IBM Research – Zurich

Hotwater Cooling for Energy-Hungry Datacenters

Dr. G. Ingmar Meijer
IBM Research – Zurich
inm@zurich.ibm.com

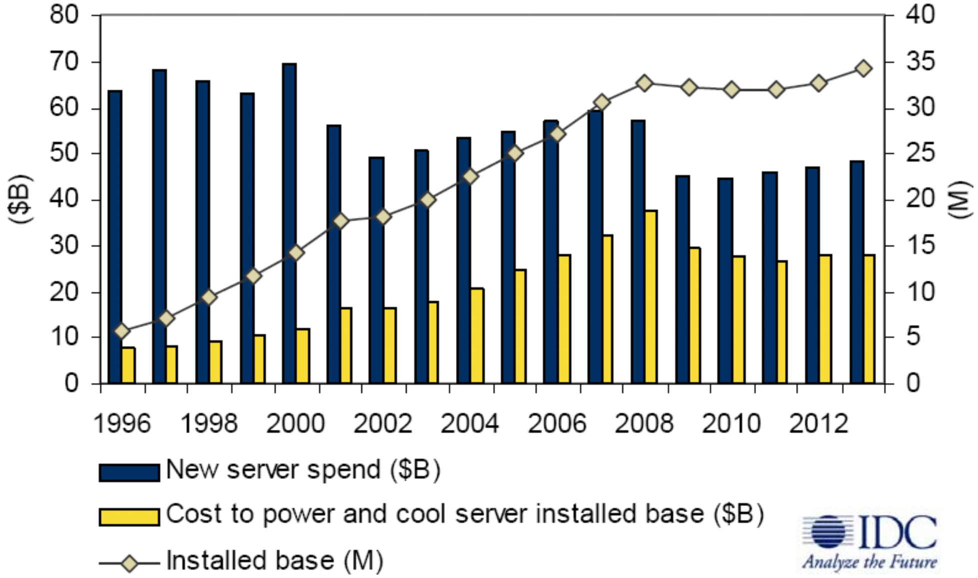
Towards a Zero – Emission Datacenter

Outline

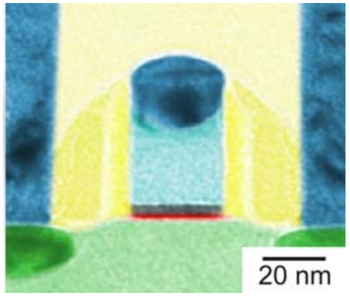
1. Introduction
   A. Datacenter Market Drivers and Trends
   B. Energy Consumption Issues
2. Hotwater Cooling; **Concept**
3. Prototype High-Performance Computer ("First-of-a-Kind")
4. **IBM System x iDataPlex Water Cooled dx360 M4**
   A. Direct Water Cooling
   B. **Power** & **Performance**
5. Summary

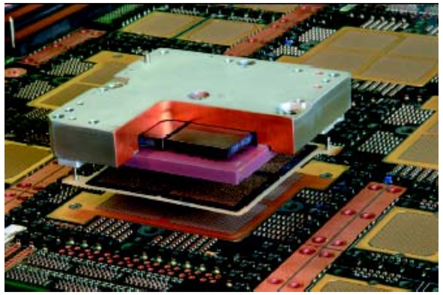
1.A. Datacenter Market Drivers and Trends

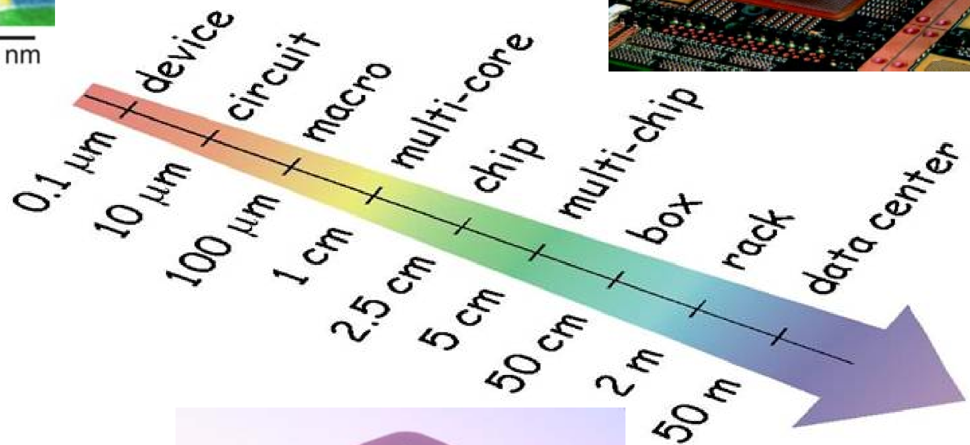
- **Total cost of ownership** and **environmental footprint**
- Servers used 330 TWh of electrical energy: **€25bn** or **2%** of the electricity production (2009).
- ICT industries **carbon emission** as **aviation**.

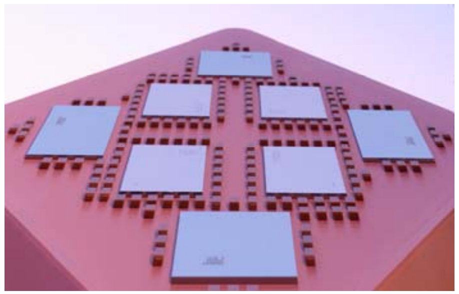


1.B. Energy-Consumption Issues (and Thermal Issues): propagate through hardware levels









Energy Consumption of Transistor (Leakage Current)

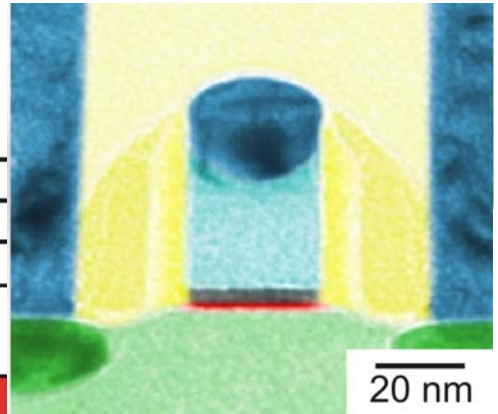
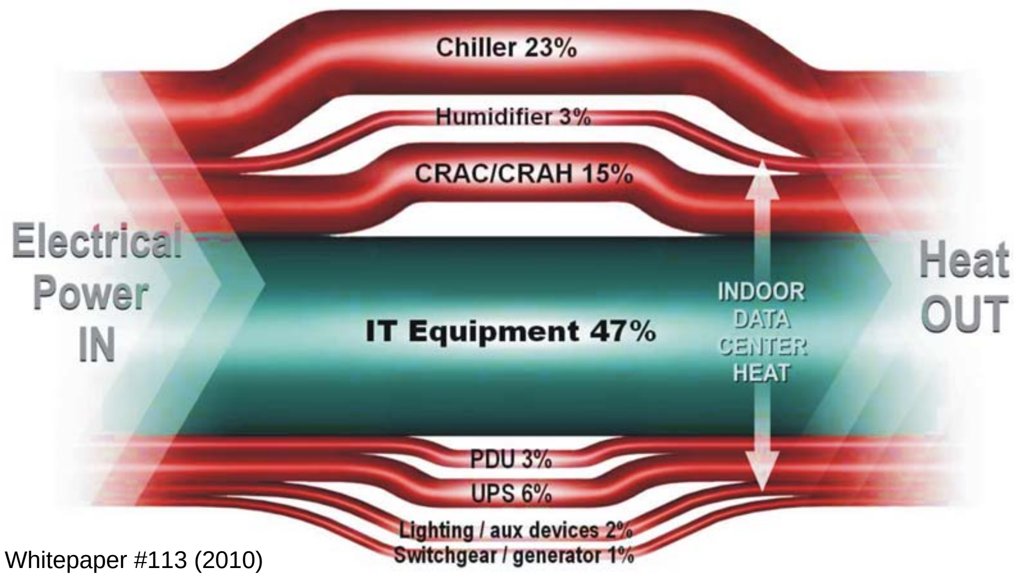


Table PIDS2 High-performance Logic Technology Requirements

Year of Production	*2009*	*2010*	*2011*	*2012*			
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)	*54*	*45*	*38*	*32*			
L_g: Physical Lgate	*29*	*27*	*24*	*22*			
Equivalent Oxide Thickness (nm)							
Extended planar bulk	**1**	**0.95**	**0.88**	**0.75**			
UTB FD					**0.7**	**0.68**	**0.6**
MG							**0.77**
Maximum gate leakage current density (kA/cm^2)							
Extended Planar Bulk	**0.65**	**0.83**	**0.9**	**1**	**1.1**	**1.2**	**1.3**
UTB FD					**1.1**	**1.2**	**1.3**
MG							**1.3**
V_{dd}: Power Supply Voltage (V) [8]							
P bulk/UTB FD/MG	**1**	**0.97**	**0.93**	**0.9**	**0.87**	**0.84**	**0.81**

International Technology Roadmap for Semiconductors

Energy Consumption of Datacenter



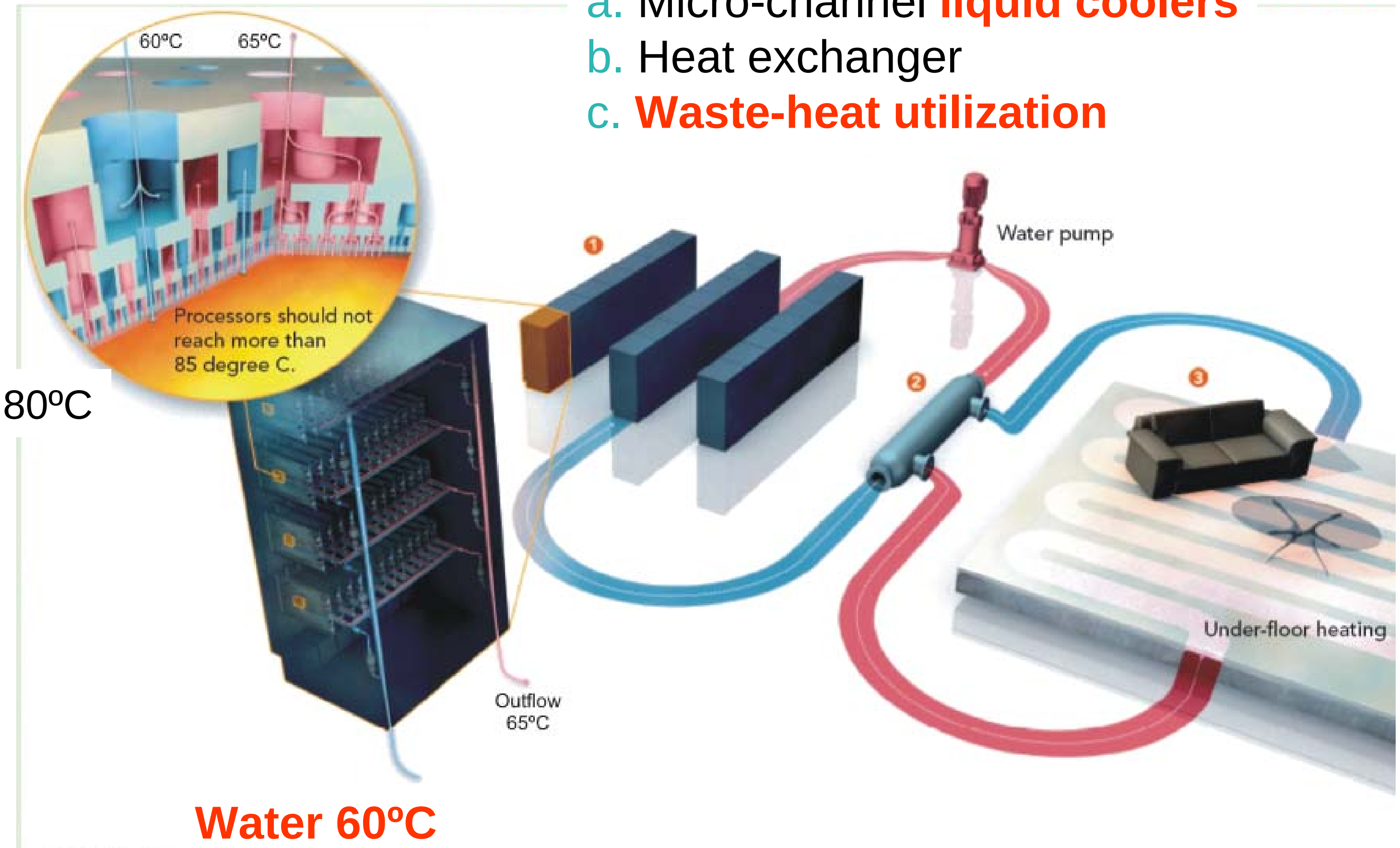
APC, Whitepaper #113 (2010)

- **Air-cooled datacenters** are **inefficient**. Typical cooling needs as much energy as IT equipment and both are thrown-away.
- Provocative: datacenter is a huge **"heater with integrated logic."**
- For a 10 MW datacenter €2m–€3m is wasted per year.

2. Zero – Emission Datacenter Concept

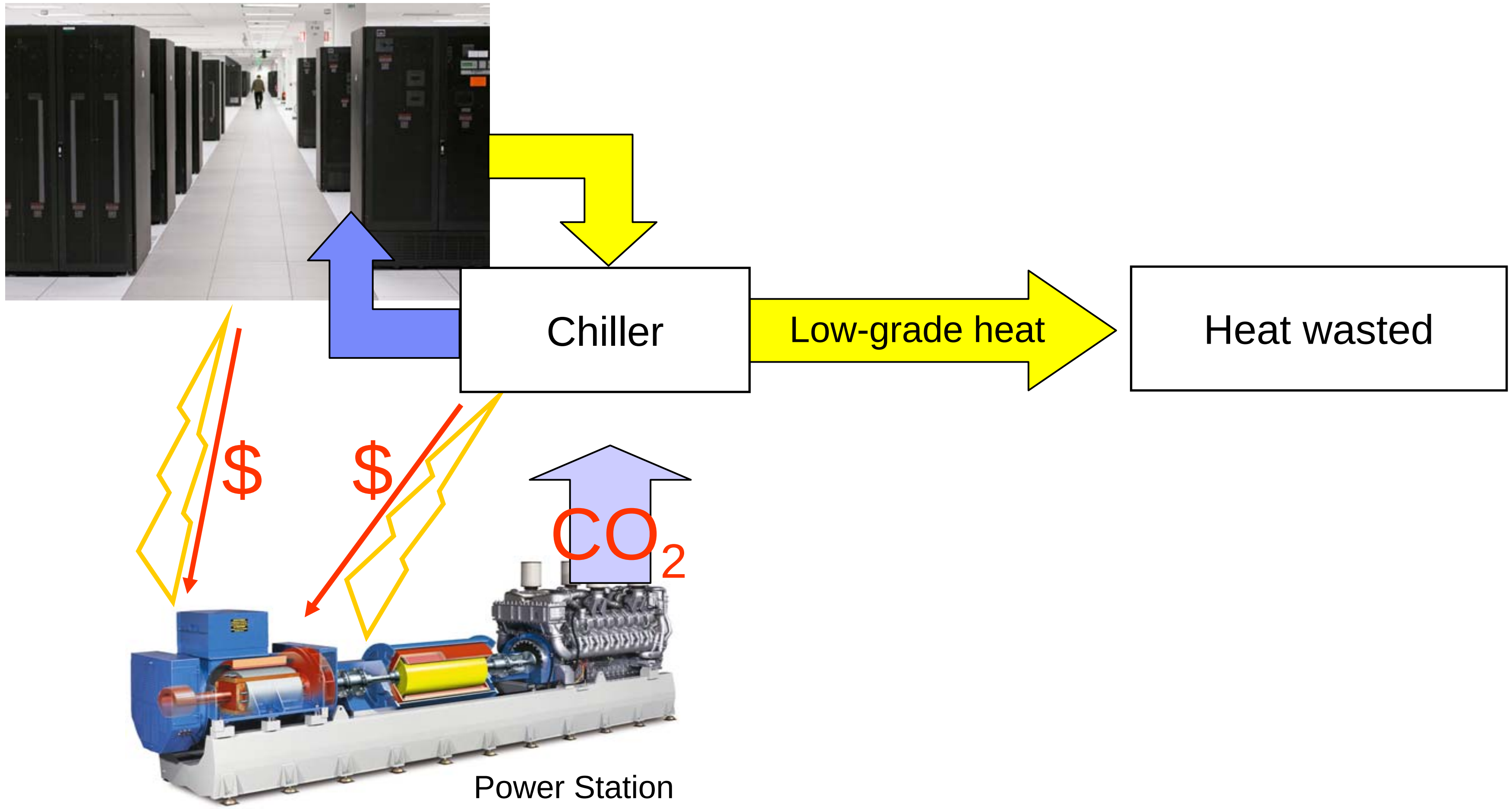
a. Micro-channel **liquid coolers**
b. Heat exchanger
c. **Waste-heat utilization**

CMOS 80ºC

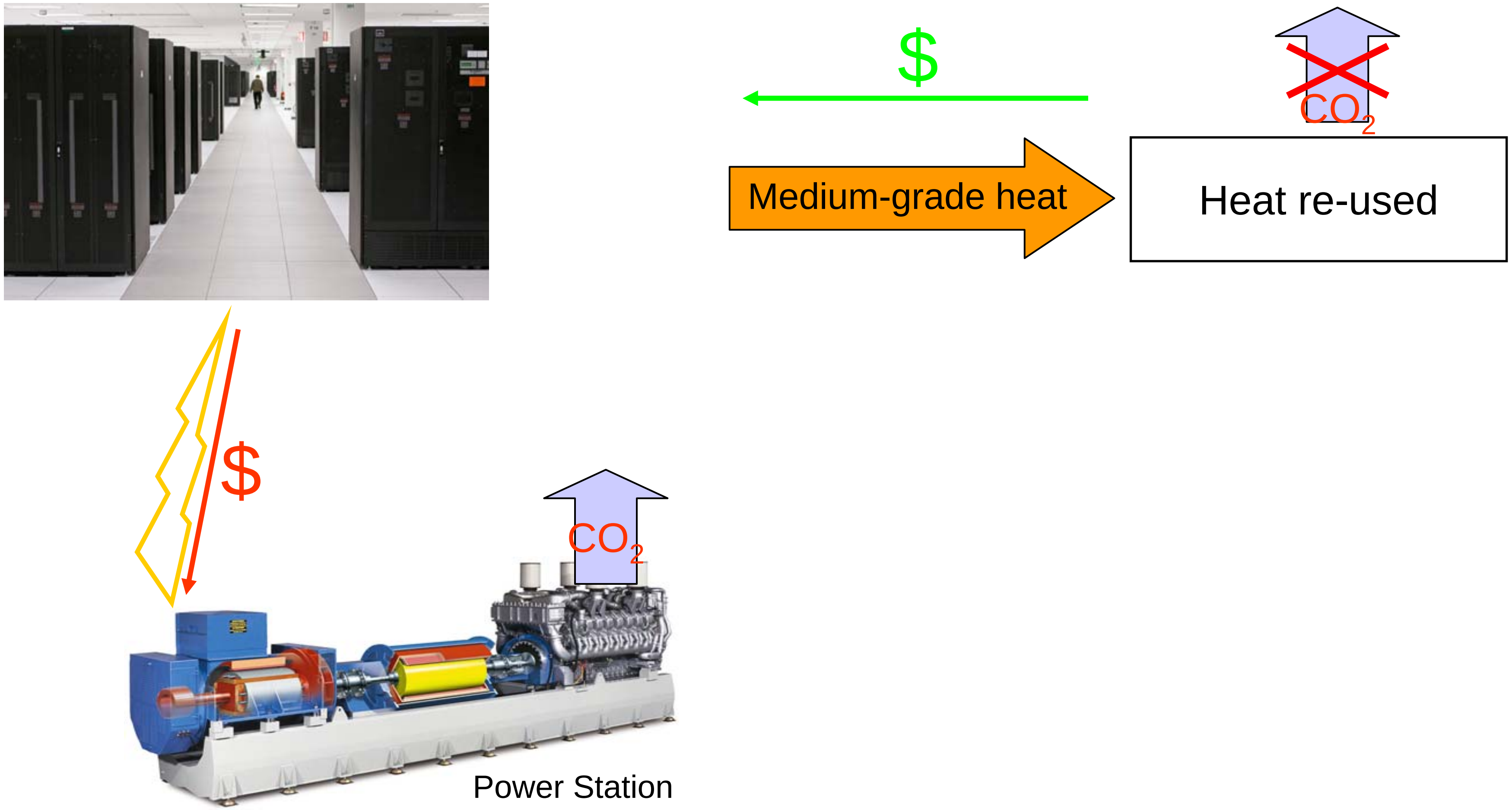


Water 60ºC

Energy and Emission of Conventional Datacenter



Zero – Emission Datacenter (CO_2 < 15 % net)



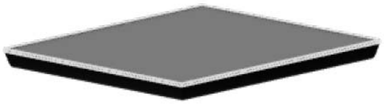
Comparison: Water Cooling vs. Air Cooling

Water Cooling:

1. High heat capacity

$c_V \approx 1$ Wh/(L·K)

2. Low thermal resistance



$\Delta T = R_{th} \cdot \dot{q}''$

$R_{th} = 0.1 \text{ K cm}^2/\text{W}$

$\dot{q}'' = 50 - 100 \text{ W/cm}^2$

$\Delta T \sim 5 - 10$ K

Air Cooling:

1. Low heat capacity

$c_V \approx 0.0003$ Wh/(L·K)

2. High thermal resistance

$\Delta T = R_{th} \cdot \dot{q}''$

$R_{th} = 1 \text{ K cm}^2/\text{W}$

$\dot{q}'' = 50 - 100 \text{ W/cm}^2$

$\Delta T \sim 50 - 100$ K

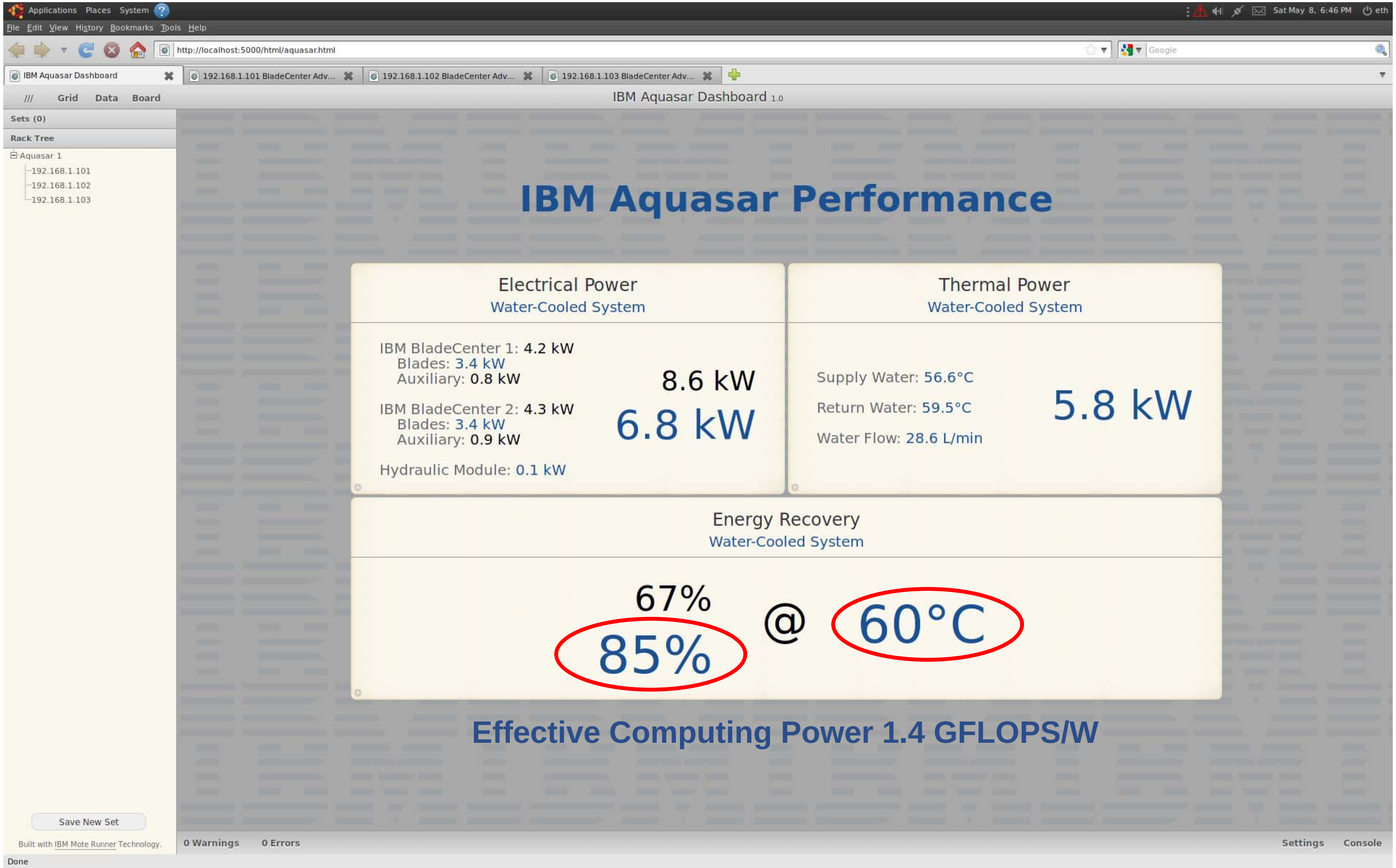
3. Zero – Emission Datacenter Prototype HPC Cluster at ETH Zurich, "Aquasar"

Energy consumption reduced by up to 40%

Direct use of waste heat cuts CO_2 emissions by up to 85%

World record in facility-level efficiency in TFLOP/gCO_2 7.9

Water-Cooled IBM BladeCenter QS22


Applications Places System
Sat May 8, 6:46 PM
File Edit View History Bookmarks Tools Help
http://localhost:5000/html/aquasar.html
IBM Aquasar Dashboard
192.168.1.101 BladeCenter Adv...
192.168.1.102 BladeCenter Adv...
192.168.1.103 BladeCenter Adv...
Grid Data Board
IBM Aquasar Dashboard 1.0
Sets (0)
Rack Tree
Aquasar 1
192.168.1.101
192.168.1.102
192.168.1.103
IBM Aquasar Performance
Electrical Power
Water-Cooled System
IBM BladeCenter 1: 4.2 kW
Blades: 3.4 kW
Auxiliary: 0.8 kW
IBM BladeCenter 2: 4.3 kW
Blades: 3.4 kW
Auxiliary: 0.9 kW
Hydraulic Module: 0.1 kW
8.6 kW
6.8 kW
Thermal Power
Water-Cooled System
Supply Water: 56.6°C
Return Water: 59.5°C
Water Flow: 28.6 L/min
5.8 kW
Energy Recovery
Water-Cooled System
67%
85%
@
60°C
Effective Computing Power 1.4 GFLOPS/W
Save New Set
Built with IBM Mote Runner Technology.
0 Warnings
0 Errors
Settings
Console
Done

4. Leibniz Rechenzentrum, Garching, Germany
SuperMUC: Warm-Water Cooled 3 PFLOPS System

1Q12—2Q12: ~10000 IBM System x iDataPlex Water Cooled dx360 M4

Hybrid Datacenter w/ Direct Water Cooled Nodes

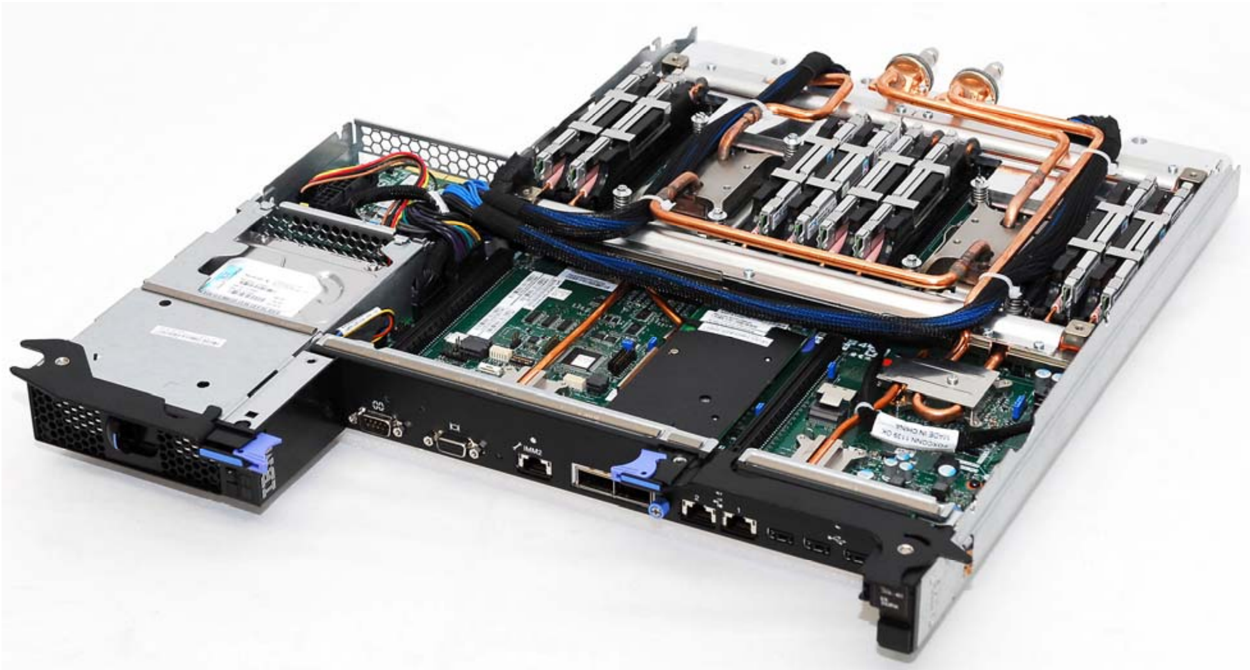
- Highly energy-efficient **hybrid-cooling** solution:
  - Compute racks
    - 90% Heat flux to warm water
    - 10% Heat flux to CRAH
  - Switch / Storage racks
    - Rear door heat exchangers

- Compute node **power consumption reduced ~ 10%** due to lower component temperatures and no fans.

- Power Usage Effectiveness P_{Total} / P_{IT}: **PUE ~ 1.1**

- **Heat recovery** is enabled by the compute node design:
  Energy Reuse Effectiveness $(P_{Total} – P_{Reuse}) / P_{IT}$: **ERE ~ 0.3**

IBM System x iDataPlex Water Cooled dx360 M4



- Heat flux > 90% to water; very low chilled water requirement.
- Power advantage over air-cooled node: warm water cooled ~10% (cold water cooled ~15%) due to lower $T_{components}$ and no fans.
- Typical operating conditions: T_{air} = 25 – 35°C, T_{water} = 18 – 45°C.

5. Main Messages / Outlook

- Technology exists to design Zero – Emission Datacenter:
  - **Hot-water cooling & Waste heat utilization**
- Business value / opportunity
  - **Energy cost reduction**
  - **Performance increase**
  - Low-carbon pioneers **"world-record TFLOP/gCO_2"**
- What is next?
  IBM System x iDataPlex Water Cooled dx360 M4

- **Supercomputing SC11 Seattle, November 12-18, 2011:**
  - Node at IBM booth

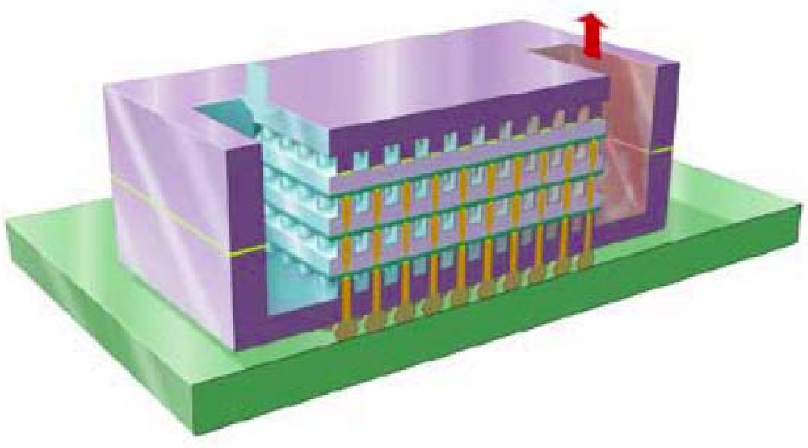
IBM Team

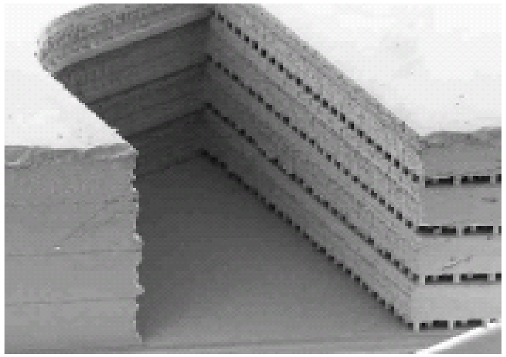
- **Aquasar**
  - IBM Research – Zurich
  - IBM Germany Research & Development
- **IBM System x iDataPlex Water Cooled dx360 M4 and SuperMUC**
  - IBM Systems & Technology Group
  - IBM Research – Zurich
  - IBM Germany

Roadmap: 3D CMOS integration and Heat Removal

3D CMOS integration requires (scalable) **interlayer liquid cooling**

- Challenge: isolate electrical interconnects from liquid
- Experimental heat removal on test vehicle 180 W/cm² per layer (extrapolated **7.2 kW from 10 layers with 4 cm²**)





ENGINEERING

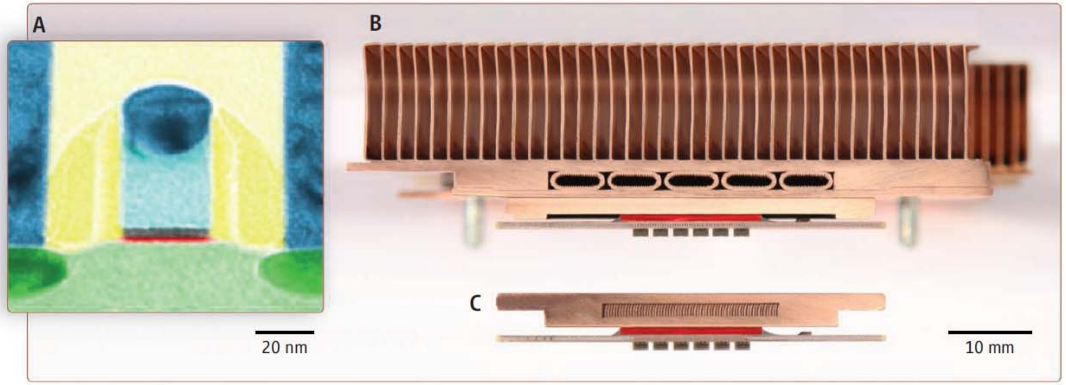
Cooling Energy-Hungry Data Centers

The information technology industry is focusing on approaches to hot-water cooling for the design of energy-efficient data centers.

G. I. Meijer

Soon after the Internet took off in the mid-1990s, enterprise computing infrastructures with warehouses full of servers, known as data centers, became commonplace. The energy consumption challenges posed by such data centers are considerable. The power dissipation of servers has to be managed skillfully. Perhaps surprisingly, the power consumption of the cooling infrastructure that is required to keep the microelectronic components from overheating is on a par with that of the servers themselves. sistors, leakage currents consume more power than the actual computational processes. To alleviate this burden, new materials were introduced in the late 2000s. Most notably, the replacement of the SiO_2 gate oxide, which is only a few atomic layers thick, with a physically thicker layer of a hafnium-based oxide enabled an appreciable reduction of the gate tunneling currents while maintaining the electrical performance of the transistor (5, 6). Nevertheless, keeping the gate leakage power per unit area below 1 kW cm^{-2} will remain a par- the electricity consumption currently goes toward powering this cooling infrastructure. It therefore appears ironic that Moore's law (a projection of microprocessor performance) is widely known and often cited, while increasingly critical physical laws of thermodynamics receive little popular attention.

The first law of thermodynamics states that energy is conserved. The electrical energy that is supplied to the computer system is eventually entirely converted into thermal energy. The standard method to remove this



Heat generation and cooling in the data center. (**A**) Transmission electron microcopy image of a transistor. Heat is primarily generated near the gate oxide (marked red). (**B**) Cross-section of microprocessor module with air-cooled heat sink. Area in red indicates location of microprocessor. (**C**) Cross section of microprocessor module with liquid-cooled microchannel heat sink.

from regulators (2, 3).

The information-technology industry therefore needs efficacious concepts to reduce the energy consumption of data centers. The key culprits are the server microprocessors, or more precisely, the transistors inside these microprocessors (see the figure, panel A). Currently, transistors with 45-nm lateral features are in volume production, and the pace of miniaturization continues unabated (4). It is a formidable challenge to keep the power dissipation of these transistors within acceptable limits. With shrinking dimensions of the tran-

IBM Research–Zurich, 8803 Rüschlikon, Switzerland. E-mail: inm@zurich.ibm.com

energy consumption and carbon footprint?

For the newest members of microprocessor families, sophisticated circuit architectures have been introduced, which allow the power associated with computational processes and also the leakage power to be adapted (7, 8). The microprocessor frequency can be adjusted and circuit blocks can be temporarily powered down completely when not in use. These innovations lead to energy savings for a computational load that comes in bursts or that is bound to memory latency or input/output operations.

An alternative approach is to tackle the problem at the cooling infrastructure (9). Approximately 50% (industry average) of energy consumption of a computing facility is liquid cooling. The reason is that thermodynamically liquid cooling is much more efficient than air cooling because the heat capacity of liquids is orders of magnitude larger than that of air (for example, for water it is 4 MJ m^{-3} K^{-1} versus 1 kJ m^{-3} K^{-1} for air). Once the heat has been transferred to the liquid, it can be handled very efficiently. Critics of liquid cooling might contend that it comes at the price of increased mechanical complexity. True, but this can be managed as computers were once equipped with liquid cooling when the power density of bipolar-transistor–based computer systems reached its peak during the 1980s. For example, the

318 16 APRIL 2010 VOL 328 SCIENCE www.sciencemag.org
Published by AAAS