Empowered by Innovation NEC

NEC SX-6 am HLRS

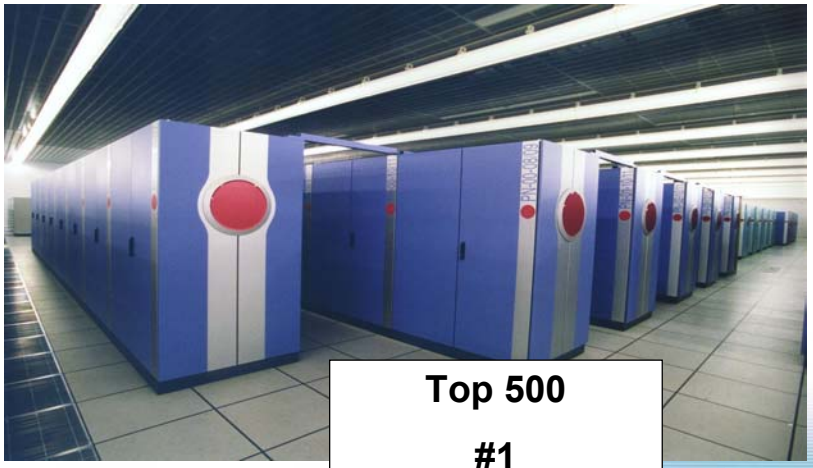
•Holger Berger, NEC HPCE,

•EHPCTC Stuttgart

•hberger@hpce.nec.com

Empowered by Innovation NEC

Earth Simulator



Top 500

#1

35.8 TFLOPS

© NEC HPCE, 2003, finishdesign

Empowered by Innovation NEC

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: $2.6/MFlops
- Sustained Price Performance: $59/MFlops
- Sustained Power Performance: 0.7 Gflops/KW

Earth Simulator

- 40.96 Tflops peak
- 5120 Processors (640, 8 processor GS40 nodes)
- 8 Gflops/processor
- Estimate 30% of peak on NCAR codes
- Est. Max sustained on workload: 12,200 GFlops
- Est. Peak Price Performance: $8.5/MFlops
- Est. Sustained Price Performance: $28/MFlops
- Est. Sustained Power Performance: 1.525 Gflops/KW

Empowered by Innovation NEC

HLRS 2005

•Quote from NECs offer to HLRS in 2003:

“ a’ viertele Earth-Simulator”

© NEC HPCE, 2003, finishdesign

Empowered by Innovation NEC

Company Overview

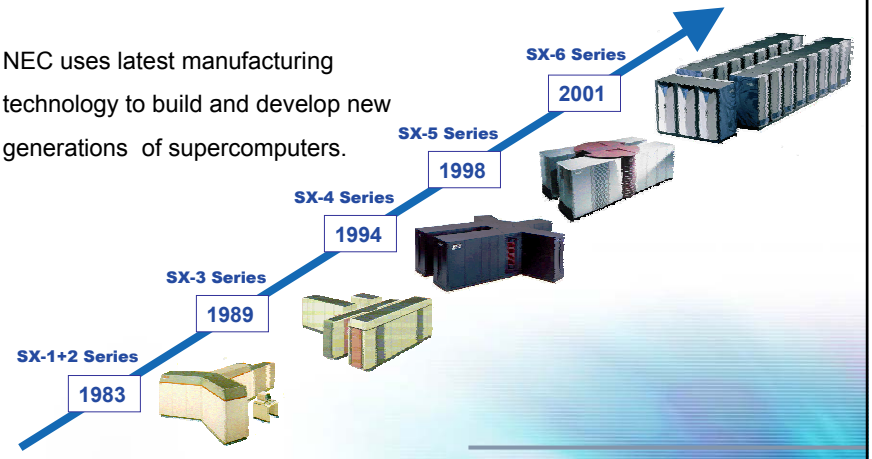
- Newly (in 2003) created NEC subsidiary
- Dedicated to HPC business
- Headquarters in Düsseldorf, Germany
- Serving the European Market
- Branch offices in France, Italy, The Netherlands, Switzerland and United Kingdom
- Application tuning & support centre in Stuttgart
- About 95 employees

Largest dedicated HPC operation in Europe!

Empowered by Innovation NEC

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

Empowered by Innovation NEC

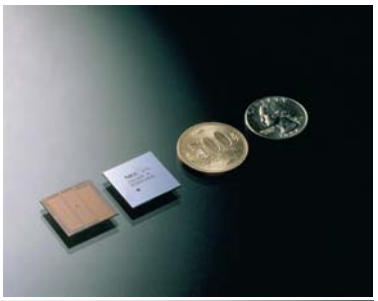
SX Technology

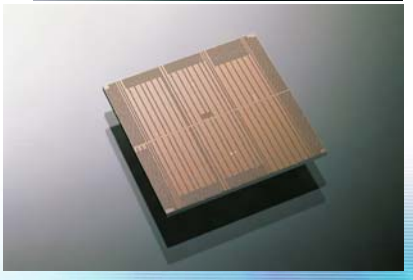
Hardware dedicated to scientific/engineering applications:

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

- Ease of Use
- Efficieny, not Macho-Flops
- single chip vector processor
- Latency hiding via vectorisation
- unprecedented memory bandwidth per chip:

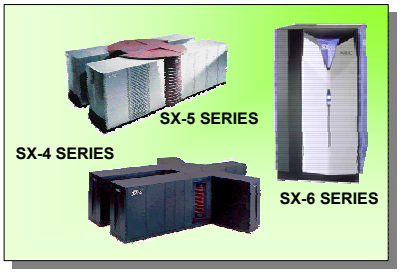
36 Gbytes/s

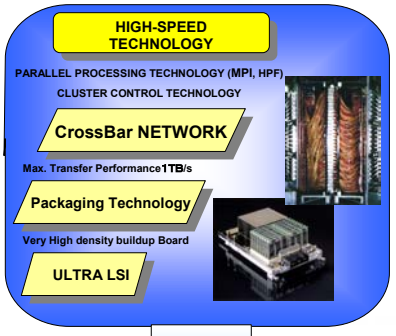




Empowered by Innovation NEC

SX Series As Technology Driver





SERVER / WORKSTATION

Driver for high-end technology development.

Express5800/1000 Series Server (a.k.a. TX7)

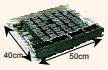
Express5800/ft Server

Empowered by Innovation NEC

Technological Progress

CPU Technology Trend

SX-4(1995)



8Wide Vector Pipe
:2GFLOPS(8.0ns)
:0.35μm CMOS
:37 Chips

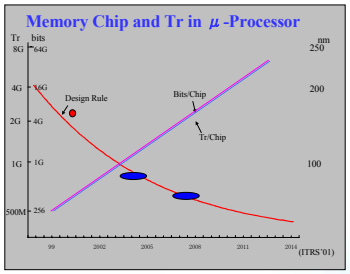
SX-5(1998)

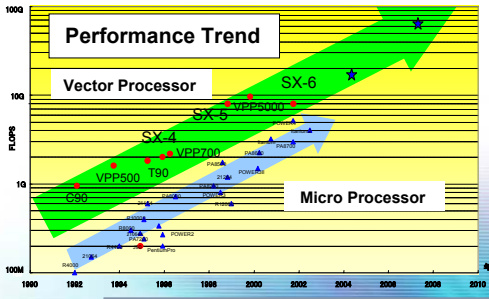
16 Wide Vector Pipe
:8GFLOPS(4.0ns)*
:0.25μm CMOS
:32 Chips

SX-6(2001)

8Wide Vector Pipe
:8GFLOPS(2.0ns)
:0.15μm CMOS
:1 Chip Vector CPU

NEC Inhouse Leading Edge Technology

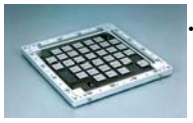




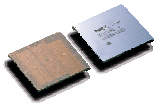
Empowered by Innovation NEC

Packaging Technology Progress

Vector CPU



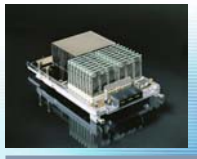
•SX-5 Vector CPU
-Multi Chip Package
▪225 mm sq
▪11,000+ Pinout
▪32 Layers



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

Main Memory

•SX-5 MMU
-457 x 386 mm^2
-4-8 GB / card
-64 -128 Mb SDRAM



•SX-6 MMU
-105 x 176 mm^2
-2 GB /card
-256 Mb DDR-SDRAM

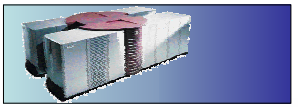
Empowered by Innovation NEC

HLRS & NEC

XXX cpus
XX Teraflops

SX-6X

48 cpus @ 9.2GF
384 GB memory



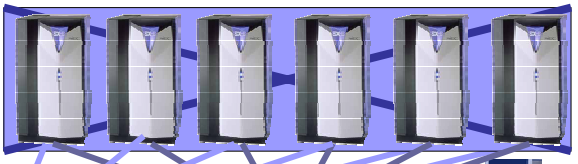
32 cpus @ 4GF
80 GB memory

32 cpus @ 2GF
8 GB memory

|1996 |2000 |2004 | 2005

Empowered by Innovation NEC

SX-6 configuration at HLRS



SX-6
compute
nodes

HWW-
Network

GbitEther-
Switch

FC-Switch

TX-7
interactive
server/file
server

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

Empowered by Innovation NEC

TX7 interactive server

- 16(32)xIntel 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

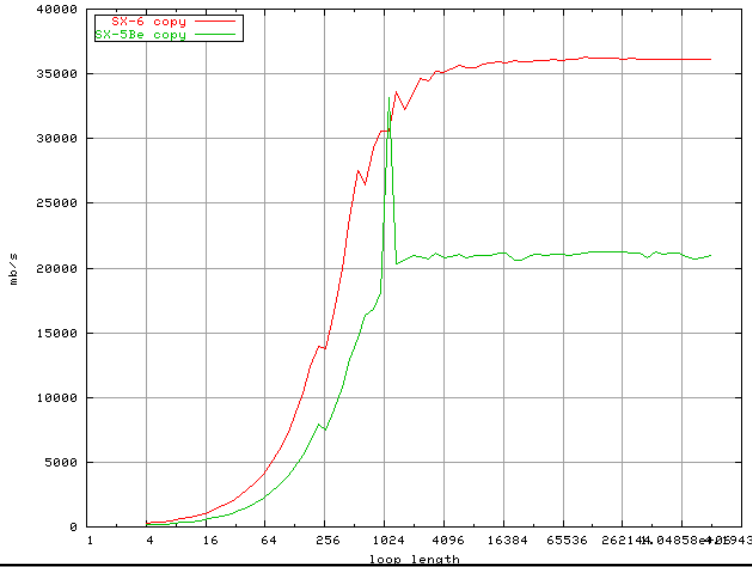
Empowered by Innovation NEC

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!

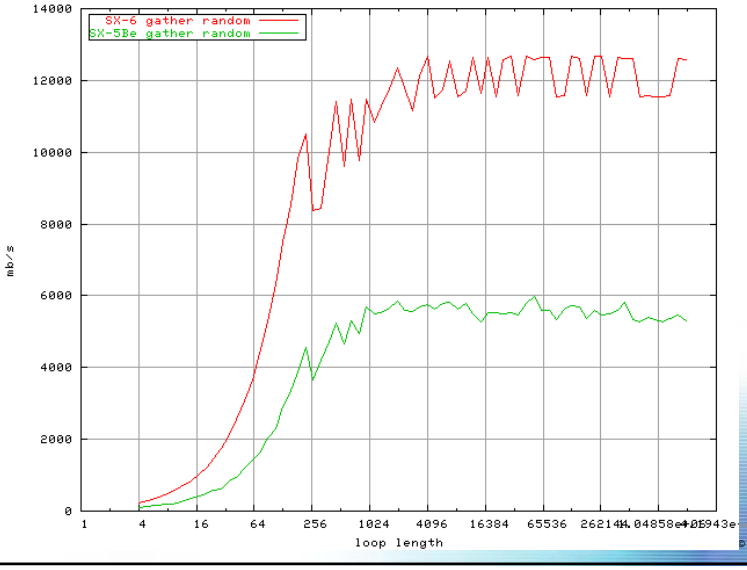
Empowered by Innovation NEC

Facts: Copy



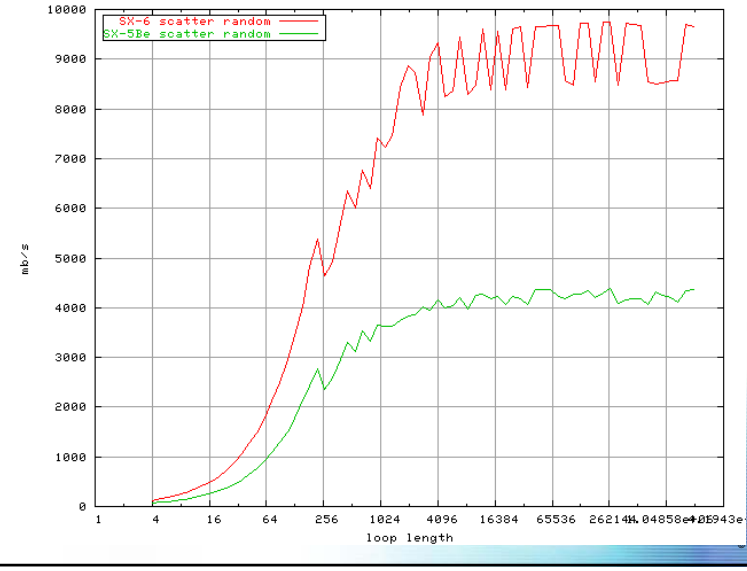
Empowered by Innovation NEC

Facts: Gather



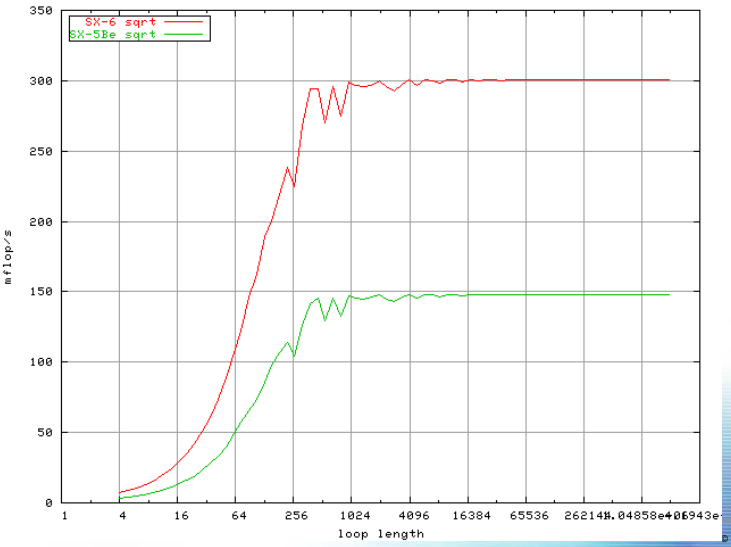
Empowered by Innovation NEC

Facts: Scatter



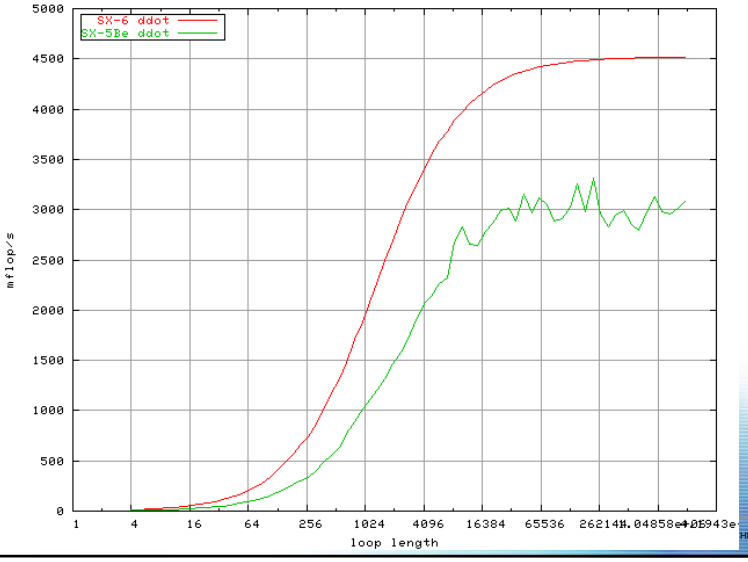
Empowered by Innovation NEC

Facts: Sqrt



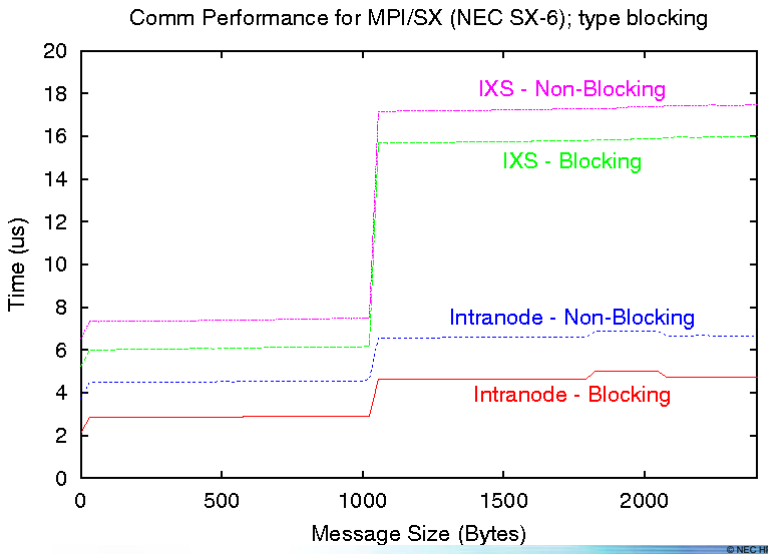
Empowered by Innovation NEC

Facts: Ddot



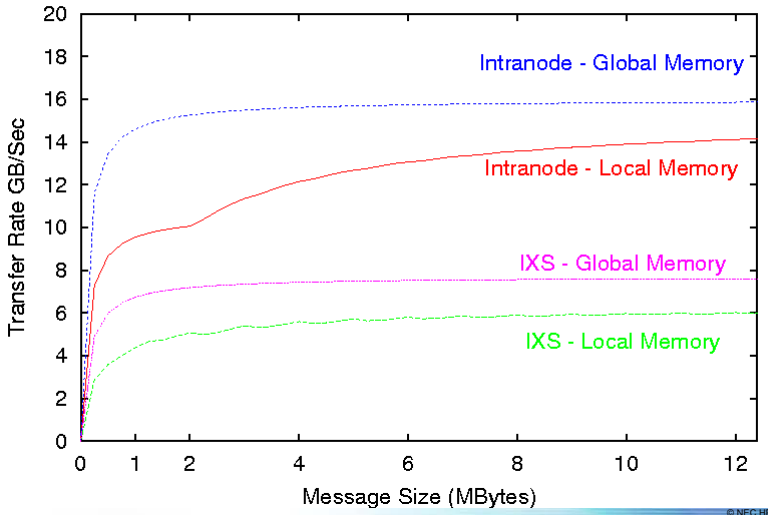
Empowered by Innovation NEC

MPI Performance: Latency



Empowered by Innovation NEC

MPI Performance: Bandwidth



Empowered by Innovation NEC

Thank you.
Questions?