Total Cost of Ownership in High Performance Computing

HPC datacenter energy efficient techniques: job scheduling, best programming practices, energy-saving hardware/software mechanisms

SoSe 2014
Dozenten: Prof. Dr. Thomas Ludwig, Dr. Manuel Dolz
Vorgetragen von Hakob Aridzanjan
03.06.2014

Outline

1. HPC energy efficiency: Current Situation
2. Calculating a datacenter‘s energy efficiency
3. Different levels of energy saving
4. Node level hardware factors
5. Node level optimization techniques
6. Grid and datacenter power management
7. Oversubscribing facility power
8. Future Predictions
9. References

HPC datacenter energy efficiency: Current Situation

- Datacenters consume more energy the more powerful they become

- US datacenter energy consumption 2006: ~61 billion kWh, total cost $4.5 billion and more than twice as high as in 2000

- Energy efficiency as a secondary priority in HPC design

HPC datacenter energy efficiency: Current Situation

- Datacenters consume more energy the more powerful they become

- US datacenter energy consumption 2006: ~61 billion kWh, total cost $4.5 billion and more than twice as high as in 2000

- Energy efficiency as a secondary priority in HPC design

HPC datacenter energy efficiency: Current Situation

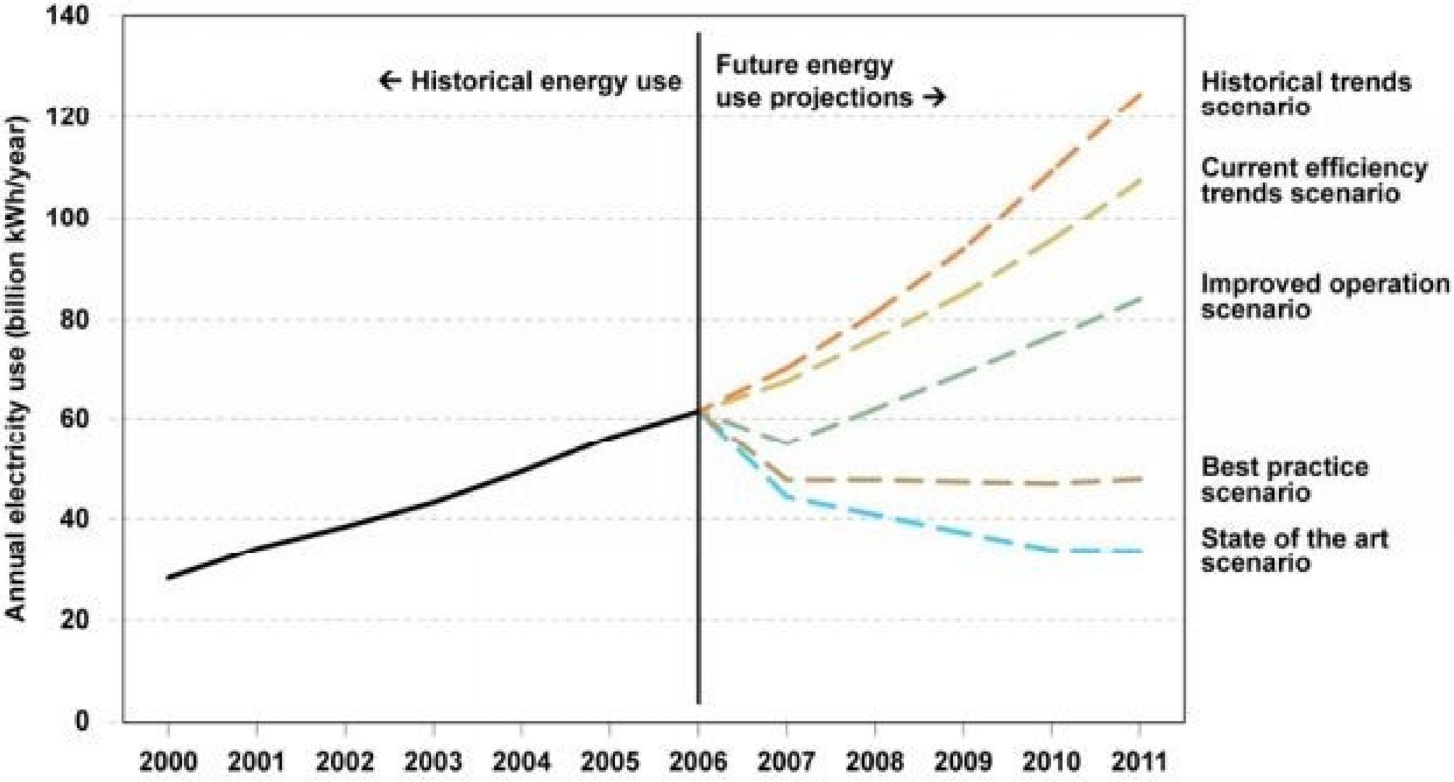
- Under current trends, US national datacenter energy consumption could double in the next five years

- Additional energy may require up to ten additional power plants in the US

- Various techniques to increase energy efficiency are readily available and in development

HPC datacenter energy efficiency: Current situation

Figure ES-1. Comparison of Projected Electricity Use, All Scenarios, 2007 to 2011



Source: U.S. Environmental Protection Agency,
“Report to Congress on server and datacenter energy efficiency,” Public Law 109-431, August 2, 2007.

Calculating a datacenter‘s energy efficiency

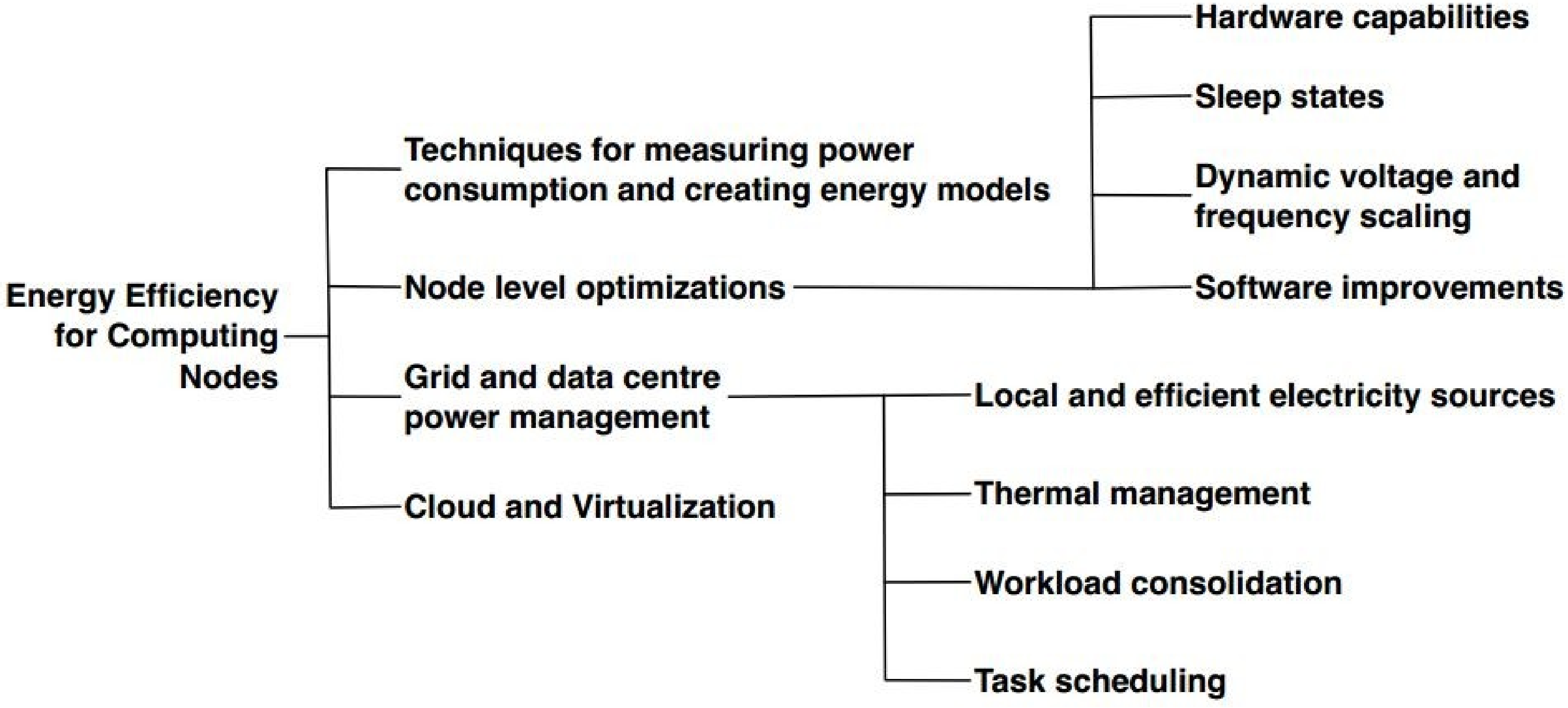
- Power Usage Effectiveness (PUE):

$$\mathrm{PUE} = \frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}}$$

- “Ratio of how much energy is used for actual computing“

- 1.0: Best possible value
- 1.2: Average Google-designed datacenter
- 1.83: US national average (Greenberg 2007)
- 3.0+: Exceptionally poor

Different levels of energy saving



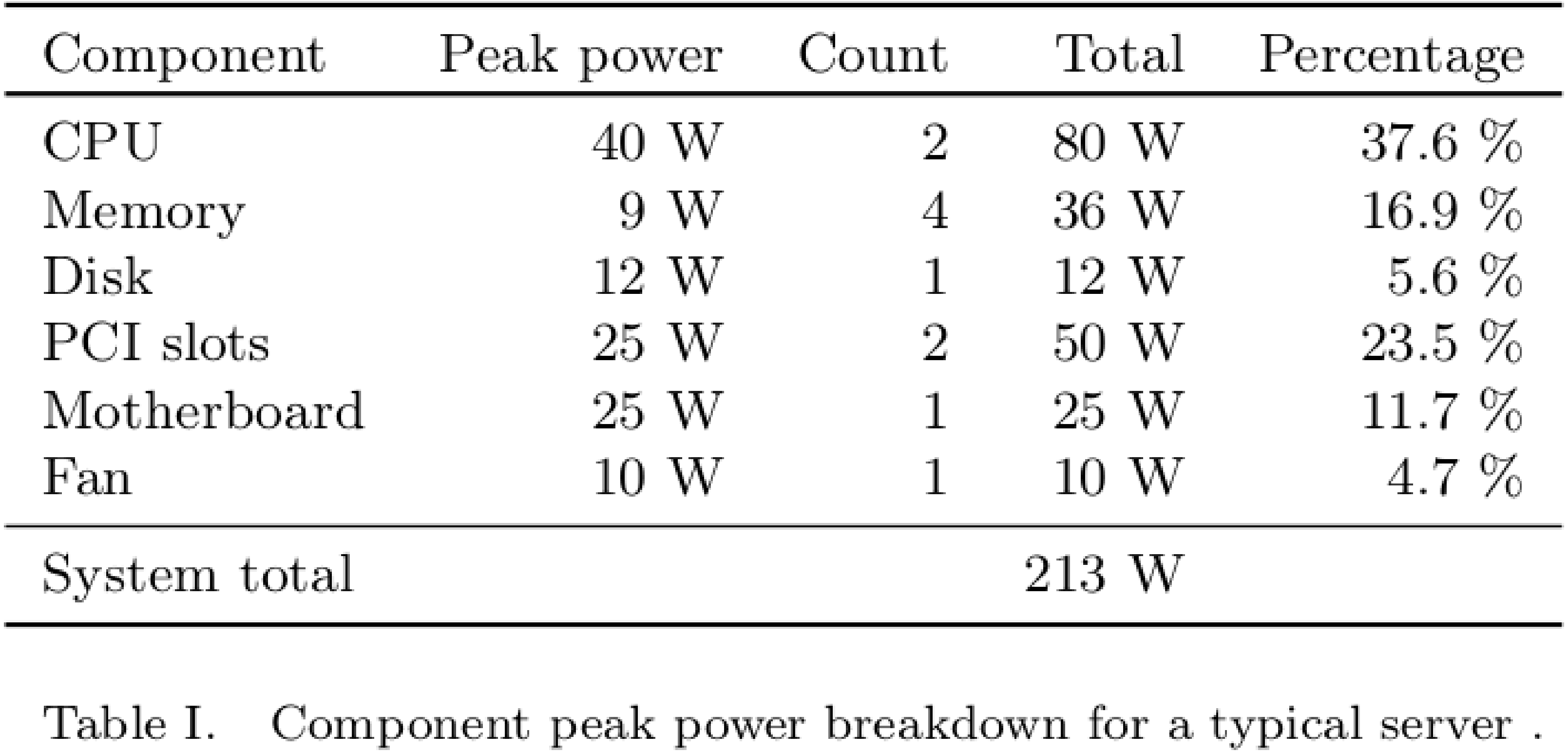
Energy Efficiency for Computing Nodes

- Techniques for measuring power consumption and creating energy models
- Node level optimizations
  - Hardware capabilities
  - Sleep states
  - Dynamic voltage and frequency scaling
  - Software improvements
- Grid and data centre power management
  - Local and efficient electricity sources
  - Thermal management
  - Workload consolidation
  - Task scheduling
- Cloud and Virtualization

Fig. 1. Overview of techniques to improve the efficiency of computing nodes.

Source: Anne-Cécile Orgerie et al.
A survey on techniques for improving the energy efficiency of large scale distributed systems
ACM Computing Surveys, Vol. TBD, No. TBD, TBD 2013.

Node level hardware factors



Component	Peak power	Count	Total	Percentage
CPU	40 W	2	80 W	37.6 %
Memory	9 W	4	36 W	16.9 %
Disk	12 W	1	12 W	5.6 %
PCI slots	25 W	2	50 W	23.5 %
Motherboard	25 W	1	25 W	11.7 %
Fan	10 W	1	10 W	4.7 %
System total			213 W	

Table I. Component peak power breakdown for a typical server .

Source: U.S. Environmental Protection Agency,
“Report to Congress on server and datacenter energy efficiency,” Public Law 109-431, August 2, 2007.

Node level optimization techniques

- Energy-aware hardware capabilities:
  Hardware that can detect access and **lower or raise activity levels**
  (f.i. HDDs, CPUs)
- Sleep state
  Hardware that **powers down when idle** and wakes up upon a signal
- Dynamic Voltage and Frequency Scaling (DVFS)
  CPUs capable of switching between **P-states of performance** and **C-states of idleness** to reduce **dynamic and static consumption** respectively
- Software improvements
  Optimization of applications, optimized OSs, disabling of unneeded modules in BIOS

Grid and datacenter power management

- Local and efficient electricity sources
  Depending on location, solar, wind or hydroelectric energy as viable **power generation** or cheap power plants nearby
- Thermal management
  Grid-based systems assigning heavy work to datacenters depending on season
- Workload consolidation
  Adjusting the **number of nodes** used per tasks efficiently
- Elevated cold aisle temperatures
  Google cools servers at **25-27°C** instead of 20°C with no sacrifices in stability or performance
- Energy-aware task scheduling
  Assigning and scheduling tasks to as **few nodes** as viable

Oversubscribing facility power

- Defined by Google* as 'hosting a number of systems (and storage, networking, etc.) where the addition of their cumulative peak power will exceed the facility's maximum IT power budget.'
- 6-month case study: Rack units used less than 65% of their peak power in 80% of the time, 93% peak power was reached
- Whole cluster units never ran above 72% of aggregate peak power, leaving 28% of the allocated power stranded

Same amount of electricity could power
40% more machines

*Source: The Datacenter as a Computer. An Introduction to the Design of Warehouse-Scale Machines.
Luiz André Barroso and Urs Hölzle, Google Inc.

Oversubscribing facility power

- Assign energy partially to non-critical processes → easy to pause/abort in case the energy demands become higher than usual
- Stability and reliability higher priority than efficiency → safer to have too much than too little power
- Oversubscribed power may decrease upon datacenter upgrade or expansion

Future Predictions

- 2025: Data center energy consumption will be 'much lower' according to Emerson Network Power (2014)
- Average datacenter power density to be between 40 and 80 kW
- Private power generation to increase with advancements in solar panel efficiency
- Solar energy predicted to replace coal as the most used energy source

Future Predictions

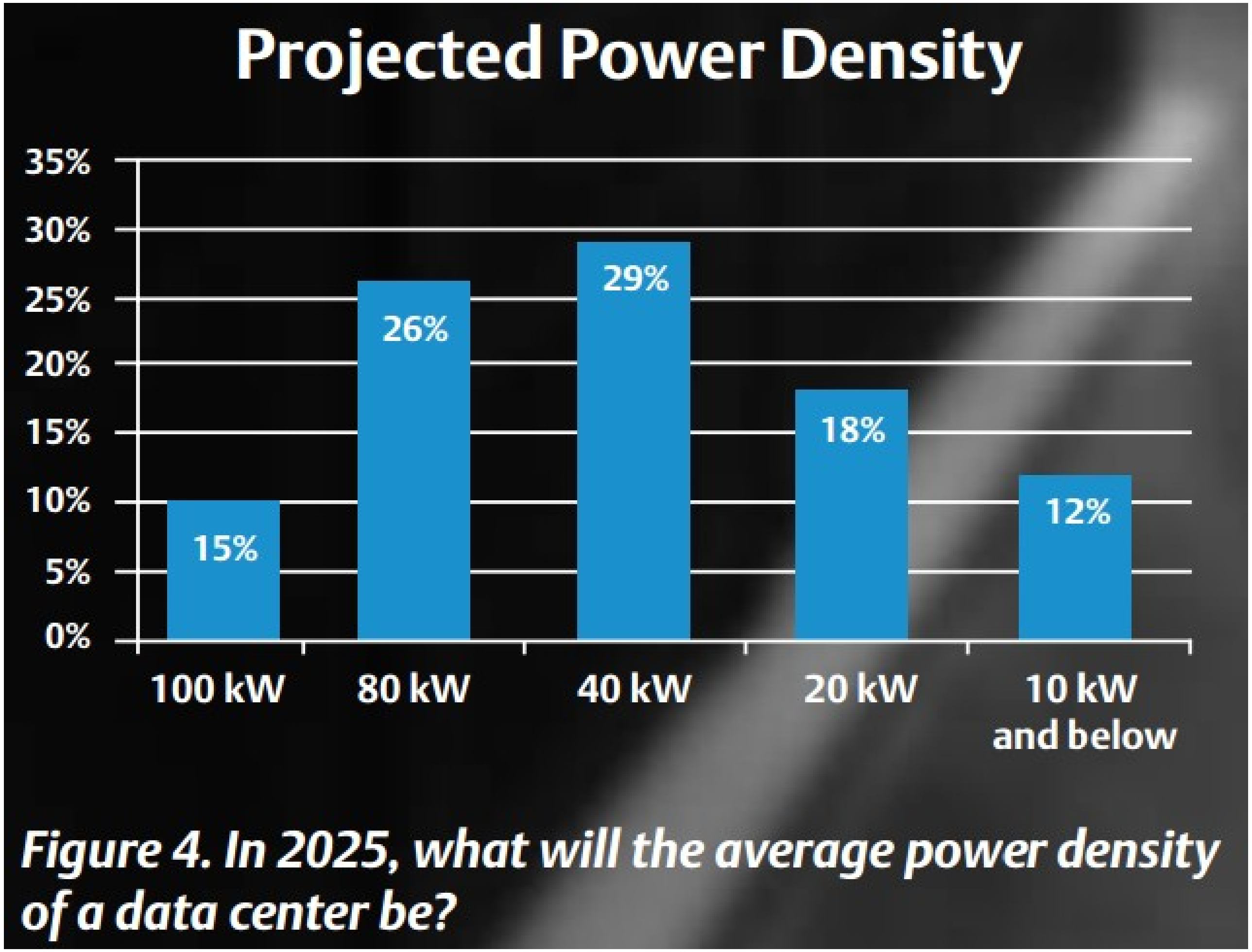


Figure 4. In 2025, what will the average power density of a data center be?

Source: Emerson Network Power. “Data Center 2025: Exploring the Possibilities” 2014

References

Emerson Network Power. “Data Center 2025: Exploring the Possibilities” 2014

The Datacenter as a Computer. An Introduction to the Design of Warehouse-Scale Machines.
Luiz André Barroso and Urs Hölzle, Google Inc.

U.S. Environmental Protection Agency,
“Report to Congress on server and datacenter energy efficiency,” Public Law 109-431, August 2, 2007.

Anne-Cécile Orgerie et al.
A survey on techniques for improving the energy efficiency of large scale distributed systems
ACM Computing Surveys, Vol. TBD, No. TBD, TBD 2013.

James H. Laros III et al. “Energy-Efficient High Performance Computing. Measurement and Tuning“
Springer 2013

Thank you for listening