising the energy consumption of data centres by:
  - **Optimisation** of their **design** and **operation**
  - **Tuning** to specific **workloads** and **conditions**
- Results available soon:
  - Next generation RECS prototype
  - Monitoring and control platform
  - Open DEBB specification

Questions?

Comments?

wolfgang.christmann@christmann.info

www.coolemall.eu