Thermal Management

-Servers and Data Centers-

Roger Schmidt, IBM Fellow
Chief Engineer - Data Center Energy Efficiency
Module Level Cooling

- Heat Sink
- External Interface (TIM2)
- Lid or Cap
- Internal Interface (TIM1)
- Substrate
- Chip

External
Internal
Heat Flow Path

- Transistor
- Chip
- Module
- Card
- Drawer
- Rack
- Room
- Atmosphere

Source: Hendrick Hamann, IBM Research
Incredible performance improvements

As the number of transistors goes UP

Energy per Transistor goes DOWN

~ 1 Million Factor Reduction In Energy/Transistor Over 30+ Years
Moore's Law: Delivering DC Performance

2004
- 79 TFlops
- 118 racks
- 10,000 ft²
- 3.5 MW

2008
- 79 TFlops
- 9 racks
- 300 ft²
- 0.5 MW
## Supercomputer Trends

<table>
<thead>
<tr>
<th>Year</th>
<th>Computer</th>
<th>Cores</th>
<th>GFlops</th>
<th>Power (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>Numerical Wind Tunnel(U.S.)</td>
<td>140</td>
<td>235</td>
<td>0</td>
</tr>
<tr>
<td>1995</td>
<td>CP-Pacs/2048(Japan)</td>
<td>2048</td>
<td>614</td>
<td>0</td>
</tr>
<tr>
<td>1997</td>
<td>ACSI Red(U.S.)</td>
<td>9152</td>
<td>1830</td>
<td>850</td>
</tr>
<tr>
<td>1999</td>
<td>ACSI Red(U.S.)</td>
<td>9632</td>
<td>3207</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>ACSI White(U.S.)</td>
<td>8192</td>
<td>12298</td>
<td>0</td>
</tr>
<tr>
<td>2002</td>
<td>Earth Simulator(Japan)</td>
<td>5120</td>
<td>40960</td>
<td>1.2</td>
</tr>
<tr>
<td>2004</td>
<td>Bluegene/L(U.S.)</td>
<td>32768</td>
<td>91750</td>
<td>0</td>
</tr>
<tr>
<td>2005</td>
<td>Bluegene/L(U.S.)</td>
<td>131072</td>
<td>367,000</td>
<td>0</td>
</tr>
<tr>
<td>2007</td>
<td>Bluegene/L(U.S.)</td>
<td>212992</td>
<td>596,378</td>
<td>0</td>
</tr>
<tr>
<td>2008</td>
<td>Roadrunner(U.S.)</td>
<td>129600</td>
<td>1,456,704</td>
<td>2.35</td>
</tr>
<tr>
<td>2009</td>
<td>Jaguar(U.S.)</td>
<td>224162</td>
<td>2,331,000</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>Tianhe-1A(China)</td>
<td>186368</td>
<td>4,701,000</td>
<td>4</td>
</tr>
<tr>
<td>2011</td>
<td>&quot;K&quot; computer(Japan)</td>
<td>548,352</td>
<td>10,510,000</td>
<td>12.6</td>
</tr>
<tr>
<td>2012</td>
<td>Titan(U.S.)</td>
<td>560640</td>
<td>17,590,000</td>
<td>8.2</td>
</tr>
<tr>
<td>2013</td>
<td>Tianhe-2 (China)</td>
<td>3120000</td>
<td>33,862,700</td>
<td>17.8</td>
</tr>
</tbody>
</table>

A 1 exaflop (1,000,000,000,000 GFlops) may require 100’s of MW to power.
Control Volume for IT Equipment

Maximize

Bits/Bytes

Heat(Q)

Useful energy(?)

Minimize

Power(kva)

Cooling(air/water)
A Data Center Crisis...or Opportunity?

Projected Data Center Energy Use Under Five Scenarios

- 2.9% of projected total U.S. electricity use
- 1.5% of total U.S. electricity usage
- 0.8% of total US electricity usage

Koomey, 2011, estimates 36% growth
Typical Data Center Energy End Use

- File server link utilization (daytime) (Bennett, 2006)

Data Center Energy Breakdown

- IT Equipment 50.0%
- Cooling 33.3%
- Power 16.7%
- X86 average utilization 5 – 12%

100 Units
35 Units
30 Units

File Server Bandwidth Utilization Profile

File Server Bandwidth Utilization Profile
IT Infrastructure

Bits/Bytes → Power(kVA) → Cooling

Heat(Q) → Power(kVA) → Cooling

Data Center Energy Breakdown
- IT Equipment: 52.2%
- Power: 16.7%
- Cooling: 33.3%
- Building chilled water loop: 16.7%
- CRAC & server fans: 16.7%

13.8 kV

480V AC

Battery cabinet(s)

UPS

AC/DC

DC/AC

Chrg

Bypass

208V AC

PDU

380-410V DC

AC/DC

PSU

DC/DC

VR

12V

Server

Rack

AC/DC

DC/AC

AC/DC

DC/DC

AC/DC

IT Equipment

Power

Cooling

Building chilled water loop

CRAC & server fans

Bypass

480V AC

13.8 kV

Chrg

B

P

C

P

B

B

Pumping power

Pumping power

Compressor power

Pumping power

Pumping power

Electrical power

NSF/IURUC

Annual Meeting

Roger Schmidt

Jan, 2014
Typical Data Centers – what really happens

Most of today’s existing data centers attempt to cool the IT equipment by flooding the air supply with as much cool air as possible. Precision air flow, as opposed to flooding the space, reduces costly as well as unnecessary of cooling air and the power to produce it.
Some Big Issues

✓ Energy efficiency at all levels of power distribution
✓ Air cooling higher powered processors
✓ Air cooling high rack powers in a DC
✓ Enhanced and cost effective liquid cooling
✓ Use of waste heat
✓ Maintaining chip temperatures for high powered chips
✓ Improved interface materials
✓ Achieving exascale computing levels within power envelopes
✓ Air moving devices for improved performance/reduced noise
Thank you for your attention