UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

URZ

EMCL

Heidelberg Institute for Theoretical Studies

HITS

- Part 1 -
High Performance Computing: Quid ?

Vincent Heuveline

From „Universitas“ to the Comprehensive University of the Future

Oldest German university (founded in 1386)

Successful in both rounds of the Excellence Initiative

Students (term 14/15)	30,898
Doctorates (2013)	1,237
Total number of staff	13,544

Energy Efficient High Performance Computing

1. **Data centre point of view – HPC installation at URZ**

   Computing Centre of Heidelberg University

2. **Numerics point of view – current activities at EMCL/DMQ**

   Engineering Mathematics and Computing Lab @IWR
   Data Mining and Uncertainty Quantification @HITS

EMCL HITS

Engineering Mathematics and Computing Lab (EMCL@IWR)

– Research group in Scientific Computing and High Performance Computing

– Link mathematical modeling, numerical simulation, optimization and hardware aware computing

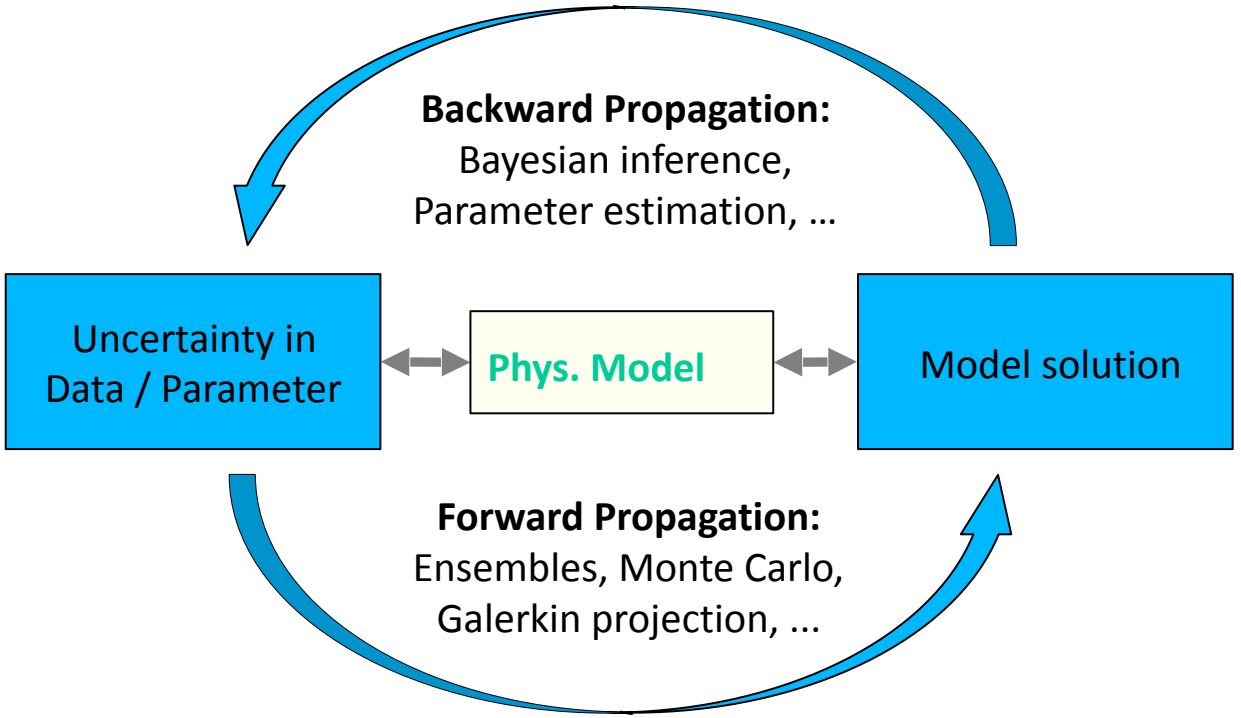
– Interdisciplinary research activities with application in
  - Meteorology, climate and environment
  - Medical Engineering
  - Energy research

– Interface with the industry, technology transfer

IWR
Interdisciplinary Center for Scientific Computing

EMCL HITS

Data Mining and Uncertainty Quantification (DMQ@HITS)



- Uncertainty in measurements, lack of knowledge in model parameters,…
- Quantify the impact of the uncertainty
- Compute mean values, variances, higher order moments, paths, …
- Large complex systems of equations requires High-Performance Computing

Exa2Green Project

Project Partners

Engineering Mathematics and Computing Lab (EMCL)
Interdisciplinary Center for Scientific Computing (IWR)
Heidelberg University
www.emcl.uni-hd.de
Germany

High Performance Computing and Architectures Group
Universitat Jaume I de Castellon
www.hpca.uji.es
Spain

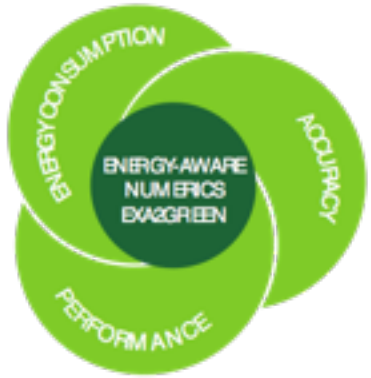
IBM Research - Zurich
www.zurich.ibm.com
Switzerland

Institute for Meteorology and Climate Research
Karlsruhe Institute of Technology (KIT)
www.imk-tro.kit.edu
Germany

Scientific Computing Group
Department of Informatics
Universität Hamburg
www.wr.informatik.uni-hamburg.de
Germany

Steinbeis-Europa-Zentrum
www.steinbeis-europa.de
Germany

Swiss Federal Institute of Technology Zurich
Swiss National Supercomputing Centre
www.cscs.ch
Switzerland



www.exa2green.eu

Contact

Project coordinator

Prof. Dr. Vincent Heuveline
Engineering Mathematics and Computing Lab (EMCL)
Interdisciplinary Center for Scientific Computing (IWR)
Heidelberg University
vincent.heuveline@iwr.uni-heidelberg.de
www.emcl.uni-hd.de
www.iwr.uni-heidelberg.de

For further information on Exa2Green please visit
www.exa2green.eu

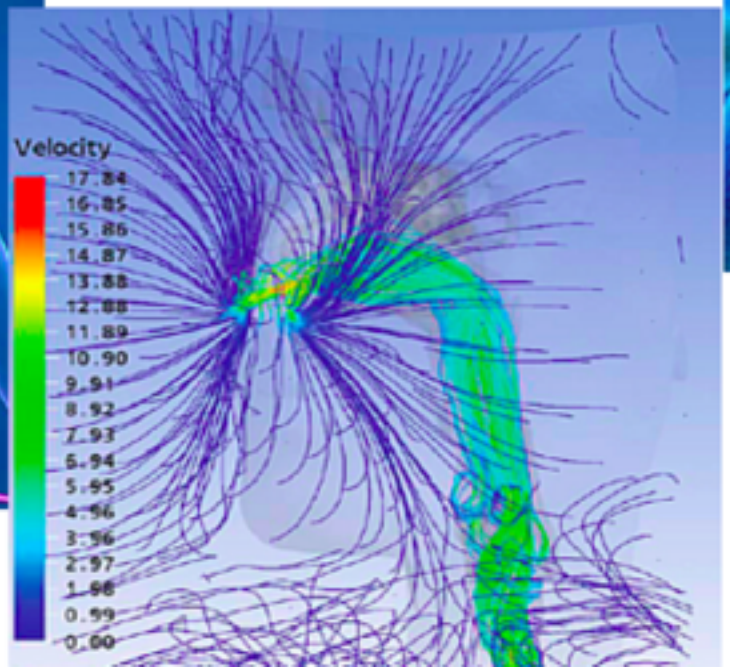
Exa2Green
energy-aware numerics

Energy-aware Sustainable Computing
on Future Technology
Paving the Road to Exascale Computing

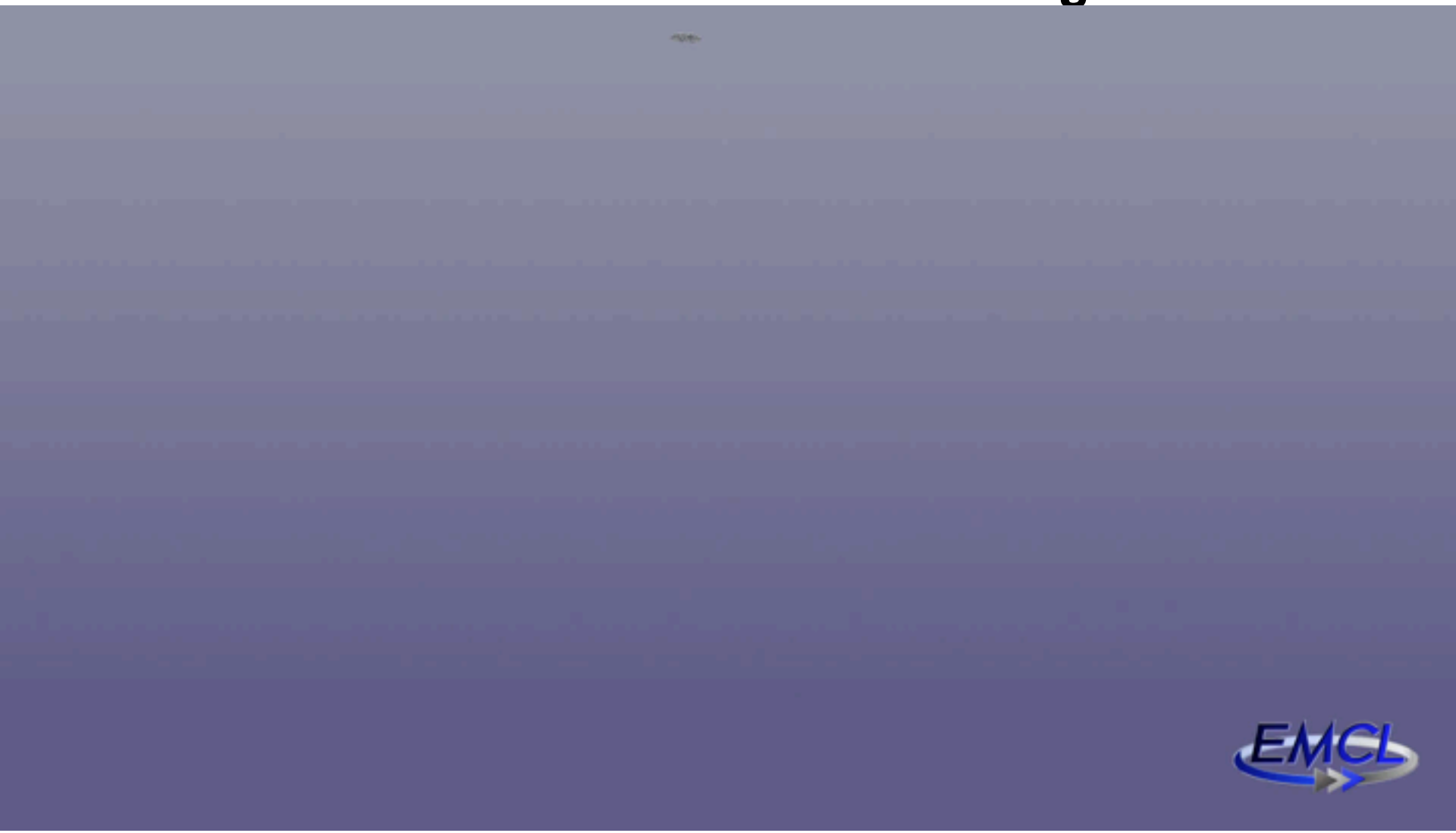
Exa2Green is co-financed by the European Commission under the 7th Framework Programme

Application in Medical Engineering

- Study of the **environmental impact** on the nose, sinuses, larynx and lungs' airways
- Quantitative analysis of the inhaled **deposition in the airways**



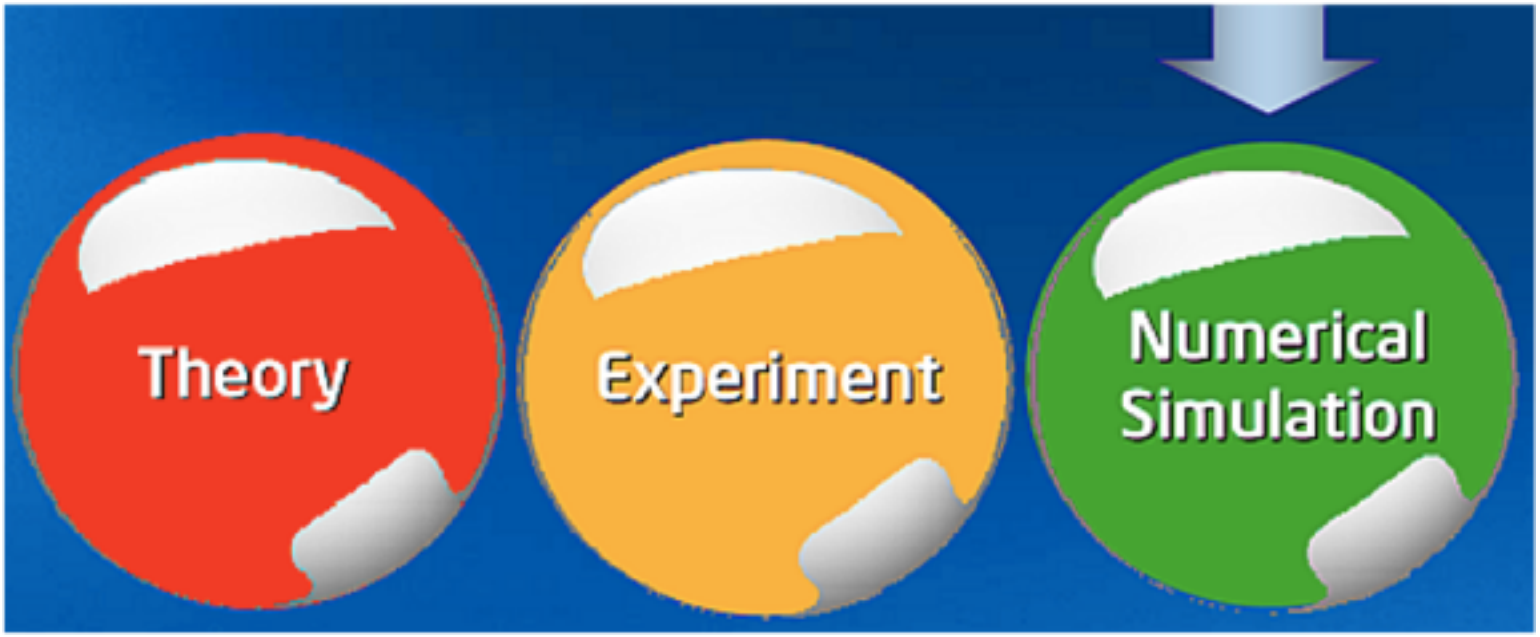
Numerical Simulation of the Human Lung



Hyperreality vs. Infrareality

Terminology: Philosopher Jean Baudrillard (1929-2007)
„Simulakrum and Simulation“

Simulation: von lat. simulare, imitate, mimic
Simulakrum: von lat. Simulo: mirror, image, illusion,chimera



… Infrareality

– Need to define the reality ...

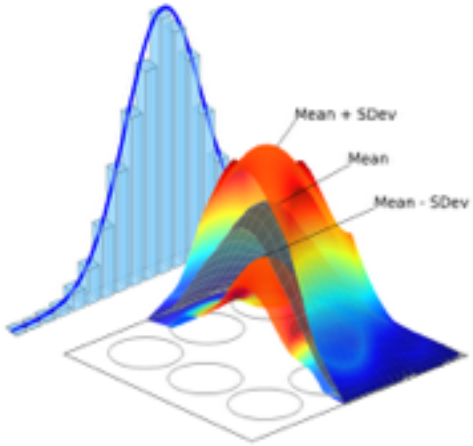
– Mathematical models (complexity not always better!)

– Lack of ressources
  - Compute power
  - Time
  - Money

– Numerical methods (convergence, condition number)

– Hardware (rounding error, ...)

– Error in measurement (Modelcalibration, modelvalidation)
– ...



Definition of Supercomputer

Thomas Sterling, 2007

A computing system exhibiting high-end performance capabilities and resource capacities within practical constraints of technology, cost, power, and reliability.

Merriam-Webster Online

a large very fast mainframe used especially for scientific computations.

Encyclopedia Britannica Online

any of a class of extremely powerful computers. The term is commonly applied to the fastest high- performance systems available at any given time. Such computers are used primarily for scientific and engineering work requiring exceedingly high-speed computations.

Measure of Performance

FLOPS = Floating Point Operation per Second

- 1 MFLOPS $= 10^6$ FLOPS
- 1 GFLOPS $= 10^9$ FLOPS
- 1 TFLOPS $= 10^{12}$ FLOPS
- 1 PFLOPS $= 10^{15}$ FLOPS

Two perspectives on Performance

Peak performance

- Maximum theoretical performance possible for a system

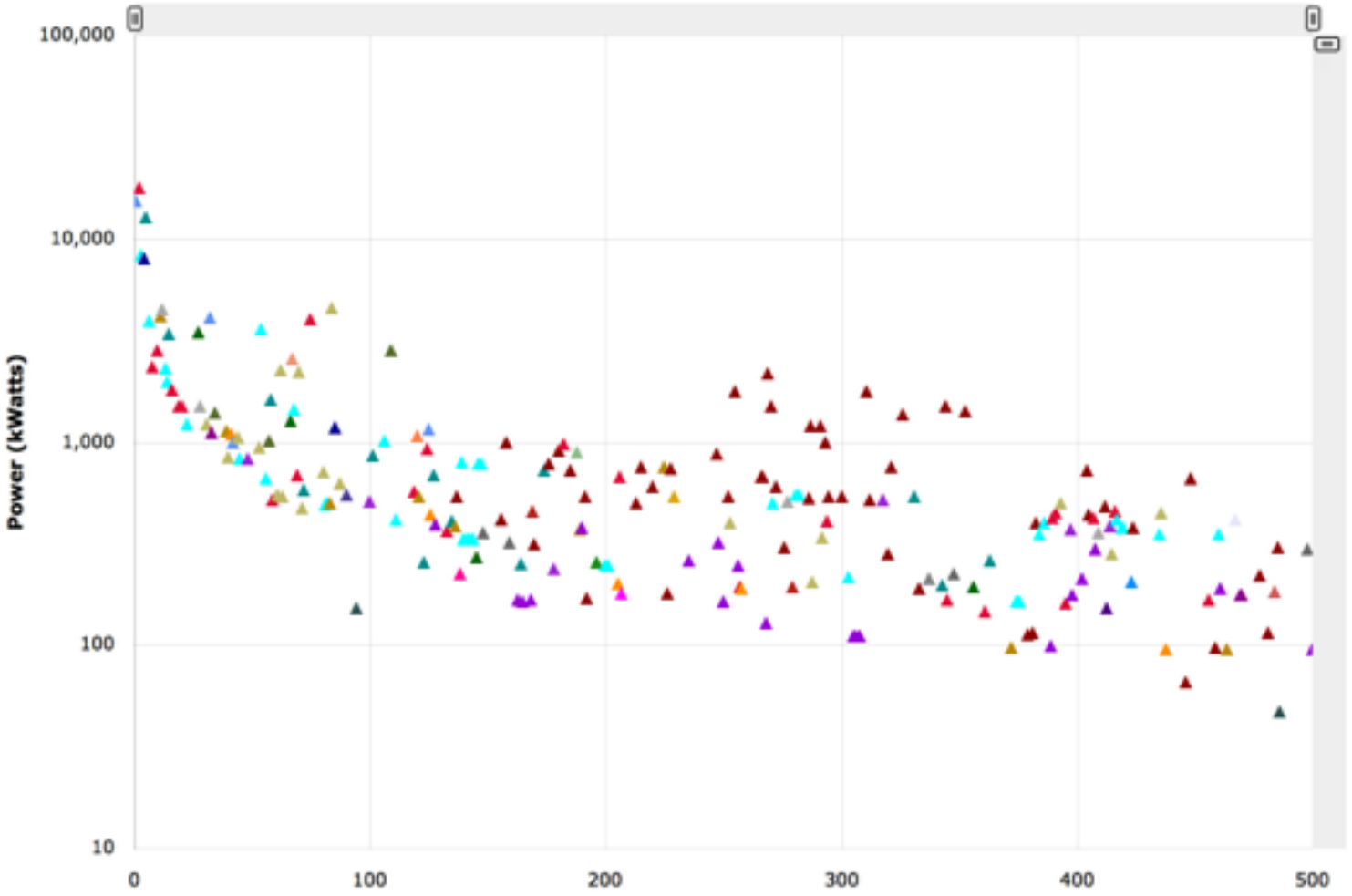
Sustained performance

- Observed performance for a particular workload and run
- Varies across workloads and possibly between runs

Top 500 - June 2016

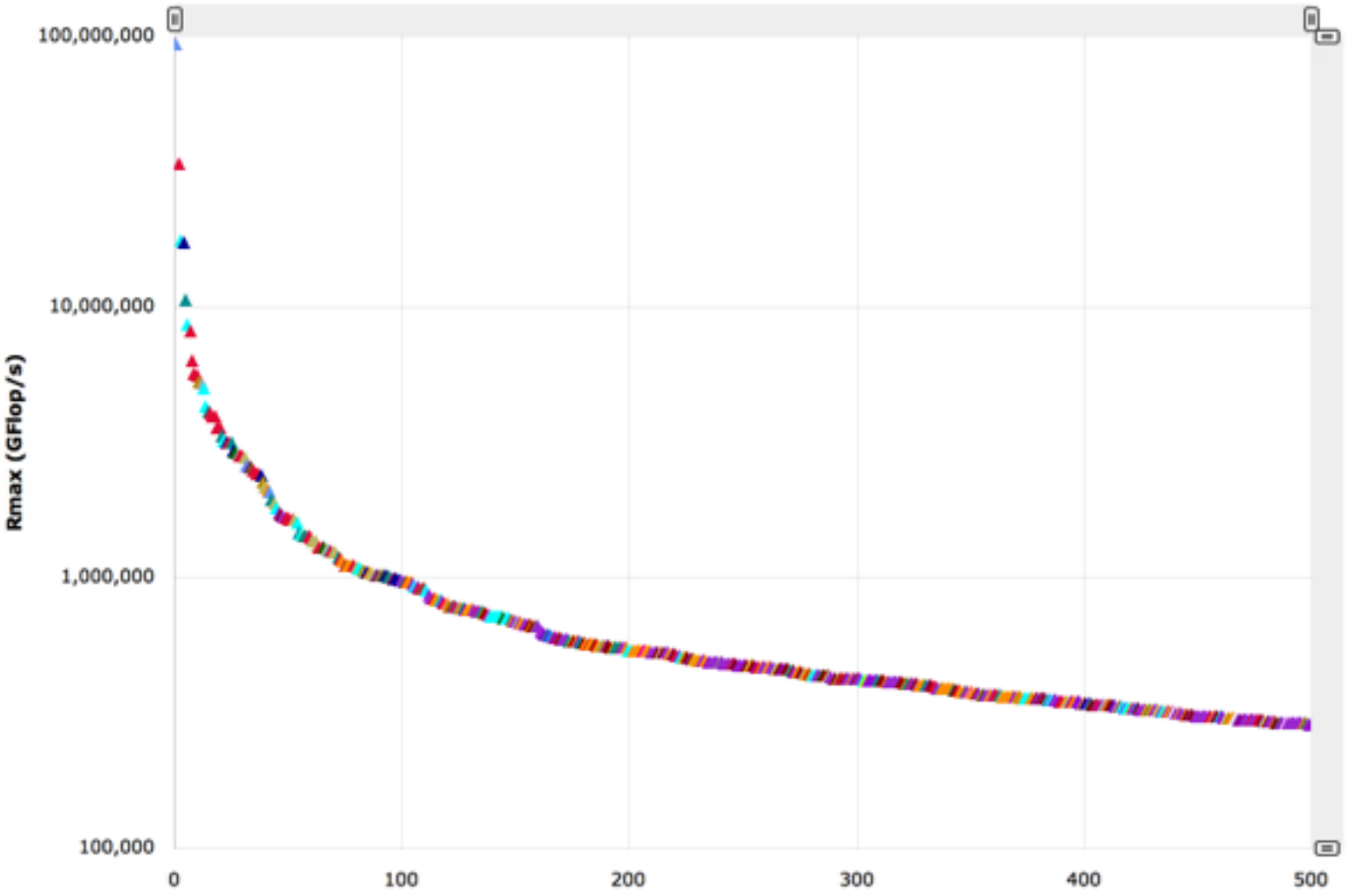
Rank	Site	System	Cores	Rpeak (TFLOP/S)	Power (kW)
1	National Supercomputing Center in Wuxi **China**	Sunway SW26010 260C 1.45GHz **NRCPC**	10,649,600	125,435	15,371
2	National Super Computer Center in Guangzhou **China**	Intel Xeon E5-2692 12C 2.200GHz **NUDT**	3,120,000	54,902	17,808
3	DOE/SC/Oak Ridge National Laboratory **United States**	Titan - Cray XK7 , Opteron 6274 16C 2.200GHz **Cray Inc.**	560,640	27,112	8,209
4	DOE/NNSA/LLNL Livermore **United States**	BlueGene/Q, Power BQC 16C 1.60 GHz **IBM**	1,572,864	20,132	7,890
5	RIKEN Advanced Institute for Computational Science **Japan**	K computer, SPARC64 2.0GHz **Fujitsu**	705,024	11,280	12,660

Top 500 - Power Consumption (kWatts)



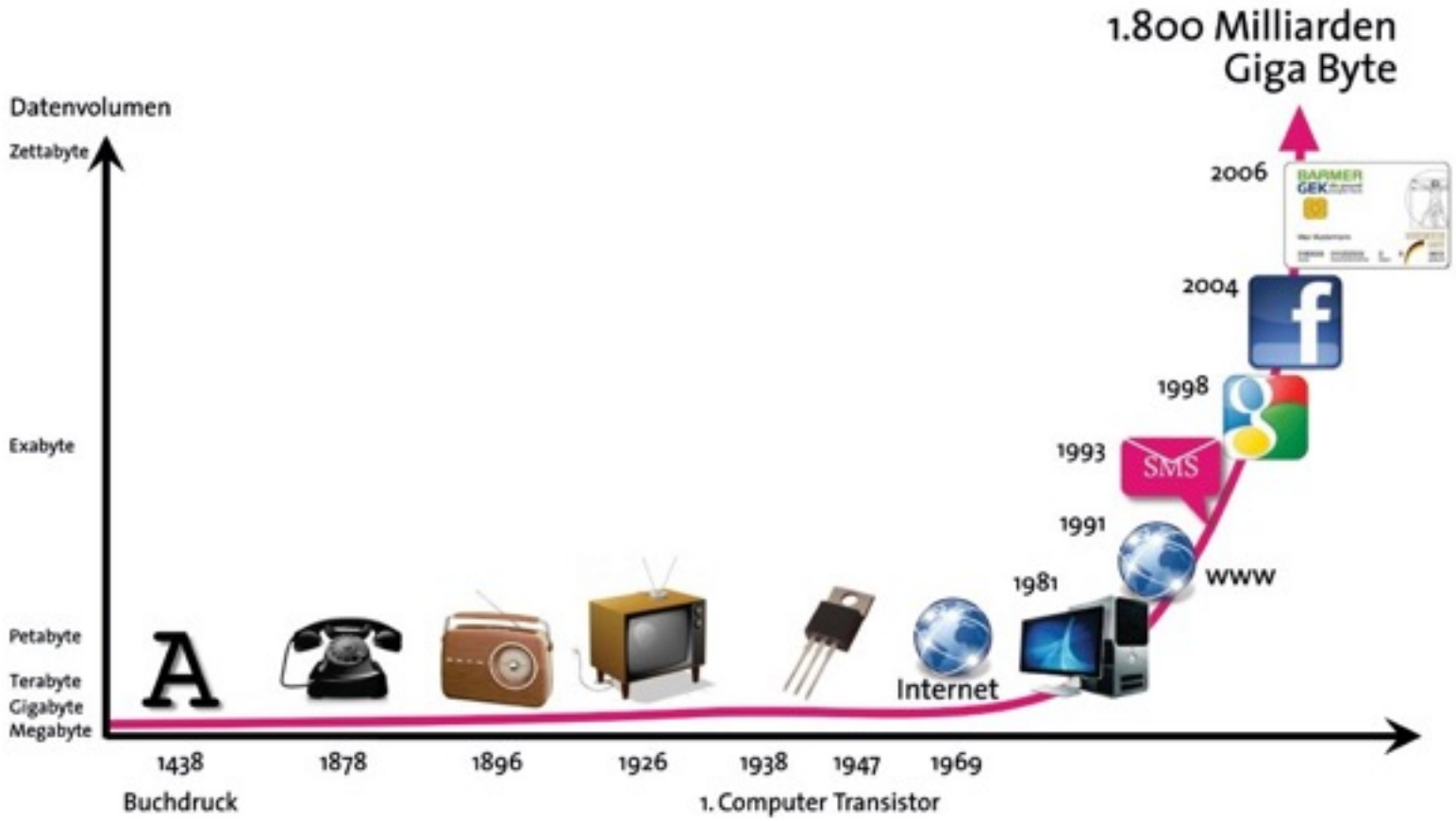
Source: top500.org

Top 500 - Performance (GFlop/s)



Source: top500.org

Exascale Computing / Datenvolumen



Source: „Big-Data im Praxiseinsatz, Leitfaden", BITKOM 2012

Exa: 10^{18}

Pokemon-Go

- Energy, Data storage?
- How do you recognize a power player?

Pokemon-Go … Powerbank

Data centre point of view – HPC installation at URZ

– Energy consumption for IT and cooling
  - Servers and network components
  - Chillers, computer room air conditioners – cold air, cold water
  - Power supplies including uninterruptible power supplies (UPS)
  - Fans

– Other energy consumption
  - Heating
  - Air conditioning for offices
  - Warm water
  - Other building equipment, e.g. light

Emergency power generator at URZ

Energy consumption in data centres

– Measure for efficiency of energy usage

Power Usage Effectiveness – PUE

$$\mathrm{PUE} = \frac{\text{Total energy data centre (kWh)}}{\text{IT energy (kWh)}}$$

Ideal: PUE = 1.0
Usual: PUE ~ 1.5 (good) ... 2.0 (bad)

Computing Centre Heidelberg: PUE < 1.1 (very good :))

HPC installation at URZ

- total of ~600 compute nodes based on Intel Haswell
- total of ~10,000 CPU cores
- includes co-processor nodes based on Intel Xeon Phi
- IT power ~250 kW

Server racks in computer room at URZ

Current situation in HPC centres

- High packing and energy density
  - **30 kW** per rack already operational
  - Envisage **100 kW** and more per rack
- High load on servers
  - Want to operate at **full capacity**
- Often homogeneous infrastructure
  - **Dedicated cooling** circuit for one machine reasonable
- **Reuse of waste heat** possible, but not constantly available
  - Unplanned (crash) or planned (maintenance) shutdowns reduce availability

What about the costs?

– example for 200 kW use

  200 kW * 24 h/day * 365 day/year * 20 Ct/kWh ≈ 350,000 EUR/year

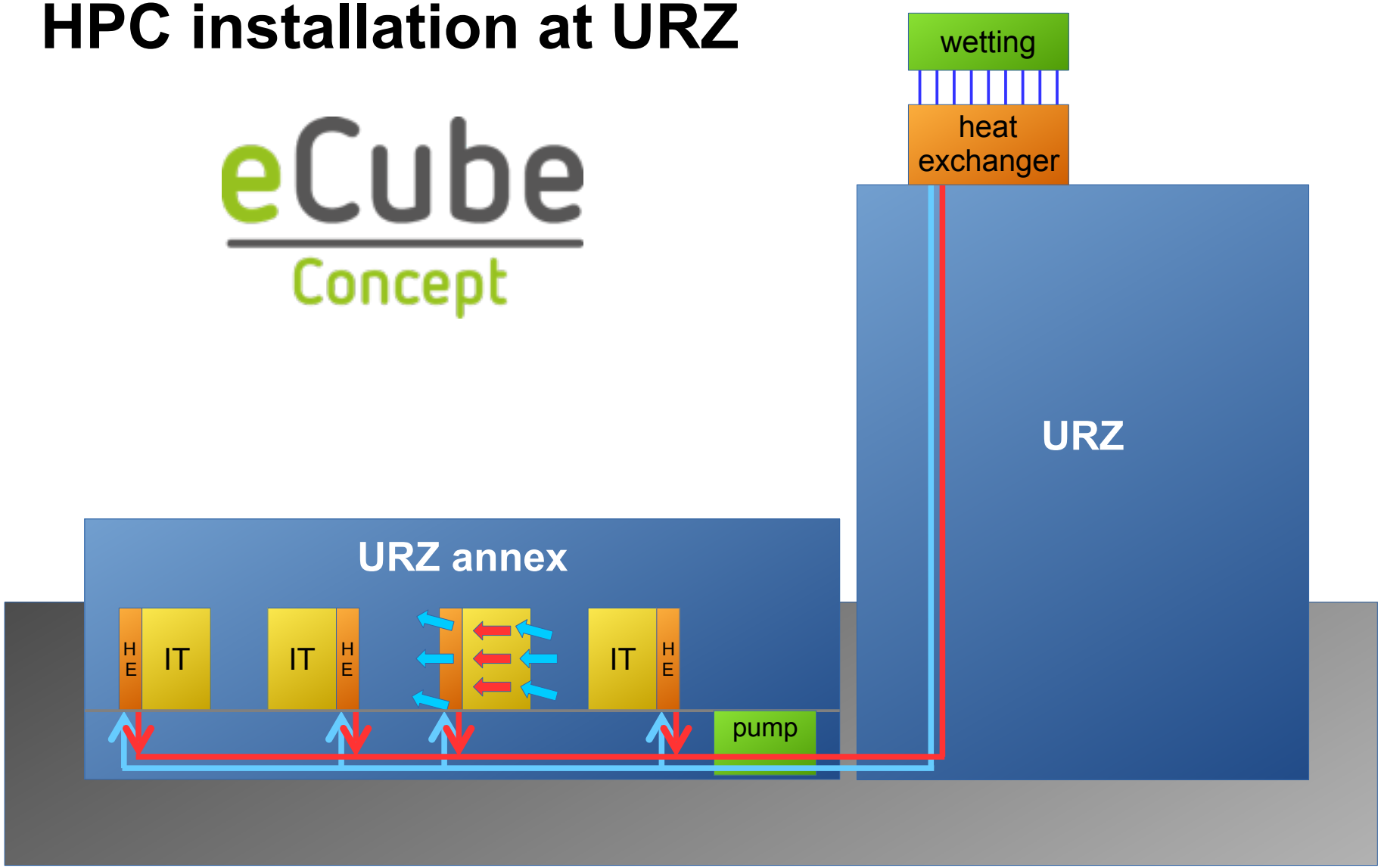
– traditional cold air cooling: PUE ≈ 2

  ► additional 0.35 Mio. EUR/year only for cooling!

– URZ uses passive cooling technology
  - outdoor environmental temperature used to cool IT
  - additional cooling by evaporation in hot summer days
  - goal: achieve PUE ~ 1.0

HPC installation at URZ

eCube Concept



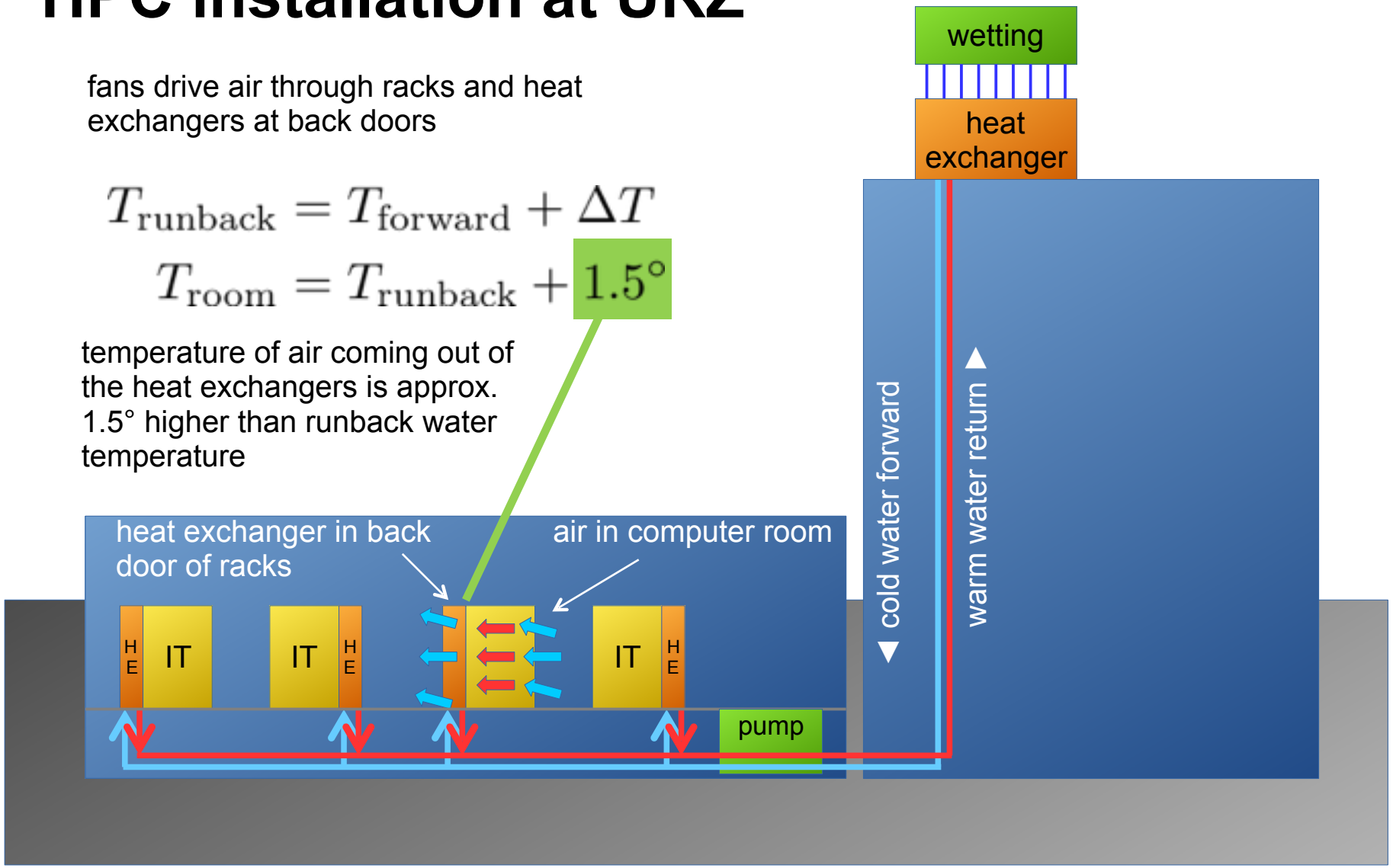
HPC installation at URZ

fans drive air through racks and heat exchangers at back doors

$$T_{\text{runback}} = T_{\text{forward}} + \Delta T$$

$$T_{\text{room}} = T_{\text{runback}} + 1.5°$$

temperature of air coming out of the heat exchangers is approx. 1.5° higher than runback water temperature



HPC installation at URZ

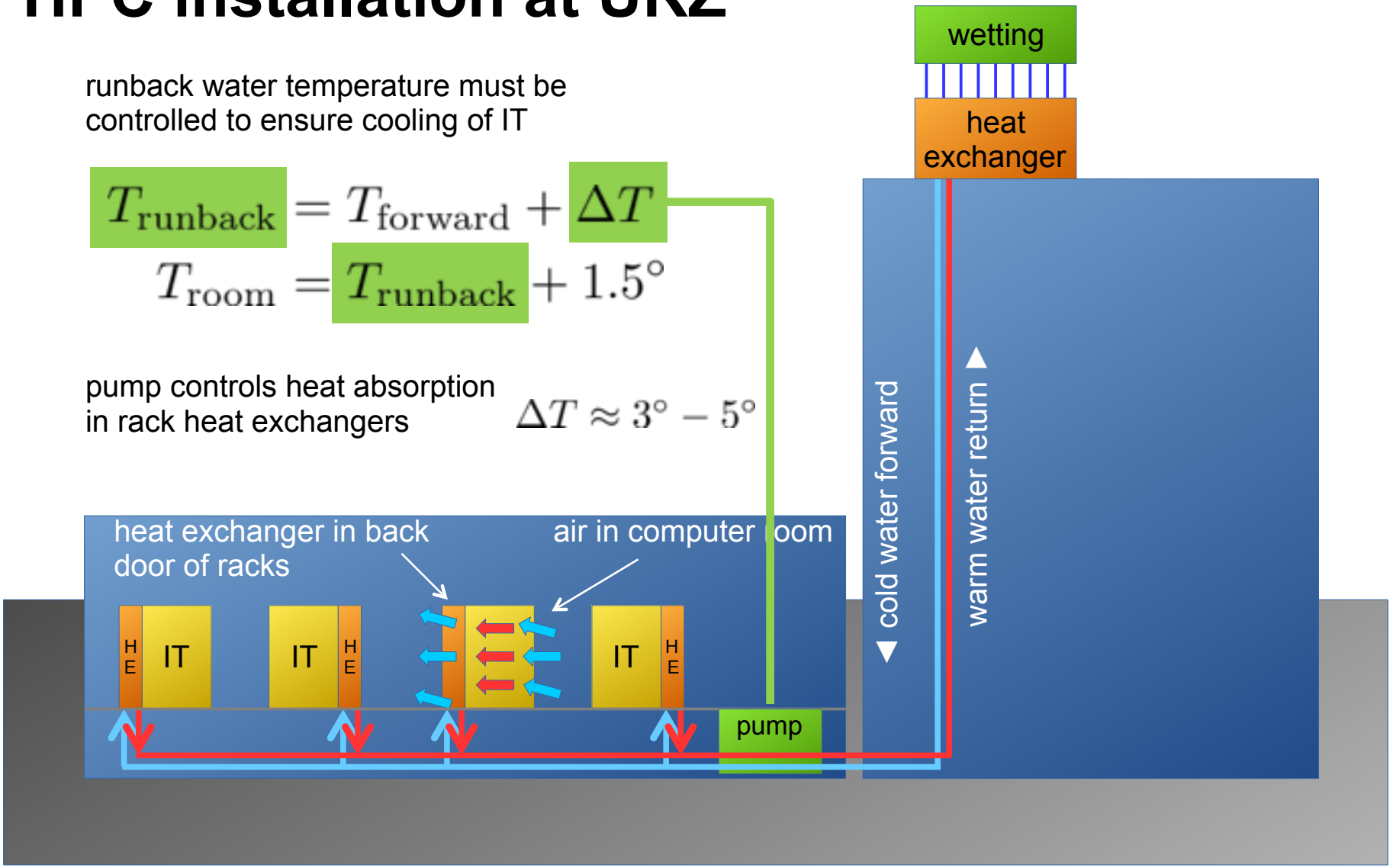
runback water temperature must be controlled to ensure cooling of IT

$$T_{\text{runback}} = T_{\text{forward}} + \Delta T$$

$$T_{\text{room}} = T_{\text{runback}} + 1.5°$$

pump controls heat absorption in rack heat exchangers

$$\Delta T \approx 3° - 5°$$



HPC installation at URZ

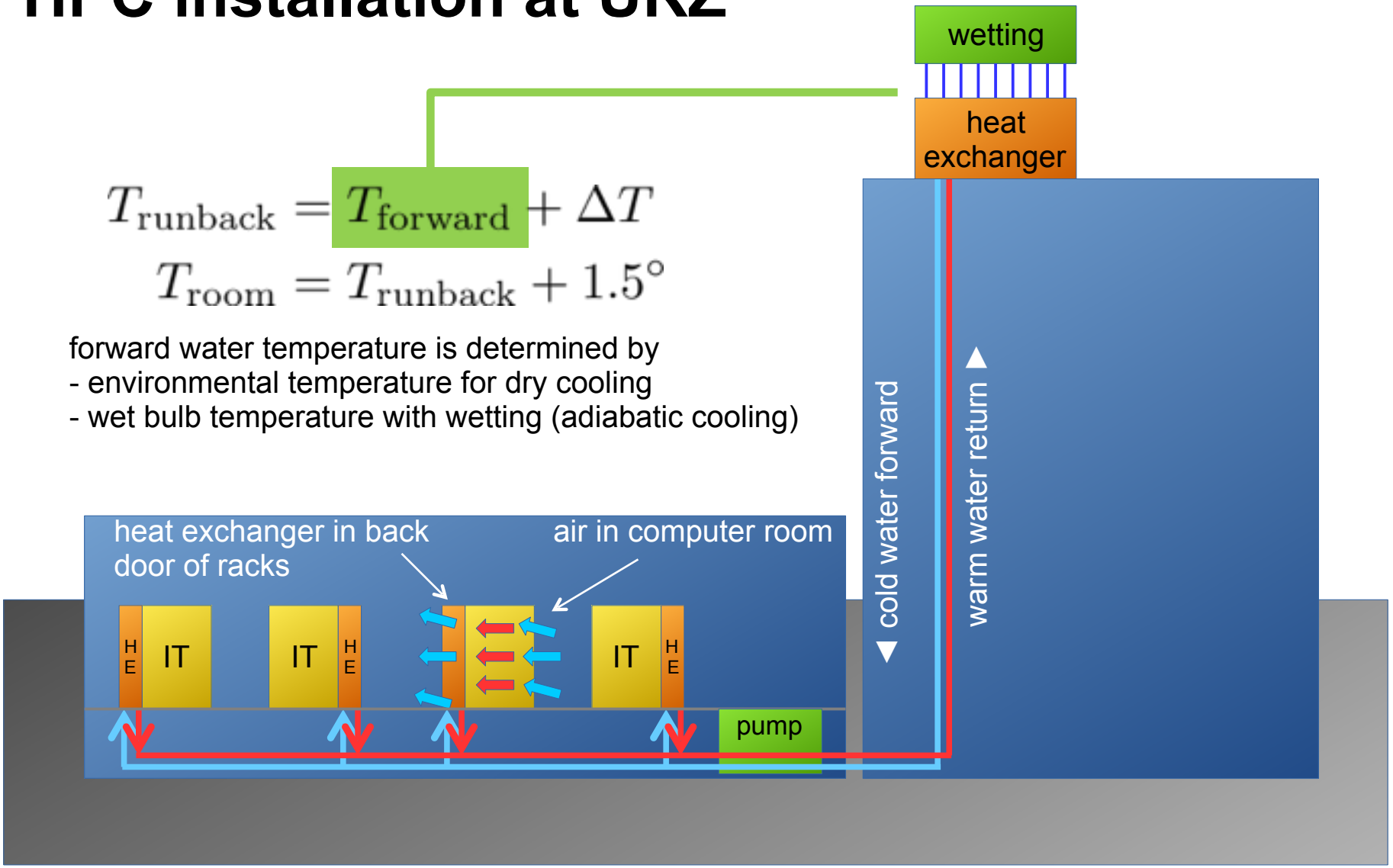
Heat exchanger and wetting on roof of URZ

HPC installation at URZ

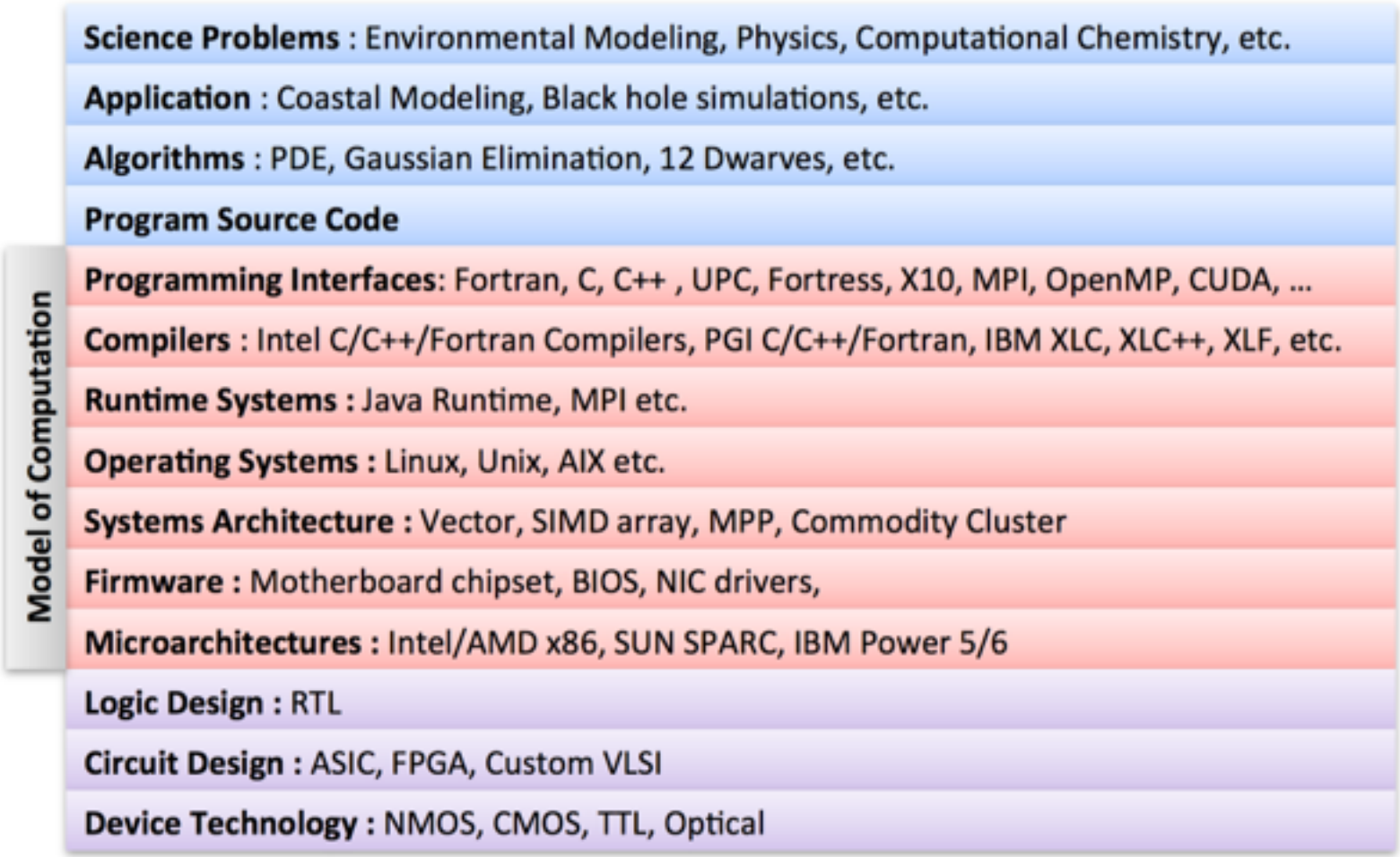
$$T_{\text{runback}} = T_{\text{forward}} + \Delta T$$

$$T_{\text{room}} = T_{\text{runback}} + 1.5°$$

forward water temperature is determined by
- environmental temperature for dry cooling
- wet bulb temperature with wetting (adiabatic cooling)



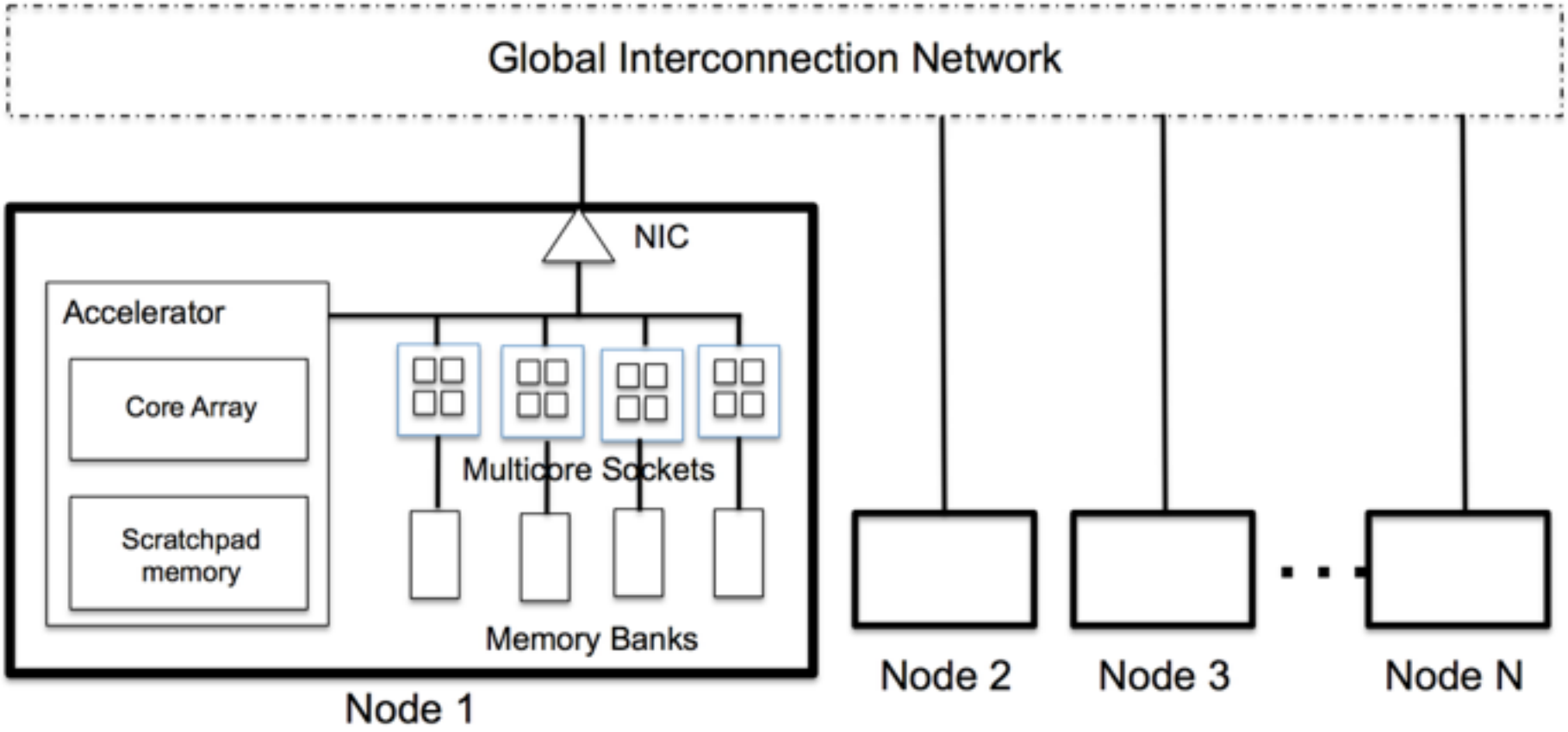
HPC System Stack



Practical Constraints

- Cost
  - ► Deployment
  - ► Operational support
- Power
  - ► Energy required to run the computer
  - ► Energy for cooling
- Size
  - ► Floor space …
- Generality
  - ► How good is it across a range of different problems
- Usability
  - ► How hard is it to program and manage

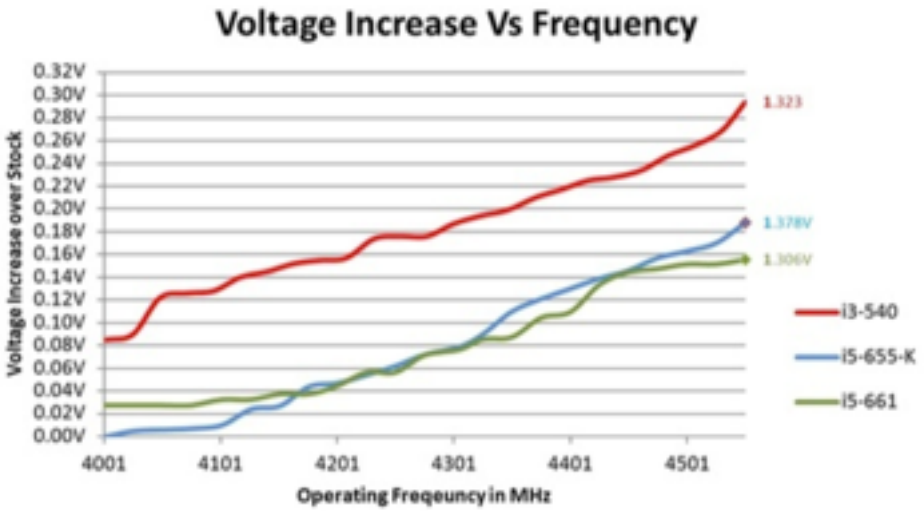
Heterogeneous Multicore System Architecture



Why Multicore?

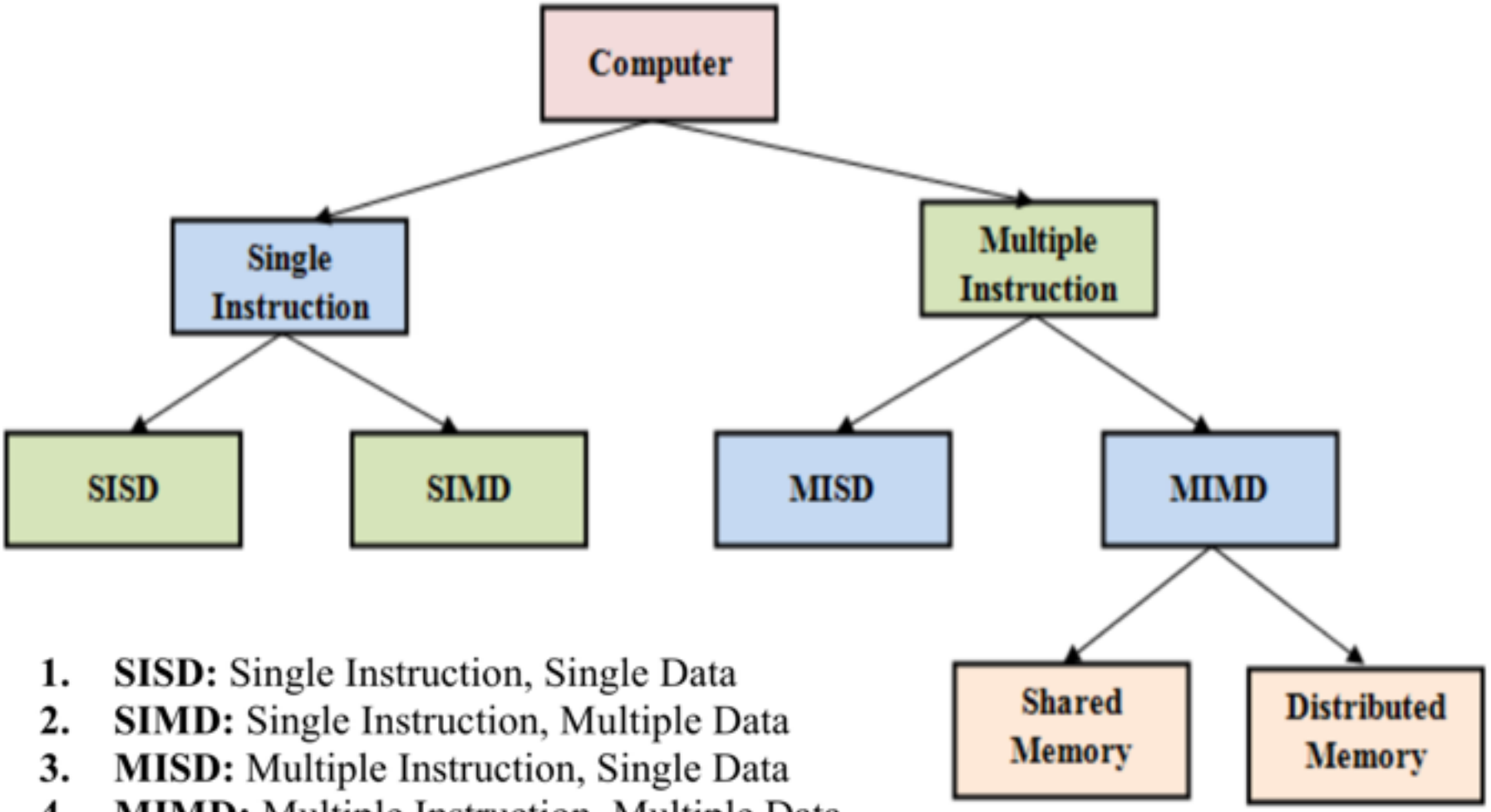
$$P=V^2/R$$

- Where P is **power**, V is **voltage** and R is **resistance**. You can see that there is an exponential in that equation (V^2). This means that power consumption grows at the square of voltage. So an increase in max frequency results in a quadratic growth in power.
- Compare this with using 2 cores, where the total power consumption for 2X the cores is 2X the power.



Source: www.anandtech.com

Flynn´s taxonomy



1. **SISD:** Single Instruction, Single Data
2. **SIMD:** Single Instruction, Multiple Data
3. **MISD:** Multiple Instruction, Single Data
4. **MIMD:** Multiple Instruction, Multiple Data

My Computer: error free?

Disk surface: $S = \pi * R * R$

π: irrational

First 500 digits of π:

3.
1415926535 8979323846 2643383279 5028841971 6939937510 5820974944
5923078164 0628620899 8628034825 3421170679 8214808651 3282306647
0938446095 5058223172 5359408128 4811174502 8410270193 8521105559
6446229489 5493038196 4428810975 6659334461 2847564823 3786783165
2712019091 4564856692 3460348610 4543266482 1339360726 0249141273
7245870066 0631558817 4881520920 9628292540 9171536436 7892590360
0113305305 4882046652 1384146951 9415116094 3305727036 5759591953
0921861173 8193261179 3105118548 0744623799 6274956735 1885752724
8912279381 8301194912 …

My Computer: error free? … obviously not!

Disk surface: $S = \pi * R * R$

Die Zahl π: irrational

First 500 digits of π:

3.
1415926535 8979323846 2643383279 5028841971 6939937510 5820974944
5923078164 0628620899 8628034825 3421170679 8214808651 3282306647
0938446095 5058223172 5359408128 4811174502 8410270193 8521105559
6446229489 5493038196 4428810975 6659334461 2847564823 3786783165
2712019091 4564856692 3460348610 4543266482 1339360726 0249141273
7245870066 0631558817 4881520920 9628292540 9171536436 7892590360
0113305305 4882046652 1384146951 9415116094 3305727036 5759591953
0921861173 8193261179 3105118548 0744623799 6274956735 1885752724
8912279381 8301194912 …

Curious properties

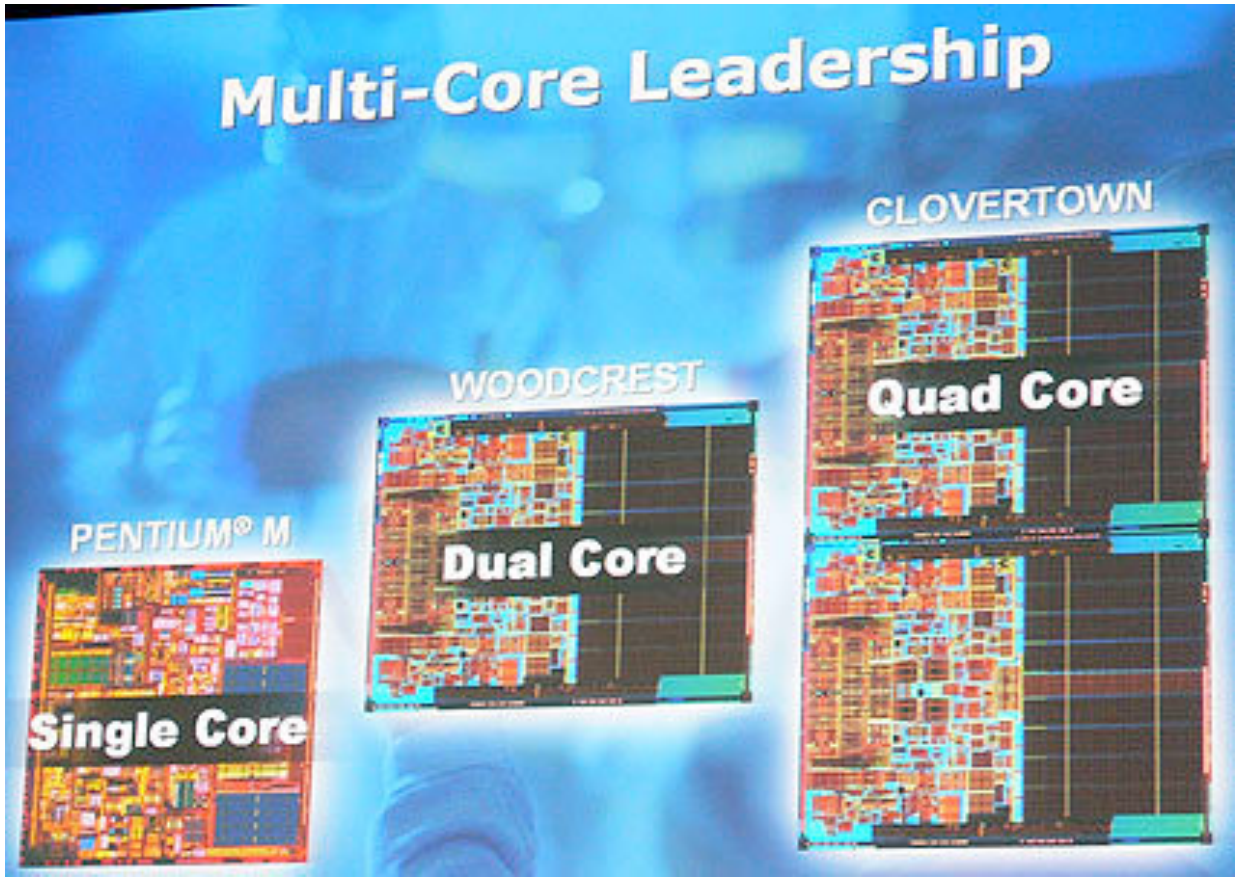
$$x > 0, \quad 1 + x = 1 \ ?$$

Yes and No …

$$1 + 0.1 \neq 1$$
$$1 + 0.01 \neq 1$$
$$1 + 0.001 \neq 1$$
$$1 + 0.0001 \neq 1$$
$$\ldots$$
$$1 + 0.0000000000000001 = 1 \ \ldots$$

Transitivity

$$(a*b)*c = a*(b*c)$$



Thank you