



## The NEC SX-Aurora TSUBASA Vector Engine

Tim Cramer

# Motivation

---

- Interesting Architecture
- Very high memory bandwidth
  - User codes of the RWTH Compute Cluster are often memory-bound  
→ might benefit from SX-Aurora capabilities
- Performance portability: Single application for multiple types of devices / architectures
  - standard-compliance, e.g., MPI, OpenMP

# SX-Aurora TSUBASA Vector Engine (VE)

- Vector processor as accelerator in x86 environment
- PCIe card
- Combines SIMD and pipelining
  - vector register length:  $256 * 64 = 16384$  bits
- OS for the VE (VEOS) runs on the vector host (VH)
- Specs



VE 1.0

|                         |                                    |
|-------------------------|------------------------------------|
| Vector Cores            | 8                                  |
| FMAs                    | 3                                  |
| Core Frequency          | 1.4 – 1.6 GHz                      |
| Theor. Peak Performance | 307 GFLOPs (DP)<br>604 GFLOPs (SP) |
| Memory Capacity         | 24/48 GB                           |
| Main Memory Bandwidth   | 1.2 TB/s (6x HBM2 stacks)          |
| TPD                     | 300 Watts                          |
| Technology              | 16 nm FinFET process               |



# SX-Aurora in a Cluster Environment

- Supercomputer Model
  - DLC with 40° C

- Rack Mount Model
  - Air cooled

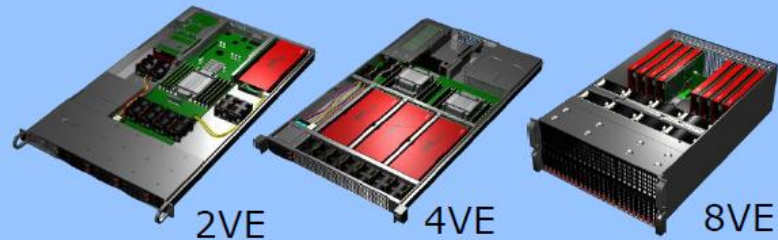
- Tower Model

## A500 series



64VE-

## A300 series



2VE

4VE

8VE

## A100 series

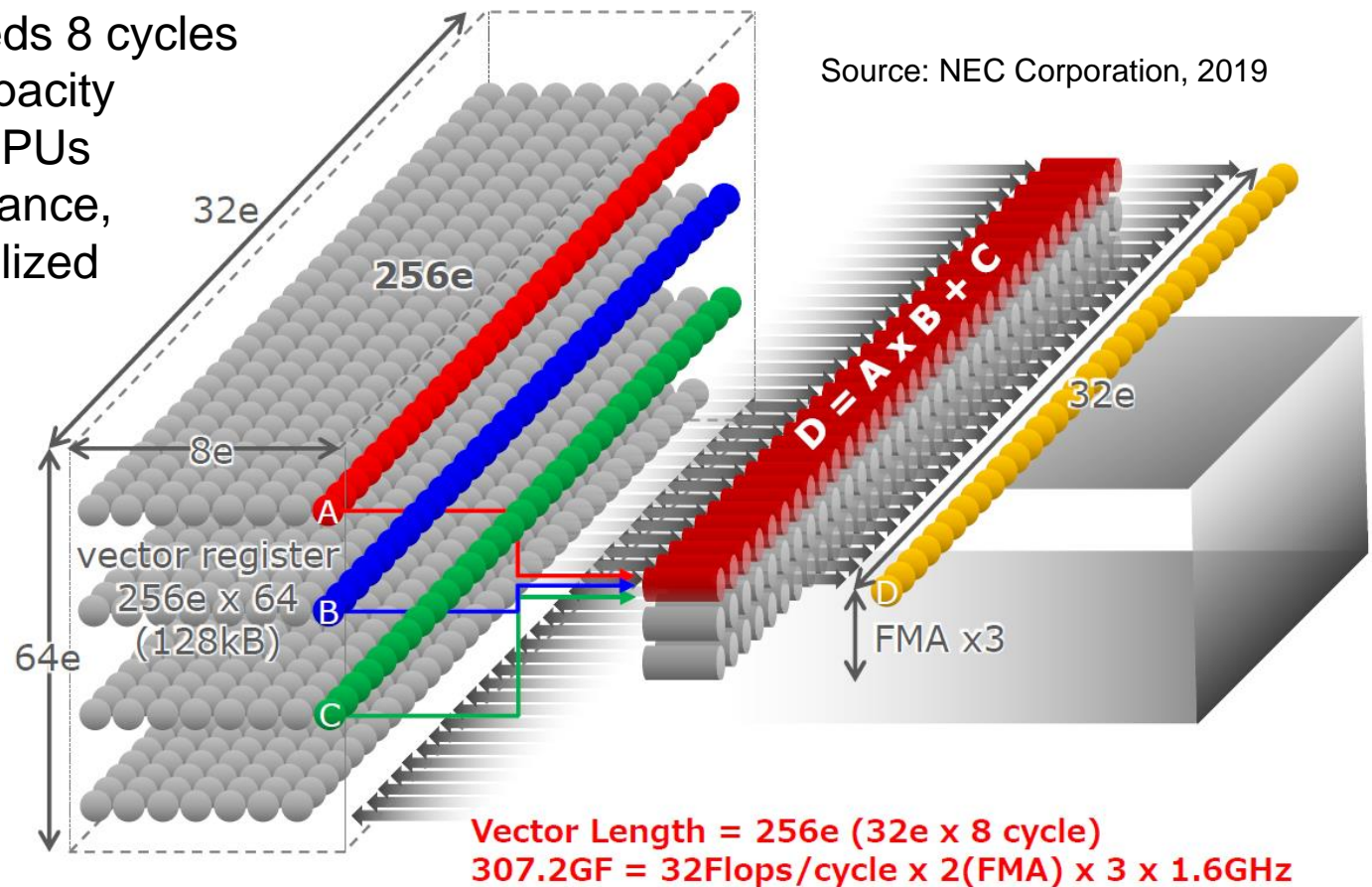


1VE

Source: NEC Corporation, 2019

# Vector Execution

- SX-Aurora combines vectors with pipelining
  - vector length = 16384 bit
  - single FMA needs 8 cycles
- Higher register capacity than L1 in many CPUs
- Very poor performance, if FMAs are not utilized



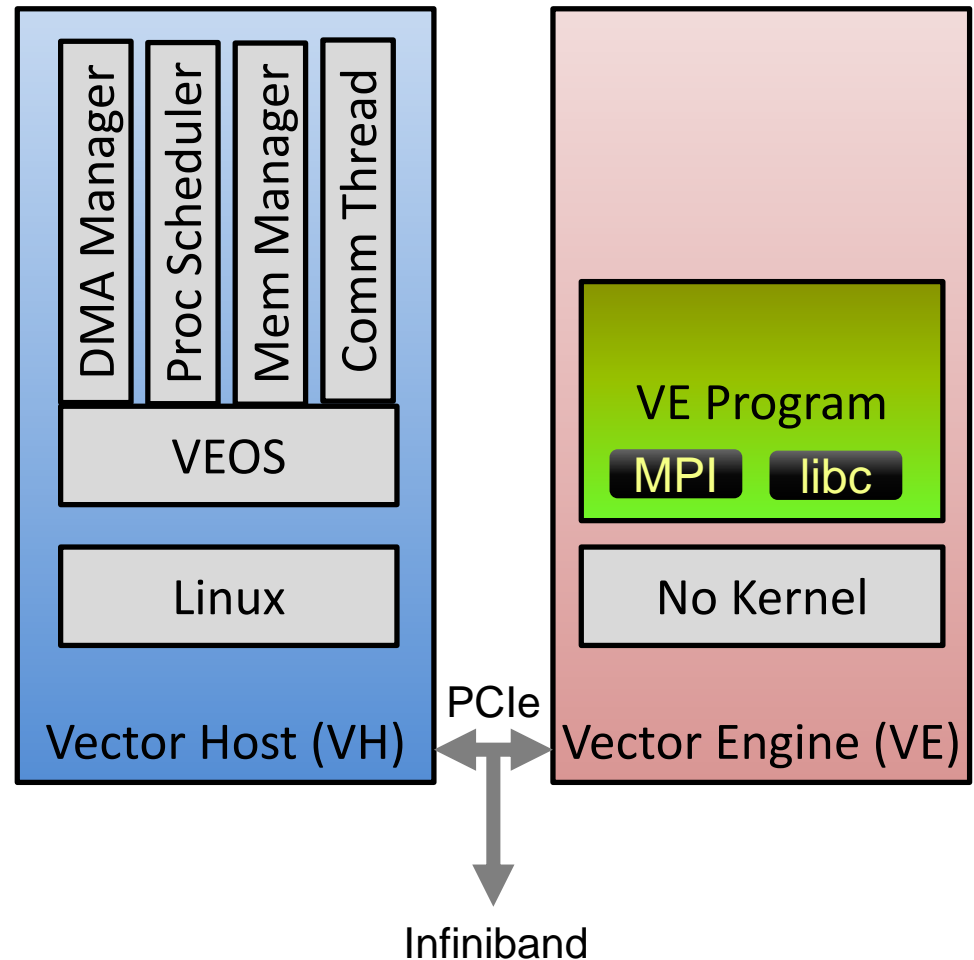
# VEOS / Installation

## VEOS

- VEOS is running on the host (Linux environment)
- Provides OS, Linux compatible and HPC functionality for VE programs
- Installation via yum repositories

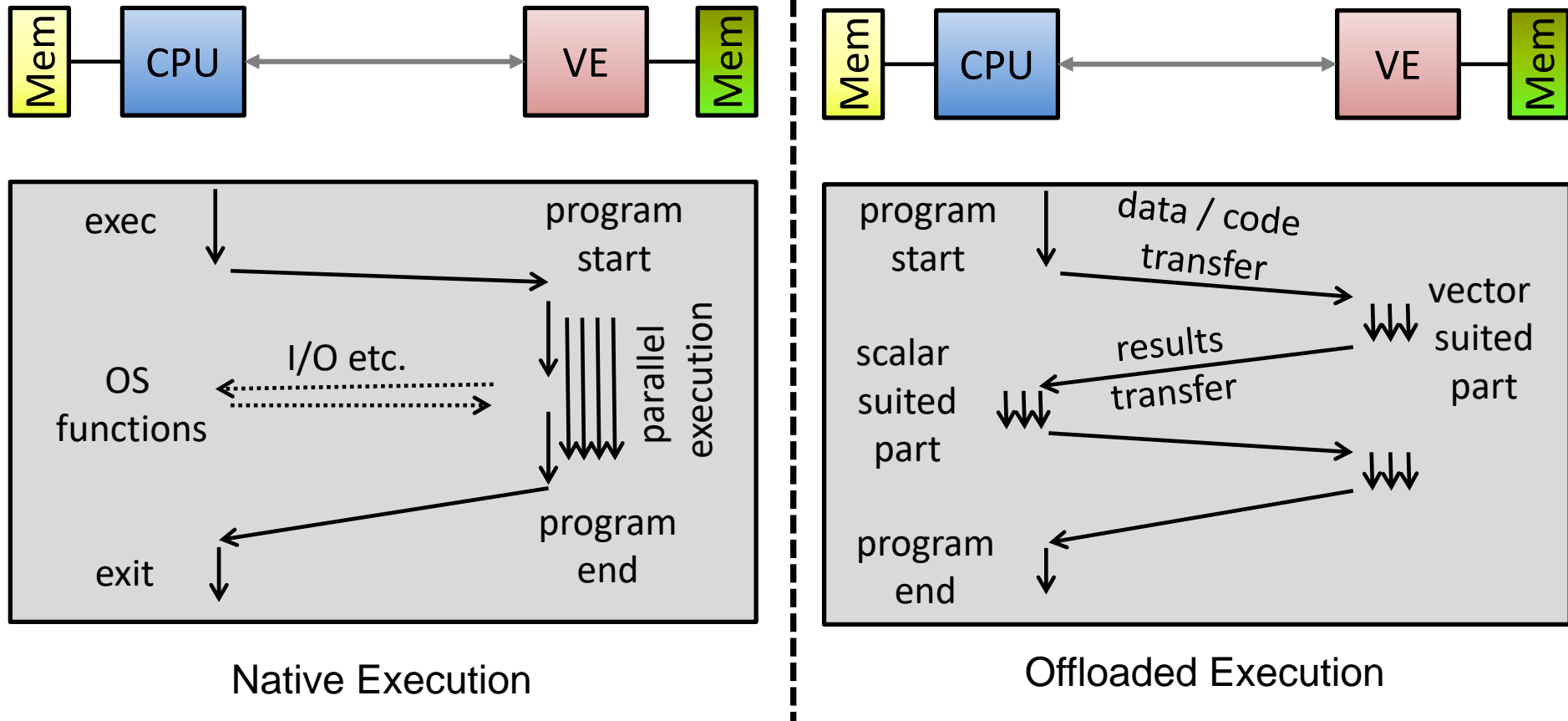
## Compiler

- C/C++ and Fortran compiler available in different versions (ncc, nfort)
- Very picky, compatible flags available
- Covers different tools (gdb, top, nm, etc.)
- Limited modules support



# Aurora Execution Models

- Supports native and (limited) offloaded execution
- “Reverse Offloading”: VHCall (limited Fortran support)



# Software Packages / MPI

---

- MPI
  - NEC MPI available (proprietary, MPI-3.1)
  - Hybrid MPI (VE + VH) possible, no experiences in Aachen yet
  - MPI communication happens on VE (in contrast to system calls)
- Numerical libraries
  - BLAS, SBLAS, LAPACK, SCALAPACK, ASL, Heterosolver



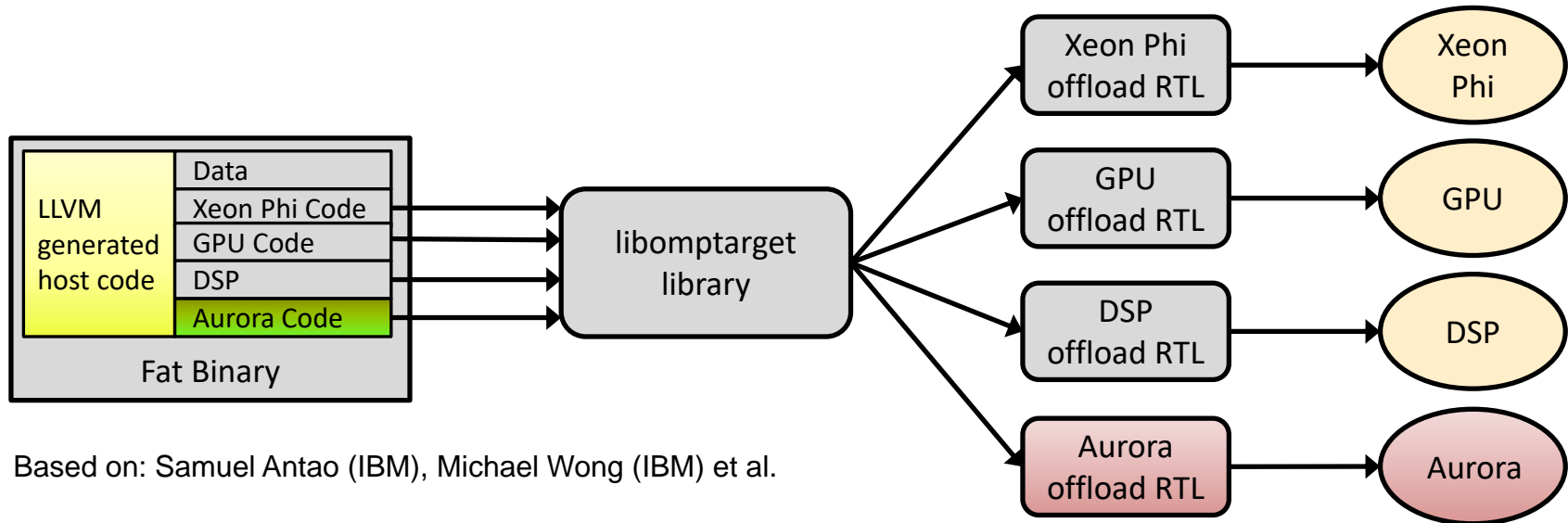
# SX-Aurora TSUBASA in Aachen

---

- A300-8: 8 VEs in a single system
- Accessible only for selected users
- Used as dialog system
  - No scheduling right now
  - Slurm plugin available in general
  - Developing, testing, performance analysis
- Although it is a quite specific system, the effort for installation / maintenance is reasonable
- Use cases in Aachen
  - Research / development of an LLVM-based offloading infrastructure
  - ML experiences (bachelor thesis)
  - Porting / Testing of applications
    - QE: Promising application
    - GROMACS: Issues with boost compilation
    - XNS (GRMES dominated): Works, performance is ok, but less than 2 Intel Xeon
    - CIAO: Work in progress
    - CP2K: Work in progress
    - LAMMPS: Some hotspots much faster than on Intel CPU, very expensive initialization

# OpenMP Offloading


- Not officially supported (yet)
- RWTH Aachen implemented a prototype based on LLVM/Clang and VEO
- Source-2-source transformation for target device code
- `$ clang -fopenmp -fopenmp-targets=aurora-nec-veort-unknown input.c`



Based on: Samuel Antao (IBM), Michael Wong (IBM) et al.


## 2<sup>nd</sup> Aurora Deep Dive Workshop

- NEC and RWTH Aachen University organizing a workshop for SX-Aurora
- Please contact me, if you interested


Bereiche ▾ Durchsuchen ▾🔍🔗 Anmelden

### Aurora Deep Dive Workshop

Programming for SX-Aurora TSUBASA



May, Thu 2 - Fr 3, 2019  
11:30 am to about 2 pm



IT Center RWTH Aachen University  
Kopernikusstraße 6  
Seminar Room 3

[Registration](#)[Agenda](#)

#### Introduction

The NEC SX-Aurora TSUBASA is a new vector processor based accelerator for HPC workloads. Since this vector engine (VE) provides a very high memory bandwidth (1.2 TB/s), it is a very promising architecture for memory bound applications. The VE is provided as a PCIe card in a x86-64 server. Each VE is equipped with 8 cores. This workshop provides detailed information about the Aurora Architecture. Experts from NEC will present the programming environment and the different execution models including hints for performance tuning. The attendees of the workshop have the opportunity to explore the system by working on provided hands-on or bringing their own code.

Gesteuert von Atlassian Confluence 5.4.3, der Kollaborationssoftware für Teams · IT-ServiceDesk · Impressum · Disclaimer · Feedback · Anfrage/Request online · SprachenChat Support

# Documentation

---

- Man Pages
  - Very limited man pages ☹️
- Compiler and library documents
  - [www.hpc.nec/documents](http://www.hpc.nec/documents)
- Aurora forum
  - [www.hpc.nec/forums](http://www.hpc.nec/forums)
- NEC Blog
  - <https://sx-aurora.github.io>
- RWTH OpenMP Offloading
  - <https://rwth-hpc.github.io/sx-aurora-offloading/>
  - <https://github.com/RWTH-HPC>