FiPS

Modeling Impact of Power- and Thermal-aware Fans Management on Data Center Energy Consumption

Wojciech Piątek[1], Ariel Oleksiak[1,2], Micha vor dem Berge[3]

1IPoznan Supercomputing and Networking Center
2IPoznan University of Technology
3IChristmann Informationstechnik & Medien

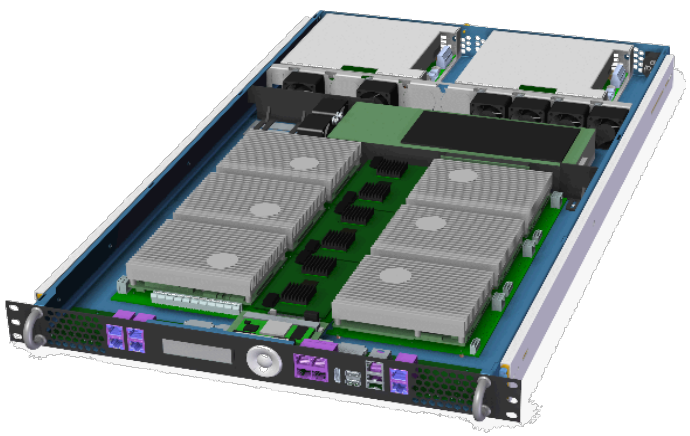
Problem and motivation

- Problem
  - Impact of micro servers on the overall data center energy consumption
  - PUE is not sufficient in minimizing data center energy consumption
- Motivation
  - Complexity and time requirements of Computational Fluid Dynamics methods
- The approach
  - Models for transient simulations of data center
  - Reflecting correlations between PUE and savings
  - Fan management to minimize overall energy consumption and high temperature excesses

FiPS Project

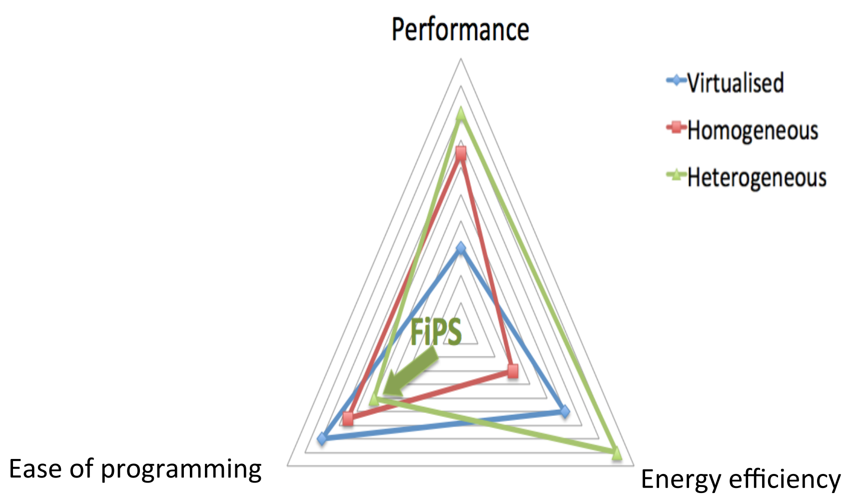
FiPS – Field-Programmable Servers

- Heterogeneous servers
  - CPU, embedded CPU, FPGA
  - Cluster architecture
  - 100 – 10k processing elements



- Application to hardware mapping methodology

- Goal: Performance/Energy +40%



Outline

- Power and thermal models of data center
- How PUE can be misleading…
- Simulation experiments - fans power management

POWER AND THERMAL MODELS OF DATA CENTER

Power use modeling

- Data Center

$$P_{DATA\_CENTER} = \sum_{i=1}^{n} P_{RACK} + P_{COOLING}$$

- Cooling*

$$P_{COOLING} = \frac{P_{IT\_LOAD}}{CoP}$$

- Rack

$$P_{RACK} = (\sum_{i=1}^{l} P_{NODE\_GROUP} + c) / \eta_{PSU}$$

*J.Moore, J. Chase, P. Ranganathan, and R. Sharma. Making scheduling "cool": Temperature – aware workload placement in data centers. In Proceedings of the 2005 USENIX Annual Technical Conference, 2005.

Power use modeling

- Node group (e.g. blade center)

$$P_{NODE\_GROUP} = \sum_{i=1}^{m} P_{NODE} + \sum_{j=1}^{k} P_{FAN}$$

- Fan

$$P_{FAN} = \frac{V * dp}{\mu} \qquad P_{FAN} = k_f * V^3$$

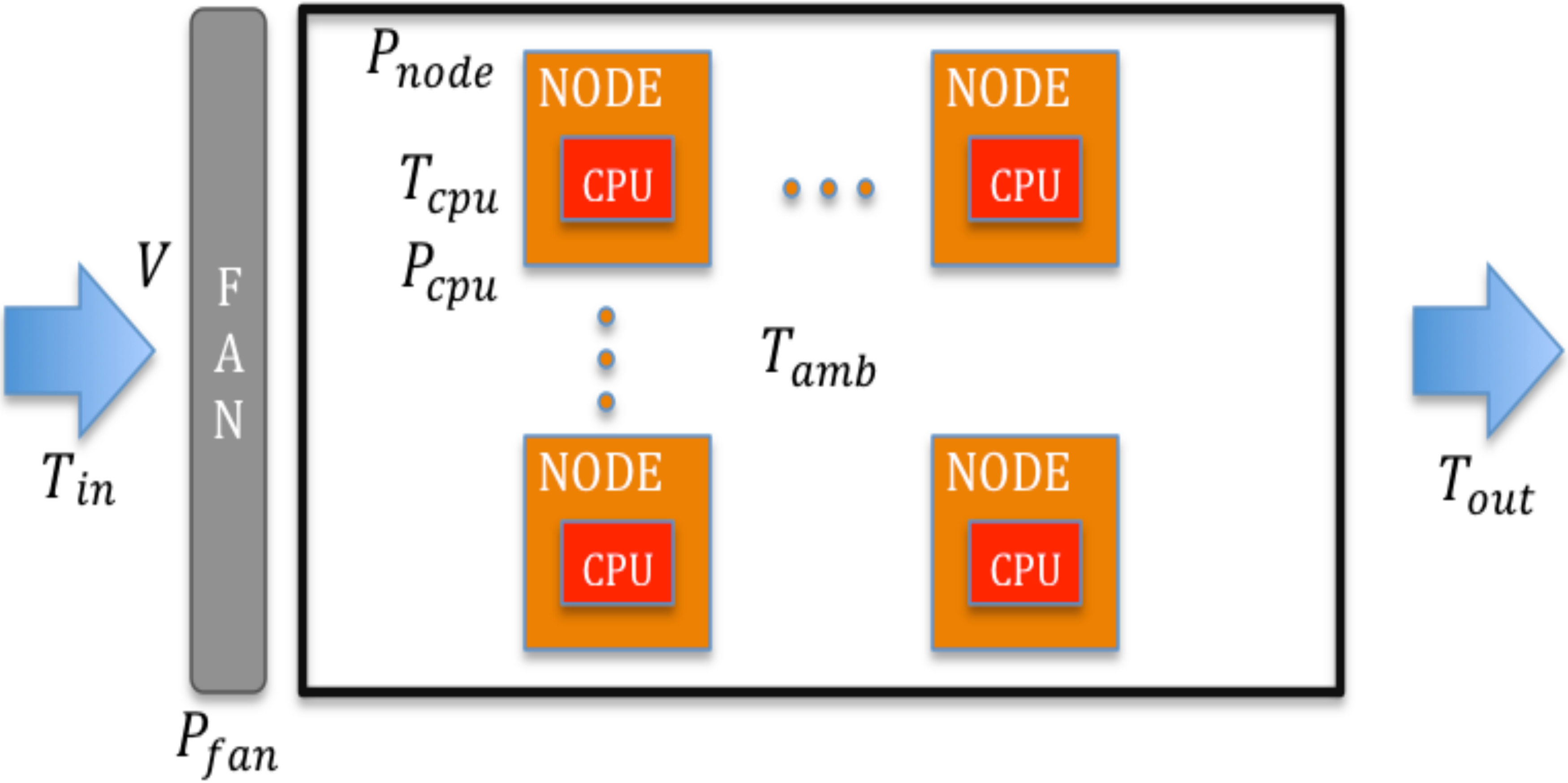
- Node (server)

$$P_{NODE} = \sum_{i=1}^{n} P_{CPU} + P_{MEM} + P_{NET} + P_{HDD}$$

- Processor

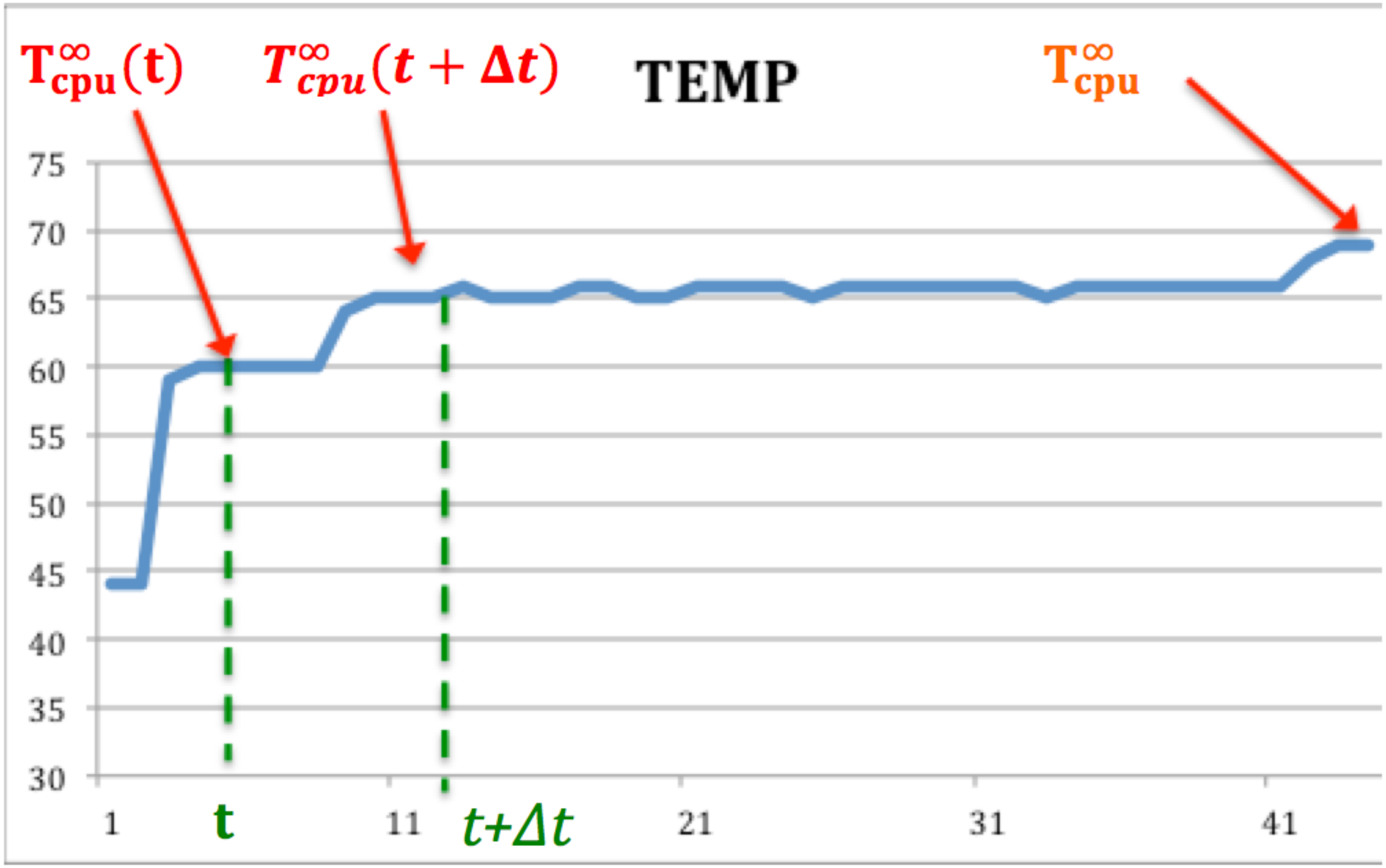
$$P_{CPU}(f,l) = (P_{CPU}(idle) + l * (P_{CPU}(f,100) - P_{CPU}(idle)) * g(T_{CPU})$$

Thermal modeling



Thermal modeling - processor

$$T_{cpu}(t+\Delta t) = T_{cpu}^{\infty} + (T_{cpu}(t) - T_{cpu}^{\infty})e^{-\frac{\Delta t}{RC}}$$



Thermal modeling - processor

$$T_{cpu}(t+\Delta t) = T_{cpu}^{\infty} + (T_{cpu}(t) - T_{cpu}^{\infty})e^{-\frac{\Delta t}{RC}}$$

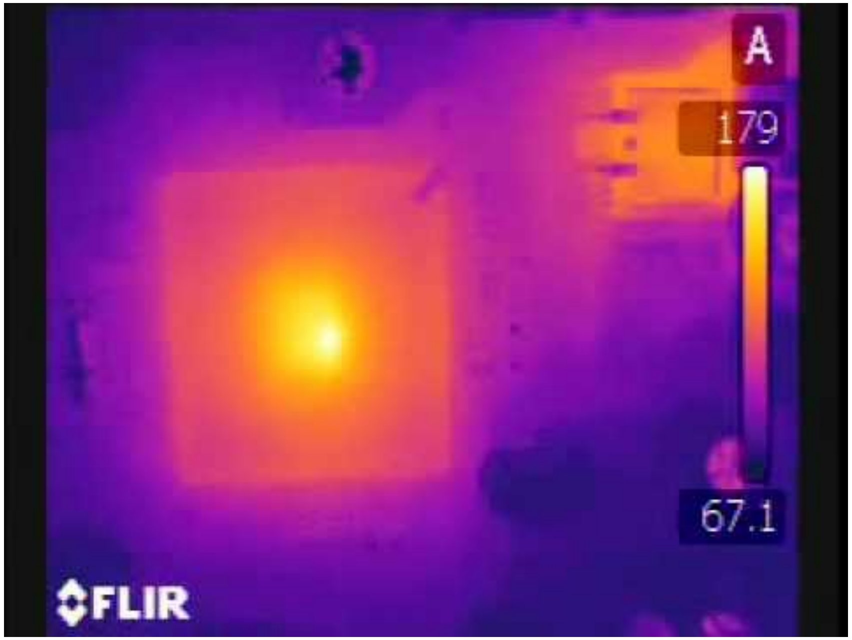
$$T_{cpu}^{\infty} = P_{cpu} * R + T_{amb}$$

$$R = R_{cond} + R_{conv}$$

$$R_{conv} = \frac{1}{k_h * V^n}$$

$$V = \sqrt[3]{k_f * P_{fan}}$$

$$T_{amb} = (T_{cpu} + 2 * T_{in}) / 3$$



Thermal modeling - outlet

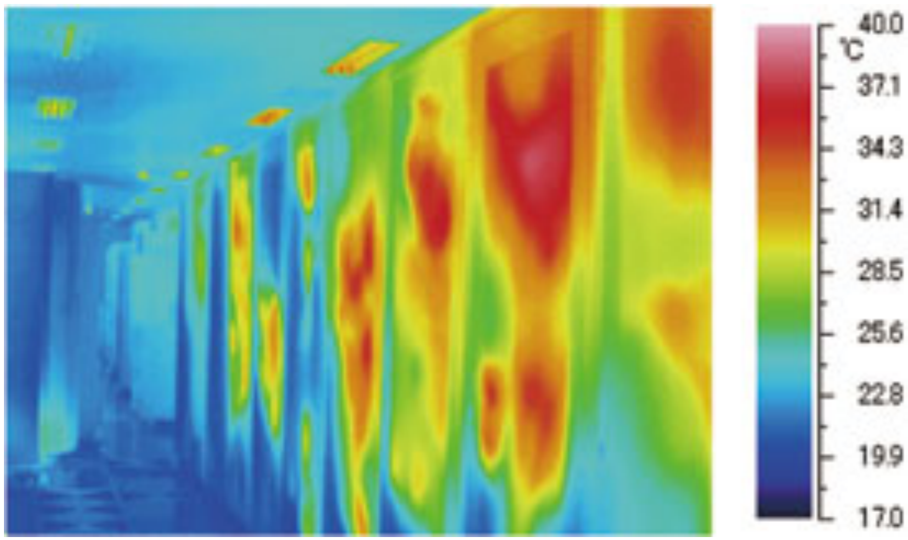
$$T_{out}(t+\Delta t) = T_{out}^{\infty} + (T_{out}(t) - T_{out}^{\infty})e^{-\frac{\Delta t}{RC}}$$

$$T_{out}^{\infty} = \frac{P_{node} + P_{fan}}{K} + T_{in}$$

$$R = R_{cond} + \frac{1}{k_h * V^n}$$

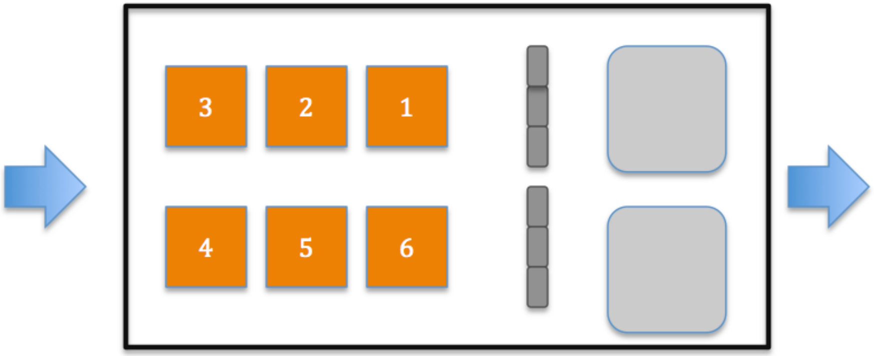
$$K = \rho * V * C_p$$

$$T_{out}^{\infty} = \frac{\sum_{i=1}^{n} P_{node_i} + P_{fan}}{K} + T_{in}$$



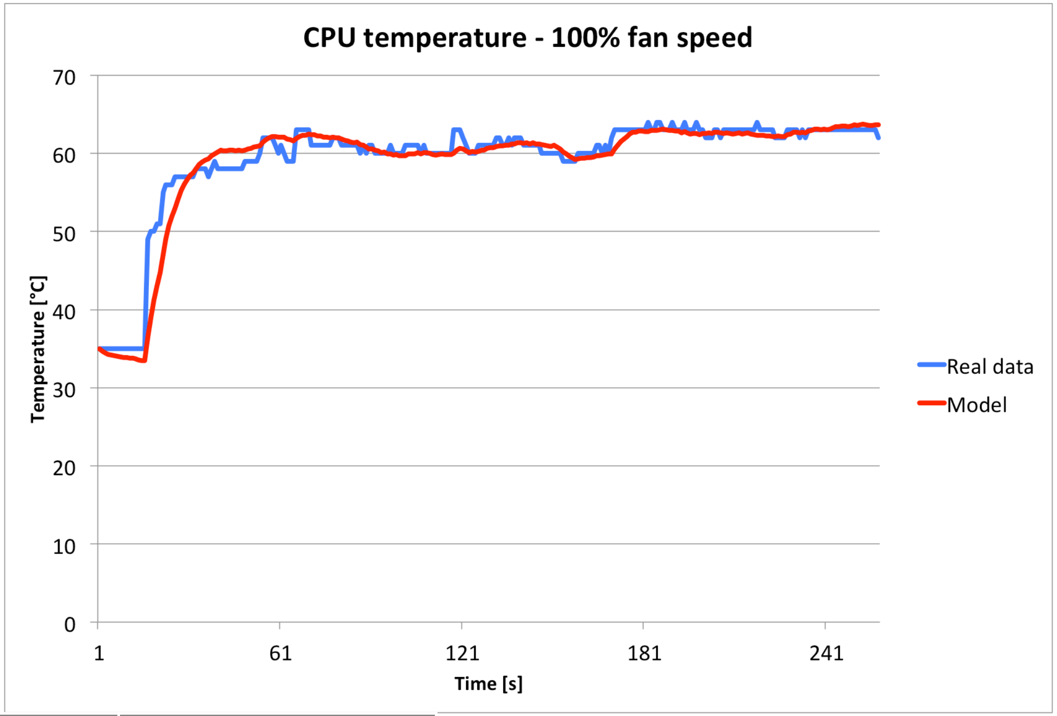
Thermal modeling - evaluation

- Tin = 20.5 °C
- Monitoring: f = 1Hz
- Intel Core i5-4400E
  - stress workload generator -> stress workload generator
  - fan speed: 30%, 70%, 100%



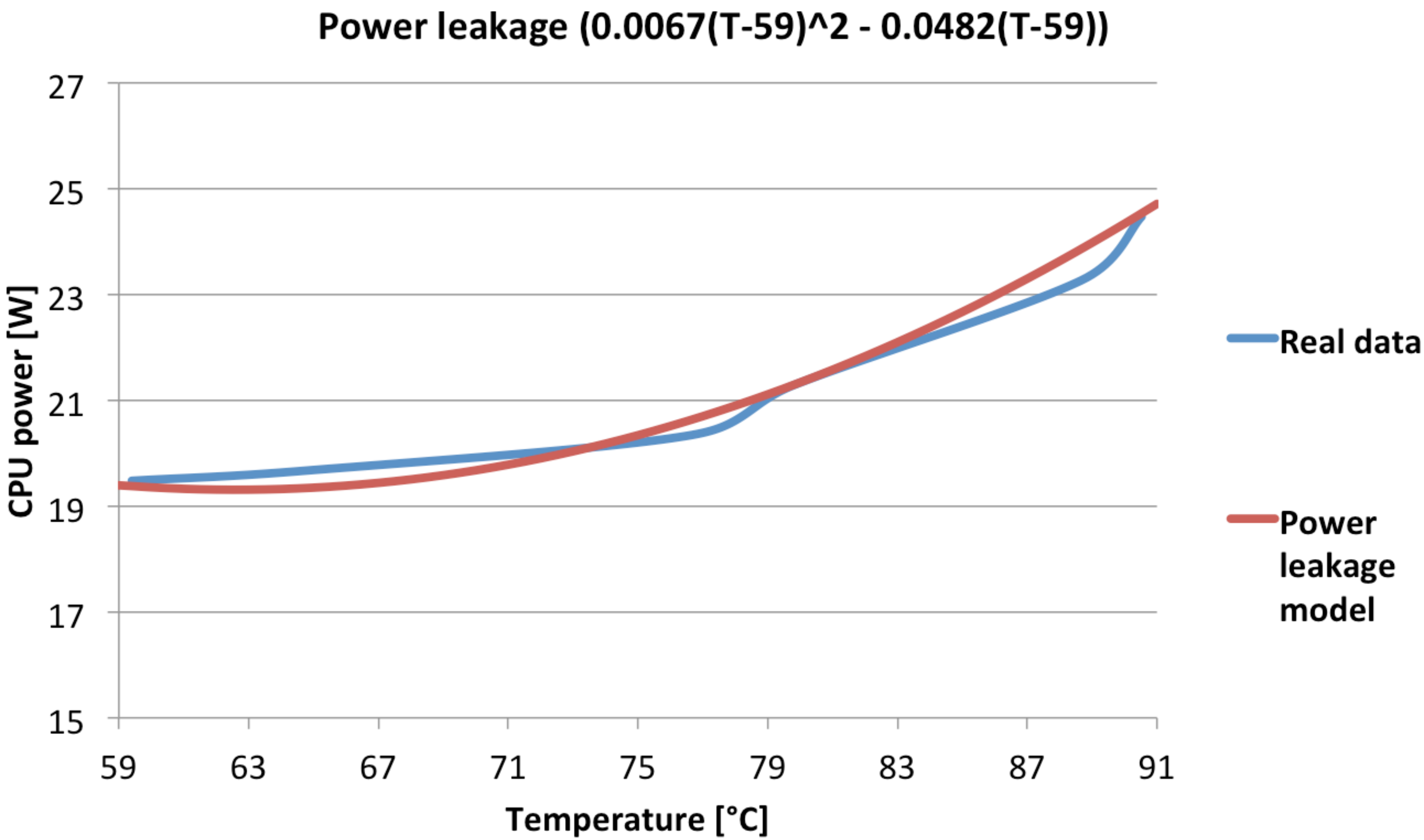
Thermal modeling - evaluation

Parameter	Value
R_cond	1,06 [°C]/[W]
R_conv	0,28-0,7 [°C]/[W]
K_h	49 [W][s]/[°C][m^3]
n	0,77
C	8 [W][s]/[°C]



Fans speed	Avg. temperature difference	Avg. relative error
30%	1,91°C	2,75%
70%	1,95°C	3,14%
100%	1,04°C	1,86%

Thermal modeling - processor power leakage



F. Fallah and M. Pedram. Standby and active leakage current control and minimization in cmos vlsi circuits. In IEICE Transactions on Electronics, pages 509-519,2005.

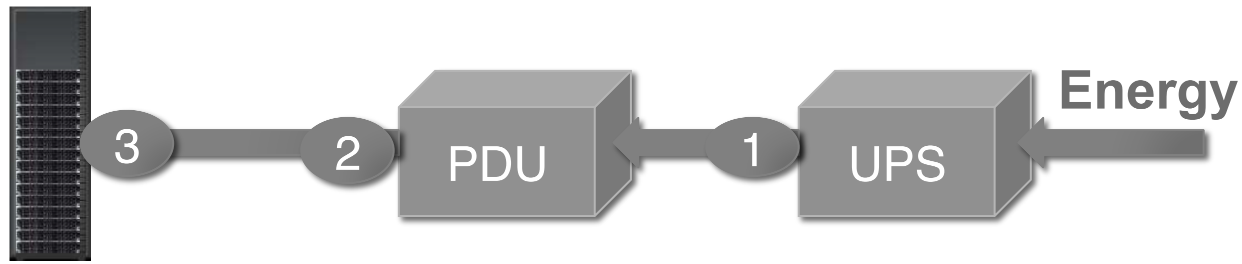
HOW PUE CAN BE MISLEADING…

How PUE can be misleading… - Levels of PUE

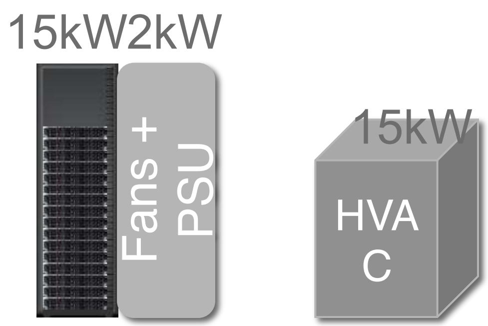
PUE Level 1

PUE Level 2

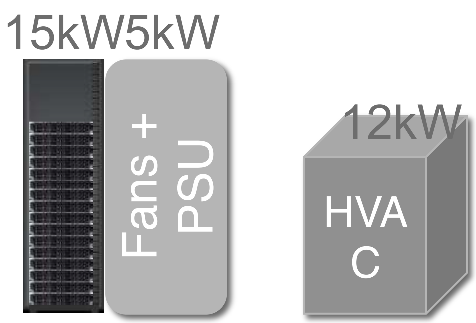
PUE Level 3



Different PUE for the same energy use!



PUE = 1.88

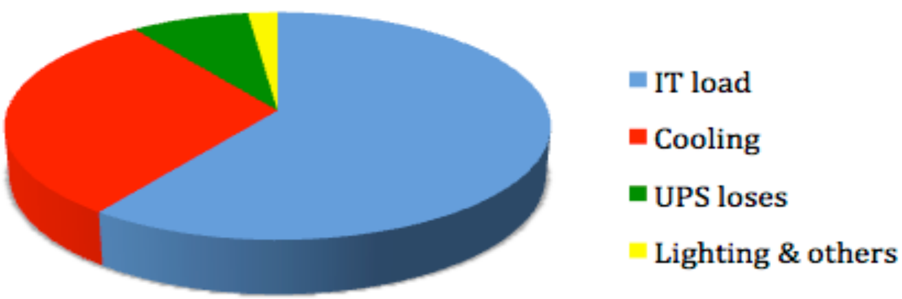


PUE = 1.6

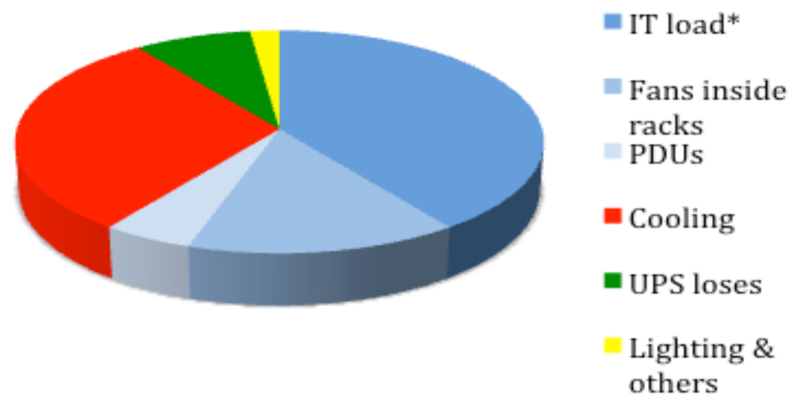
How PUE can be misleading… - PUE vs PUE4

- PUE-Level 4:
  - Subtracting fans and PSU consumption
  - To quantify efficiency at IT level

$$PUE = \frac{Total\ load}{IT\ load}$$



$$PUE\ Level\ 4 = \frac{Total\ load}{IT\ laod - Fans\ inside\ racks - PDUs}$$

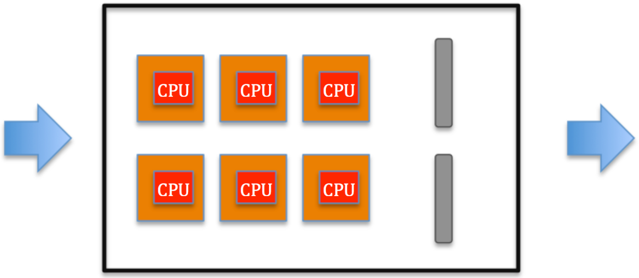


SIMULATION EXPERIMENTS

Simulation experiments

- Simulations with SVD Toolkit
- Resource: 10 racks, 42 1-unit enclosures with Intel Core i5-4400E
- Cooling model*

$$CoP = 0.0068 * T_{in}^2 + 0.0008 * T_{in} + 0.458$$



- Load: 7200s, 10 periods (uniform distribution) with 0-100% utilization -> 46%

*J.Moore, J. Chase, P. Ranganathan, and R. Sharma. Making scheduling "cool": Temperature – aware workload placement in data centers. In Proceedings of the 2005 USENIX Annual Technical Conference, 2005.

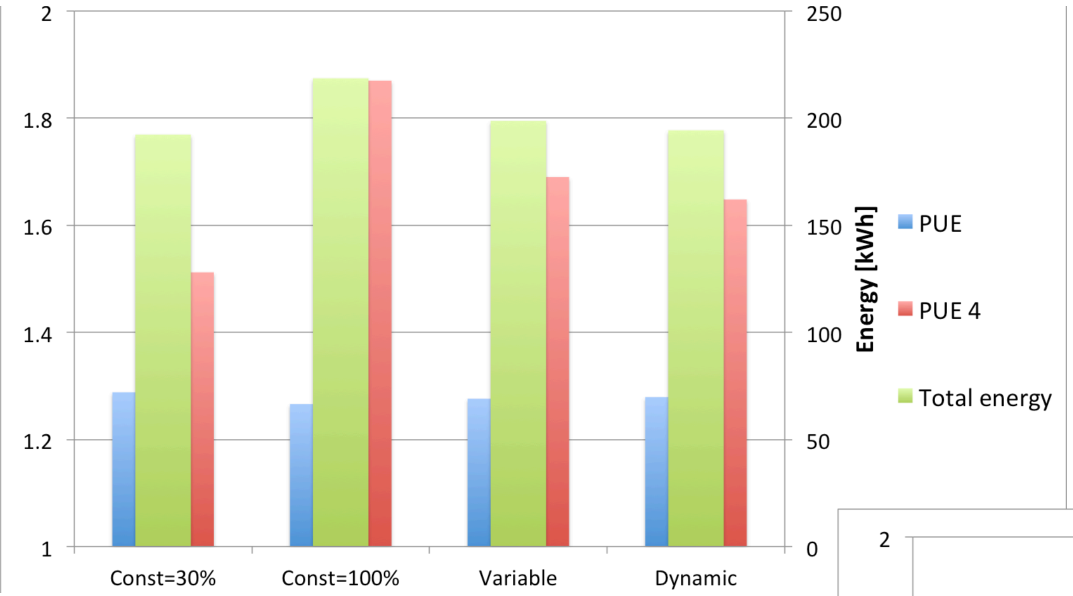
Simulation experiments

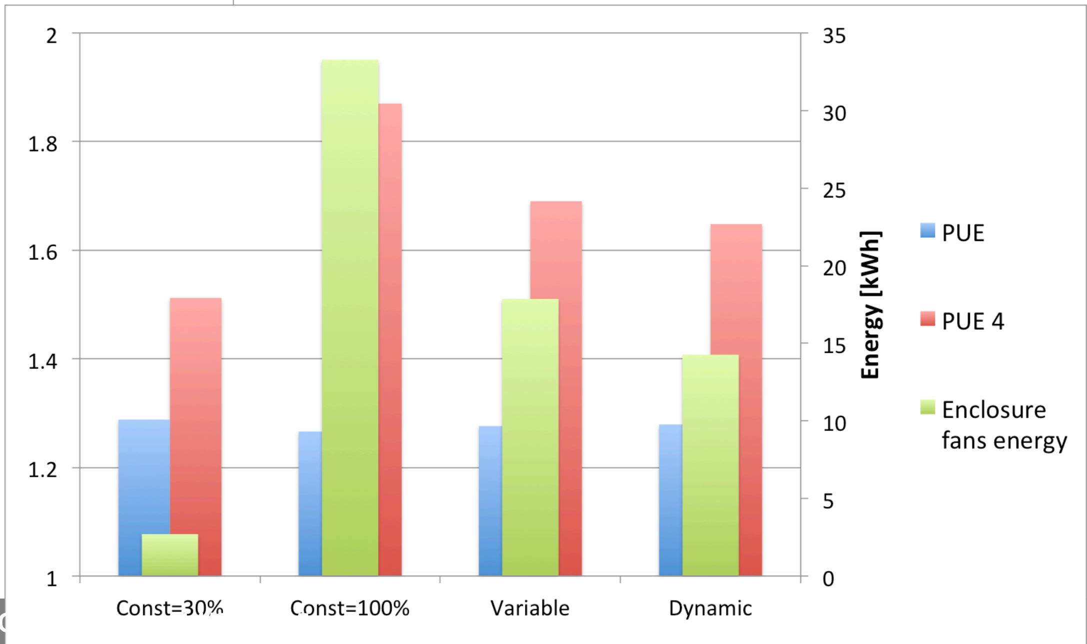
- Constant fan speed - 30%
- Constant fan speed - 100%
- Variable fan speed - 3 speed levels: 30%, 70% and 100%
- Dynamic fan speed control - 8 speed levels; tuning the fan speed in advance

Simulation experiments Tin=22C

Metrics	Const = 30%	Const = 100%	Variable	Dynamic
Energy Leakage [kWh]	14,49	4,17	4,91	5,22
Avg temp, exceeded time (90 C) [s]	1311	0	0	0
Processors energy [kWh]	87,78	77,46	78,2	78,51
Enclosure fans energy [kWh]	2,7	33,27	17,86	14,26
Racks energy [kWh]	149,23	172,55	155,69	151,91
Cooling energy [kWh]	40,08	42,58	39,93	39,35
Total energy [kWh]	192,3	218,58	198,73	194,3
PUE	1,288	1,266	1,276	1,279
PUE Level 4	1,512	1,87	1,69	1,648

Simulation experiments Tin=22C





Simulation experiments Tin=24C

Metrics	Const = 100%	Variable	Dynamic
Energy Leakage [kWh]	4,91	5,37	5,88
Avg temp, exceeded time (90 C) [s]	0	0	0
Processors energy [kWh]	78,2	78,66	79,17
Enclosure fans energy [kWh]	33,28	21,37	15,25
Racks energy [kWh]	173,41	160,25	153,80
Cooling energy [kWh]	36,7	34,91	34,07
Total energy [kWh]	213,58	198,37	190,95
PUE	1,231	1,237	1,241
PUE Level 4	1,631	1,566	1,531

Conclusions

- Holistic model of data center including power and thermal models (accuracy: 1,86-3,14%)
- PUE Level 4 as a way of representing fan impact on DC effectiveness
- Thermal-aware fans management led to:
  - Up to 8% reduction on cooling
  - Up to 12% savings on total energy
  - … all with keeping the processor temperature below the given threshold
- Next steps
  - Model improvement, analysis of denser and heterogeneous architectures

Questions?

FiPS

SEVENTH FRAMEWORK PROGRAMME

Ministry of Science and Higher Education
Republic of Poland