HOTCHIPS26

SX-ACE Processor: NEC’s Brand-New Vector Processor

Shintaro Momose, Ph.D.
NEC Corporation
IT Platform Division
Manager of SX vector supercomputer development

August 11th, 2014
NEC has always provided the high sustained performance by Vector Super-Computer SX series.

**SX History and Technical Evolutions**

- **SX-2**: Bipolar Water Cooling
- **SX-3**: Multi node CMOS Air Cooling
- **SX-4**: Multi node CMOS Air Cooling
- **SX-5**: 1 Chip Vector Processor
- **SX-6**: 3D node module
- **SX-7**: Support Over 100 nodes Cluster
- **SX-8/8R**: Support >1000 nodes ECO
- **SX-9**: Multi-core All in One Chip ECO

**Software innovations**
- Auto Vectorization Compiler
- Distributed Parallelization (MPI-SX)
- Automatic Parallelization Function & SUPER-UX
- MPI support to Multilane IXS
- 100GF Processor
- ECO

**Hardware innovations**

- 1990
- 2000
- 2010
Table of Contents

■ Introduction
  • HPC Technical Trends and Issues
  • SX-ACE Processor Design Direction

■ Processor Architecture
  • Processor/Core Architectures and Implementations
  • Memory Subsystem

■ Performance Evaluations
  • Experimental Results of Several memory intensive Benchmarks

■ Conclusions
Introduction
Growing of LINPAC performance has been provided by system enlarging
User must spend their time to extract massive parallelism
Smaller # of cores with big cores can reduce the difficulty

Trend of TOP500 (1st ~ 10th system)

- Linpack [TF]
- Lipack ave. [TF]
- # of cores
- # of cores ave.
- # of nodes
- # of nodes ave.
- Core performance [GF]
- Core performance ave. [GF]
- Frequency [GHz]
- Frequency ave. [GHz]

Linpack [GF] 198%/year
Moore's Law CPU performance 142%/year
Core performance [GF] 115%/year
Frequency [GHz] 115%/year

Increasing fewer cores
big core
Nearly constant
According to Japanese Government (MEXT) working group report for a wide variety of strategic segment applications, diverse characteristics are observed.

B/F requirement from each application differs greatly. Any single architecture cannot cover all application areas.

Concepts of SX-ACE

The best solution for memory intensive APs against scalar processors trend

**Big Core**
Reducing Massive Parallel Difficulty with fewer cores

*World top-level performance:* 64GF
*The largest memory bandwidth:* 64~256GB/s

**Low Power Consumption**
The best memory bandwidth solution

GB/s / Watt compared x86 CPU 1.8x

**Hybrid Solution**
Vector / Scalar tightly coupled environment

Specialized SWs

© NEC Corporation, 2014 / HOTCHIPS26
## Processor Overview

**Architecture**: Vector

**Clock Frequency**: 1.0GHz

**SPU decode rate**: 4 instructions

**VPU Performance**: 64GFlops

**ADB size**: 1MB

**ADB bandwidth**: 256GB/s

**Memory bandwidth**: 64GB/s~256GB/s

**Core Byte/Flop**: 1.0 ~ 4.0

### CPU

- **Cores**: 4
- **Performance**: 256GFlops
- **Memory bandwidth**: 256GB/s
- **CPU Byte/Flop**: 1.0
- **Memory capacity**: 64GB

---

**Crossbar**

**Interconnect**

8GB/s x2

**Remote access Control Unit (RCU)**

**Scalar Processing Unit (SPU)**

**Vector Processing Unit (VPU)**

**Assignable Data Buffer (ADB)**

256GB/s

**Memory (DDR3)**

256GB/s

**Memory controller**

256GB/s

© NEC Corporation, 2014 / HOTCHIPS26
The SX-ACE core can provide the world top-level performance and the largest memory bandwidth.
Floor Plan of the CPU

- **“Memory access” focused layout**

- **Specifications**
  - Process rule: 28nm
  - Clock speed: 1GHz
  - Die size: 23.05 x 24.75mm
  - # of transistors: 2BTr.

- **I/F**
  - 16ch DDR3 I/F
  - IXS 8GB/s x 2
  - 2ch PCIe x8 I/F
Core Architecture

256 operations = 16 parallel x 16 clock cycles

Vector Data Registers
- 64 sets
- 64 Vector Data Registers
- 256 elements
- 256 bits

Vector Arithmetic Registers
- 16 Vector Mask Registers
- 256 elements
- 8 Vector Arithmetic Registers

Forwarding
- 64 sets

Add pipeline

Multiply pipeline

Divide/Sqrt pipeline

Logical pipeline

Mask pipeline
Memory Network Integration

- Large SMP configuration can provide high sustained performance
- But, over 70% power was consumed by the memory network
- SX-ACE processor integrates the memory network into LSI

SX-9 1node 1.6TF

- CPU (16LSI, 16cores)
- Memory network switch (32LSI)
- Printed wiring
- Memory controller (512LSI)
- RAM

SX-ACE 6nodes 1.5TF

- Many LSI consuming more than 70% of the power
- High performance maintained
- Power efficient
- Number of LSI reduced to 1/100

560LSI
30KW

6LSI
2.8KW

© NEC Corporation, 2014 / HOTCHIPS26
# Reducing DRAM Energy

## Cache line size
- DRAM activation powers are depending on
- Sustained memory bandwidth is strongly affected by adopted cache line size

## Variable cache line size feature
- Supporting 64B/128B memory access granularity
- 128B as a default to reduce power
- 64B for a sparse memory access such as stride/indirect memory accesses

---

RD:WR 1:1, Micron DDR3 power calculator 0.96
Assignable Data Buffer (ADB)

- **On-chip Cache for Vector**
  - Private, 1MB, 4-way, 16-bank
  - 256GB/s bandwidth per core
  - Software controllable cache
  - Customized for fast random access

- **Assignable Feature**
  - A bypass flag in each instruction
  - Compiler/User can control
  - Avoiding cache pollution

- **MSHR Feature**
  - Redundant memory requests same as an inflight memory request are held to reduce memory transactions
Out-of-Order Vector Memory Access

Vector memory access instruction

Consecutive memory access (by HW)
1: VST
2: VLD A
3: VLD B

Checking lower/upper bound

Stride memory access (by HW)
1: VST
2: VLD A
3: VLD B

Checking start address and stride

Indirect memory access (by SW)
1: VLD A
2: VSC
3: VLD B

© NEC Corporation, 2014 / HOTCHIPS26
Node Packaging

**CPU**

256GFlops = 64GFlops/core x 4
256GB/s memory bandwidth

**Memory**

16 DIMMs (DDR3 2000)
256GB/s, 64GB

Rated power consumption = 469W
Optimization of cooling efficiency and rack weight

- **CPU:** water cooling
- **Other components:** air cooling
Performance Evaluation
Evaluation programs

<table>
<thead>
<tr>
<th>Evaluate point</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-chip memory bandwidth</td>
<td>STREAM (TRIAD)</td>
</tr>
<tr>
<td>Off/On-chip memory bandwidth</td>
<td>Himeno Benchmark (High memory intensive)</td>
</tr>
<tr>
<td>Indirect memory access performance</td>
<td>Legendre transformation</td>
</tr>
</tbody>
</table>

Each evaluation is carried out by only using compiler optimizations without code modifications for individual systems

Performance comparison

<table>
<thead>
<tr>
<th>CPU</th>
<th>Performance</th>
<th>Memory bandwidth</th>
<th>Rated system Watts/CPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>SX-9</td>
<td>102GF = 102GF x 1c</td>
<td>256GB/s</td>
<td>1875W</td>
</tr>
<tr>
<td>SX-ACE</td>
<td>256GF = 64GF x 4c</td>
<td>256GB/s</td>
<td>469W</td>
</tr>
<tr>
<td>IVB(Xeon)</td>
<td>230GF = 19GF x12c</td>
<td>60GB/s</td>
<td>200W</td>
</tr>
<tr>
<td>Power7</td>
<td>245GF = 31GF x 8c</td>
<td>128GB/s</td>
<td>656W</td>
</tr>
<tr>
<td>FX10(Sparc)</td>
<td>234GF = 15GF x16c</td>
<td>85GB/s</td>
<td>281W</td>
</tr>
</tbody>
</table>

Power7 and FX10 are measured through a joint research with Tohoku University
Evaluation of Off-chip memory bandwidth

- Benchmark code: STREAM (TRIAD)

Sustained memory bandwidth

- Only the SX-ACE single core can use full memory bandwidth
- This can accelerate memory-intensive serial parts in parallel processing

Power efficiency (SX-ACE=1)

- SX-ACE provides the best memory bandwidth per watt
Memory Bandwidth 2

**Evaluation of Off/On-chip memory bandwidth**

- **Benchmark code:** Himeno benchmark (highly memory intensive) solving the Poisson equation with the Jacobi iterative method

**Sustained performance (SX-ACE=1)**

- ADB and MSHR improve sustained memory bandwidth compared with its predecessor
- SX-ACE is the best

**Power efficiency (SX-ACE=1)**

- SX-ACE is assumed to provide 2~25x higher power efficiency in the case of memory intensive APs having off/on chip memory accesses
Indirect Memory Access

Evaluation of Indirect memory access performance
- Benchmark code: Legendre transformation
- Cache effective BM (4.4MB data)

Sustained performance (SX-ACE=1)

- Memory bandwidth [GB/s]
- Power efficiency ratio

<table>
<thead>
<tr>
<th></th>
<th>SX-9</th>
<th>SX-ACE</th>
<th>IVB</th>
<th>Power7</th>
<th>FX10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>0.15</td>
<td>1.00</td>
<td>0.68</td>
<td>0.69</td>
<td>0.17</td>
</tr>
<tr>
<td>Power</td>
<td>1.00</td>
<td>0.68</td>
<td>0.69</td>
<td>1.5~6x</td>
<td>25x</td>
</tr>
</tbody>
</table>

- Cache is effective
- ADB, MSHR, OoO, and short memory access latency work well

Power efficiency (SX-ACE=1)

<table>
<thead>
<tr>
<th></th>
<th>SX-9</th>
<th>SX-ACE</th>
<th>IVB</th>
<th>Power7</th>
<th>FX10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>0.04</td>
<td>1.00</td>
<td>0.50</td>
<td>0.50</td>
<td>0.28</td>
</tr>
<tr>
<td>Power</td>
<td>1.00</td>
<td>1.58</td>
<td>0.6~3x</td>
<td>0.50</td>
<td>0.28</td>
</tr>
</tbody>
</table>

- SX-ACE improvement provides 25x higher power efficiency than SX-9
- But, IVB is the best due to a larger cache and a lower power consumption
Conclusions

- **Issue of modern scalar/accelerator processors**
  - Massive parallel with small cores
  - Low memory bandwidth

- **SX-ACE direction**
  - Providing the big core with large memory bandwidth
  - Improving proven vector architecture

- **SX-ACE processor**
  - 4 cores vector processor
  - 64GF core performance with 64-256GB/s memory bandwidth
  - Efficient memory subsystem for higher sustained memory bandwidth

- **Performance**
  - High sustained performance and power efficiency for memory intensive benchmarks
I would like to express my gratitude to Cyber Science Center at Tohoku University for the intensive performance evaluation of the SX vector supercomputers as a part of the joint research project with NEC Corporation.

**Tohoku University, Cyber Science Center**
- Professor Hiroaki Kobayashi
- Associate Professor Ryusuke Egawa
- Assistant Professor Kazuhiko Komatsu

**NEC Corporation**
- Senior Manager Takashi Hagiwara
- Manager Yoko Isobe
Empowered by Innovation

NEC