A64FX SVE evaluation and building up the Arm HPC Ecosystem

Shinji Sumimoto, Ph.D.
FUJITSU LIMITED
Nov. 18th, 2019
Outline of This Talk

- **Arm HPC Ecosystem Development**
  - Brand New Product from Fujitsu: Come and Ask at Fujitsu Booth #1517
  - Some Application Performance and SPACK

- **Performance Evaluation of SVE Enabled Arm Processor A64FX using Variable Vector Length**
Arm HPC Ecosystem Development

- Brand New Product from Fujitsu
- Some Application Performance and SPACK
New PRIMEHPC Lineup

PRIMEHPC FX1000
Supercomputer optimized for large scale computing

- High Performance
- High Scalability
- High Density

A64FX processor
384 nodes/Rack
Tofu Interconnect D
Water Cooling
Fujitsu Software Stack
for Supercomputing

PRIMEHPC FX700
Supercomputer based on standard technologies

- Fugaku Technology
- Ease to use
- Installation

A64FX Processor
8 nodes/2U Rackmount
InfiniBand
Air Cooling
Utilize ISV and Open Source Software Stack
PRIMEHPC FX1000

- High-performance Arm HPC server based on Fugaku
  - SVE: ISA extension for HPC
  - HBM2: High memory bandwidth
  - FP16: Enhancement for AI workloads

- High scalability
  - Tofu Interconnect D supports over 100,000 nodes

- High-density mounting
  - Over 1 PFLOPS per 1 rack (100x compute density over the K computer)
  - Efficient heat removal by water cooling
PRIMEHPC FX700

- High-performance Arm server featuring the A64FX CPU
  - Same CPU as Fugaku and FX1000

- Easy deployment and flexible configuration
  - Air-cooled, 2U rack-mountable chassis
  - From 2 to 8 CPUs per chassis

- Utilize open and standard technologies
  - InfiniBand
  - RHEL 8, OpenHPC, Bright Cluster Manager, etc.
## Comparison between the FX1000 and FX700

<table>
<thead>
<tr>
<th>Feature</th>
<th>FX1000</th>
<th>FX700</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPU</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>A64FX</td>
<td>A64FX</td>
</tr>
<tr>
<td>ISA</td>
<td>Armv8.2-A SVE</td>
<td>Armv8.2-A SVE</td>
</tr>
<tr>
<td>Cores</td>
<td>48 computation cores + 2-4 assistant cores</td>
<td>48 computation cores</td>
</tr>
<tr>
<td><strong>Frequency</strong></td>
<td>2.2GHz</td>
<td>2.0GHz / 1.8GHz</td>
</tr>
<tr>
<td><strong>Node</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPUs</td>
<td>1 CPU</td>
<td>1 CPU</td>
</tr>
<tr>
<td>Memory capacity</td>
<td>32GiB (HBM2)</td>
<td>32GiB (HBM2)</td>
</tr>
<tr>
<td>Memory bandwidth</td>
<td>1,024GB/s</td>
<td>1,024GB/s</td>
</tr>
<tr>
<td>Interconnect</td>
<td>Tofu Interconnect D</td>
<td>InfiniBand</td>
</tr>
<tr>
<td><strong>Enclosure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Form factor</td>
<td>Dedicated rack</td>
<td>2U rack-mountable chassis</td>
</tr>
<tr>
<td>Node per rack or chassis</td>
<td>384 node/rack</td>
<td>8 node/chassis</td>
</tr>
<tr>
<td>Heat-removal method</td>
<td>Water cooling + Air cooling</td>
<td>Air cooling</td>
</tr>
</tbody>
</table>
Software Stack on FX1000/FX700

- Fujitsu software stack provides powerful support for the A64FX CPU
- Fujitsu also plans to support OpenHPC and Bright Cluster Manager for FX700
High Performance on Real Applications

The performance on 1 node is evaluated for seven OSS applications:

- Measured on PRIMEHPC FX1000, A64FX 2.2GHz
- Up to 1.8x faster over Xeon Platinum 8268 x2
- High memory B/W and long SIMD length

Applications work effectively:

- OpenFOAM
- FrontISTR
- ABINIT
- SALMON
- SPECFEM3D
- WRF
- MPAS

Bar chart comparing FUJITSU A64FX to Xeon Platinum 8268 2CPU.
High Performance in Power Efficiency

The power efficiency on 1 node is evaluated for seven OSS applications:

- Measured on PRIMEHPC FX1000, A64FX 2.2GHz
- High efficiency is achieved by energy-conscious design and implementation

<table>
<thead>
<tr>
<th>Application</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>OpenFOAM</td>
<td>3.0</td>
</tr>
<tr>
<td>FrontISTR</td>
<td>2.0</td>
</tr>
<tr>
<td>ABINIT</td>
<td>3.0</td>
</tr>
<tr>
<td>SALMON</td>
<td>2.0</td>
</tr>
<tr>
<td>SPECFEM3D</td>
<td>3.0</td>
</tr>
<tr>
<td>WRF</td>
<td>3.0</td>
</tr>
<tr>
<td>MPAS</td>
<td>3.0</td>
</tr>
</tbody>
</table>

FUJITSU A64FX       Xeon Platinum 8268 2CPU
OSS Build Verification on the aarch64

RIKEN and Fujitsu verify Spack’s 3000+ recipes with multiple compilers

Spack is a package manager for supercomputers

- The verification results are available on the RIKEN’s web page
  https://postk-web.r-ccs.riken.jp/oss/public/

- Summary of results
  - Spack version used: September 6, 2019
  - Pull Requests made: RIKEN 47, Fujitsu 46

<table>
<thead>
<tr>
<th></th>
<th>GNU GCC</th>
<th>Fujitsu Compiler (under development)</th>
</tr>
</thead>
<tbody>
<tr>
<td>x86</td>
<td>2479/3451 (71.8%)</td>
<td>N/A</td>
</tr>
<tr>
<td>aarch64</td>
<td>2387/3451 (69.1%)</td>
<td>2072/3451 (60.0%)</td>
</tr>
</tbody>
</table>
Performance Evaluation of SVE Enabled Arm Processor A64FX using Variable Vector Length

Presented at Arm Research Summit 2019
Overview

- Background
- A64FX Overview
- Application Characteristics: Compute Intensive vs. Memory Intensive
- Preliminary Performance Evaluation
Background

- A64FX is the First SVE enabled Arm Processor in the world.
  - SVE realizes single binary for multiple vector length environment in order to support application binary portability.
  - A64FX supports 512, 384, 256, 128 bit vector length.

- HPC applications have several characteristics such as compute intensive and/or memory intensive.
  - SVE enabled processor can control execution vector length at runtime.
  - A64FX has a memory bandwidth controlling feature at runtime.
  - No re-compilation is needed for the above executions.

- Therefore, we have evaluated several application benchmarks in order to clarify application characteristics
A64FX: High Performance Arm CPU

- Inheriting Fujitsu HPC CPU technologies with commodity standard ISA
High Performance Arm CPU “A64FX”

**Architecture features**

<table>
<thead>
<tr>
<th>Specification</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISA</td>
<td>Armv8.2-A (AArch64 only) SVE (Scalable Vector Extension)</td>
</tr>
<tr>
<td>SIMD width</td>
<td>512-bit</td>
</tr>
<tr>
<td>Precision</td>
<td>FP64/32/16, INT64/32/16/8</td>
</tr>
<tr>
<td># of cores</td>
<td>48 computing cores + 4 assistant cores (4 CMGs)</td>
</tr>
<tr>
<td>Memory</td>
<td>HBM2: Peak B/W 1024 GB/s</td>
</tr>
<tr>
<td>Interconnect</td>
<td>TofuD: 28 Gbps x 2 lanes x 10 ports</td>
</tr>
</tbody>
</table>

Copyright 2019 FUJITSU LIMITED
A64FX Memory System

- Extremely high bandwidth
  - Asynchronous Processing in cores, caches and memory controllers
  - Maximizing the capability of each layer’s bandwidth

Performance: >2.7TFLOPS
L1 Cache: >11.0TB/s (BF= 4)
L2 Cache: >3.6TB/s (BF = 1.3)
Memory: 1024GB/s (BF =~0.37)
Application Characteristics: Compute Intensive vs. Memory Intensive
SVE: Scalable Vector Length

- ISA does not fix Vector Length
  - SVE supports VL from 128 to 2048 bit with multiples of 128 bit
  - VL is set by processor before executing a binary dynamically

- Single execution binary can be executed on processors with multiple VLs
  - Vector-Length Agnostic (VLA) programming enables ABI Compatibility

Execution Binary Portability

Execution Binary does not depend on processor’s VL

- Increasing dynamic instruction steps to double
- Reducing dynamic instruction steps to half

512bit SIMD → Execution Binary/a.out → 256bit SIMD
Application Characteristics: How to investigate?

- Application Characteristics
  - Computing Intensive
  - Memory Bandwidth Intensive

- How to investigate?
  - Using performance profiling tools: Arm Allinea Studio, OSS tools
  - Re-compiling with different compiler option and running

- A64FX can help to evaluate the characteristics easily
  - Compute Intensive Analysis: Changing Vector Length
  - Memory Intensive Analysis: Changing Memory Access Gap
Preliminary Performance Evaluation
Benchmark Applications, Evaluation Environment and Evaluation Parameters

**Benchmark Applications**
- STREAM (TRIAD)
- DGEMM
- Himeno Benchmarks
- NAS Parallel Benchmarks OMP Class C (EP, CG, LU, FE, IS, MG, BT, SP)

**Evaluation Environment**
- Hardware: Fugaku-Prototype System with A64FX (single node, 2.0GHz)
- Compiler: Fujitsu Compiler (development version)

**Evaluation Parameters**
- SVE Vector Length: 128 - 512
  - Using a command with prctl(PR_SVE *)
- Memory Bandwidth: 100%-20%
  - Using an option of submission job

<table>
<thead>
<tr>
<th>Application</th>
<th>VL 128</th>
<th>VL 256</th>
<th>VL 512</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBM 100%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBM 80%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBM 60%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBM 40%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBM 20%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* https://lwn.net/Articles/717804/
STREAM TRIAD: 1 Thread vs. 12 Threads (1CMG)

Relative Performance VL512-HBM100% = 100
STREAM TRIAD: 1 Thread vs. 12 Threads-SIMD Effects

- Relative Performance VL128-HBM 100% = 100
DGEMM: 12 Threads (1CMG)

- Relative Performance VL512-HBM100% = 100
Himeno Benchmark: 1 CMG vs. 4 CMG

Relative Performance VL512-HBM 100% = 100

Himeno OpenMP 12 thread

Himeno OpenMP 48 thread
Himeno Benchmark: 1 CMG vs. 4 CMG-SIMD Effects

Relative Performance VL128-HBM100% = 100
NPB: with No SMID Effect and Compute Intensive

- Relative Performance VL512-HBM100% = 100
NPB: with SMID Effect and Compute Intensive

Relative Performance VL512-HBM100% = 100
NPB: with SMID Effect and Compute Intensive and Memory Intensive(1)

Relative Performance VL512-HBM100% = 100
NPB: with SMID Effect and Compute Intensive and Memory Intensive(2)

- Relative Performance VL512-HBM100% = 100
Summary

- Arm SVE provides Application Binary Portability
  - Variable Vector Length is defined at runtime

- SVE feature can be used application performance analysis
  - Reduction of vector length shows application is compute intensive or not

- Benchmark Evaluation on A64FX Enables:
  - Compute Intensive Analysis: Changing Vector Length
  - Memory Intensive Analysis: Changing Memory Access Gap
shaping tomorrow with you