Orchestrating a brighter world

NEC

Vector Engine Processor of NEC's Brand-New Supercomputer SX-Aurora TSUBASA

Yohei Yamada, NEC Corporation
Shintaro Momose, Ph.D., NEC Deutschland GmbH
21 August 2018

Agenda

- Introduction
- SX-Aurora TSUBASA
- Vector Engine
- Benchmarks
- Conclusion

Introduction

History of NEC's Vector Supercomputer

35 years
Experience
For
High Sustained
Performance

SX-Aurora TSUBASA

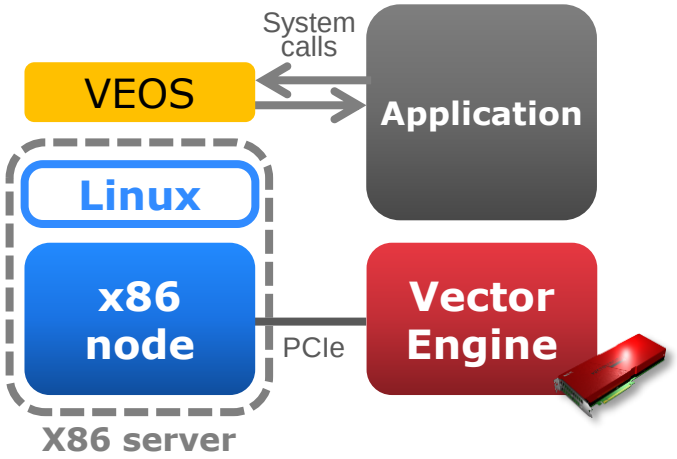
SX-Aurora TSUBASA

Vector computing in a standard environment

- High sustained performance vector processing
- Vector capability is transparently provided on x86/Linux

Aurora architecture

Applications are entirely executed on VE side



Hardware

- Vector Engine (VE) + x86 node
- High memory bandwidth
- Flexible configuration

SW Environment

- x86 / Linux OS
- Fortran/C/C++ standard programming
- Automatic vectorization and parallelization by proven vector compiler

Scalable Vector Supercomputer

Supercomputer Model
- For large scale configurations
- DLC with 40°C/104°F water

Rack Mount Model
- Flexible configuration
- Air Cooled

Tower Model
- For developer/programmer
- Personal supercomputer

Vector Engine Card

Air Cooled Card

- Two types of packages

Passive Cooling Type
For Server

Active Cooling Type
For Tower/Workstation

Water Cooled Card

- Direct liquid cooling
- Hot water cooling available

40°C/104°F water

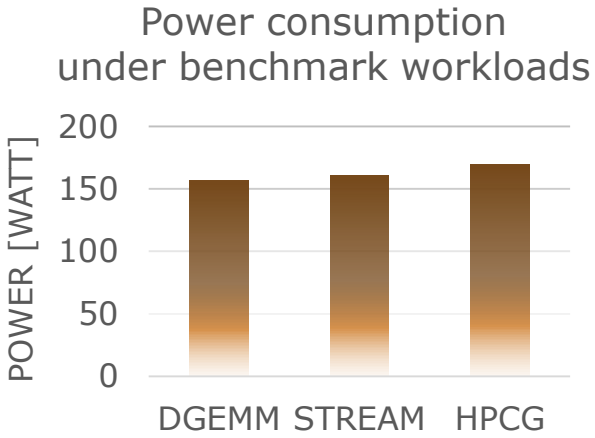
Direct Liquid Cooling Type
For Supercomputer

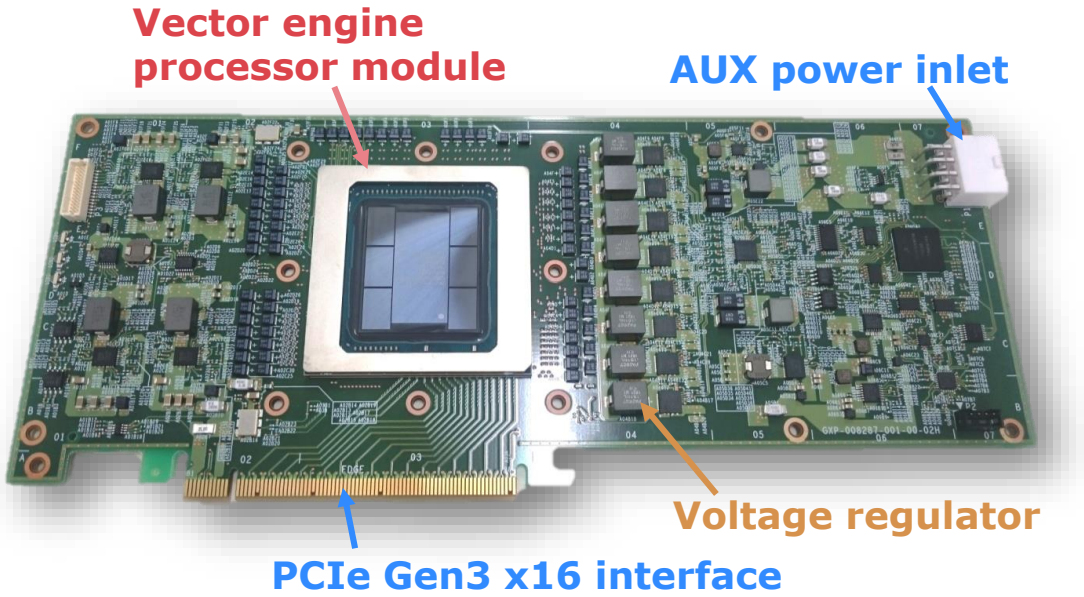
Vector Engine

Vector Engine Card Implementation

Standard PCIe card

- PCIe Gen3 x16 interface
- Full-length full-height card
- Dual slot
- <300W power

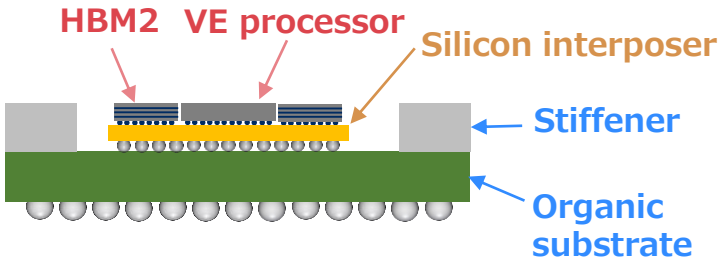


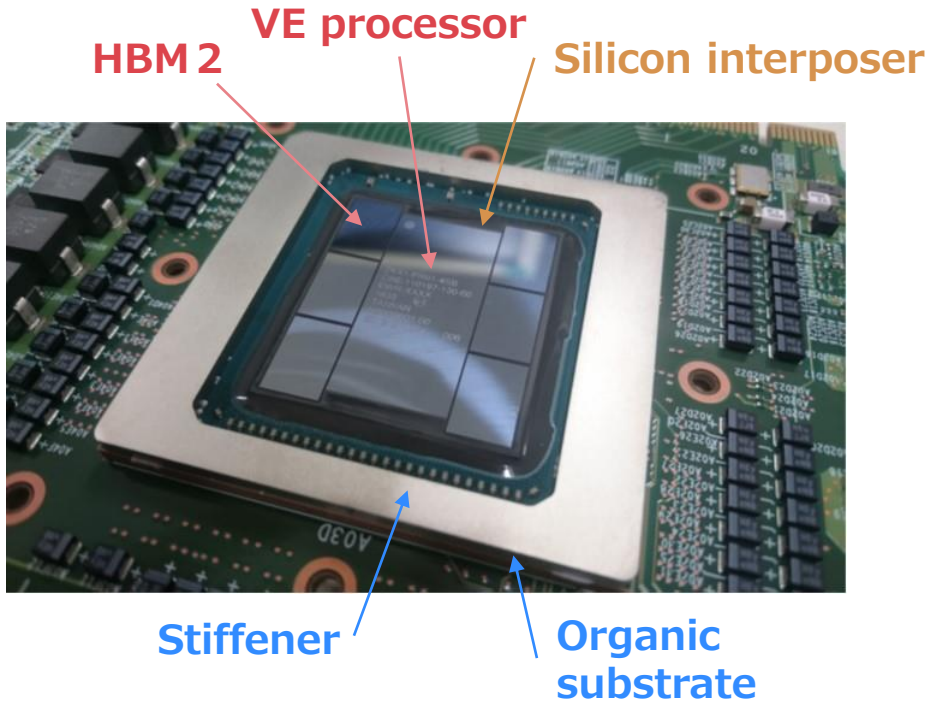


Vector Engine Processor Module

2.5D implementation

- A VE processor and six 8Hi or 4Hi HBM2 modules on a silicon interposer
- Lidless package to minimize thermal resistance
- Package size: 60mm x 60mm
- Interposer size: 32.5mm x 38mm
- VE processor size: 15mm x 33mm





World's first implementation of a processor with 6 HBM2s

Vector Engine Processor Overview

Components

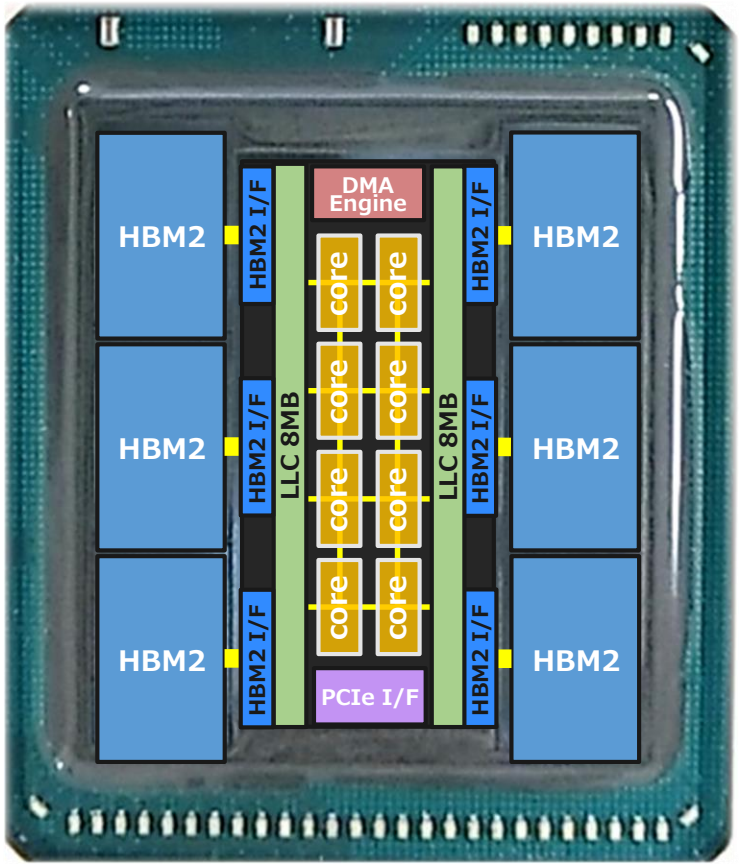
- 8 vector cores
- 16MB LLC
- 2D mesh network on chip
- DMA engine
- 6 HBM2 controllers and interfaces
- PCI Express Gen3 x16 interface

Specs

Core frequency	1.6GHz
Core performance	307GF(DP) 614GF(SP)
CPU performance	2.45TF(DP) 4.91TF(SP)
Memory bandwidth	1.2TB/s
Memory capacity	24/48GB

Technology

- 16nm FinFET process



Vector Core

Vector Processing Unit (VPU)

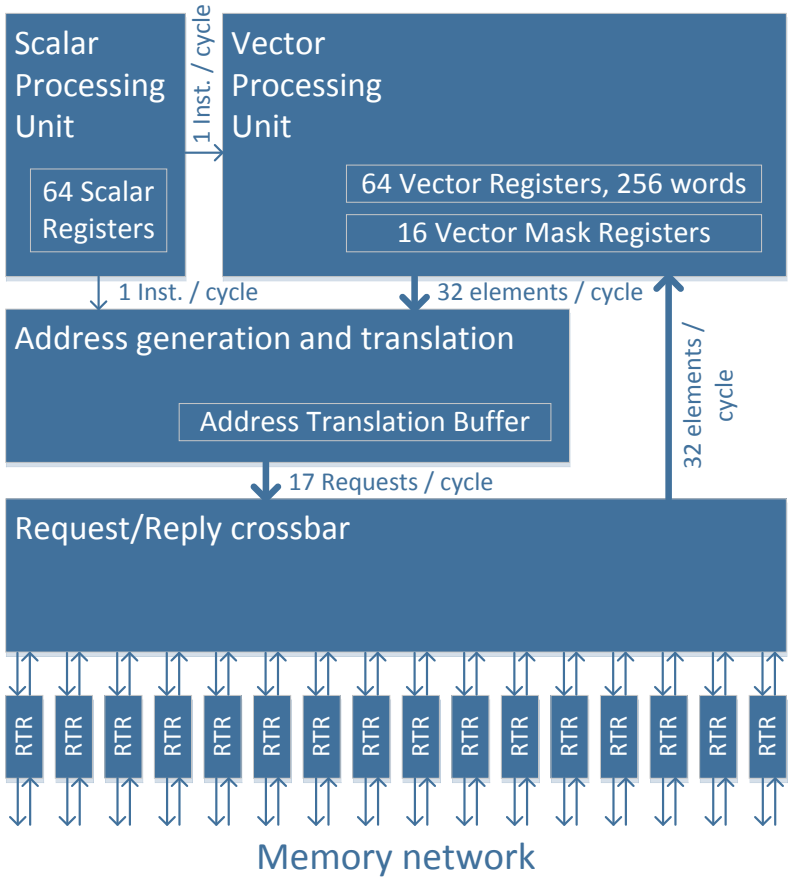
- Powerful computing capability
  - 307.2GFLOPS DP / 614.4GFLOPS SP performance
- High bandwidth memory access
  - 409.6GB/sec Load and Store

Scalar Processing Unit (SPU)

- Provides the basic functionality as a processor
  - Fetch, decode, branch, add, exception handling, etc…
- Controls the status of complete core

Address translation and data forwarding crossbar

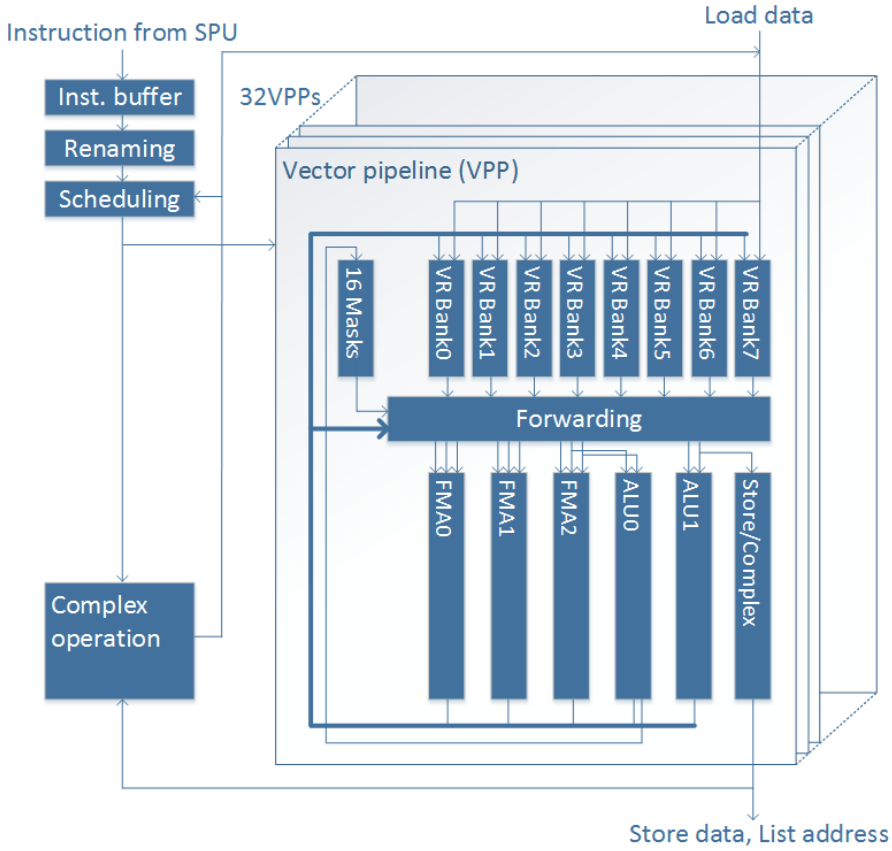
- To support contiguous vector memory access
  - 16 elements/cycle vector address generation and translation, 17 requests/cycle issuing
  - 409.6GB/sec load and 409.6GB/sec store data forwarding



Vector Processing Unit

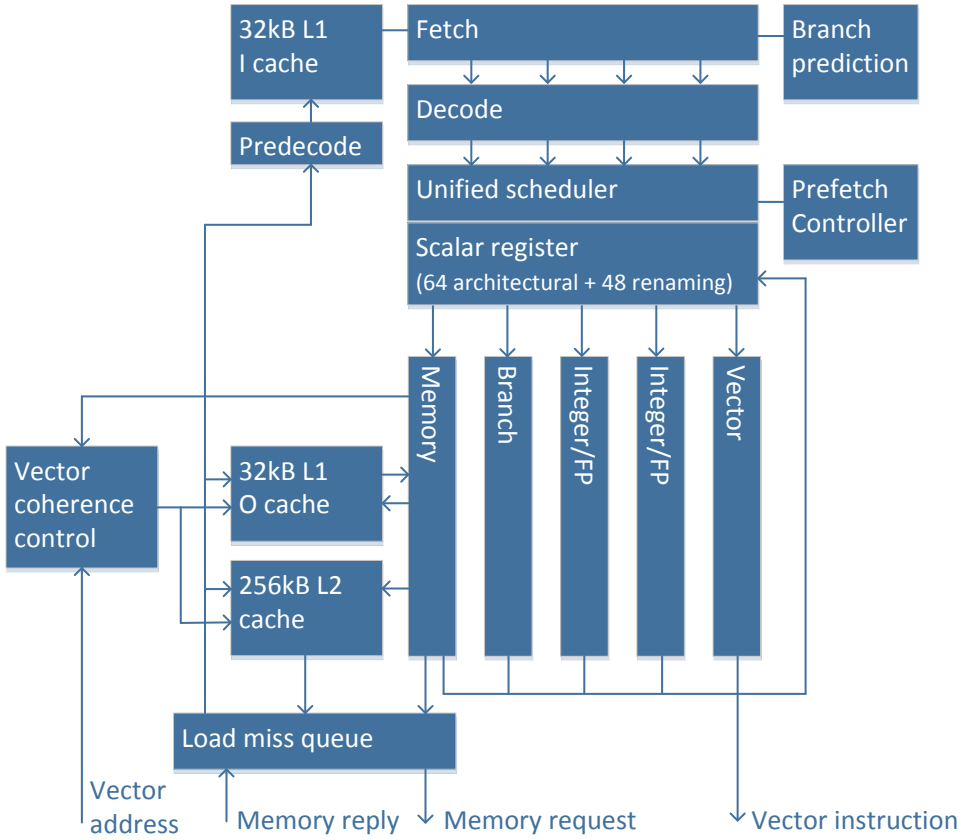
- Four pipelines, each 32-way parallel
  - FMA0: FP fused multiply-add, integer multiply
  - FMA1: FP fused multiply-add, integer multiply
  - ALU0/FMA2: Integer add, multiply, mask, FP FMA
  - ALU1/Store: Integer add, store, complex operation

  } Total 96 FMAs
- Doubled SP performance by 32bit x 2 packed vector data support
- Vector register (VR) renaming with 256 physical VRs
  - 64 architectural VRs are renamed
    - Enhanced preload capability
    - Avoidance of WAR and WAW dependencies
- OoO scheduling
- Dedicated complex operation pipeline to prevent pipeline stall
  - Vector sum, divide, mask population count, etc.



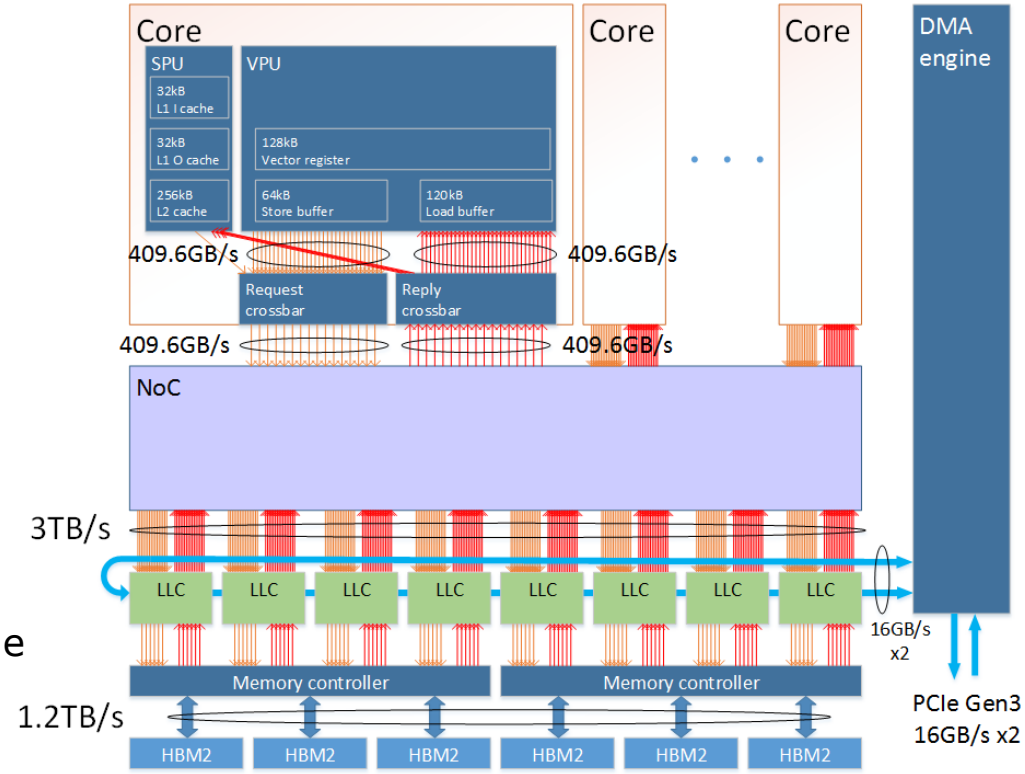
Scalar Processing Unit

- General enhancements
  - 4 instructions / cycle fetch and decode
  - Sophisticated branch prediction
  - OoO scheduling
  - 8-level speculative execution
  - Four scalar instruction pipes
  - Two 32kB L1 caches + unified 256kB L2 cache
  - Hardware prefetch
- Support for contiguous vector operation
  - Dedicated vector instruction pipe
  - 16 elements / cycle coherency control for vector store



Memory Subsystem

- High bandwidth
  - 409.6GB/s x2 core bandwidth
  - Over 3TB/s LLC bandwidth
  - 1.2TB/s memory bandwidth
- Caches
  - Scalar L1/L2 caches in each core
  - 16MB shared LLC
- Two memory networks
  - 2D mesh NoC for core memory access
  - Ring bus for DMA and PCIe traffic
- DMA engine
  - Used by both vector cores and x86 node
  - Can access VE memory, VE registers, and x86 memory



Network on chip (NoC)

2D mesh network

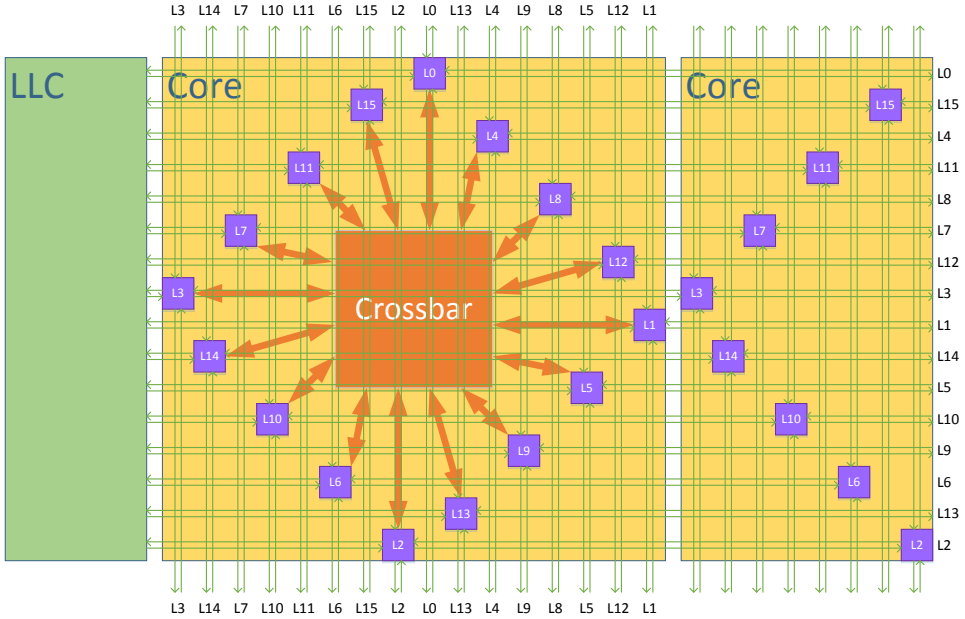
- Maximize bandwidth with minimal wiring
- Minimizing data transfer distance
- 16 layered mesh

Deadlock avoidance

- Dimension-ordered routing
- Virtual channels for request and reply

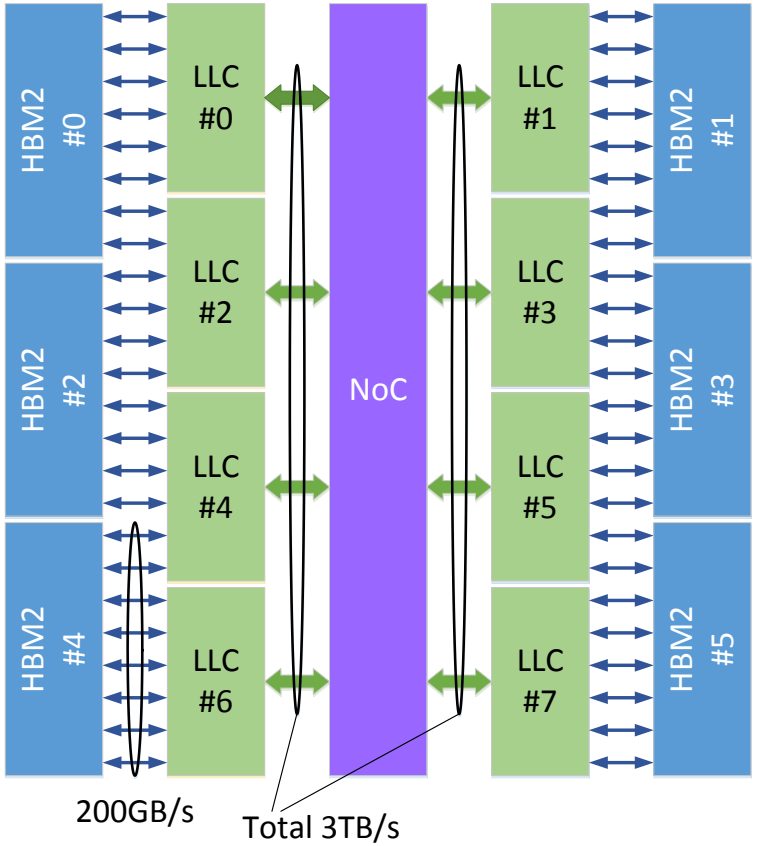
Adaptive flow control

Age based QoS control



Last Level Cache (LLC)

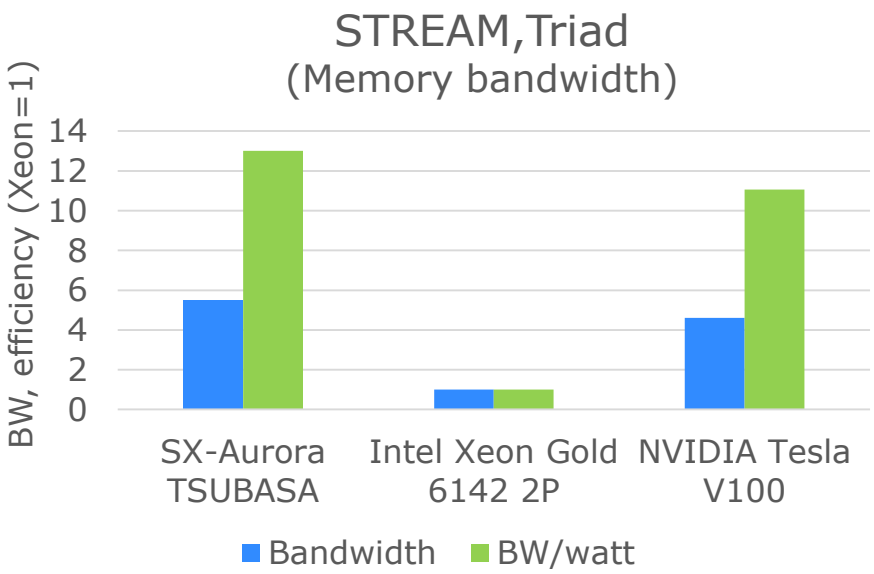
- Memory side cache
  - Avoiding massive snoop traffic
  - Increasing efficiency of indirect memory access
- 16MB, write back
- Inclusive of L1 and L2
- High bandwidth design
  - 128 banks, in total more than 3TB/s bandwidth
- Auto data scrubbing
- Assignable data buffer feature
  - Priority of data can be controlled by a flag for vector memory access instructions



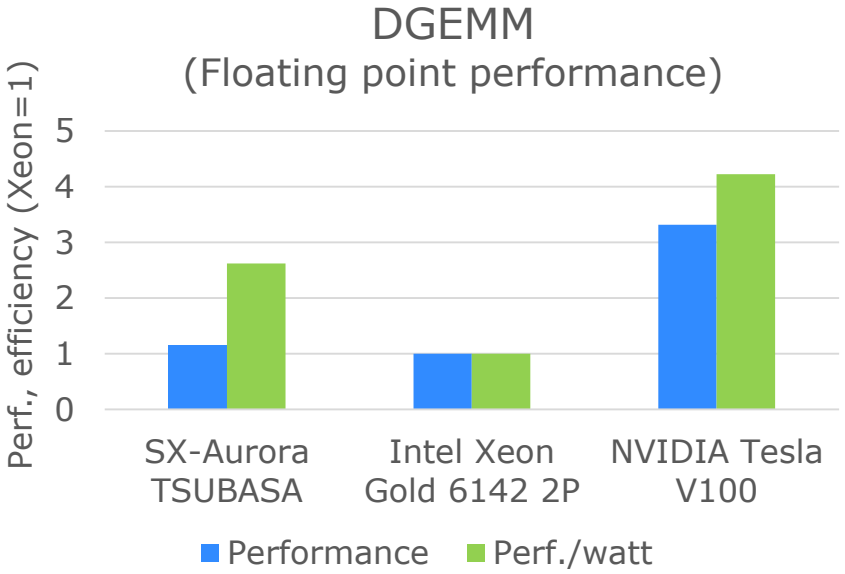
Benchmarks

Benchmark conditions
SX-Aurora TSUBASA: SX-Aurora TSUBASA A500 model
Intel Xeon: Intel Xeon Gold 6142 2 sockets, 192GB DDR4-2666
NVIDIA Tesla V100: Intel Xeon CPU E5-2630v4 2 sockets, 128GB DDR4-2400, NVIDIA Tesla V100 16GB

Floating point calculation and memory bandwidth



Industry leading memory access performance and efficiency

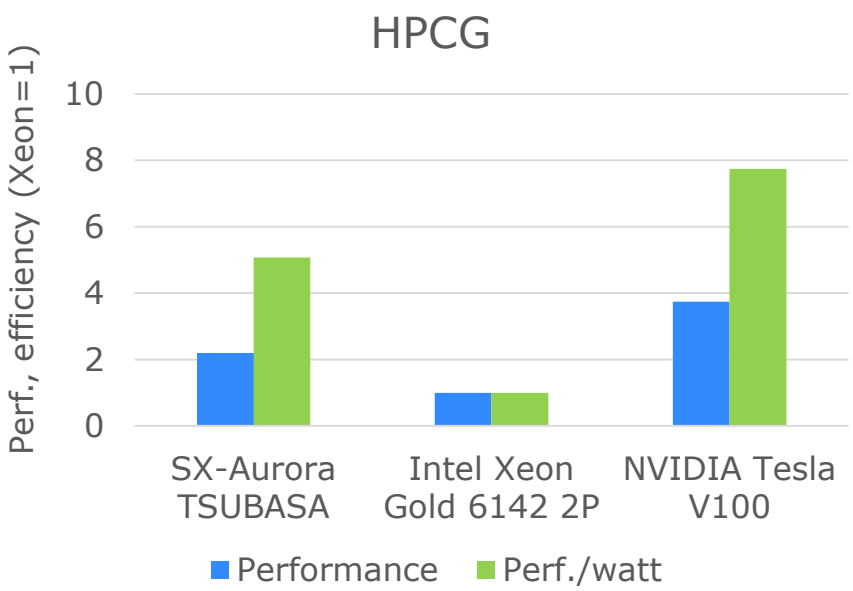


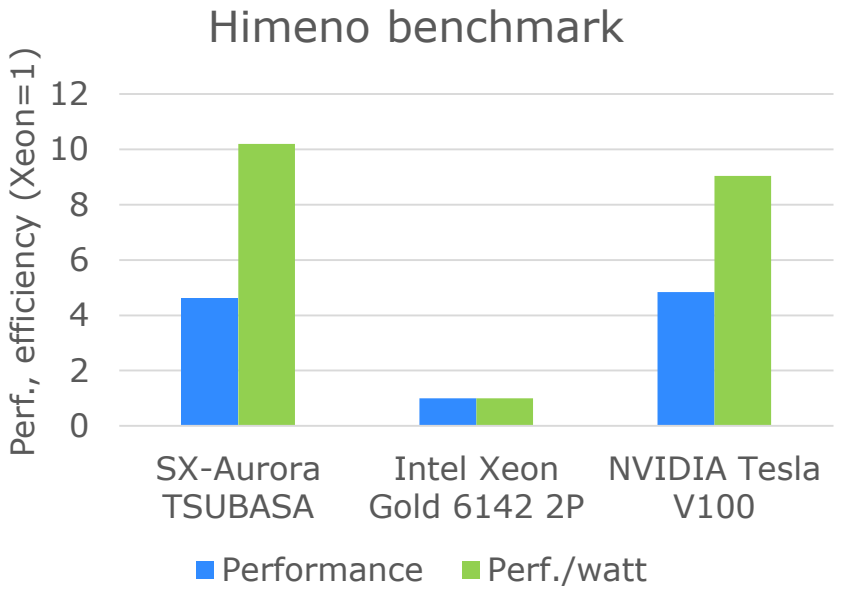
Comfortable enough compute capability for memory intensive workloads

Note: VE price is much cheaper than V100

Performance on General Benchmarks

HPCG and Himeno benchmark (Poisson equation solver)





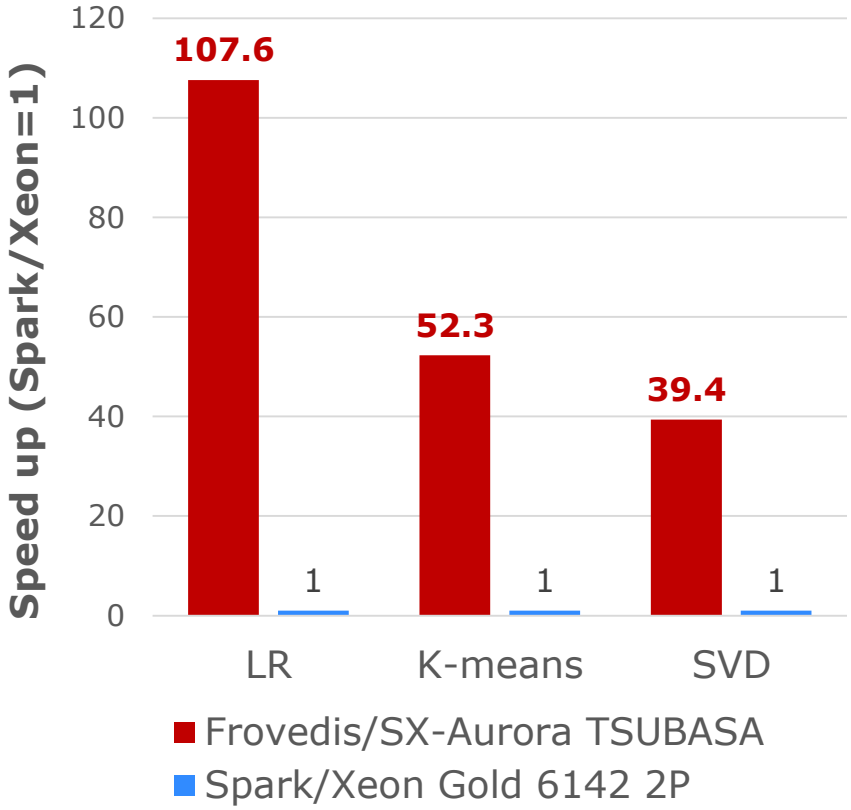
Competitive performance and power efficiency
available using standard programming paradigms

Note: VE price is much cheaper than V100

Performance on Machine Learning

Statistical machine learning

- Workloads
  - Web ads optimization (Logistic regression)
  - Document clustering (K-means)
  - Recommendation (Singular value decomposition)
- NEC's *Frovedis*™ framework for AI/BigData processing
  - Apache Spark MLlib compatible API
  - Open source
    - https://github.com/frovedis



Summary

SX-Aurora TSUBASA

- A new product line of vector supercomputers based on Aurora architecture
- Vector capability is provided in a standard x86/Linux environment

Vector Engine

- High memory bandwidth by six HBM2s configuration
- Enhancements of the vector microarchitecture to provide high sustained performance and power efficiency

Benchmarks

- Very competitive performance and power efficiency using standard programming paradigms
- Outstanding performance on statistical machine learning workloads with Frovedis framework

Thank you!


Orchestrating a brighter world
NEC