NEC SX-6 am HLRS

• Holger Berger, NEC HPCE,
• EHPCTC Stuttgart
• hberger@hpce.nec.com

Earth Simulator

Top 500
#1
35.8 TFLOPS
Richard Loft, NCAR @ CAS Conference

IBM P690 Cluster
- 5.3 TFlops peak
- 1024 processors (32, 32 way P690 nodes)
- 5.2 GFlops/processor
- Observed 4.1-4.5% of peak on NCAR codes
- Max sustained on workload: 213.5 GFlops
- Est. Peak Price Performance: 516 GFlops
- Sustained Price Performance: 9.7 GFlops/MP

Earth Simulator
- 40.96 TFlops peak
- 120 Processors (640, 8 processor GS40 nodes)
- 8 GFlops/processor
- Estimate 30% of peak on NCAR codes
- Est. Max sustained on workload: 12,000 GFlops
- Est. Peak Price Performance: 8.3 GFlops/MP
- Est. Sustained Price Performance: 1.2 GFlops/MP

HLRS 2005

Quote from NECs offer to HLRS in 2003:

“a' viertele Earth-Simulator”
Company Overview

- Newly (in 2003) created NEC subsidiary
- Dedicated to HPC business
- Headquarters in Düsseldorf, Germany
- Serving the European Market
- Branch offices in France, Italy, The Netherlands, Switzerland and United Kingdom
- Application tuning & support centre in Stuttgart
- About 95 employees

Largest dedicated HPC operation in Europe!

NEC's SX-Series is a consistent innovation driver and today’s leading high performance platform

NEC uses latest manufacturing technology to build and develop new generations of supercomputers.
**SX-6 specifications**

- Vector Supercomputer
- 9-11 GF / CPU
- Maximum 128 nodes
- Maximum 1024 CPUs, max 11 TFLOPS
- Internode crossbar Switch
- 8 GB/s (bidirectional) interconnect bandwidth per node
- 1 TB/s maximum interconnect bandwidth per system in a flat crossbar

**SX-6 midlife kicker changes**

- 565 Mhz instead of 500 Mhz
- 9.2 Gflops instead of 8 gflops
- Faster IO processor – fast-iop
- PCI 64/66
SX Technology

Hardware dedicated to scientific/engineering applications:

Vector is the only architecture to solve the current and future bandwidth, latency and efficiency issues!

- Ease of Use
- Efficiency, not Macho-Flops
- Single chip vector processor
- Latency hiding via vectorisation
- Unprecedented memory bandwidth per chip:

36 Gbytes/s

SX Series As Technology Driver

Driver for high-end technology development.
Technological Progress

CPU Technology Trend

  - 8 Wide Vector Pipe
  - 2G FLOPS (8.0 ns)
  - 0.35 μm CMOS
  - 37 Chips

  - 16 Wide Vector Pipe
  - 8G FLOPS (4.0 ns)
  - 0.25 μm CMOS
  - 32 Chips

- SX-6 (2001)
  - 8 Wide Vector Pipe
  - 8G FLOPS (2.0 ns)
  - 0.15 μm CMOS
  - 1 Chip Vector CPU

Performance Trend

- SX-5
- SX-6
- Micro Processor

Packaging Technology Progress

Vector CPU

- SX-5 Vector CPU
  - Multi Chip Package
  - 225 mm sq
  - 11,000+ Pinout
  - 32 Layers

- SX-6 single chip CPU
  - 8 way parallel vector pipes
  - 8 Gflops (2.0 ns)
  - 0.15 μm CMOS
  - 5200 IO pads

Main Memory

- SX-5 MMU
  - 457 x 386 mm²
  - 64 - 128 Mb SDRAM

- SX-6 MMU
  - 105 x 176 mm²
  - 2 Gb / card
  - 256 Mb DDR-SDRAM
<table>
<thead>
<tr>
<th>Year</th>
<th>SX-6X Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>32 cpus @ 2GF, 8 GB memory</td>
</tr>
<tr>
<td>2000</td>
<td>32 cpus @ 4GF, 80 GB memory</td>
</tr>
<tr>
<td>2004</td>
<td>48 cpus @ 9.2GF, 384 GB memory</td>
</tr>
</tbody>
</table>

**SX-6 configuration at HLRS**

- SX-6 compute nodes
- TX-7 interactive server/file server
- HWW Network
- GbitEther-Switch
- FC-Switch
- iStorage 4TB disk space
SX-6 vector compute nodes

- 6 x 8 CPU nodes, 9.2/11Gflops performance, 64 GB shared Memory
- 36 GB/s memory bandwidth
- Connected by 8GB/s IXS crossbar
- 6 2-Gbit FC-HBAs, 4 for global filesystems
- Software:
  - NQSII: new version of batch system, syntax is different
  - Fortran/SX Compiler as so far
  - C++/SX Compiler as so far
  - HPF/SX Compiler as so far
  - MPI/SX as so far

TX7 interactive server

- 16(32)x Intel Itanium2 1.5Ghz/6M
- 256 GB Memory
- Partitioned, so not all CPUs will be visible to users
- 14 2-gbit FC-HBAs/12 Gbit ethernet interfaces
- Running NEC Linux based on RH Advanced Server
- SX cross compilers will be installed
- Filesystem will be shared with SX
- Integrated in NQSII
- Usage: Pre- and Postprocessing, compiling for SX
Changes for SX-5 users

- New batch system with new scheduler
  - Syntax oriented on POSIX standard, similar to PBS
  - Examples will be provided
- Multinode jobs will be common for many users
  - No problem for MPI users, just a question of mpirun parameters
  - Examples will be provided
- More work will be offloaded to the frontend TX7, like batch system
- Other environment is improved, but basically same as on SX-5

TX-7 frontend

- ‘AsAmA’
- Large Memory: 256 GB
- 16 CPUs in 4 cells, similar to Azusa
- Faster Memory, 6.4 GB/cell
- Better Snoop filters on the cells
- Faster crossbar
- Partitioned, there are two partitions with 16 CPUs and own memory not visible to you
SX-6 Memory

- **Capacity**
  - SX-5: avg. 2.5 GB per CPU, 16K banks
  - SX-6: 8 GB per CPU, 64GB per node, 4K banks
- Larger memory footprint for 'small' jobs is possible
- For MPI multinode jobs, ~240 GB will be available
- **Bandwidth**
  - SX-5: 32 GB/s
  - SX-6: 36 GB/s (clock-up model)
- But: Improved gather/scatter!!!
- If you encounter low performance for table-lookup: contact NEC!

Facts: Copy

NEC – Introduction
0–10
MPI Performance: Latency
Comm Performance for MPI/SX (NEC SX-6); type blocking

MPI Performance: Bandwidth
Comm Performance for MPI/SX (NEC SX-6); type blocking
Thank you.
Questions?