« Green Computing » implique la minimisation du coût énergétique et de l'empreinte sur l'environnement des systèmes informatiques

Prof. David Atienza Alonso
Embedded Systems Laboratory (ESL)
Institute of EE, Faculty of Engineering
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

INFORUM
Nouvelles technologies de l'information

INFORUM Keynote, 23 mars 2010

What is “Green Computing”? The ICT Ecosystem

It is a global target, to be addressed at all ICT levels

[Courtesy: Prof. M. Pedram, USC]

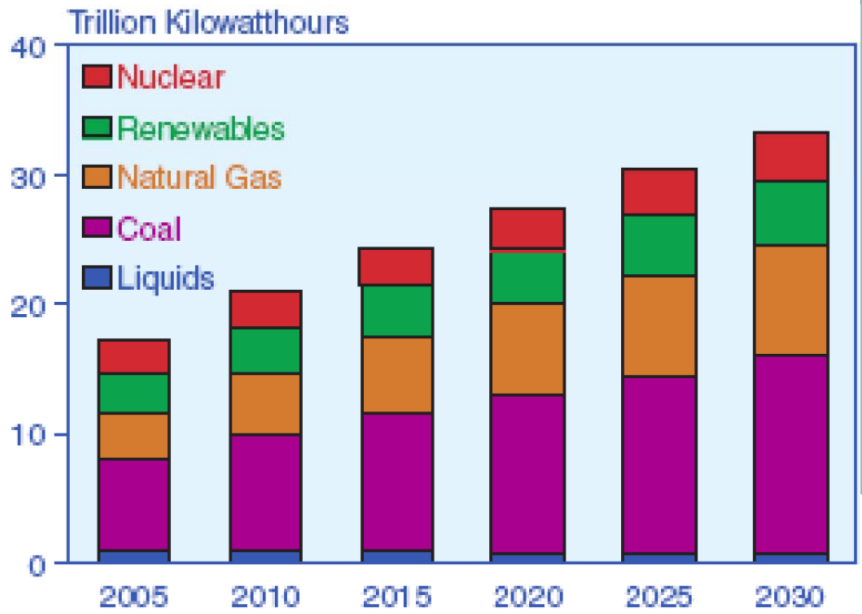
- The Information Communication Technology (ICT) ecosystem encompasses [Harvard Law]:
  - Policies, strategies, information, technologies, applications & stakeholders
  - Defines global technology environment for a country, government or enterprise

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

What is “Green Computing”? (for Electrical Engineers and IT Professionals)

- Looking for a meaning: “Green” as physical magnitude?
  - Performance?
  - Area?
  - Power?
  - Energy?
  - Style?

= Multi-objective target
Energy-minimal computing with certain performance and area constraints



[Sources: Energy Information Administration (EIA), International Energy Annual and Analysis of Global Energy Markets and Electricity Module]

Exponential increase in ICT energy demands...
2 Terawatts (TW) today to 5 TW by 2050, why?

Number of Electronic Systems: Global ICT Carbon Footprint by Subsector

Number of PCs (desktops and laptops) expected to increase from 592 Million in 2002 to more than four Billion in 2020

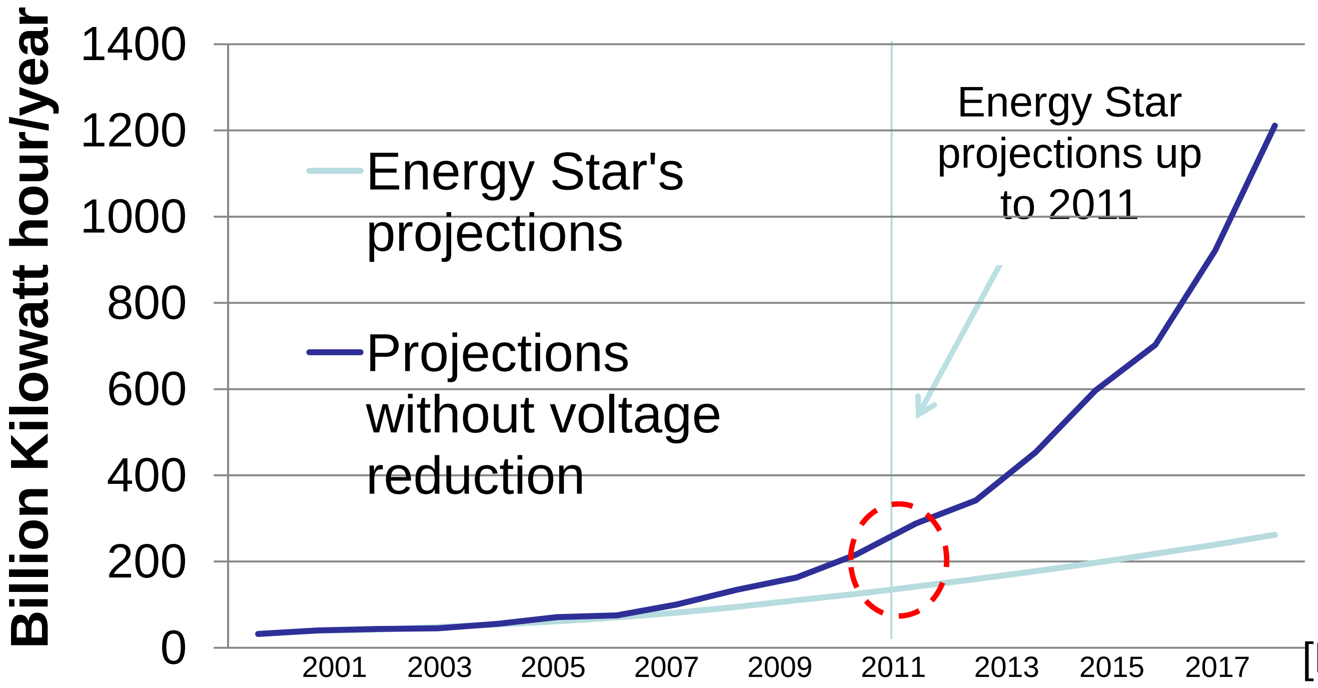
* Printers were 11% of the total ICT footprint in 2002, 8% in 2007 and will be 12% in 2020.

www.smart2020.org

THE CLIMATE GROUP GeSI

Power Density in Computing Systems: "Economic Meltdown of Moore's Law"

- 40 years of IT industry & Moore's law being discontinued
  - 2x more transistors every 2 years, quadratic energy reduction (voltage)
  - Voltages have started to level, ITRS projections in 2000 for voltage levels in 2009 were 30% lower

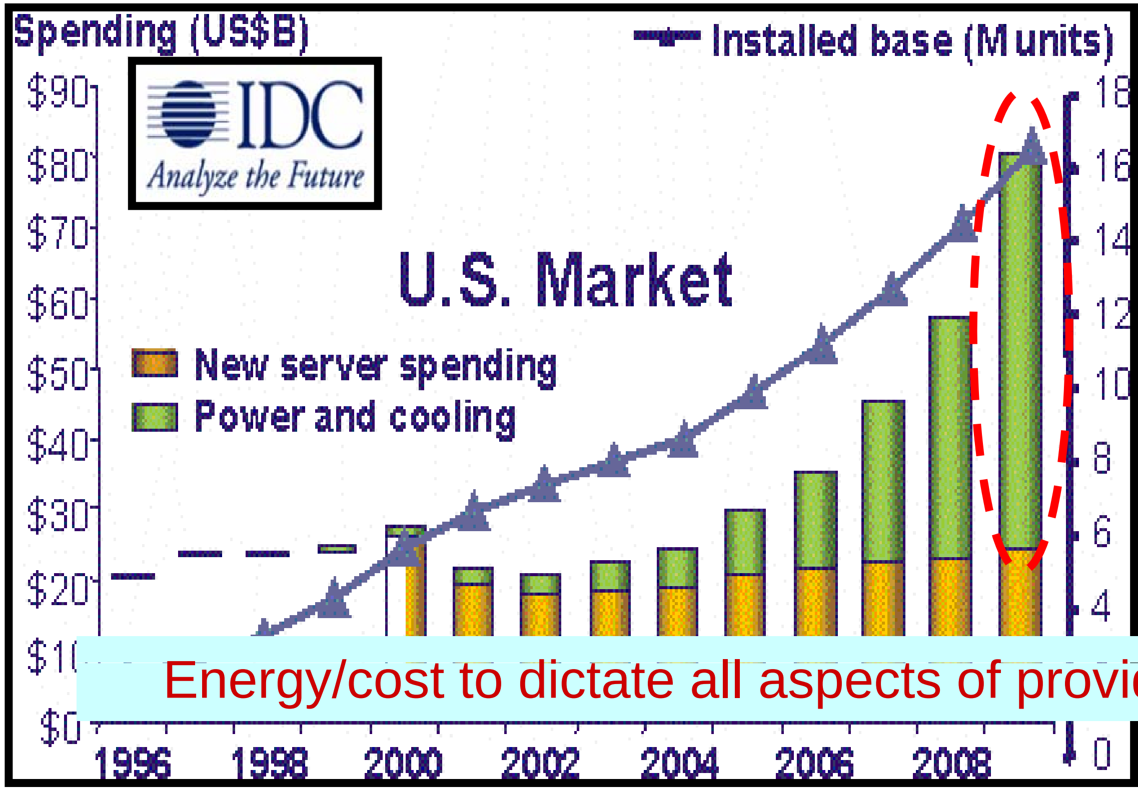


[Kenneth Brill, Uptime Institute]

Exponential increase in energy usage every new processing generation

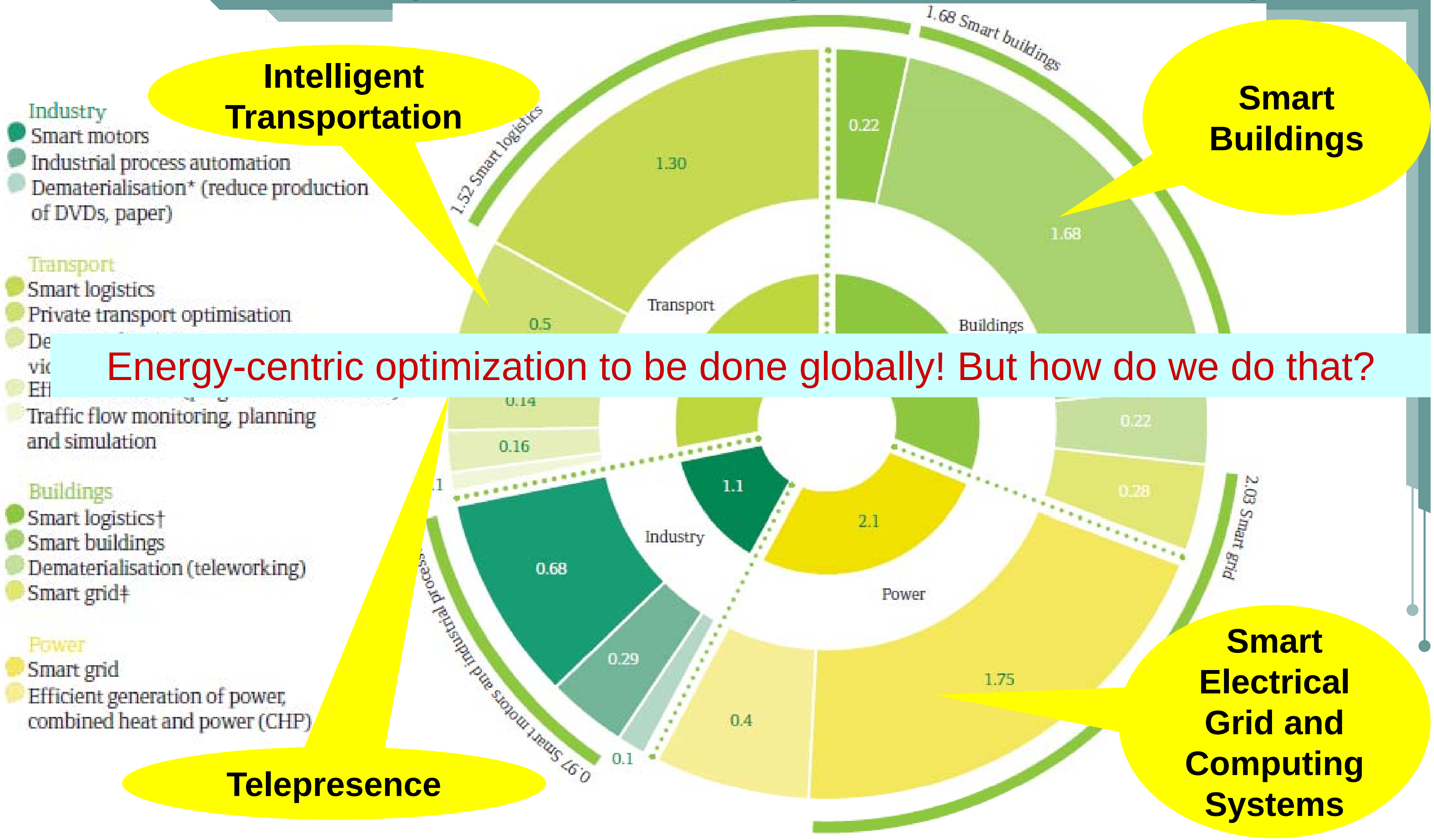
Datacenters Global Trend and Market Drivers for Energy-Centric Computing

- IT demand outpaces technology improvements
  - *Server* ***energy increase 15% / year****; not sustainable.*
  - *Servers use 1.2% of U.S. energy (Koomey Study);*
- Energy-centric globally needed: “Green painting” not enough!



Energy/cost to dictate all aspects of providing computing services

Energy-Centric Computing: Smart 2020 Challenge for Reducing ICT Emissions by 50%

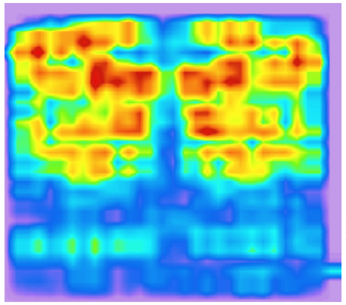


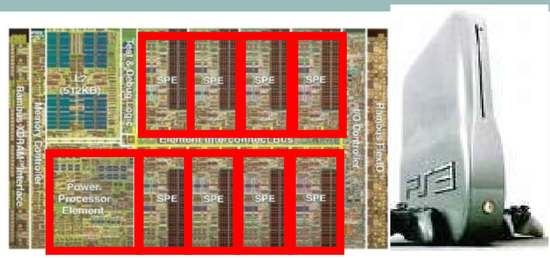
Energy-centric optimization to be done globally! But how do we do that?

Expected Total ICT 2020 Emissions are 1.43 $GtCO_2e$

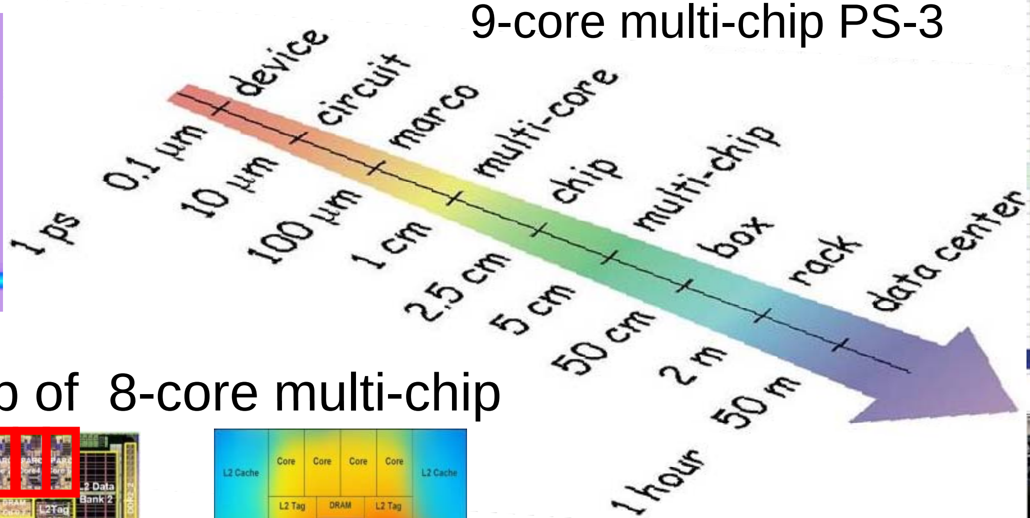
Energy and Thermal Optimizations at all Scales for Green Computing

Temperature map of dual core microprocessor

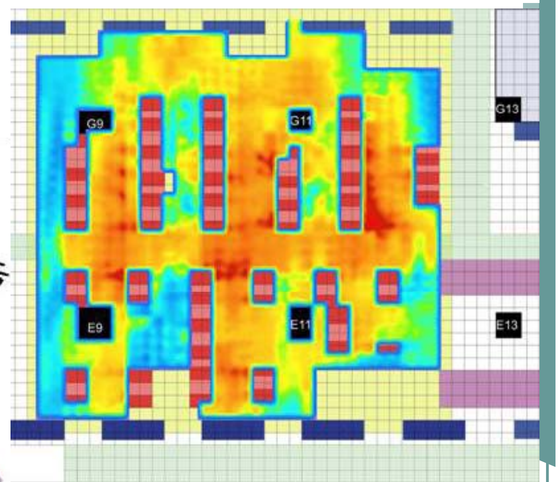




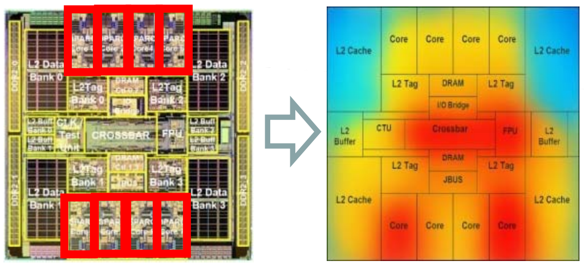
9-core multi-chip PS-3



Temperature map of data center



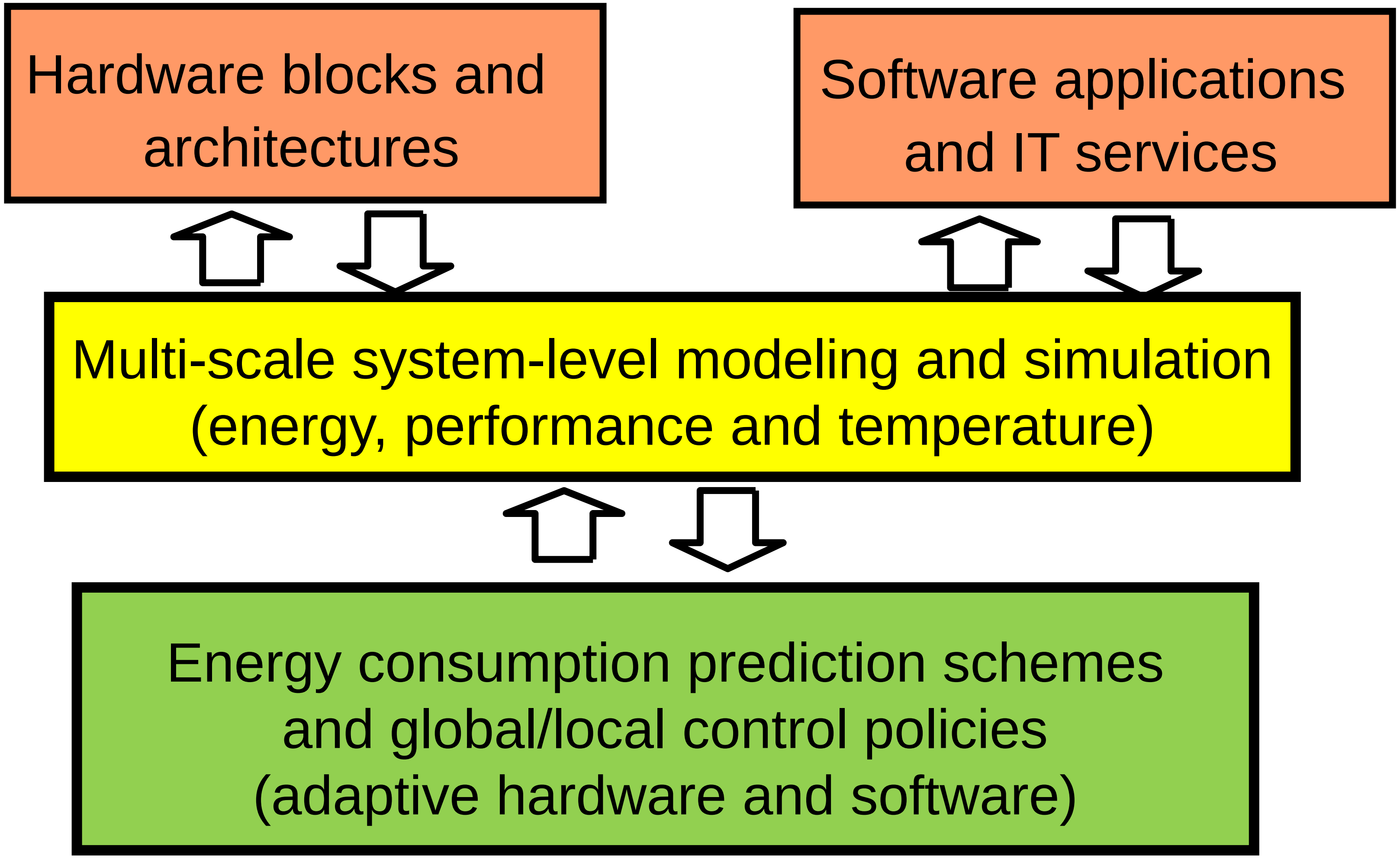
Temp. map of 8-core multi-chip



Holistic (system-level) view needed for energy-centric computing

Converging Challenges: Energy-Centric Design to Achieve Green Computing

- System-level design: optimizing at multiple levels



New Research Programs Needed for Green Computing: Nano-Tera.ch

nano — manufacturing challenges

tera — design challenges

Physical level

- Materials
- Devices
- Components
- System Structure

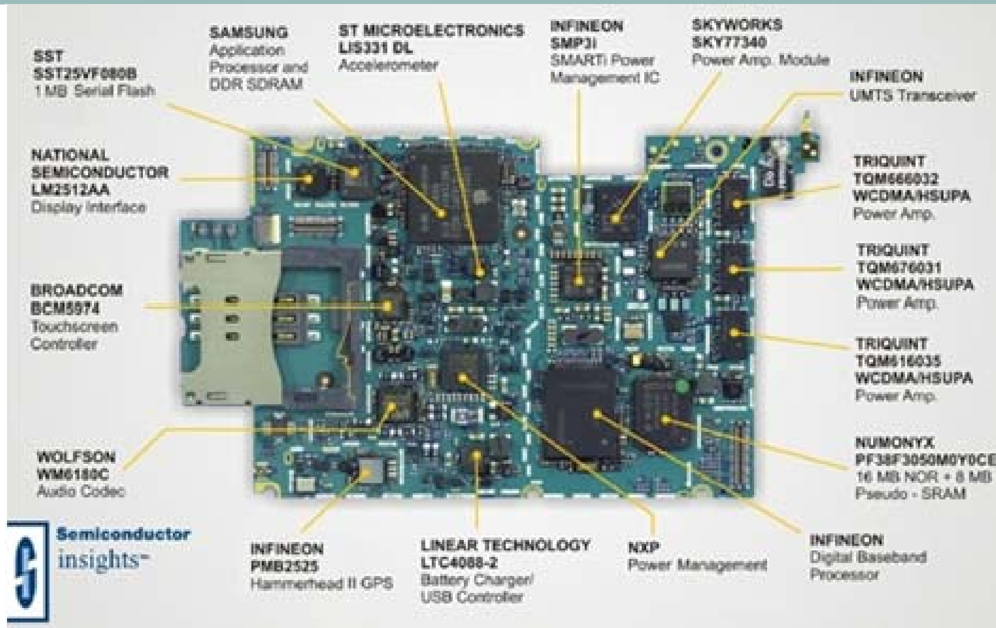
System level

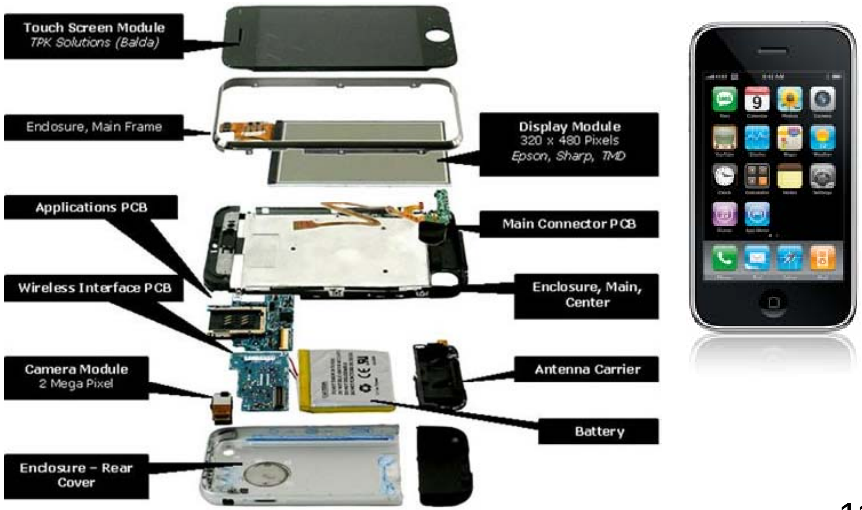
- Environmental monitoring
- Personalized health care
- Remote networking
- Energy scavenging
- Security
- Distributed intelligent agents
- Circuit design

Inter-disciplinary efforts in engineering, combining industrial and academic effort

Small-Scale Computing Systems: Multimedia Embedded Systems

- Multiple components and vendors involved
  - Processor, memory, RF…
- Value due to integration of components in final single system
- Energy optimization must be also applied in an integral way, 2 levels:
  - Architecturally
  - Algorithmically





Architectural Design to Reach Green IT in Embedded Systems

- Hardware
  - What is the right choice and combinations of components? Processors, radios, storage, networking…
- Software
  - How to manage/extend power-related decisions across abstraction layers (more in software than hardware)? Metadata methods, characterization, introspection
- Power-states and energy-saving transitions
  - What is the right number of power states and methods to move among them? Multiples voltage and frequency scaling modes, different on-off transition times

Algorithmic Improvements for Energy-Centric Design: The More You Know, The Less You Do

1. Heterogeneity is a fact of life

2. Increasing bandwidth of power consumption

3. Abstraction stacks have a high cost

Know More → System-Level design and optimization: software applications guide hardware power modes

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

1. Heterogeneity is a Fact of Life

Wireless Radios

Wi Fi

GSM

UWB, NFC

802.11a/b/g/n

Display System

ATI RADEON GRAPHICS

intel G45 EXPRESS CHIPSET

Storage System

Multiband radios: over wide frequency range and bandwidth

Multimode radios: different protocols in a single baseband

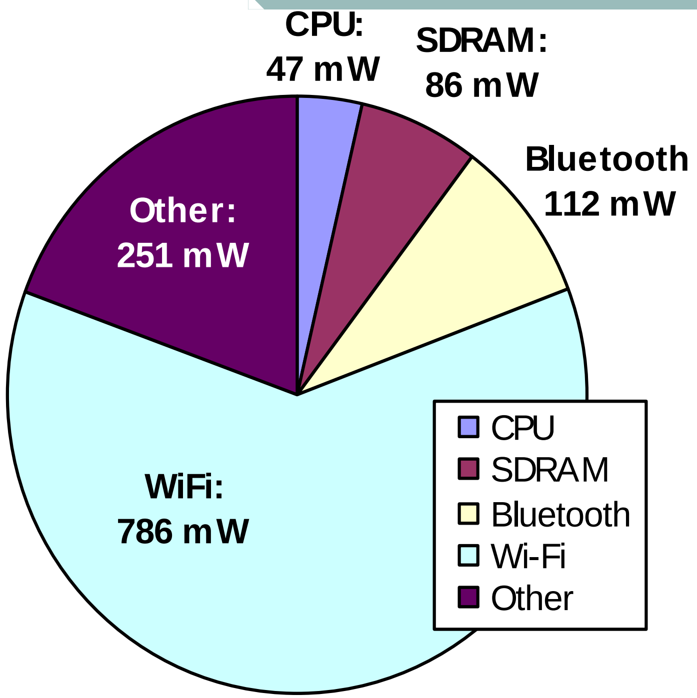
Wi Fi Medium range, High power (400mW-1W), Higher bit-rate (54Mbps)

Short range, low power (20mW-100mW), lower bit rate (2Mbps)

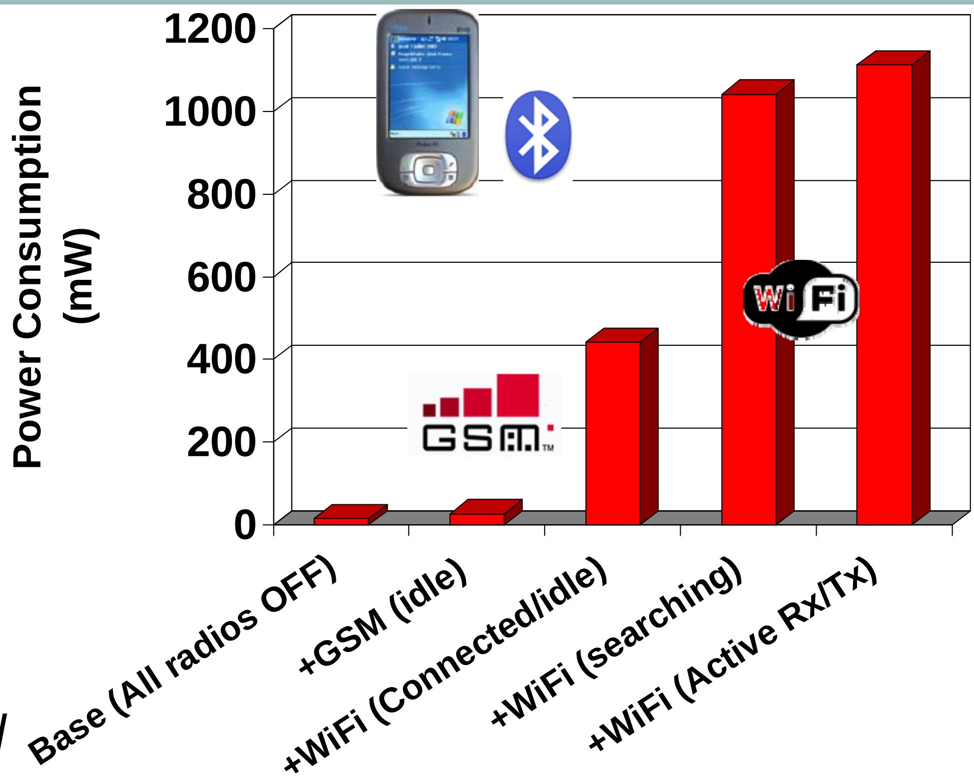
GSM **Long Range, very low power (<10mW), voice only**

From radios to storage to processors, heterogeneity is already in embedded systems

2. Increasing Bandwidth of Power Consumption



Power breakdown for a *fully connected* mobile device in *idle* mode, with LCD screen and backlight turned off.

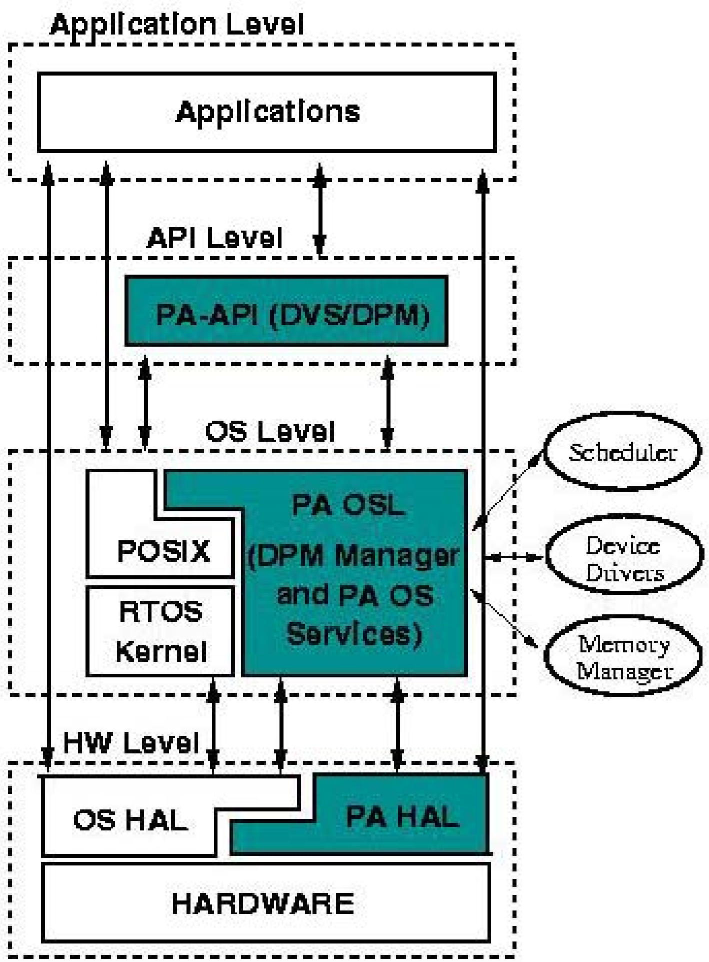


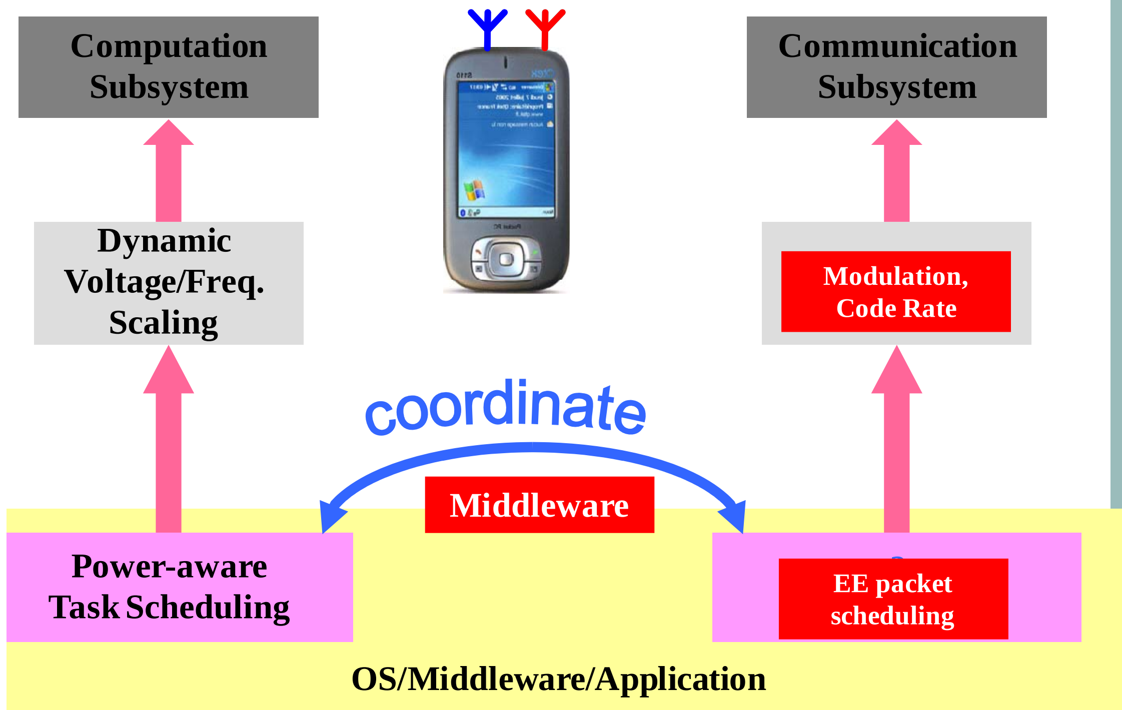
6-10x variations in power are common in each component

- Cellular voice radio (GSM) highly optimized for low idle power
  - Cingular 2125: GSM radio consumes 38 times less power than Wi-Fi!

3. Abstraction Stacks have a High Cost: Cross-Layer Optimizations

Introspect





Collaborate

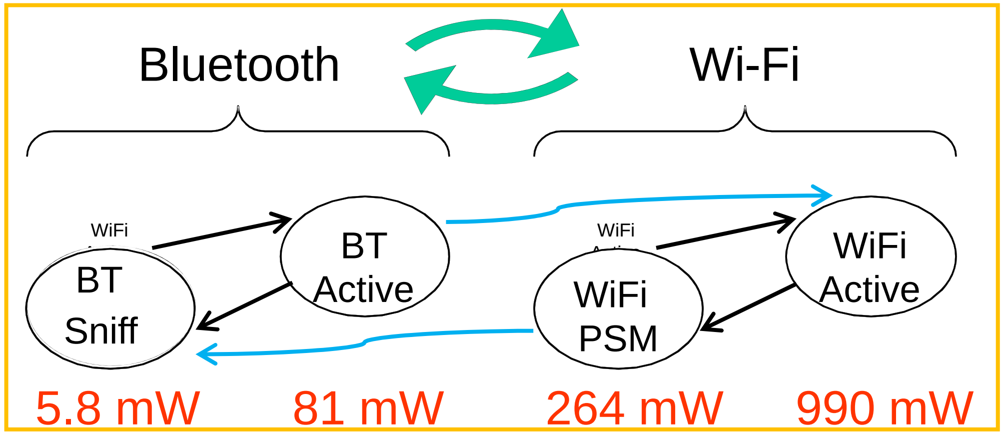
Across Radios, Processor, Radios & Processors

[Courtesy: Prof. R. Gupta, UCSD]

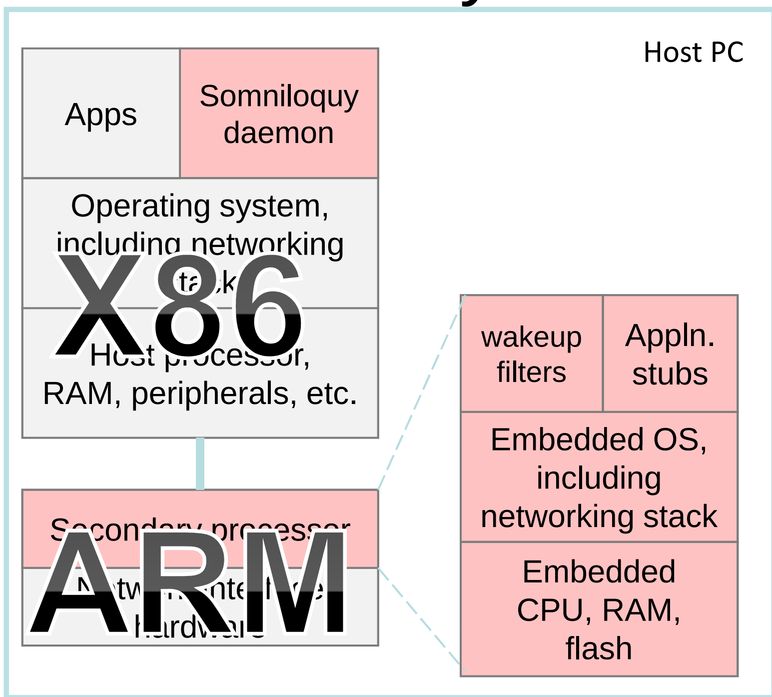
Example 1) Network Latency-Aware Energy Control

- Radio-switching hierarchies
- Wake-up based approaches in multi-core systems

Paging Radios



1.7-6.4x extension in battery life



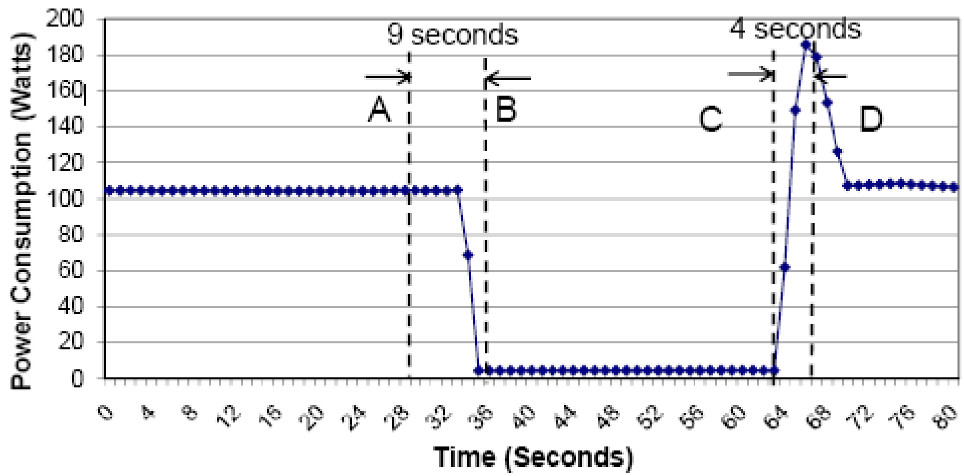
Secondary Processors

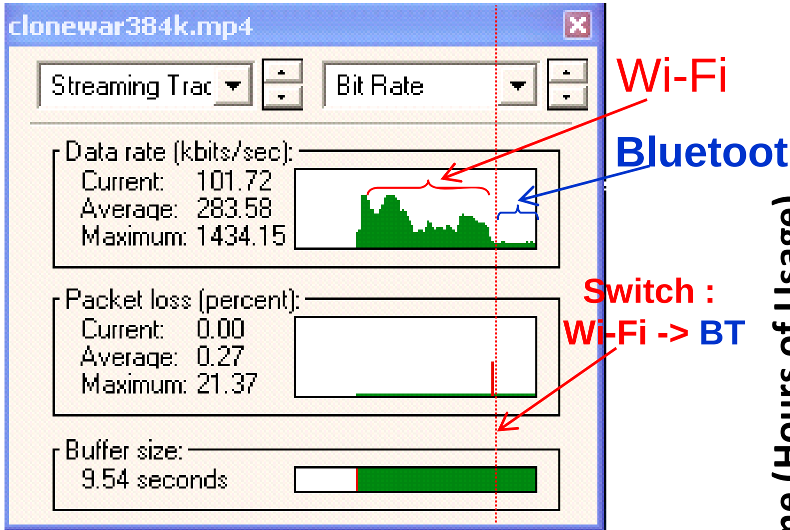
92% less power than one core alone

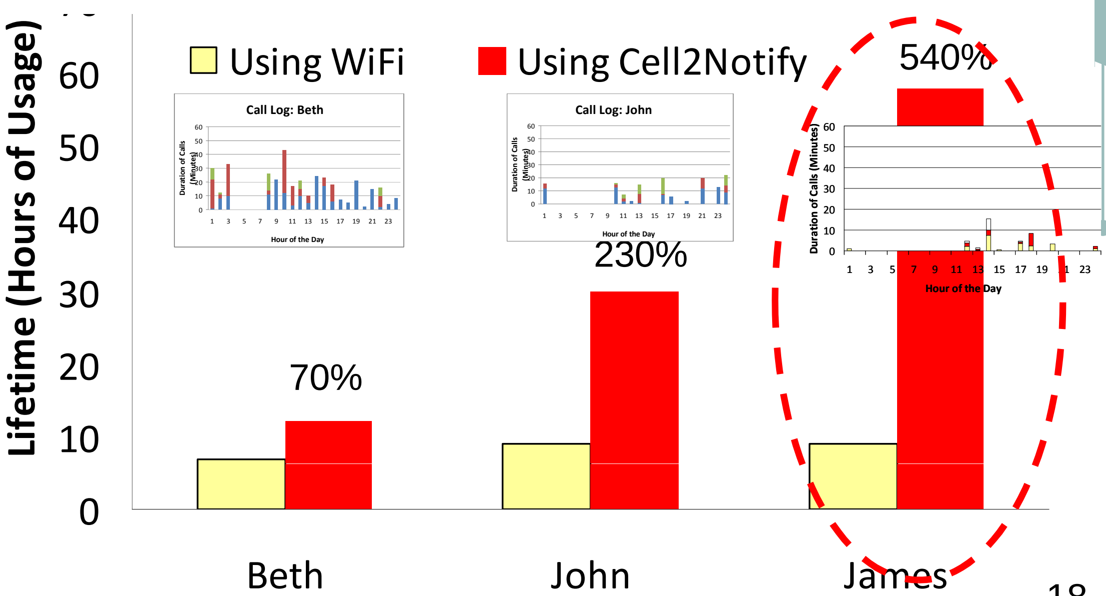
[Courtesy: Prof. R. Gupta, UCSD]

Example 1) Network Latency-Aware Energy Control: Somniloquy Prototype

[Courtesy: Prof. R. Gupta, UCSD, with Microsoft Research]

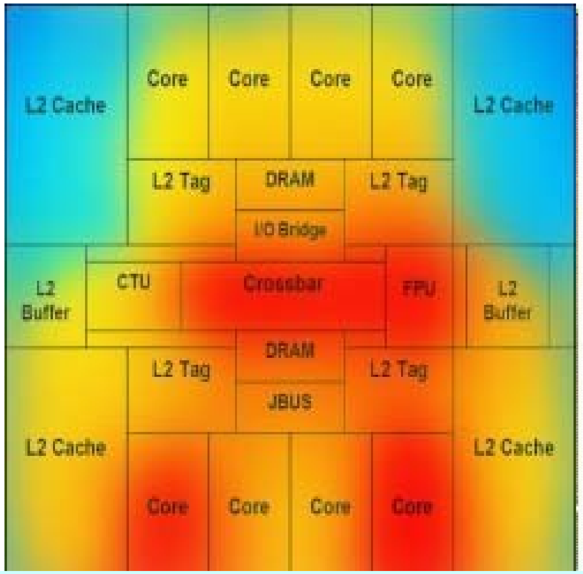




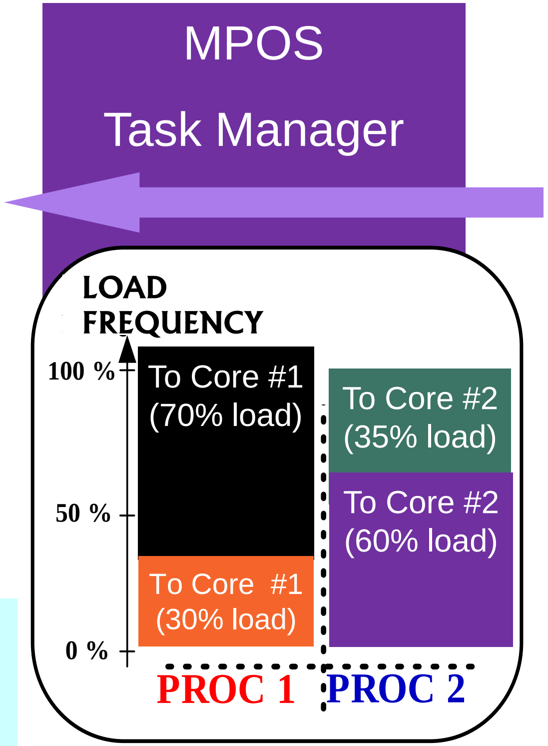


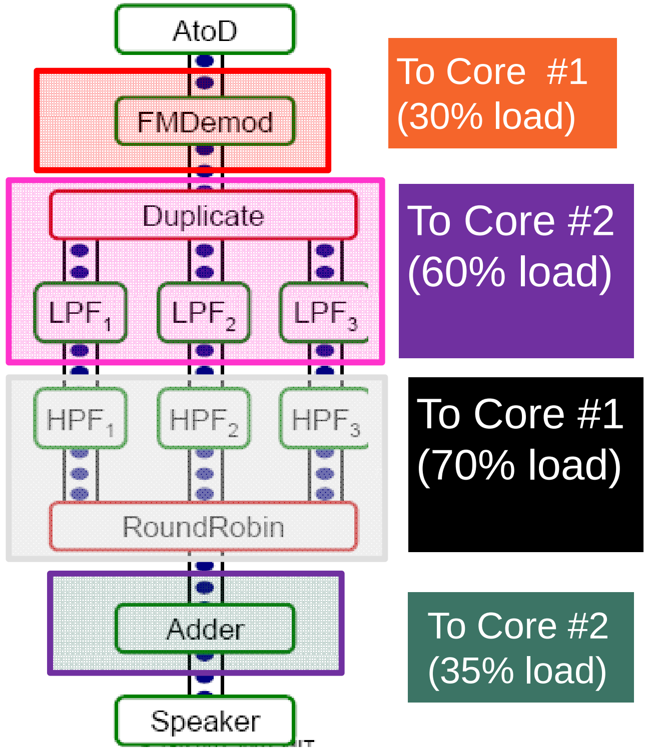
Cell2Notify up to 6.4x lower energy than WiFi

Example 2) Hardware/Software Thermal Management for Multi-Processor Systems



Can we control cooling costs as well as performance?



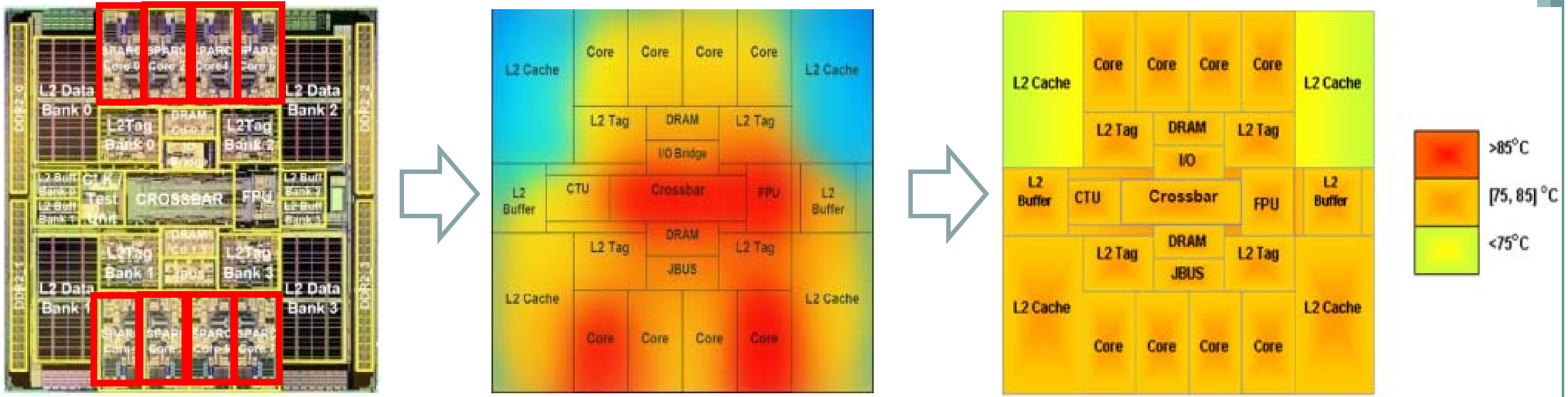


[Software FM Defined Radio]

- SW layers introduced to exploit the HW of multiprocessors
  - Applications divided in <u>tasks/threads</u>: blocks of operations to run
  - Multi-processor Operating System (MPOS) distributes the tasks
    - <u>Load balancing:</u> equal distribution of work between processors
    - Thermal sensors are already included in latest multi-core systems

Example 2) Hardware/Software Thermal Management for Multi-Processor Systems

Initial: Large gradients **New: Thermal balancing**



- HW: 8-core Sun Niagara
- MPOS: Solaris Multi-Core
- SW: Web and video processing

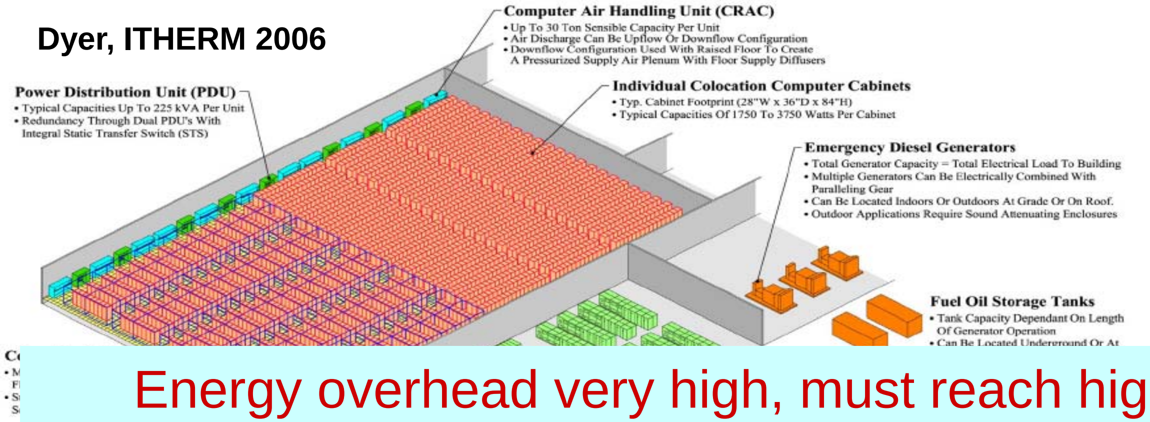
HW: convex-based dynamic voltage and frequency scaling and
SW: predictive task scheduling

Thermal balancing reduces operating temperature by 13 degrees: 40% cooling energy reduction

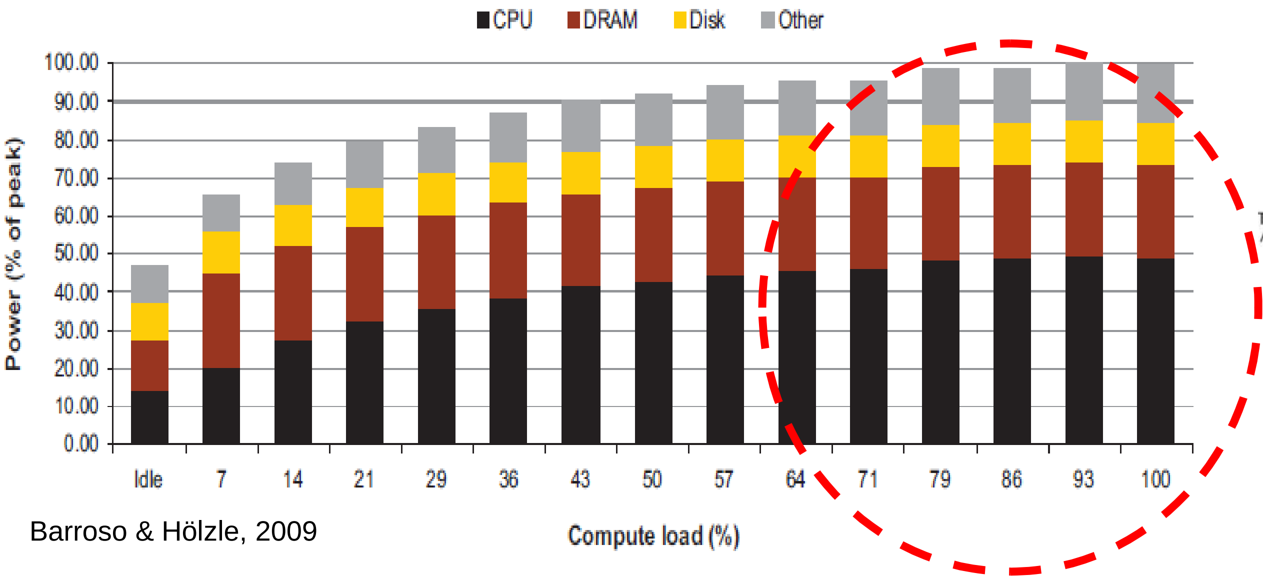
Large-Scale Computing Systems: Datacenters

- Area is expensive, always denser infrastructures
  - New containers: 2500 servers each, >10x density vs. old datacenters

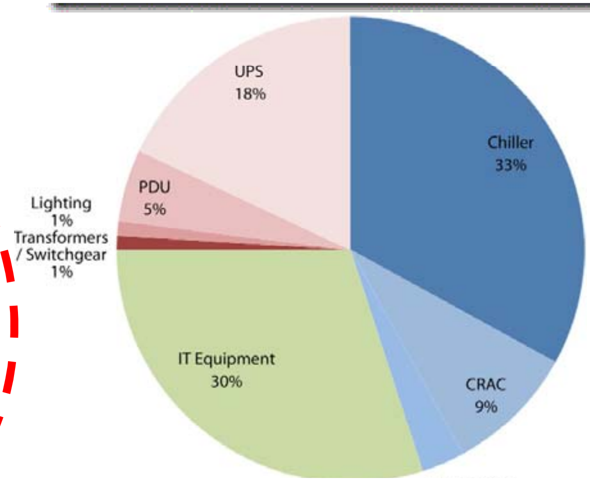
Dyer, ITHERM 2006



Energy overhead very high, must reach high computational loads

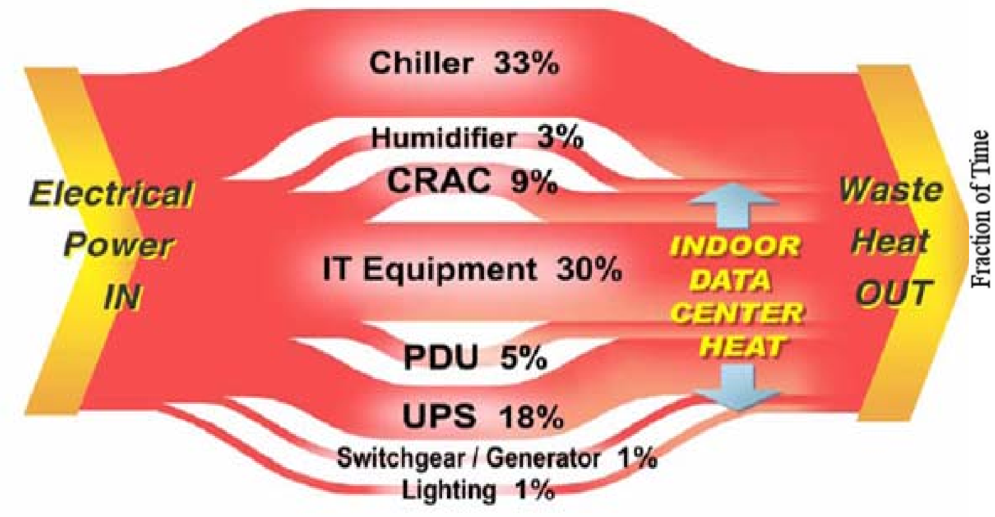


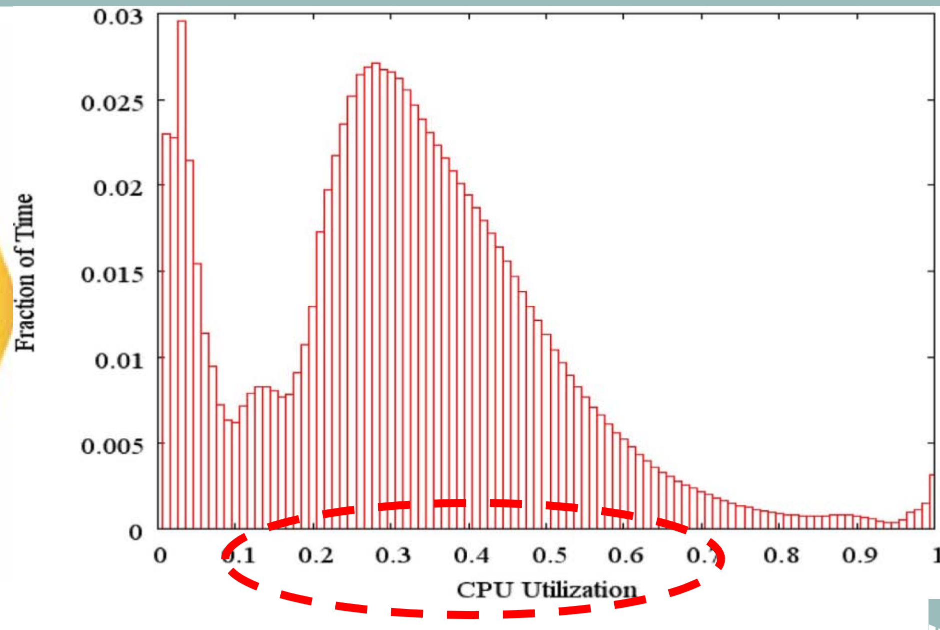
Barroso & Hölzle, 2009



Datacenter energy overhead, ASHRAE

Typical Computing Power and Energy Efficiency of Air-Cooled Datacenters





- Current **air-cooled datacenters** are **very inefficient**
  - Cooling needs as much energy as IT… and thrown-away
- Utilization of IT-hardware: 20-50% in general
- For a 10 MW datacenter : **~US$ 4M wasted** per year

Energy Proportional IT in Datacenters: Power and Thermal Management Together

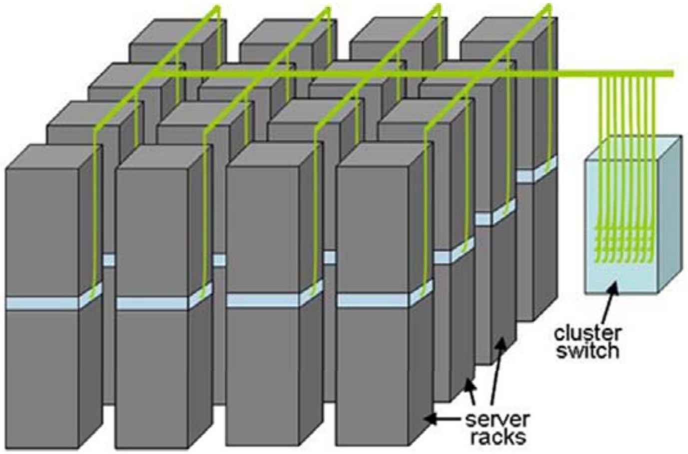
- Key challenges:
  - Effective energy management across layers with easy to develop software
  - Building energy balanced systems from imbalanced components (operation and cooling)
  - Graceful degradation of

Energy-proportional datacenters: Integrated layers of software-hardware to minimize joules/work

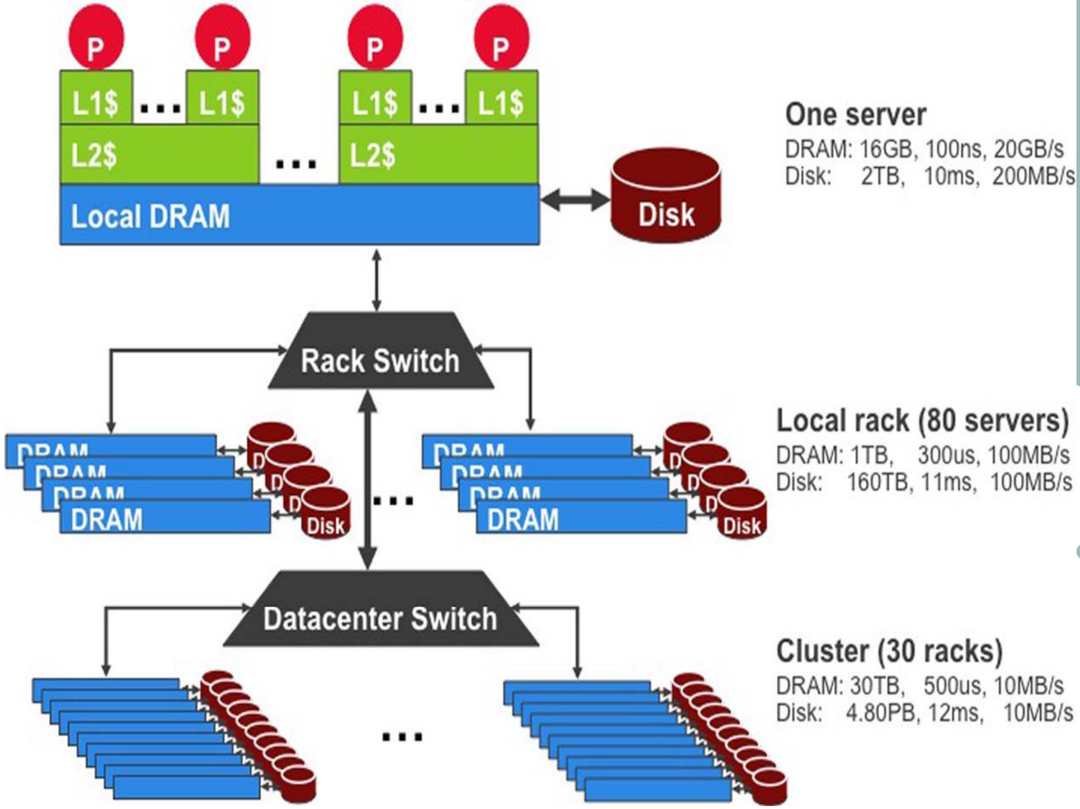
Source: R. Stroud, Google, 2009

Energy-Centric Design for Datacenters: System-Level View

- Multi-scale energy management solution needed



  - Architecture: middleware, network, applications mapping, etc.



Barroso & Hölzle, 2009

  - Technology: cooling, processing system, etc.

[Courtesy: Prof. T. Simunic, UCSD]

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Zero-Emission Datacenter: Liquid Cooling Technology and Predictive Energy Management

Datacenters are "intelligent" heaters

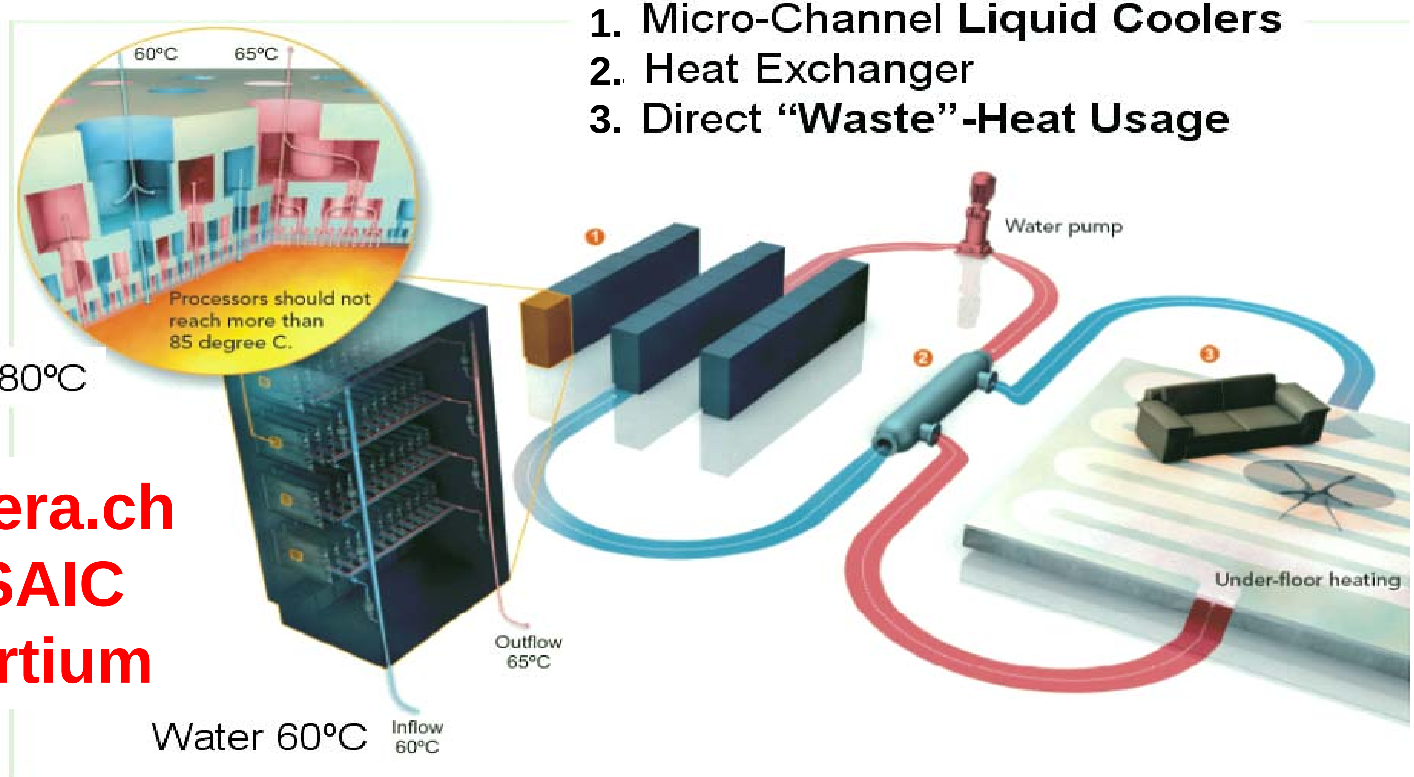
- 30-40% of carbon footprint in Europe using district heating networks

Direct re-use of heat

- "Hot"-liquid outlet temperature of more than 60°C
- Effe

Potential payback: 60-80% of electricity costs

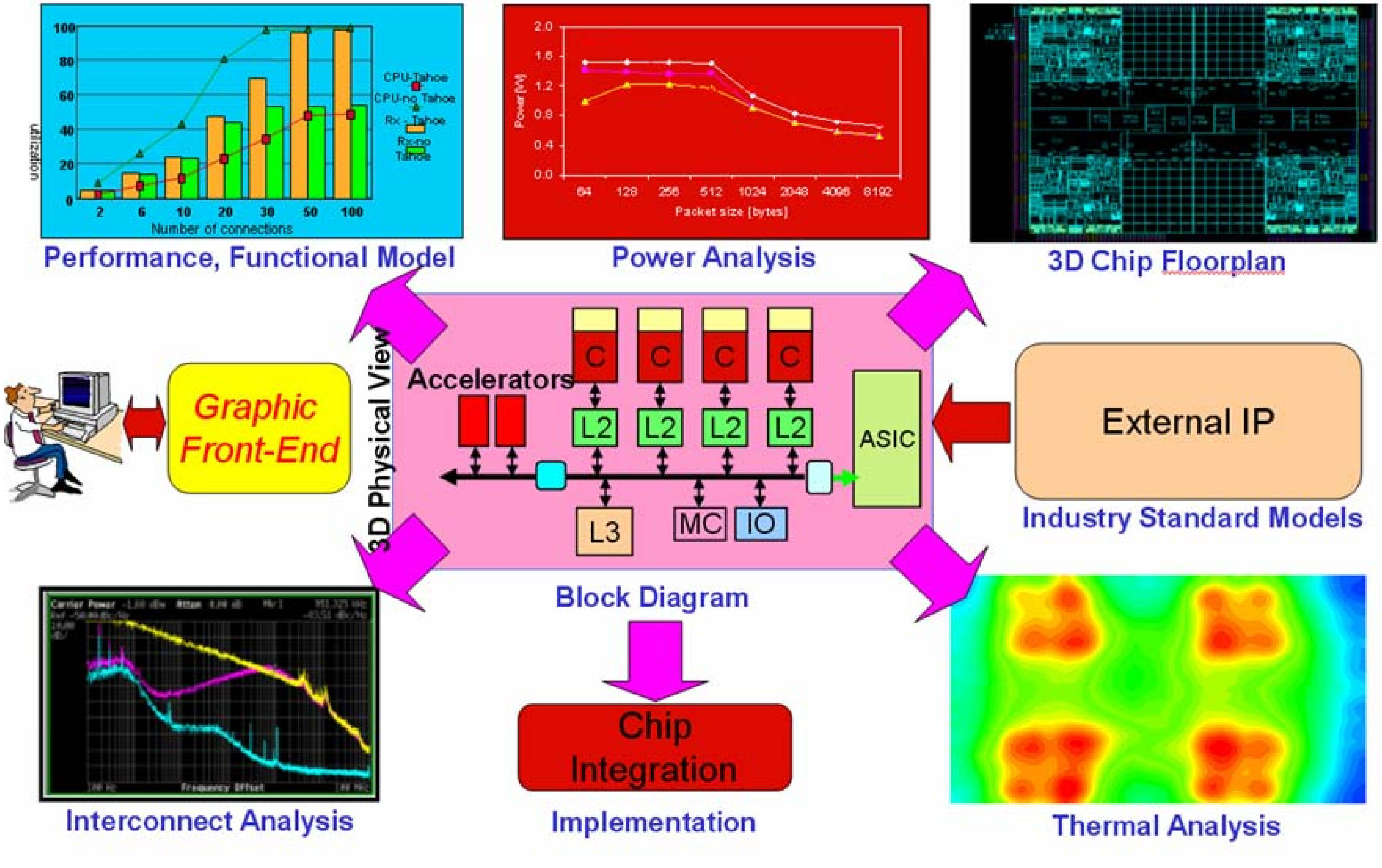
1. Micro-Channel **Liquid Coolers**
2. Heat Exchanger
3. Direct **"Waste"-Heat Usage**



Nano-Tera.ch CMOSAIC Consortium

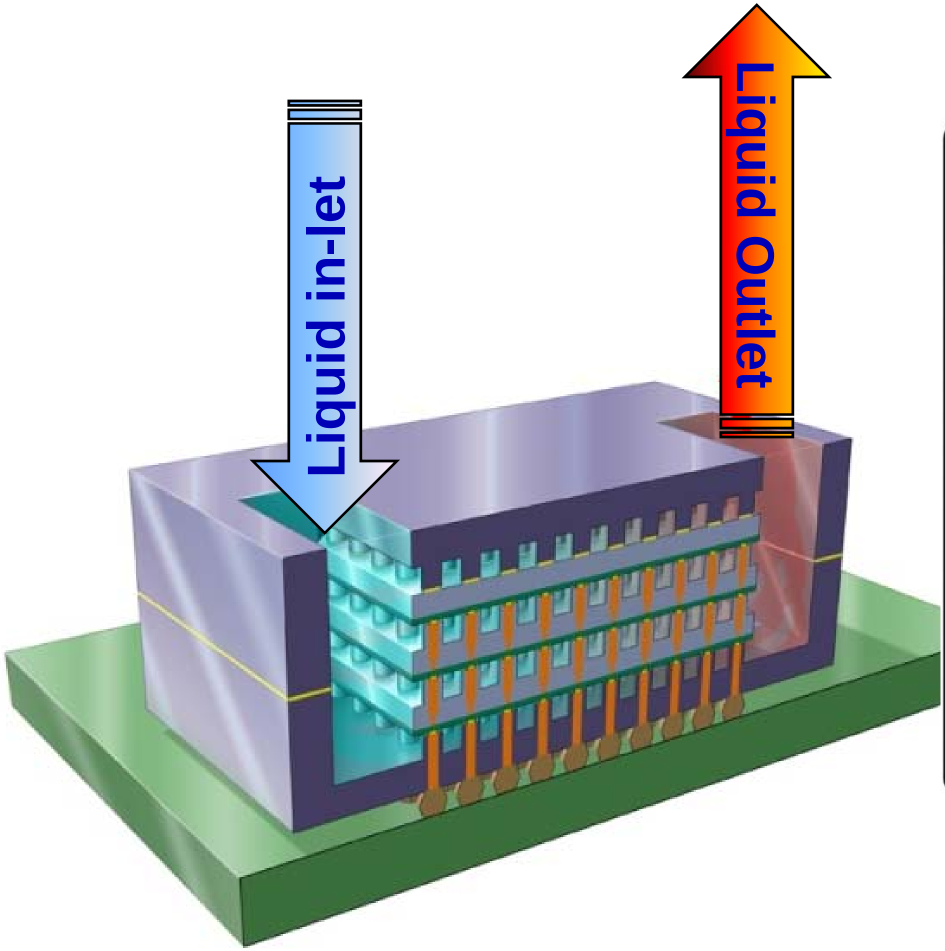
NanoTera CMOSAIC Project: Design of 3D Chips with Advanced Cooling

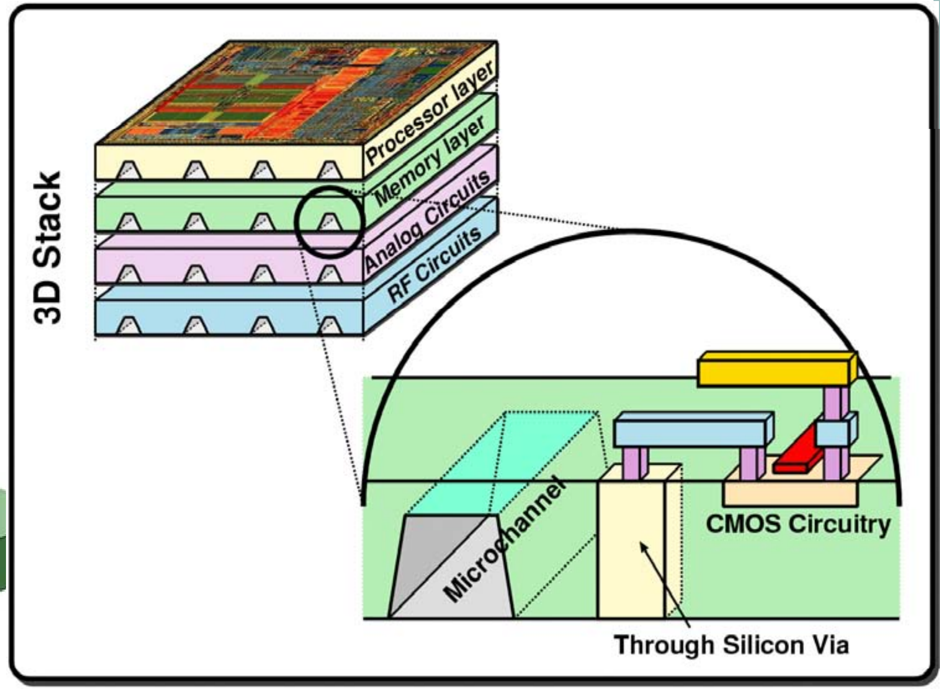
- 3D systems require novel electro-thermal co-design
  - Academic partners: EPFL and ETHZ
  - Industrial: IBM Zürich



NanoTera CMOSAIC Project: Design of 3D Chips with Advanced Cooling

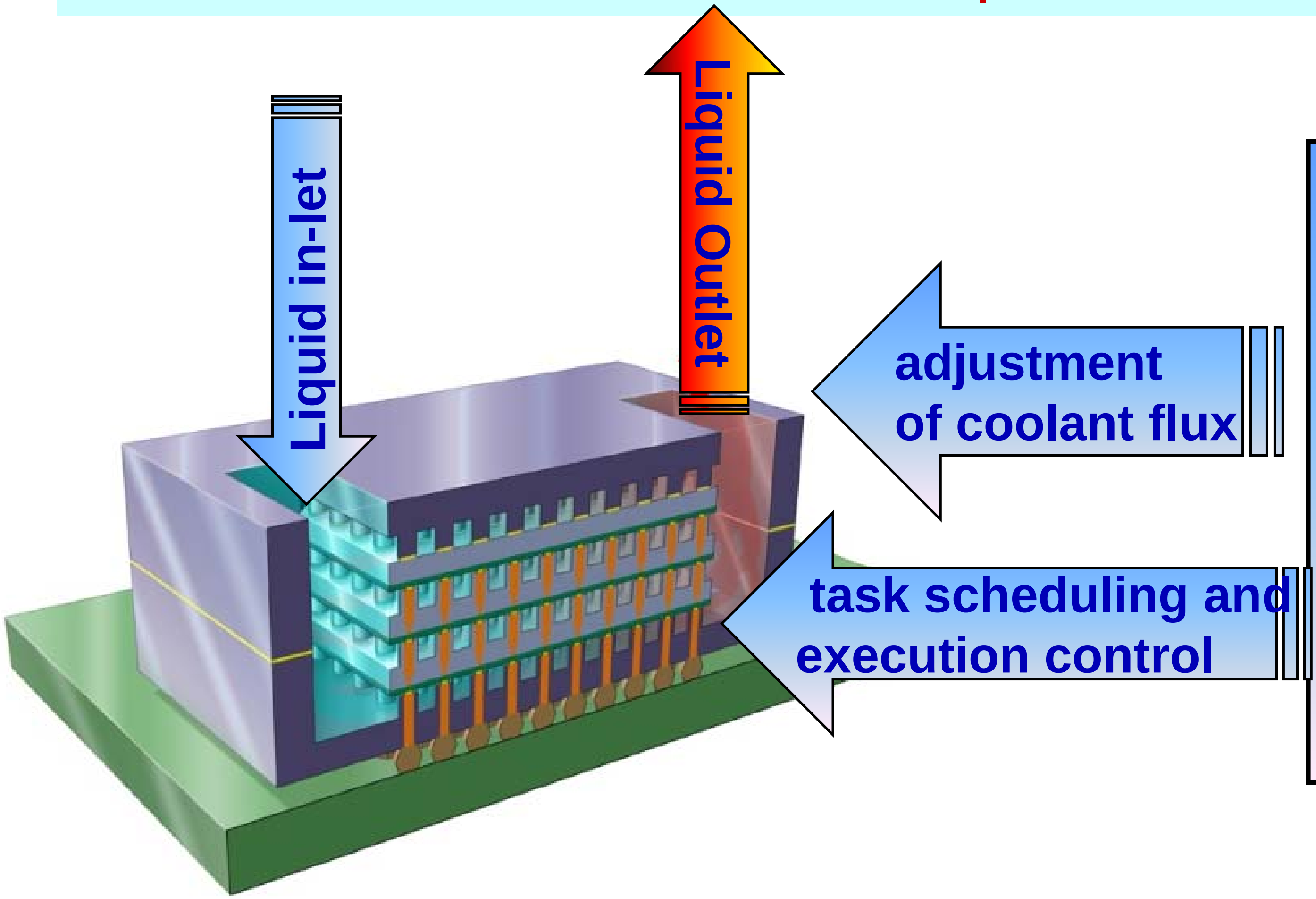
3D stacked multi-core chips: microchannels etched on back side to circulate liquid coolant





NanoTera CMOSAIC Project: Design of 3D Chips with Advanced Cooling

System-level energy management: cooling and execution cost prediction

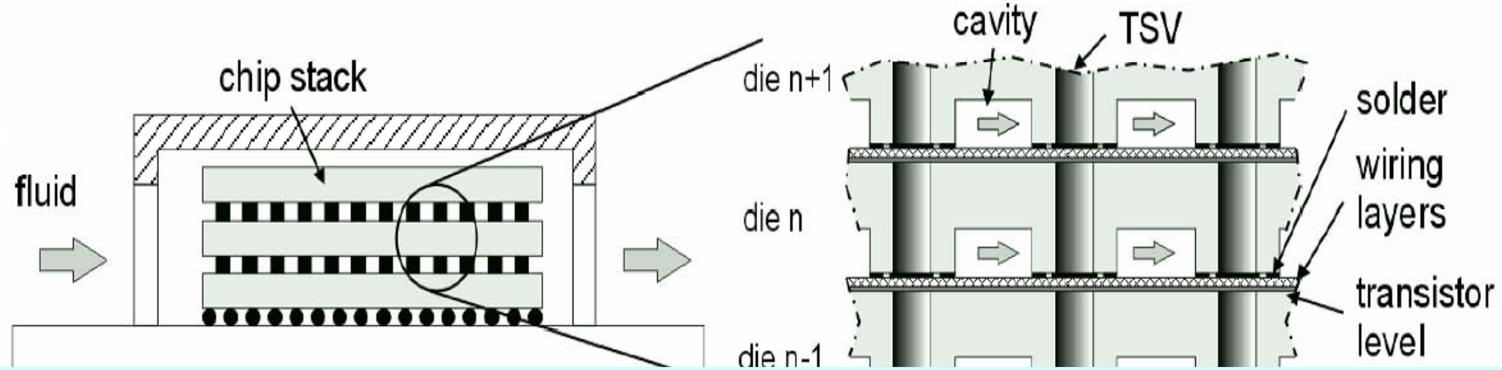


System Level Active Cooling Manager

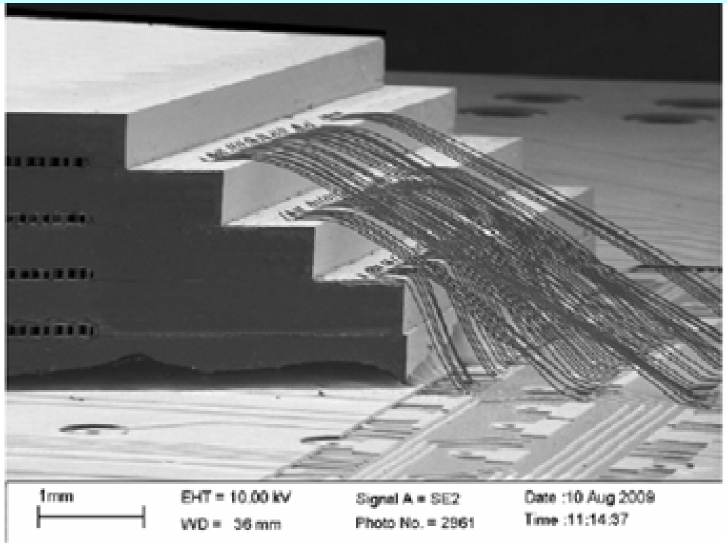
(3D heat flow prediction)

CMOSAIC: System-Level Control of Heat and Liquid Cooling Effects in 3D Stacks

- Combining liquid control, task scheduling and task migration



Initial energy savings of 90% at chip level with air-cooling techniques



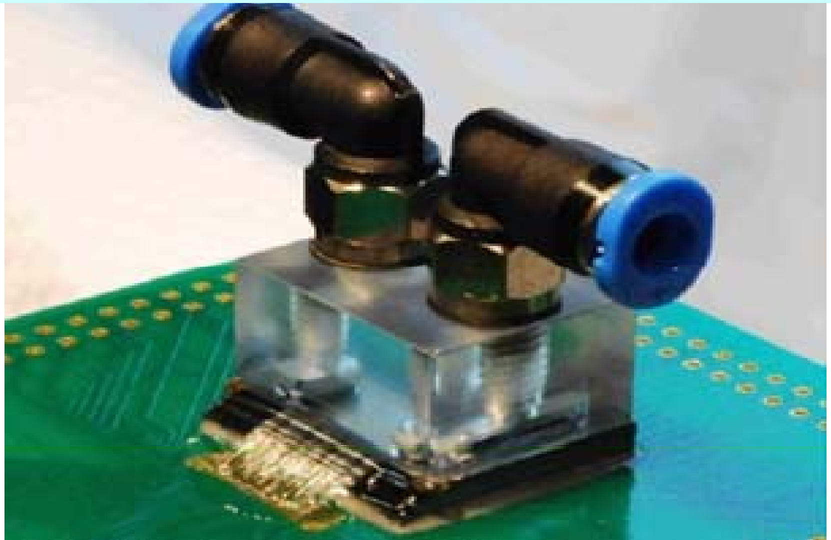
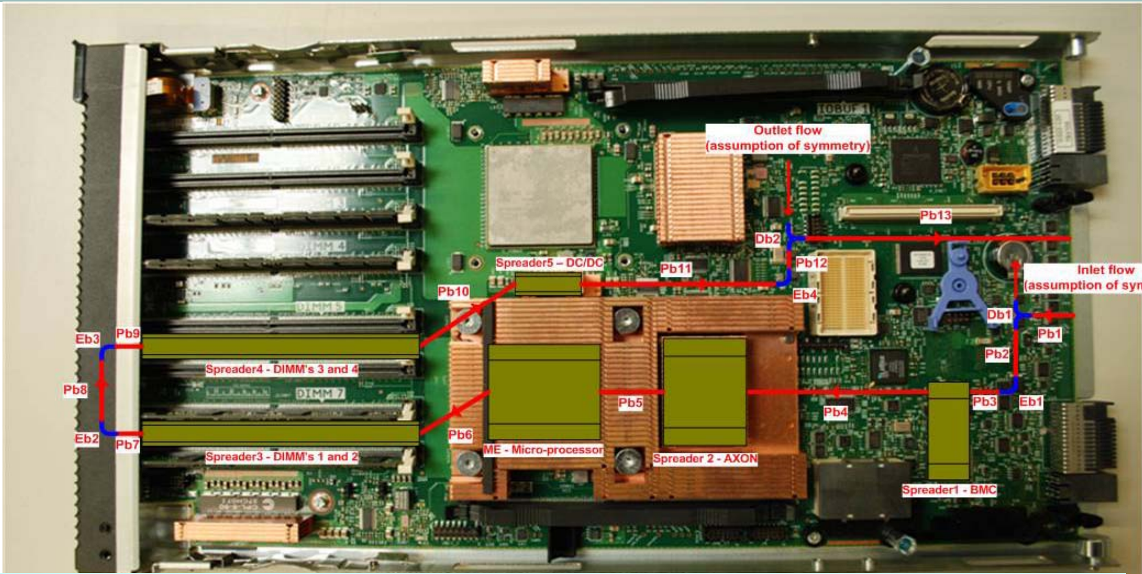


Figure: IBM Zürich & EPFL

CMOSAIC: Conversion of Air-Cooled to Two-Phase Cooled Blade and Datacenter

Integrated micro-channel cold plate in single chip



Initial energy savings in IT infrastructure of datacenters of more than 30% with respect to air-cooling techniques (no chillers needed)

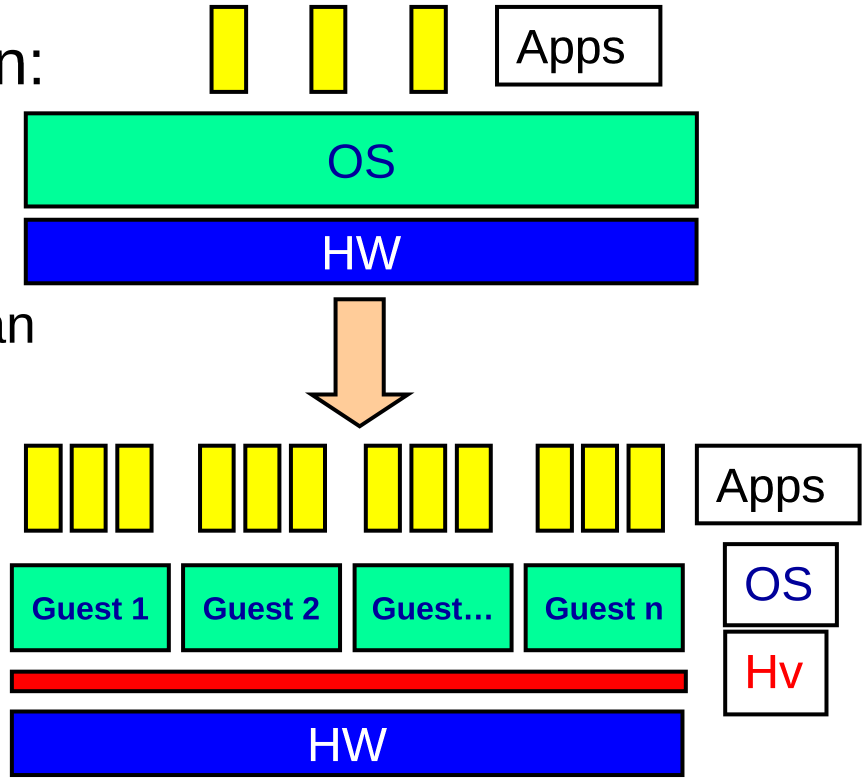
Rack-level and 2-phase datacenter tests

[Source: Prof. J. R. Thome, EPFL]

Energy-Centric IT: Virtualization and Zero-Emission Datacenter Together

- Virtualization increases utilization:
  - Multiplexing a number of existing servers onto a single physical machine
  - Mission-critical software systems can have performance isolation

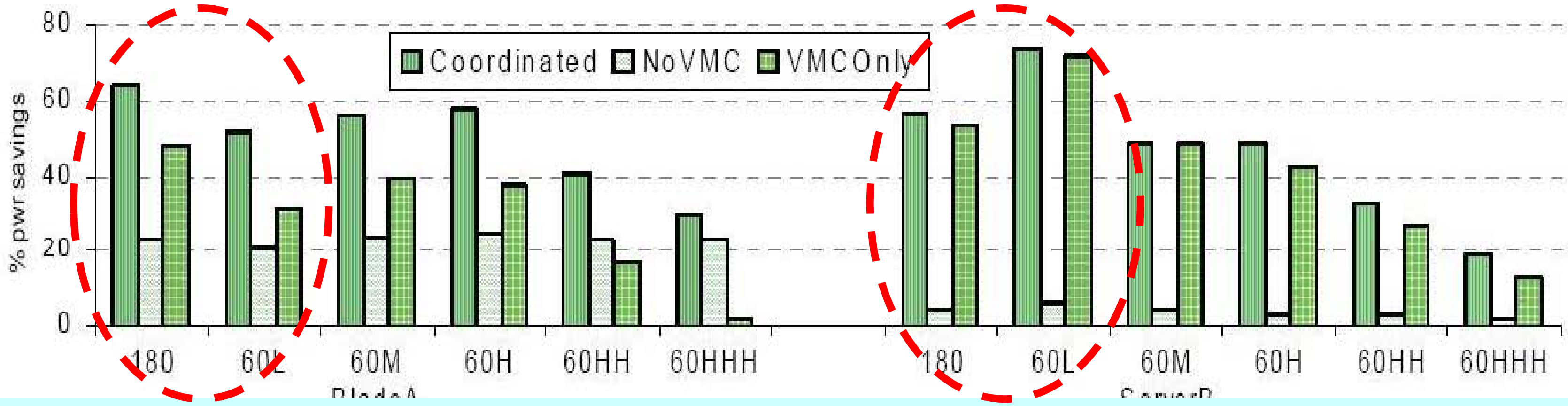
- SW-based monitoring and management: Integrated hypervisor/OS (*Hv*)
  - Distributed energy cost/task monitoring and management (including cooling and execution costs)



Exploitation of virtualization for local and global energy control

Energy-Centric IT: Virtualization and Zero-Emission Datacenter – Results

- 180 real-world traces: L -> Low workload ; H->High workload
- Two types of servers/blades (Blade A has 5 power settings and Server B has 6 power settings)
- **Global energy control: higher potential gains**
- **Higher power savings: low utilization workloads**

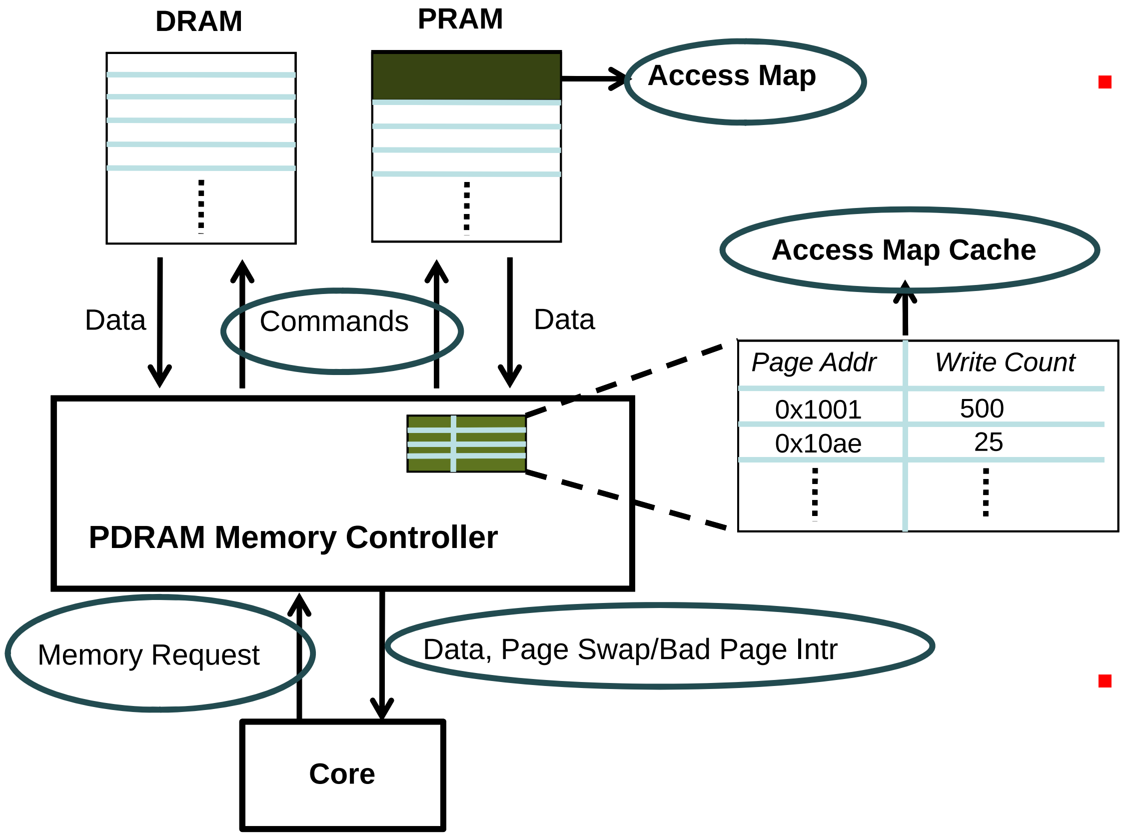


Good adaptation: No work done, no energy spent!

Energy-Efficient Storage for Datacenters

- Shift from hard disk drive (HDD) storage to solid state disks or memories (SSD)

New scalable storage managers and data layout schemes needed

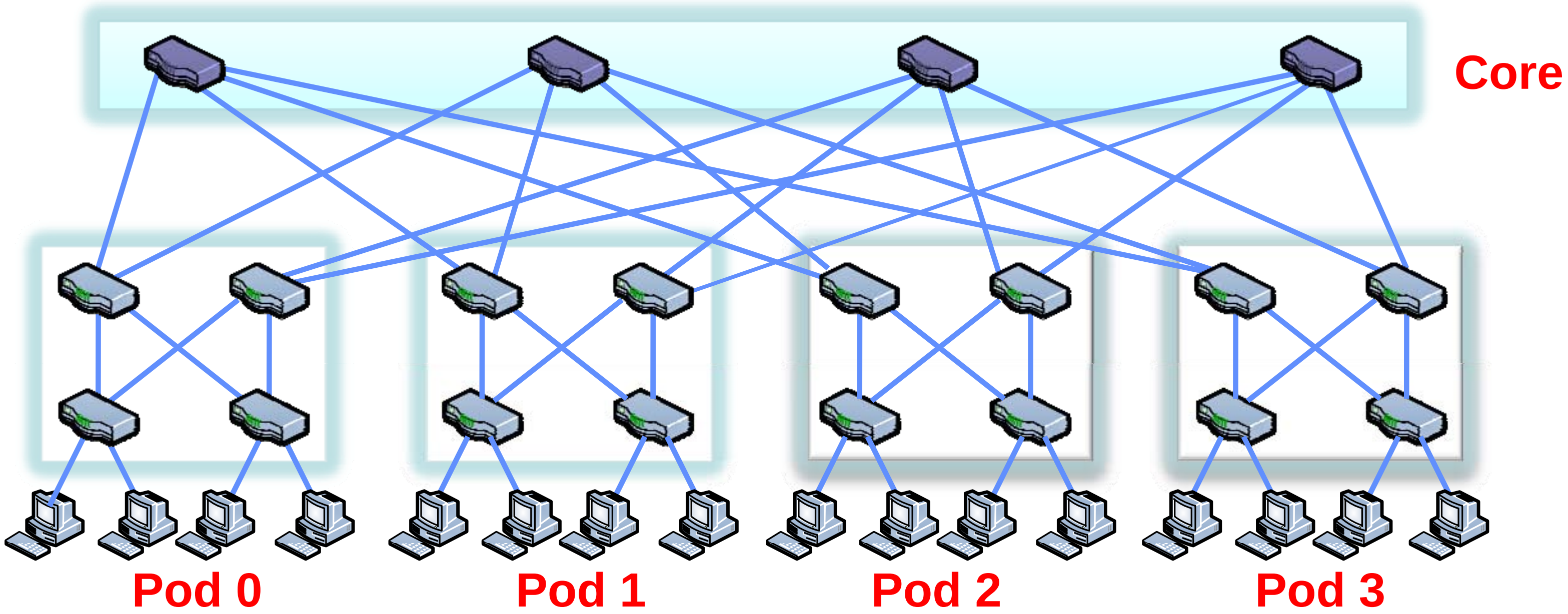


- Benefits:
  - New and fast low- power states
  - Higher energy efficiency
  - Lower latency to storage
- Challenges:
  - HW/SW interdependence

[Source: Prof. T. Simunic Rosing, UCSD]

Energy-Efficient Network for Datacenters

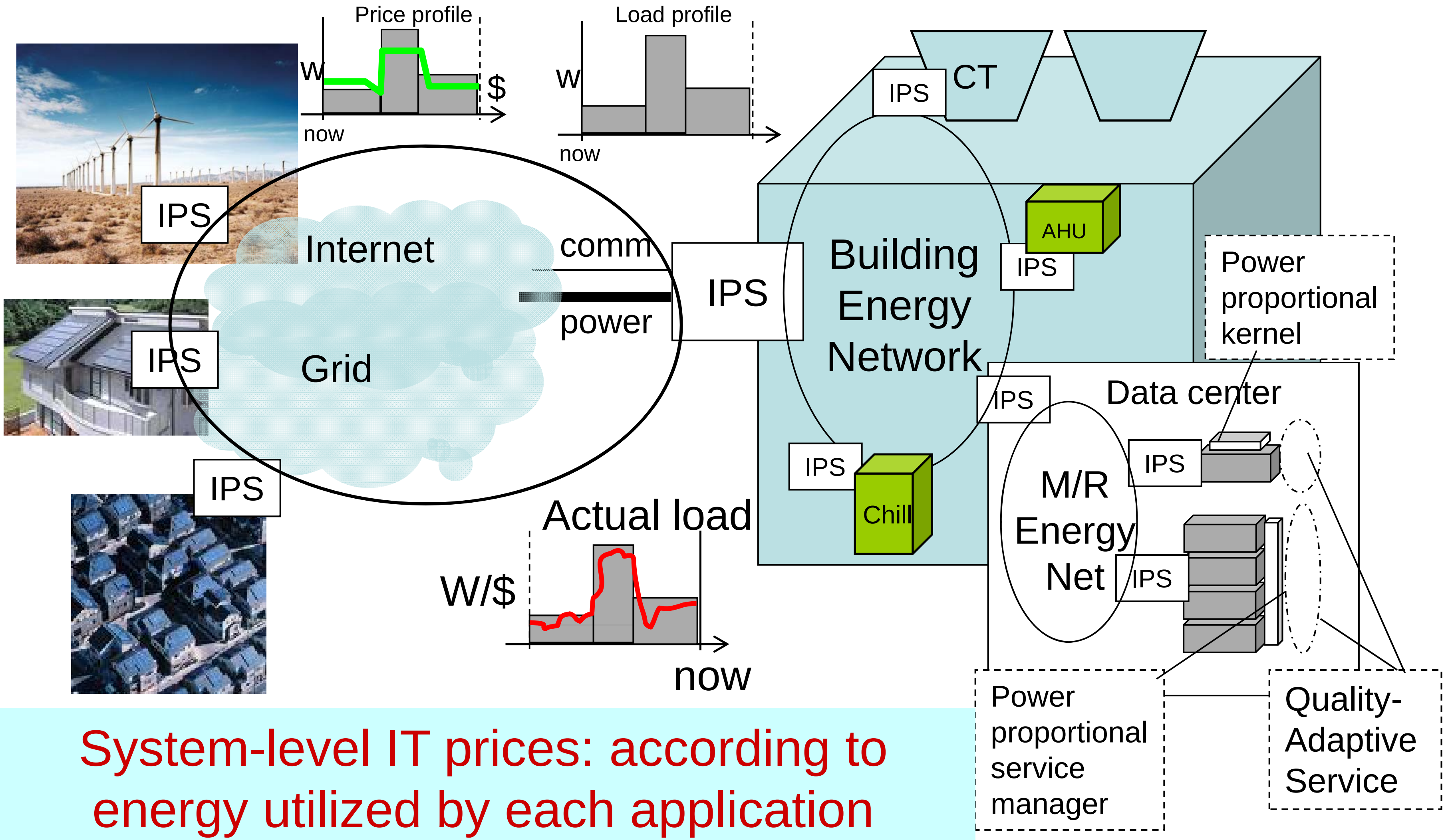
- Develop power-efficient network components: switches, routers and network processors with power-saving modes
- New end-to-end protocols to build energy-balanced systems
  - Scheduling algorithms to leverage path diversity
  - Dynamic energy management; optics for energy-efficient networks



[Source: Prof. D. Kostic, EPFL]

Long-Term Goal: Energy-Proportional IT Services in the Future

[Source: Prof. T. Simunic Rosing, UCSD]



System-level IT prices: according to energy utilized by each application

[Source: Prof. T. Simunic Rosing, UCSD]

EcoCloud Center @ EPFL
(soon to come @ ecocloud.ch)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Collaboration of multiple laboratories and industrial affiliates to reach long-term goal
- Research lines covered:
  - Energy-proportional data processing, communication and storage
  - Scalable cloud applications and services
  - Vertically-integrated computing and cooling

Making tomorrow's clouds and IT services green and sustainable

Conclusions

- Green (or energy-centric) computing is necessary due to worldwide environmental impact of ICT technologies
- System-level management of available energy is required in order to achieve cost efficient IT services
  - New IT costs metrics: IT services per joule per CHF/EUR/$
- Cross-layer optimization applies to both small- and large-scale computing systems:
  - Execution cost and cooling costs are key for energy-efficient systems
  - Combination of architectures and algorithms for system-level design
  - Inter-disciplinary design and optimization approaches required
- The overall goal must be reaching energy-proportional IT services in the “Green IT” future

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Merci

Thank You