Energy-Aware Service Execution

LCN 2011, Bonn, Germany

October 4, 2011

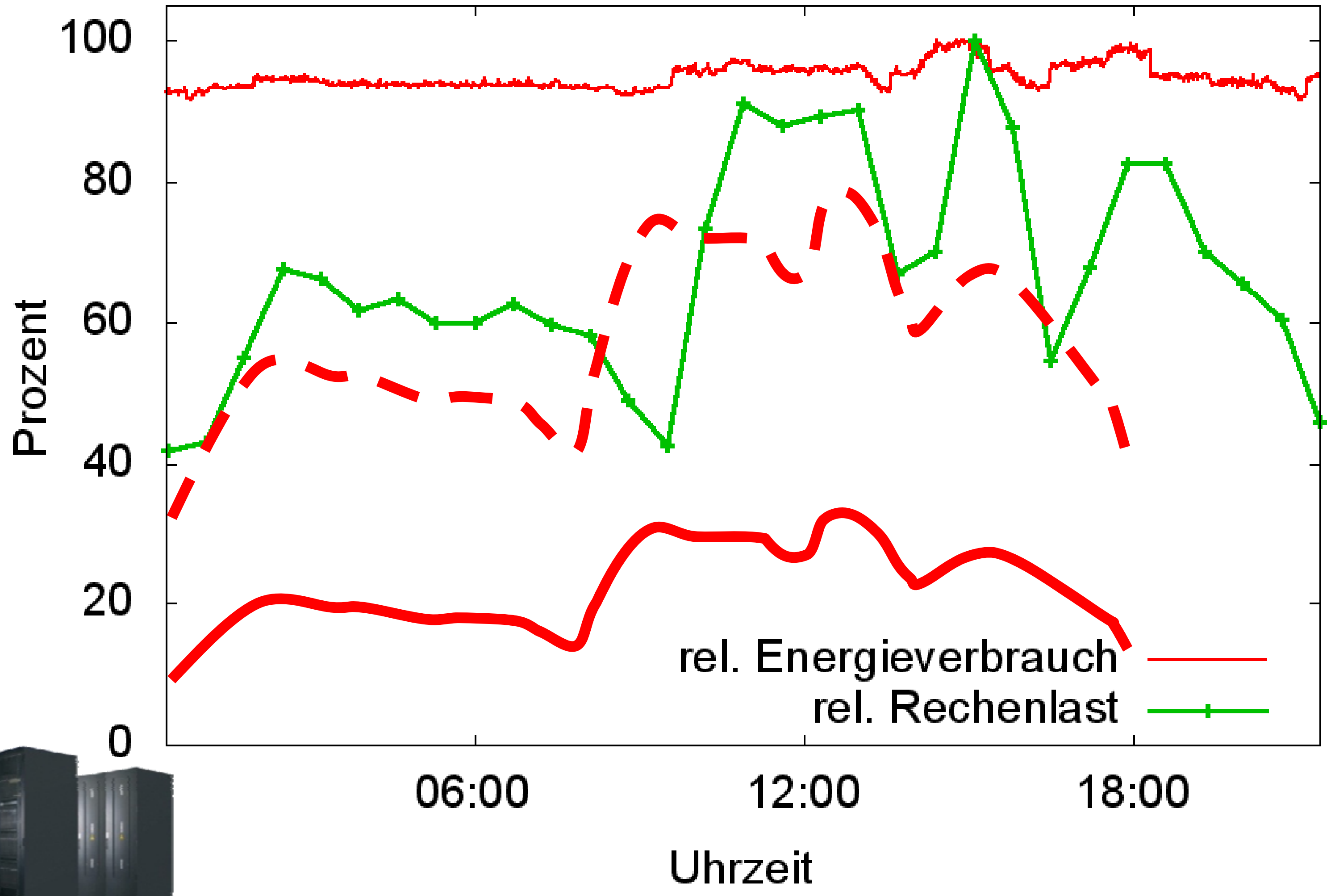
Waltenegus Dargie
TU Dresden
Chair of Computer Networks

TECHNISCHE UNIVERSITÄT DRESDEN

Outline

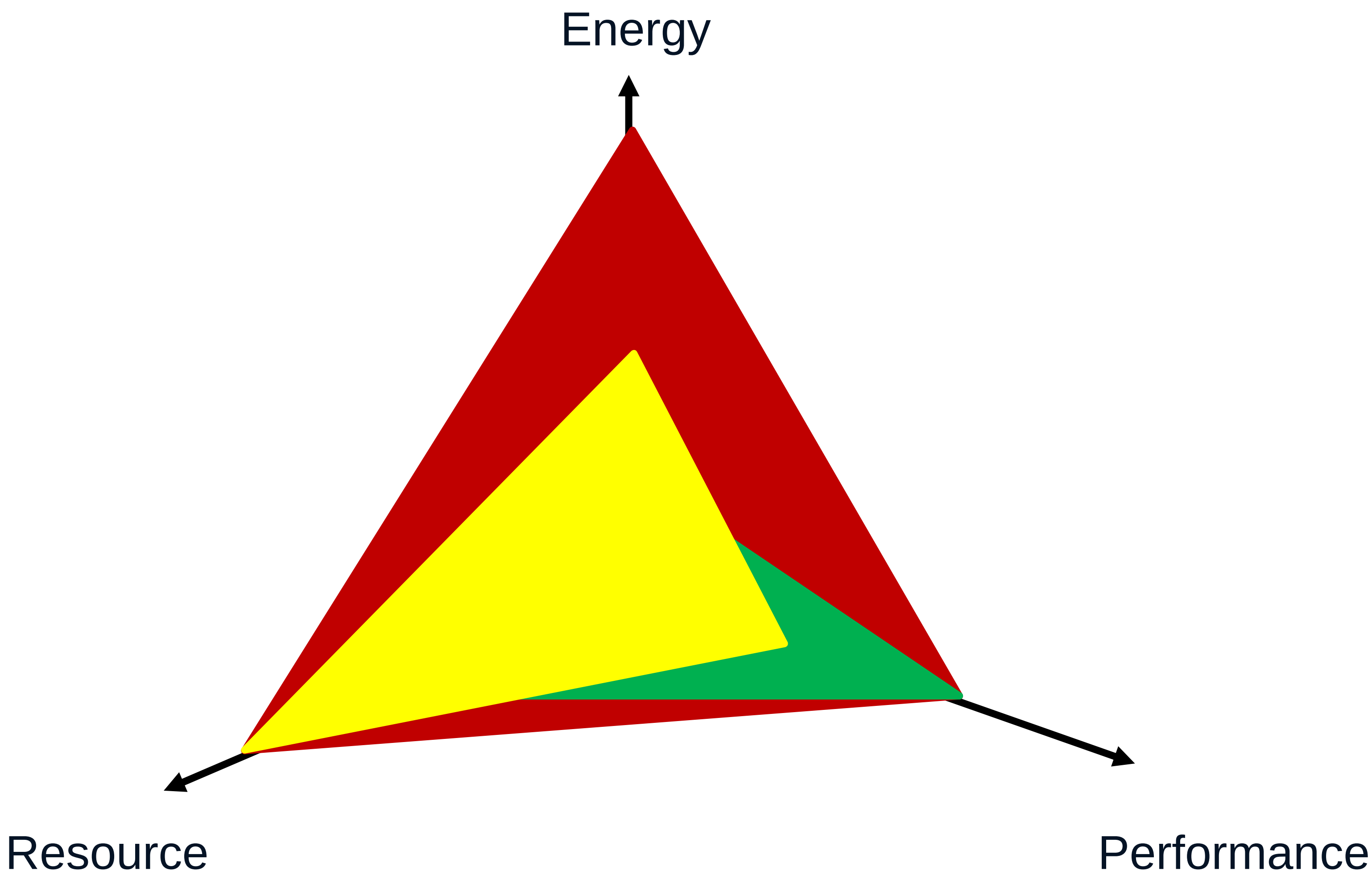
- **Motivation**
- **Related Work**
- **Proposed Architecture**
- **Quantifying the energy cost of service execution**
- **Probabilistic Models**
- **The Cost of Service Execution**
- **Open issues**
- **Conclusion**

Energy Proportionality



Center for Information Services and High Performance Computing, TU Dresden)
Measurement on June 20, 2008

Motivation



TECHNISCHE UNIVERSITÄT DRESDEN

Related Work

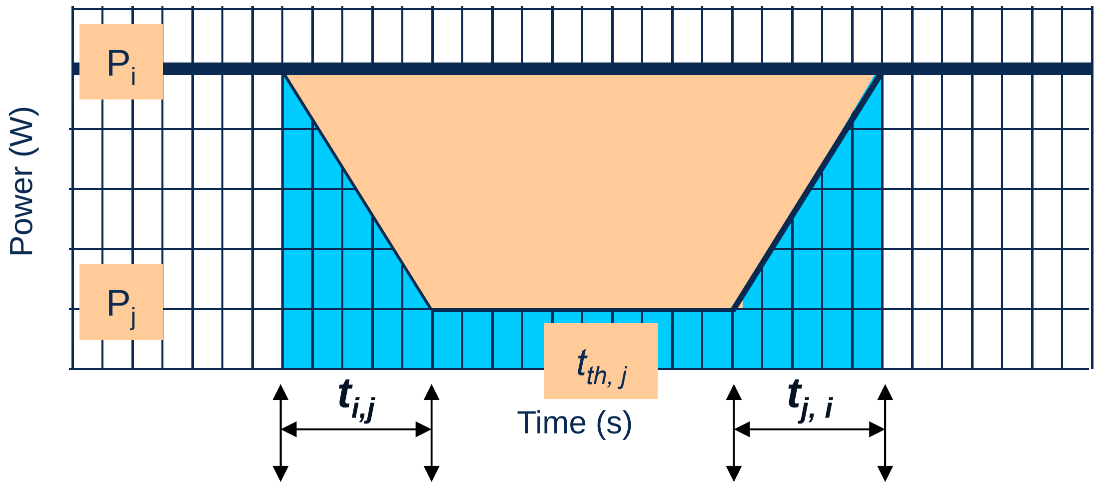
- **Qualitative description of the energy consumption of hardware devices**
  - **CPU, memory, high speed data busses, hard disk, communication**
- **Dynamic power management algorithms**
  - **Service consolidation, selective switching, dynamic voltage and frequency scaling**

TECHNISCHE UNIVERSITÄT DRESDEN

Related Work

- Our focus is mainly on *when* instead of *how* to undertake dynamic power management.
- We undertake adaptation based on *fine grained estimation* of the energy (power) consumption of individual service requests.
  - In other words services are being adapted instead of virtual machines

Related Work

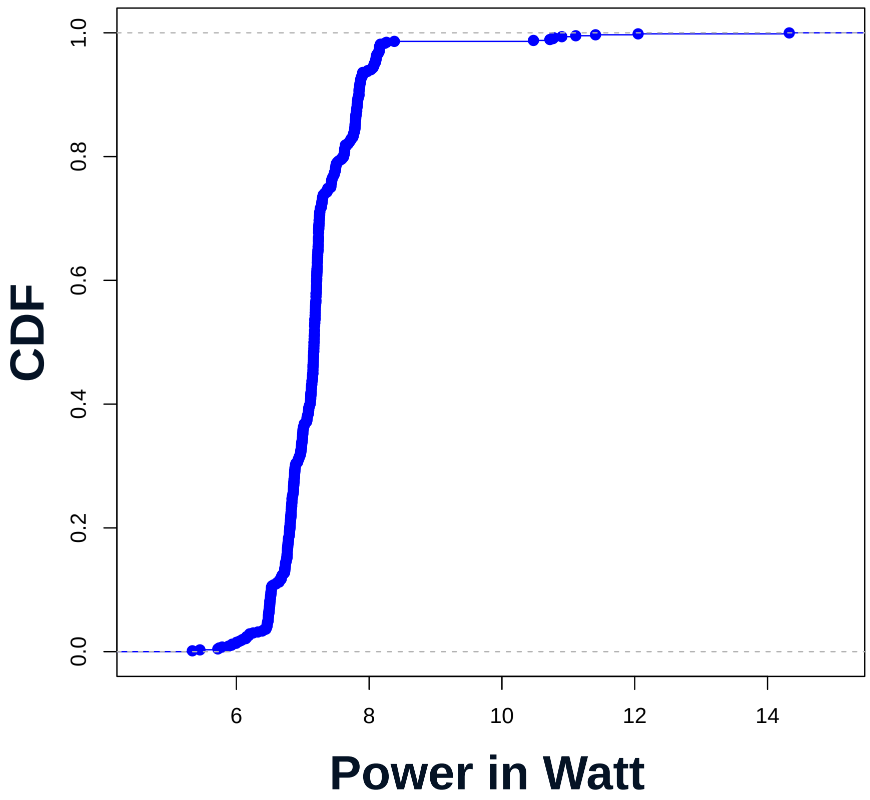


$$E_j^{saved} = P_i \left(t_j + t_{i,j} + t_{j,i} \right) - \left[P_{i,j} \times t_{i,j} + P_{j,i} \times t_{j,i} + P_j \times t_j \right]$$

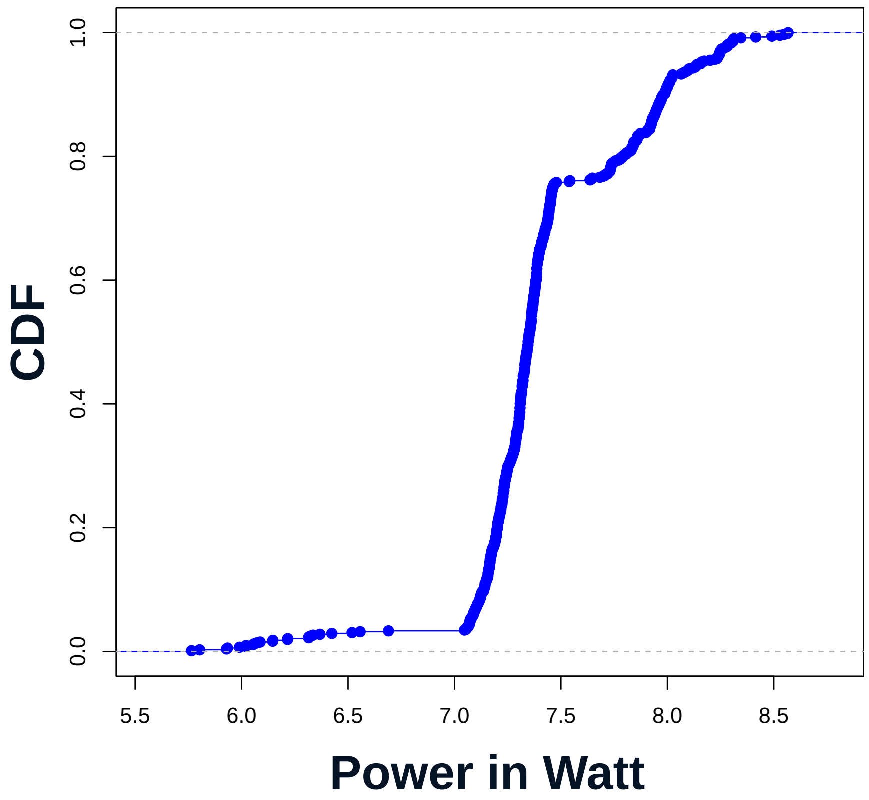
$$t_{th,\,j} \geq \max\left(0, \frac{\left(P_i - P_{i,j}\right) t_{j,i}}{\left(P_j - P_i\right)} \right)$$

Transition Cost (Power)

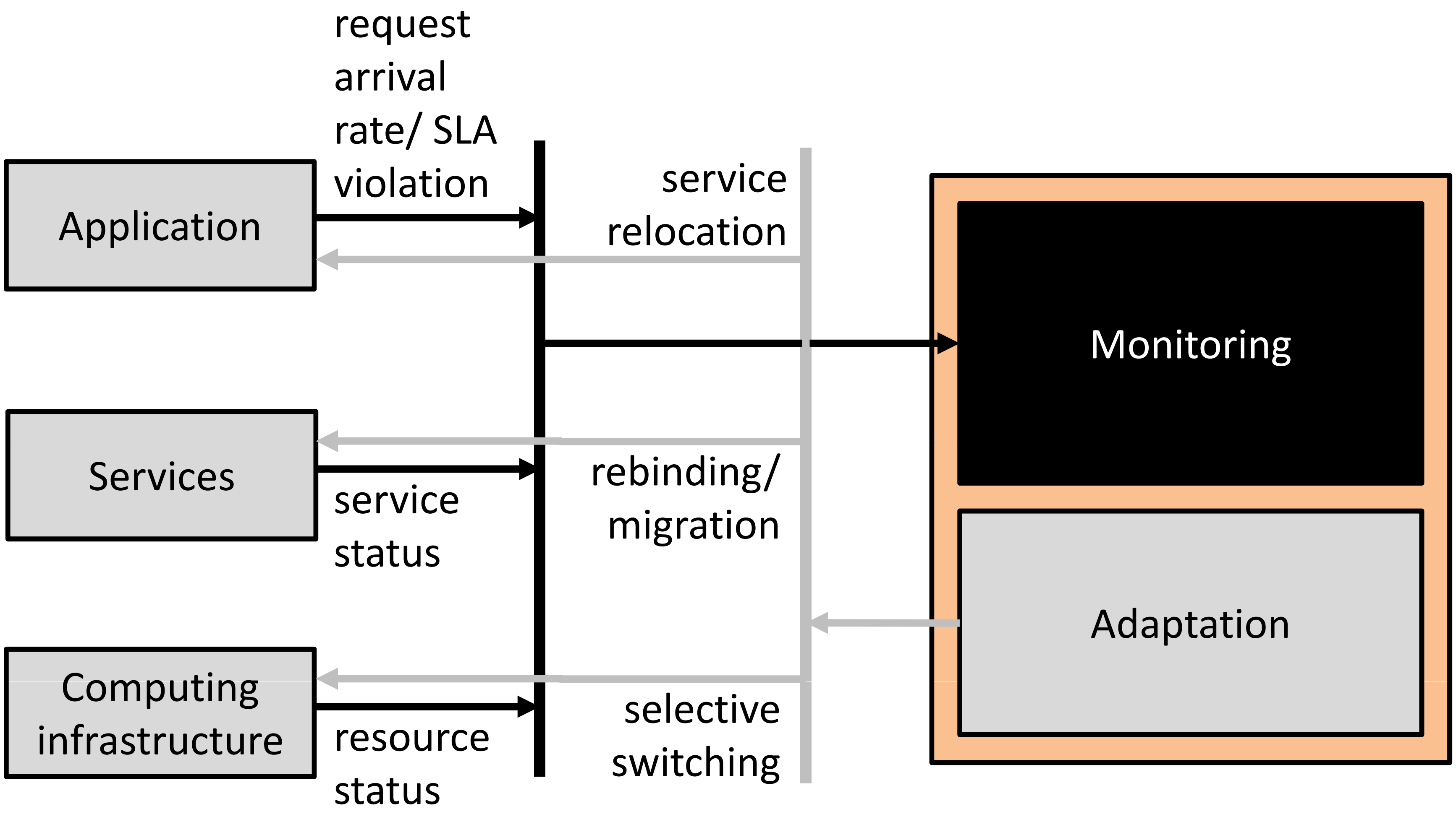
Transition Power: Motherboard: 12V



Transition Power: CPU

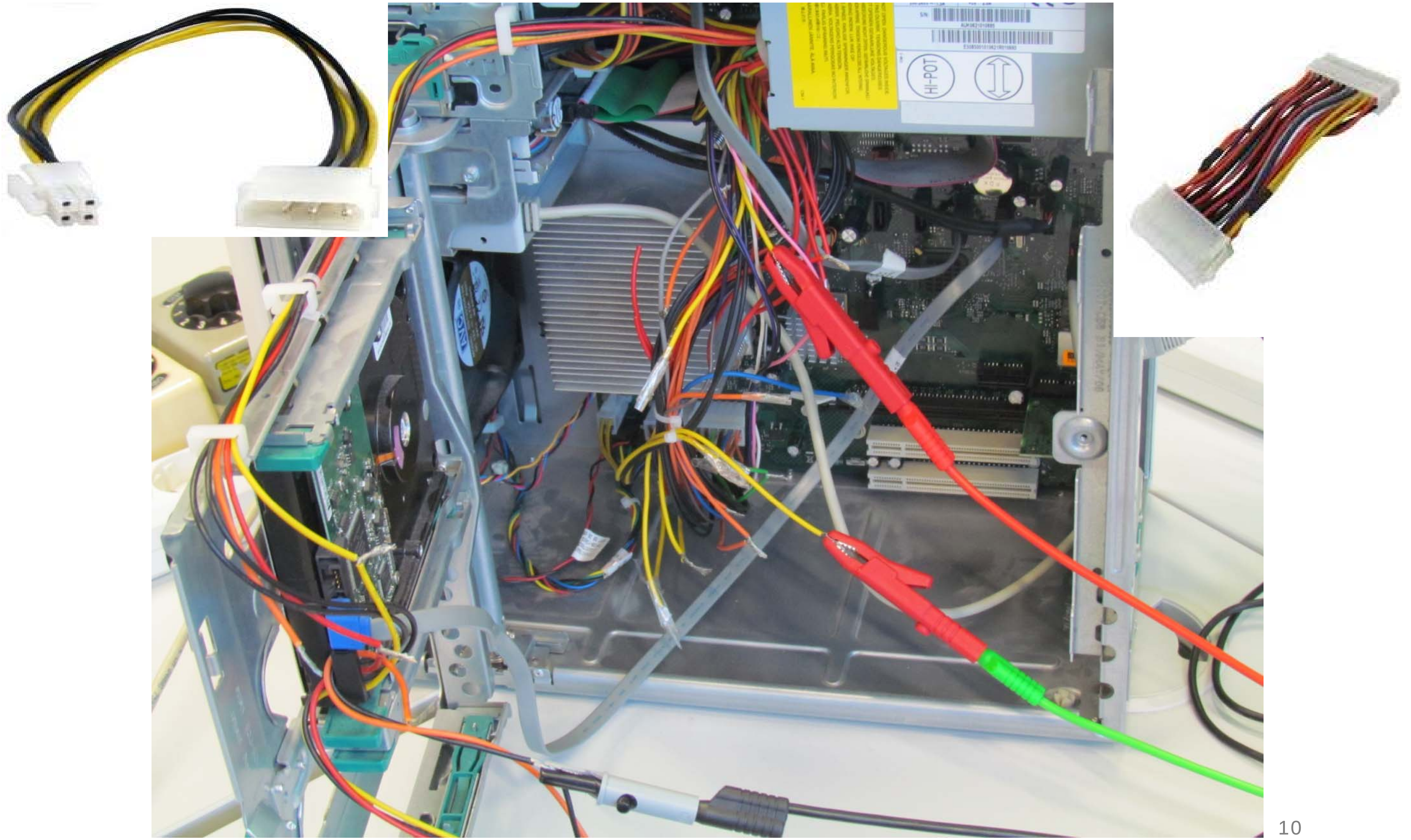


Proposed Architecture



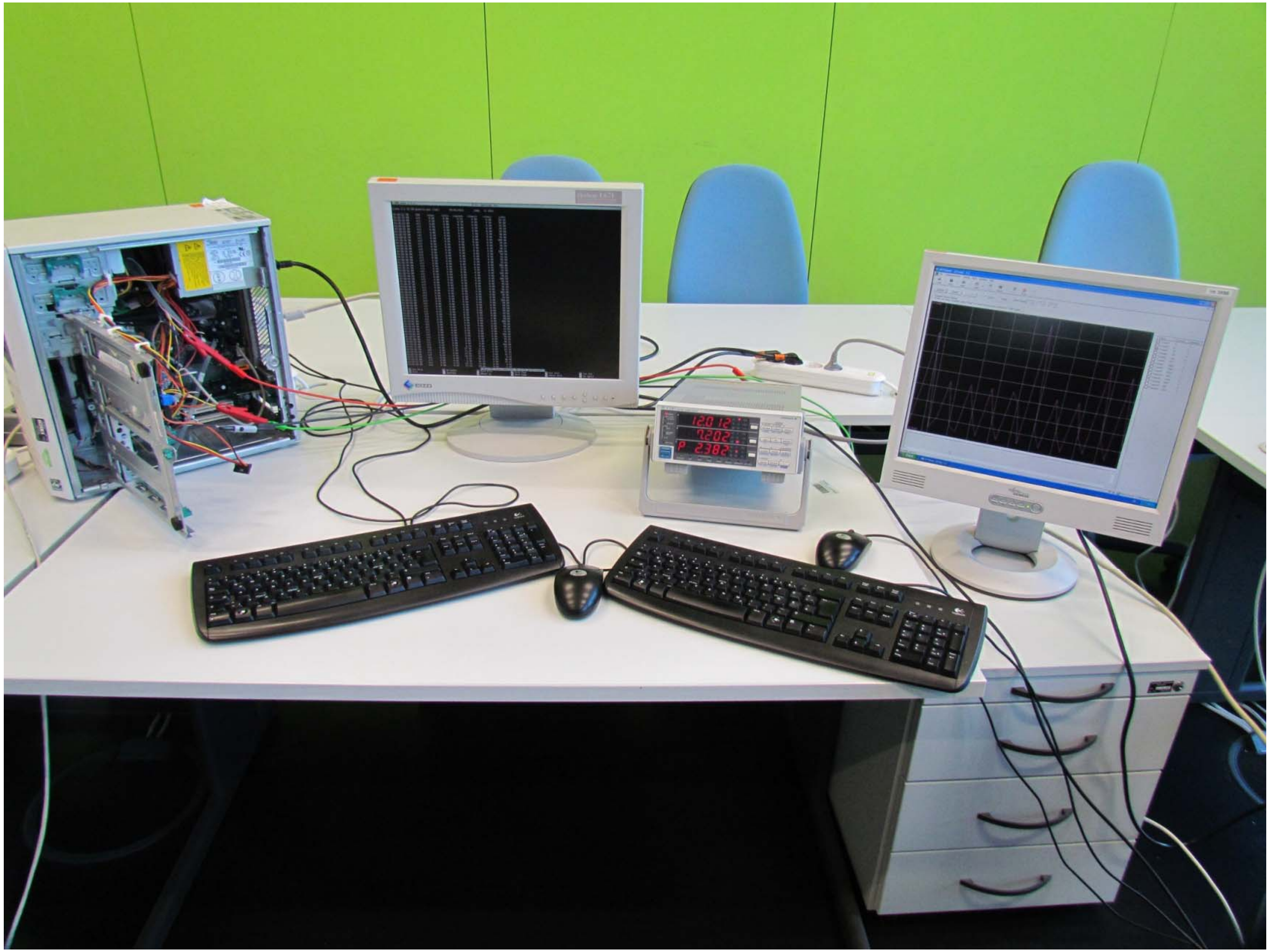
TECHNISCHE UNIVERSITÄT DRESDEN

Measurement Setting

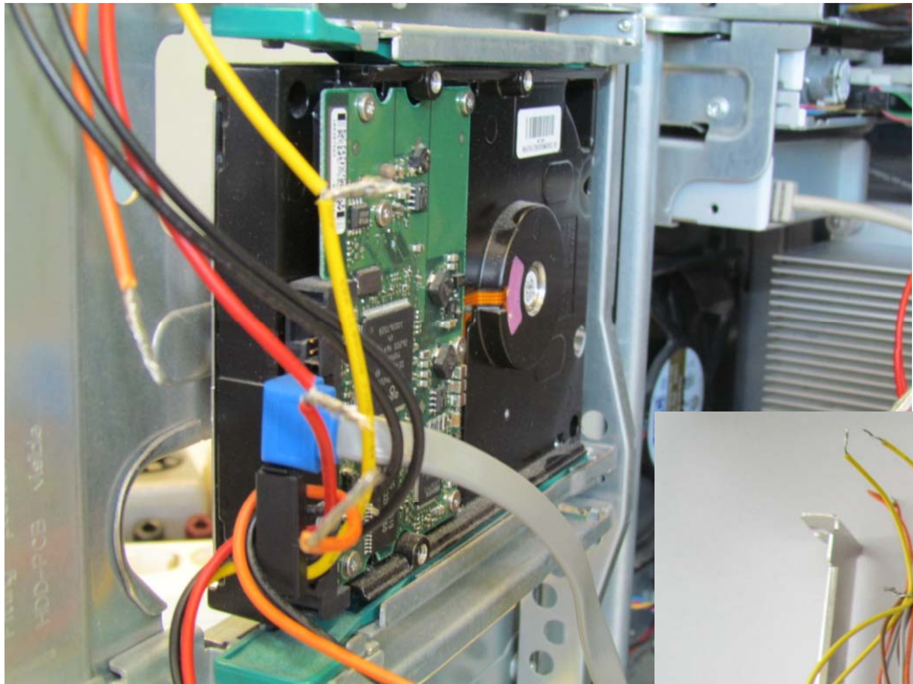


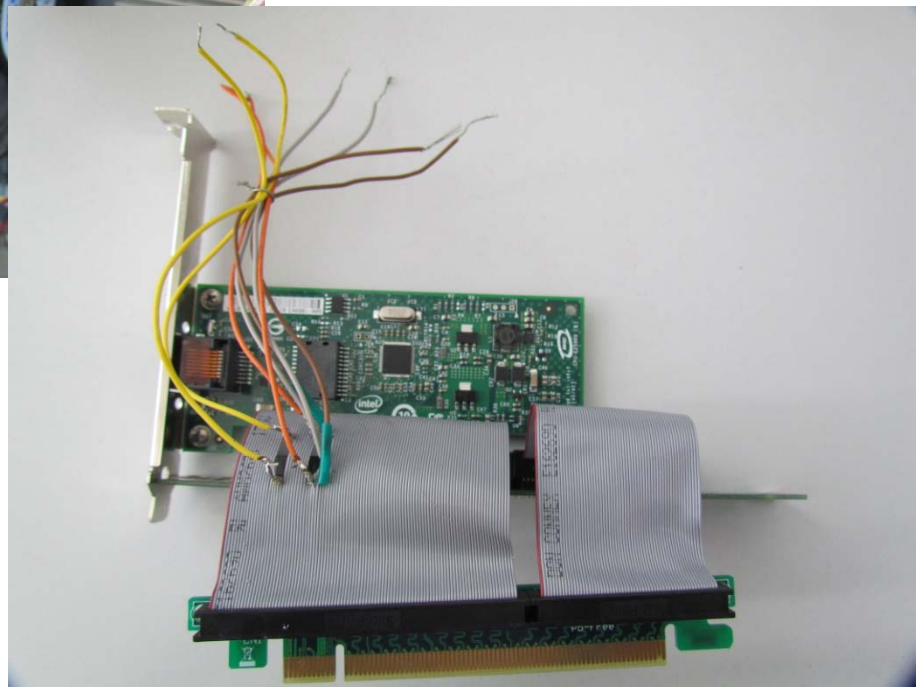
TECHNISCHE UNIVERSITÄT DRESDEN

Measurement Setting

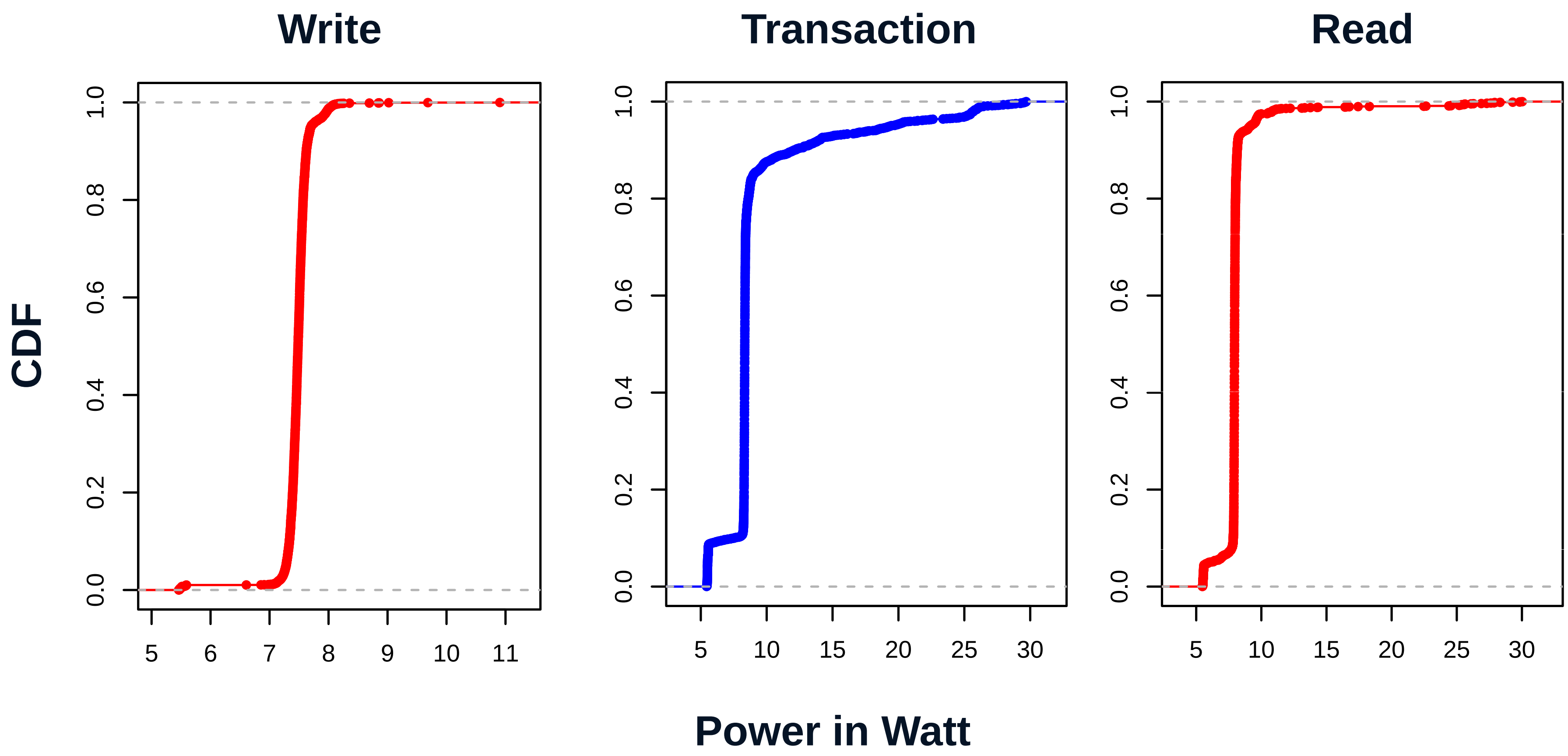


Measurement Setting

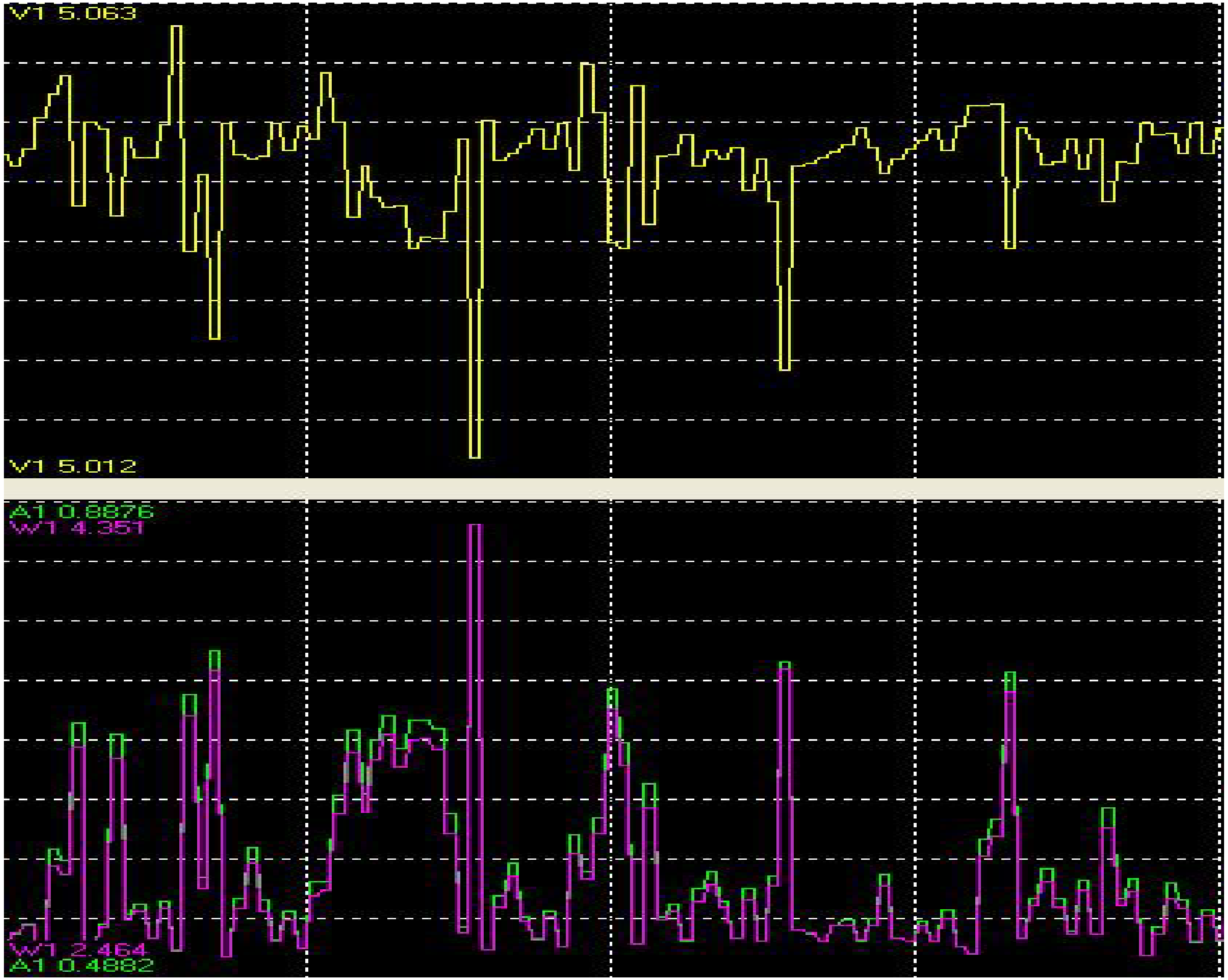




Execution Cost (Power)



Power Supply Efficiency

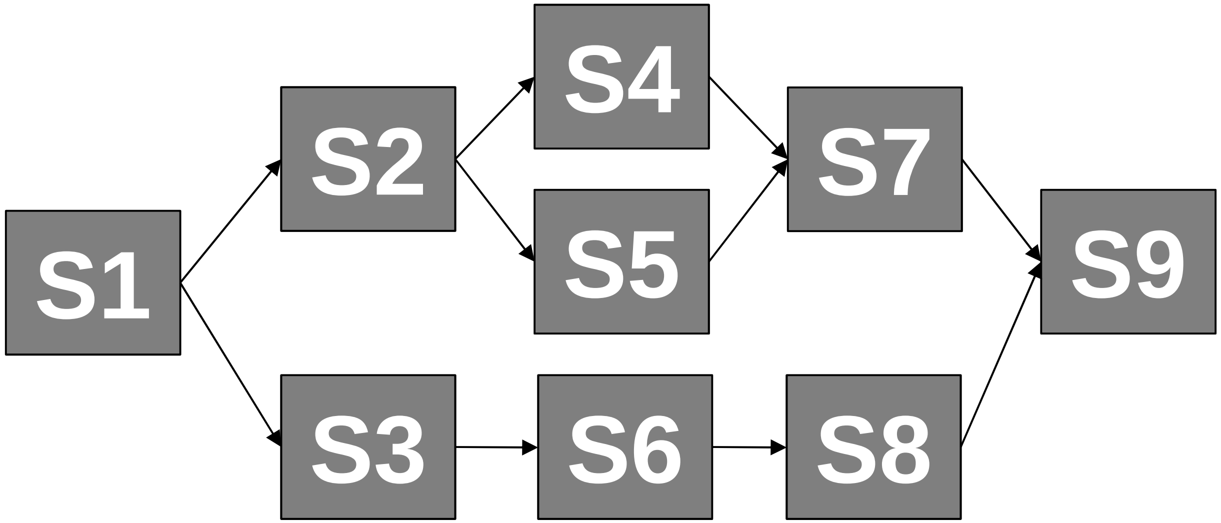


TECHNISCHE UNIVERSITÄT DRESDEN

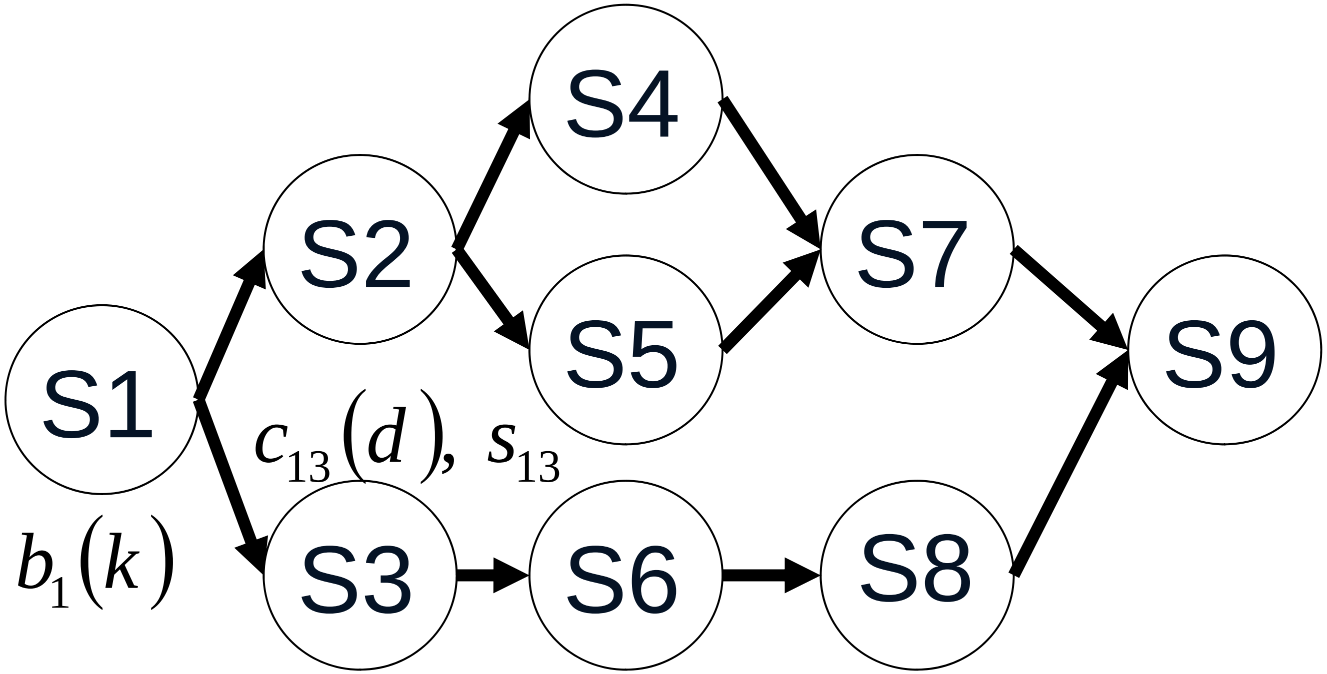
Execution Dependency

- Understanding the way contemporary applications are built offers a new insight into application based energy management
- Consider what a simple request to a YouTube application involves:
  - Search operation
  - Request related filter
  - Context related filter
  - Video streaming
  - Images and their descriptions
  - Ranked comments

Execution Dependency



Execution Model



Execution Cost

$$b_j(k) = P\left(computation\ cost\ \leq k | s_j, r\right)$$

The probability that the computation level of service s_j is k, given a request class, r

$$A = \left\{s_{ij}\right\} = P\left[s_{ij} \mid r, \lambda\right] = \begin{bmatrix} s_{11}\ s_{12}\ \dots\ s_{1n} \\ s_{21}\ s_{22}\ \dots\ s_{2n} \\ \vdots \\ s_{n1}\ s_{n2}\ \dots\ s_{nn} \end{bmatrix}$$

The probability that a communication between service i and j exists, given a request class, r, and the model λ

$$c_{ij}(d) = P\left(data\ size \leq d \mid s_{ij}, r\right)$$

The probability that d amount of data is exchanged between service s_i and s_j given that a communication exists between service i and j

TECHNISCHE UNIVERSITÄT DRESDEN

Execution Cost (Energy)

$$PE_j(r) = \sum_{k=1}^{K} b_j(k) E_j(k)$$

$$PE(r) = \sum_{j=1}^{n} PE_j(r)$$

The expected processing cost

$$CE_{ij}(r) = s_{ij} \sum_{d=1}^{D} c_{ij}(d) E_{ij}(d)$$

$$CE(r) = \sum_{i=1}^{n} \sum_{j=1}^{n} CE_{ij}(r)$$

The expected communication cost

TECHNISCHE UNIVERSITÄT DRESDEN

Execution Cost (Energy)

$$PE_j(r) = \sum_{k=1}^{K} b_j(k) E_j(k)$$

$$PE(r) = \sum_{j=1}^{n} PE_j(r)$$

The expected processing cost

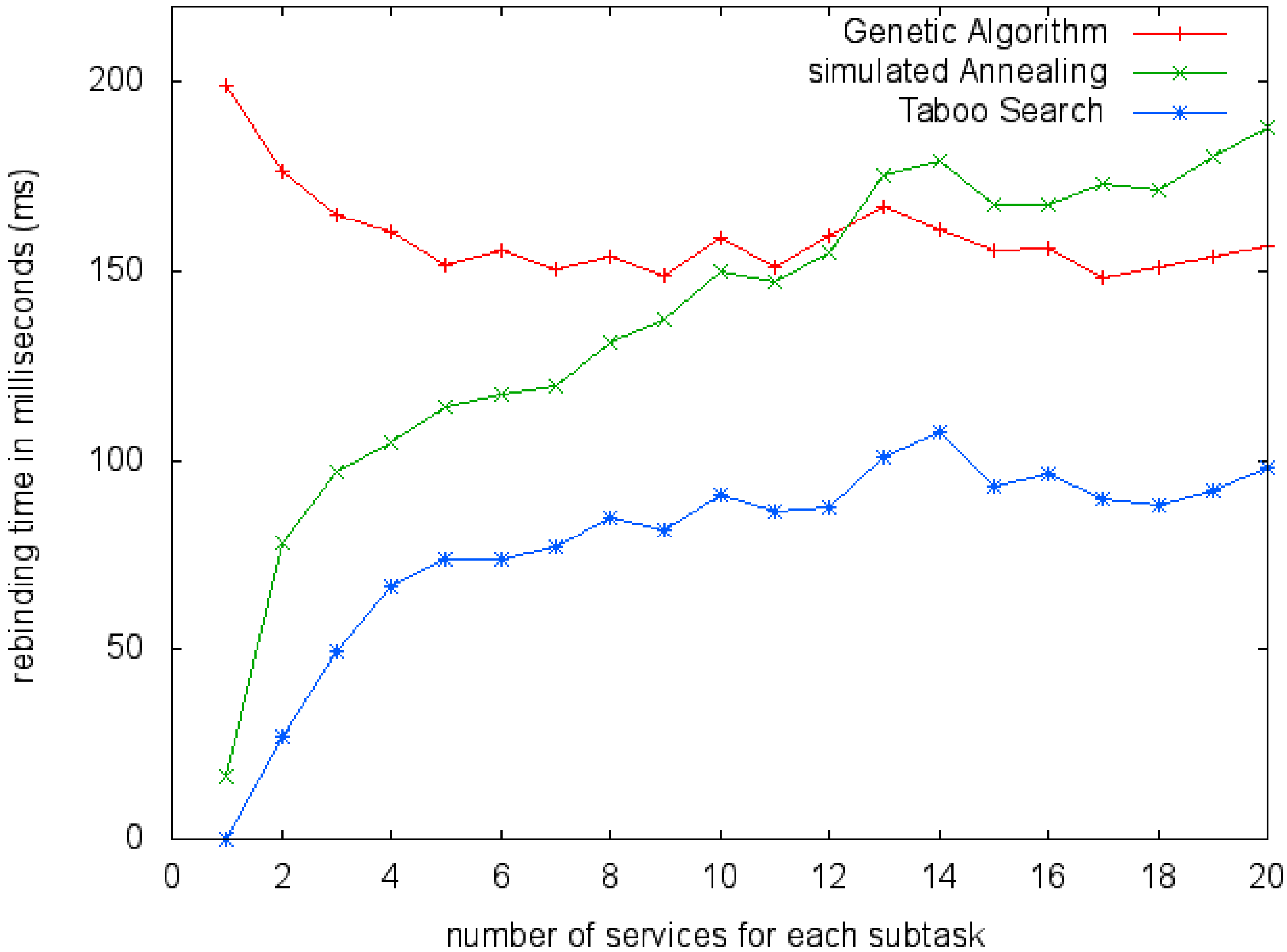
$$CE_{ij}(r) = s_{ij} \sum_{d=1}^{D} c_{ij}(d) E_{ij}(d)$$

$$CE(r) = \sum_{i=1} \sum_{j=1} CE_{ij}(r)$$

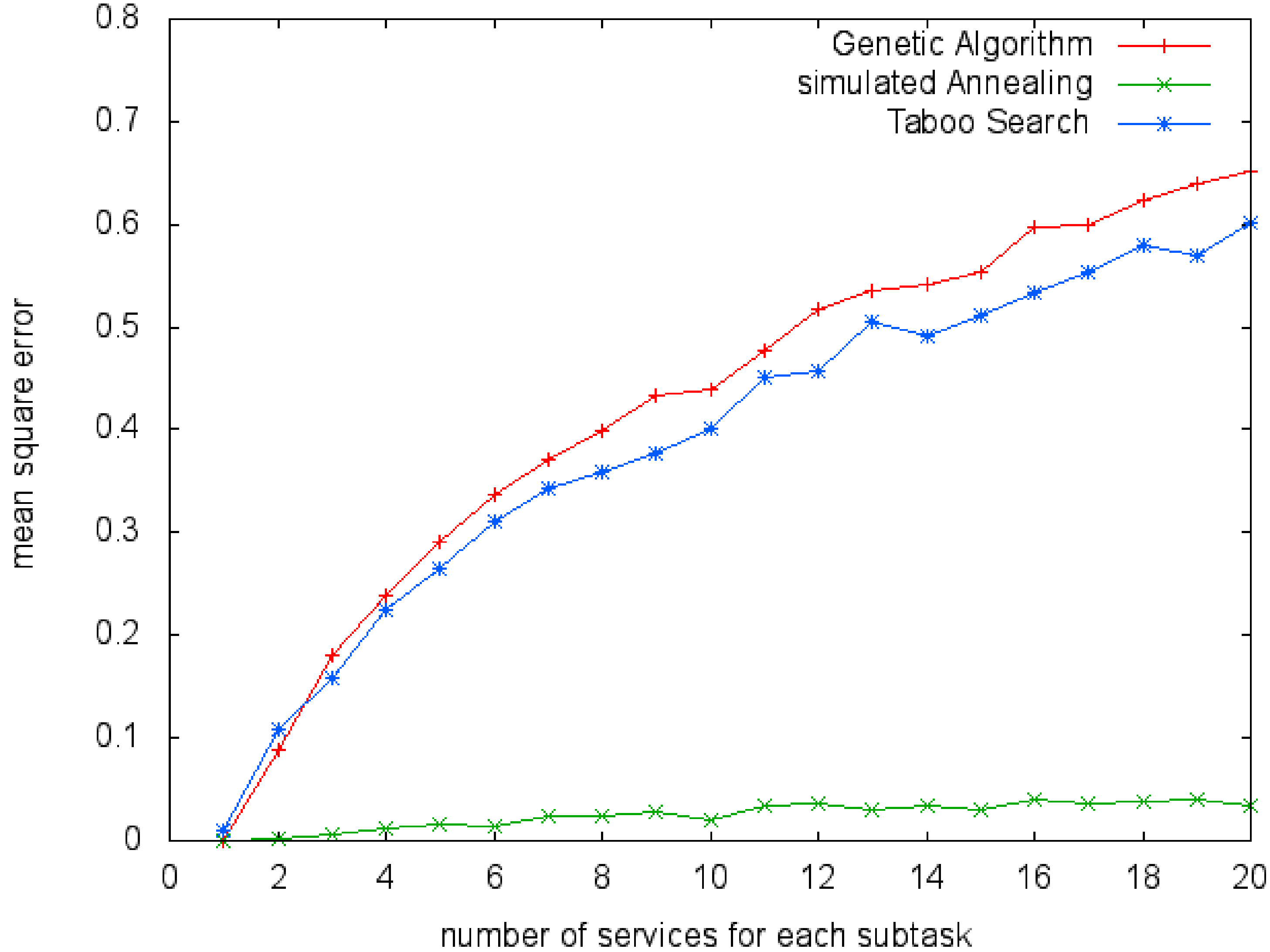
The expected communication cost

Execution Cost (Delay)

Execution Cost (Delay)



Execution Cost (Delay)



TECHNISCHE UNIVERSITÄT DRESDEN

Conclusion

- The energy consumption of servers or their subsystems (CPU, disk, network, etc.) is never deterministic.
- Probabilistic adaptation strategies enable greater flexibility of trade-off between performance and energy consumption
- We believe that understanding the dependency between distributed services can be useful to design DPM strategies
- The cost of adaptation is not properly studied so far and needs further investigation.

TECHNISCHE UNIVERSITÄT DRESDEN

Thanks for Listening