Packaging of New Servers
- energy efficiency aspects-

Roger Schmidt, IBM Fellow
Chief Engineer, Data Center Energy Efficiency

Many slides provided by Ed Seminaro
Chief Architect of HPC Power6 & 7

1st Berkeley Symposium on Energy Efficient Electronics, June 11 -12, 2009
Agenda

✓ Liquid cooling
✓ Power6 packaging
✓ Power7 packaging
✓ Data Center issues
Back to the future - Rear Door Heat Exchanger

Leveraging Mainframe Heritage

Capabilities

- A water-cooled, heat exchanger designed to remove heat generated from the back of your computer systems before it enters the room
- Patented "rear heat exchanger" can reduce heat at the source
- Utilizes chilled water to dissipate heat generated by computer systems while requiring no additional fans or electricity!

Customer Benefits

- Increase server density without increasing cooling requirements.
- Effective solution for a data center wishing to deal with computer room “hot spots,” or is at the limit of its cooling capacity, but still has useable floor space to add racks of systems.

Water cooling is 3500x more energy efficient than air

IBM uses it today in System x, iDataplex and System p
Processor Liquid Cooling for Increased Performance

System z10 EC is cost effective and can help you Go Green by delivering highly energy efficient technologies

**Customer Benefits**
- Highest server utilization
- High performance per watt
- Extremely efficient power and cooling

Power 575 turns on the water for Green Supercomputer

**Customer Benefits**
- 80% heat load to water
- 40% energy savings via improved cooling and power efficiencies
- Up to 8. trillion floating point operations per second + 3.58 TB of memory in a single rack
Today’s IBM Mainframe – z10EC

- Processor Books, Memory, MBA and HCA cards
- Ethernet cables for internal System LAN connecting Flexible Service Processor (FSP) cage controller cards
- InfiniBand I/O Interconnects
- 2 x Cooling Units
- 3x I/O cages
- 2 x Support Elements
- Power Supplies
- Internal Batteries (optional)
System Heat Load Trends 1977 - present

Bipolar

CMOS

Heat Load - kW

Water
Air

Refrigeration

3033 3081 3090 3090EX ES9000 G2 G3 G4 G5 G6 z900 z990 29 30 40 50 60 70 80 90

Heat Load Trends 1977 - present

CMOS

Bipolar

Refrigeration

Water
Air
IBM Computer Platforms

Focus of Talk = System-p

- System-x Servers
  - DS3K/4K Midsize Storage
    - x86 uP & chip set
    - Linux & Windows OS’s
    - Industry Standard Interconnect
    - Key Competitors: HP, Dell, etc.
    - Principal DE site: Raleigh, NC

- System-pi Servers
  - System-z Mainframes
  - DS8K High-End Storage
    - System-pi: “Power uP” based
    - DS8K: “Power uP” based
    - IBM chip set
    - AIX, Linux, i5 OS’s, IBM FW
    - Industry Standard Interconnect
    - Competitors: HP, Sun, Cray, SGI
      - EMC, Hitachi, etc.
    - Principal DE sites: Austin, Tx.
      - MHV, N.Y.
      - Rochester, Mn.
      - Tucson, Az.
    - System-z: “370 uP” based
    - IBM chip set throughout
    - IBM cluster interconnect
    - Key Competitors: HP, Sun, etc.
    - Principal DE site: MHV, N.Y.

- Blue Gene System
  - Cell Blade
    - Special Purpose HPC
    - BG: “Power uP” based
    - Cell: “Power uP & SPU” based
    - Customized Interconnect
    - Competitors: Cray, etc.
    - Principal DE site: Yorktown, N.Y.
    - Rochester, Mn.

- Blue Gene System
  - Cell Blade
    - Special Purpose HPC
    - BG: “Power uP” based
    - Cell: “Power uP & SPU” based
    - Customized Interconnect
    - Competitors: Cray, etc.
    - Principal DE site: Yorktown, N.Y.
    - Rochester, Mn.
## IBM Power Server Systems™

(Shown without lids and heat sinks)

<table>
<thead>
<tr>
<th>P5 Systems</th>
<th>Blade JS2x</th>
<th>520 HV</th>
<th>550 &amp; 560 HV8</th>
<th>570 L4 / ML</th>
<th>575 IH</th>
<th>595 HE</th>
</tr>
</thead>
<tbody>
<tr>
<td>![P5 System Image](p5 systems image)</td>
<td>![Blade JS2x Image](blade js2x image)</td>
<td>![520 HV Image](520 hv image)</td>
<td>![550 &amp; 560 HV8 Image](550 560 hv8 image)</td>
<td>![570 L4 / ML Image](570 l4 ml image)</td>
<td>![575 IH Image](575 ih image)</td>
<td>![595 HE Image](595 he image)</td>
</tr>
</tbody>
</table>

## P6 Systems
(Transition to organic for some systems)

<table>
<thead>
<tr>
<th>P6 Systems</th>
<th>550 HV</th>
<th>570 L4 / ML</th>
<th>575 IH</th>
<th>595 HE</th>
</tr>
</thead>
<tbody>
<tr>
<td>![P6 System Image](p6 systems image)</td>
<td>![550 HV Image](550 hv image)</td>
<td>![570 L4 / ML Image](570 l4 ml image)</td>
<td>![575 IH Image](575 ih image)</td>
<td>![595 HE Image](595 he image)</td>
</tr>
</tbody>
</table>

*P6 titles linked to press announcements*
The Following Slides Describe the P6 Power 575
High Density Mainstream HPC Server

Announced at SC07
Began Shipping to Customers in 1H08

We went all out on the Power-575 to create an Extreme Server that represents the future of Energy efficient Mid-to-Large Scale Technical or Commercial Compute Intense processing Today

- It has a system structure that eliminates large amounts of unnecessary components
- The rack is directly customer chilled water cooled eliminating most customer on-floor cooling requirements
- The processor is directly water cooled, which reduces it operating temperature and therefore power draw
- It packs 2.5x the compute power of a full rack of blade centers into a single cabinet with very flexible I/O
- It combines the advantages of a large 32w SMP with the cost effectiveness of smaller 4w servers

UI NCSA Blue Waters is based on the P7 Follow-on Product to P6/Power-575
P6-p575

121.6 GFLOPs
602 GFLOPs

1.46 TFLOPs

8.42 TFLOPs
**Water Cooled**

32W @ 4.7GHz

**Air Cooled**

32W @ 3.5GHz
Assembled Node
Water Cooled Cold Plate
Secondary Loop

- A Single WCU can Cool all 14 Nodes
- WCU’s are in parallel
- Each WCU has a door control Valve
- Door is removed from loop if temperature can’t be maintained

![Diagram of the Secondary Loop with WCU, Processor Heat Load to Water (51%), RDHX Heat Load to Water (27%), and Air Heat Load to Data Center (22%)](image)
**WCU (Water Conditioning Unit)**

- 4U Tall by Half of a 24” Rack Unit
- Contains Control Valve, Heat Exchanger, Pump, Reservoir, Sensors, Motor Drive & Interface/Control Card
Signal Cable Troughs

Raised Floor

457.2mm (18") Min. Ht.
The Power6 575: How We Designed it for Energy Efficiency

Create an Energy Efficient Solution

- Maximize the Energy Efficiency of the Data Center Power & Cooling Infrastructure

• Power 575’s Power and Cooling Input Enables High Data Center Facility efficiency
  • 480 VAC balanced 3-phase direct power connection – with line cord redundancy
    • PDU step-down losses eliminated
    • Lower building distribution losses and smaller wire sizes
    • In environments that do not require Battery Back-up facility efficiency can be 98% with 480VAC
      • Outstanding Power Line Disturbance Immunity of the Power-575 makes this mode very viable
        • 100% single phase line disturbance immunity
        • Protection against Under-voltages of up to 70%
        • Exceptional protection against Over-Voltage conditions (Line Overshoot, lightning strike, ring)
  • Direct building chilled water connection – with full redundancy
    • Saves on the order of 20% energy over conventional air cooling
    • Eliminates CRAC hardware on the floor
    • Enables tight packing of server racks
      • Without a room temperature penalty
P5-P575  
12 Node / 2 Switch  

~35 KW max  

- Footprint 137 Racks / 100TF (ASCI Purple)  
- 0.73 TFlops / Rack  
- 45 mW / MFlop  
- 100% Thermal Load to Air  
- 6 Node Line Cord Redundancy  

---  

P6-P575  
14 Node / 2 WCU  

~85 KW max  

- Footprint 12 Racks / 100TF (ASCI Purple)  
- 8.4 TFlops / Rack  
- 8 mW / MFlop  
- 80% Thermal Load to H₂O / 20% to Air  
- 14 Node Line Cord Redundancy  

---  

Same size Rack  
< 1/10th space  
< 1/5th power  
~ 1/6th energy  

- (10% Cost Savings minimum)
Customer Benefits
- 80% heat load to water
- 40% energy savings via improved cooling and power efficiencies
Systems on the Horizon

• Sequoia-system developed and manufactured by IBM will be installed at LLNL and will start running in 2012. It will have 1.6 Million Power-Processors and 1.6 Pbytes of main memory leading to a peak performance of more than 20 Pflops

• Blue Waters at U. of Illinois/NCSA - The Blue Waters project will provide a computational system capable of sustained petaflop performance on a range of science and engineering applications. The system is built by IBM on the Power7 processor
Blue Waters will be an unsurpassed national asset, delivering sustained performance of 1 petaflop (that's 1 quadrillion calculations every second) for many science and engineering applications that researchers use every day—not just benchmarks. Blue Waters is being developed and built with a $208 million grant from the National Science Foundation and will be a resource for scientists and engineers across the country.

Blue Waters will deploy IBM's POWER7 architecture, which is based on advanced multicore processor technology.
UI-NCSA Blue Waters Computing System & the Supporting Facilities Infrastructure
Data Center Efficiency

- Minimize Power Consumption and Floor-space of IT Equipment
  - Servers
  - Storage
  - Networking

- Maximize Efficiency of Data Center Power & Cooling Infrastructure
  - Power Conversion & Distribution from 13KV Building Input to Equipment Input
  - Efficiency of Computer Room Heat Removal
    - Power Required to Remove each KW of Heat from the Room
  - High Availability Design of infrastructure
  - Minimize Waste of Floor Space
Potential Future Directions to Improve Data Center Efficiency

- Move to More Integrated Large Scale IT Equipment
  - Integrate the Network into the Server
  - Unify Storage & Servers
  - Eliminate unnecessary Data Flow & Protocol Conversions that Burn Power
  - Integration Reduces IT Equipment hardware content yielding better reliability, lower cost, & potentially better performance.
Potential Future Directions to Improve Data Center Efficiency

- Improve Heat Transfer from Electronic Components in IT Equipment to the Outside Ambient Environment
  - Direct Water Cooling of Electronic Components
  - Transfer of Heat Load from IT Equipment Cabinets to Outside Air via Water with the Highest Possible Temperature (15 deg-C / 60 deg-F target)
  - Eliminate the majority of Room Level Air Cooling Needs, and provide the air conditioning needed via Air-to-Water Heat Exchange in the IT Equipment

- Improve Power Conversion & Distribution Efficiency in the Data Center
  - Distribute Higher Voltages Directly to the IT Equipment Cabinets
    - 400-480 VAC directly from the output of the building transformer
    - 400-600 VDC directly from the output of a UPS without the inverter stage
    - No on Floor Step Down Transformers
    - Assure Low Loss in Power Distribution Cabling .
  - Use efficient building transformers (98.5%) and efficient UPS’s (94% AC output / 97% DC output). Assure efficiency at actual normal operating load.
Potential Future Directions to Improve Data Center Efficiency

- Provide Customer Choice in Input Operating Voltage
  - 200-480 VAC 3-phase nominal & 400-600VDC, Enhanced Universal Input
    - Configurations may be limited for 200-240 VAC input on the largest systems
  - 277 VAC add to smaller equipment; enables 480V phase-to-neutral operation
  - Attempt to enable best operating efficiency for each customer

- Design IT Equipment Power & Cooling for True End-to-End Efficiency
  - Power efficiency optimization from 13KV all the way to 1V
    - 13KV to the IT equipment input (PUE) is not a complete metric
  - Cooling efficiency from the outside ambient all the way to the silicon junction
    - Lower Temperature CMOS silicon dissipates less power
    - Electronics packed closer together may dissipate less power
Data Center Energy Efficiency Metric

PUE: Power Usage Effectiveness
DCE: Data Center Efficiency

Building Load
Demand from grid

Total Facility Power

Power
- Switchgear
- UPS
- Battery backup
- Etc.

Cooling
- Chillers
- CRACs
- Etc.

IT Equipment Power

IT Load
- Servers
- Storage
- Telco equipment
- Etc.

PUE = \( \frac{\text{Total Facility Power}}{\text{IT Equipment Power}} \)

DCE = \( \frac{1}{\text{PUE}} = \frac{\text{IT Equipment Power}}{\text{Total Facility Power}} \)
Delivered Annual electricity consumption by sector, 1980-2030

![Graph showing delivered annual electricity consumption by sector (Commercial, Residential, Industrial) from 1980 to 2030. The graph includes historical data and projections, with a note on Data Center Electric Growth.](image)
Data Center Infrastructure

- Cooling tower air loop
- Cooling tower water loop
- Refrigeration chiller loop
- Building chilled water loop
- CRAC & server fans

Power sources:
- Pumping power
- Compressor power
- Electrical power
IT Equipment is HOT...And HEAVY

P-Series 595
31”W x 66”D x 80”H
77500 BTU
3014 Lbs

P-Series 690
62” W x 50” D x 80”H
108,000 BTU
4233 Lbs
Possible Ventilation Schemes

Diagram showing possible ventilation schemes, including Hot Aisle, Cold Aisle, Racks, CRAC Unit, Ceiling Plenum, Egg Crate Return, Ducted Return, and Power/Cabling.
Hot Air Recirculation Animation - Raised Floor
Summary

✓ Liquid cooling is back – improves overall energy efficiency
✓ Integrate functions more effectively to drive efficiency
✓ Viewing designs more holistically – end to end power and cooling solutions
Thank You