HPS

Tim Dettmar

ARM Evaluation in HPC

Project Preliminary Results

Table of contents

Outline

Evaluation Hardware

ARM Nodes

- Huawei TaiShan 200 (28x)
  - Dual Kunpeng 920 ARMv8 CPUs (2.6GHz, 2x64C)
  - 128GB RAM
  - 100Gbit/s InfiniBand
- NVIDIA ARM Developer Kit (2x)
  - Ampere Altra Q80-30 (3.0GHz, 80C)
  - 512GB RAM
  - 200Gbit/s BlueField-2 DPU with InfiniBand
    - ARM Cortex A72 (2.0GHz, 8C)
    - 16GB RAM
- Limited access time

Reference GWDG Nodes

x86 Nodes

- Xeon Server
  - Dual Xeon Platinum 9242 (2x48C)
  - 384GB RAM
  - 100Gbit/s Omni-Path

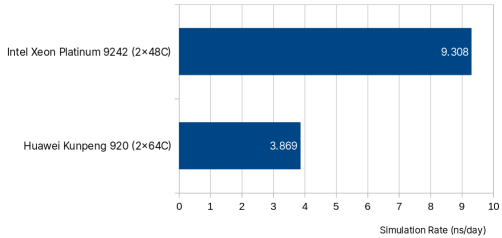
Preliminary Benchmarks

- Preliminary benchmarks were run using Phoronix Test Suite
  - HPL Linpack
  - GROMACS
- Can compare with other CPUs easily, but more tuning is needed
- Benchmark configuration is not too well optimized
- Core-to-core latency benchmark was also run

https://www.phoronix-test-suite.com/

GROMACS Performance

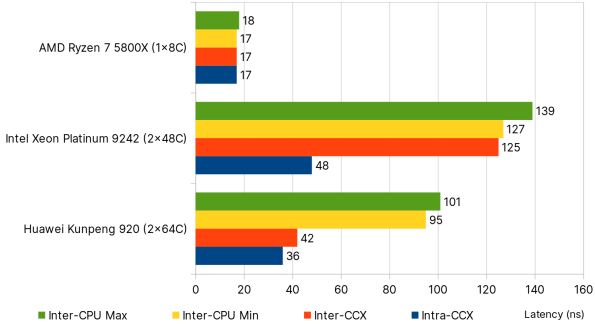
- Large discrepancy between x86 and ARM platform
- More investigation needed
  - Possible BLAS library optimization
  - Memory limitations
  - Hybrid OpenMP/OpenMPI



Preliminary Data; PTS GROMACS Benchmark 1.70, GROMACS 2022.1, MPI CPU, water_gmx50_bare, All Cores across All NUMA Nodes

Core-Core Latency

- Average Latencies
  - Consistently better latency on ARM platform
  - Much better inter-CPU latency
  - Generally: More cores, higher latency between them



CCX = Core Complex, Preliminary Data;
`https://crates.io/crates/core-to-core-latency`,
exclusive node access

Core-Core Latency

- Depends on the CPU design, not necessarily ARM vs x86
  - Monolithic die: traditional
  - Multi-die
  - Chiplet
- What the manufacturer chose and how it’s implemented matters

Outline

Existing Benchmarks

- Mostly focused on the older generation ThunderX2
- Some primary drawbacks
  - Narrow vector unit (128-bit) $\rightarrow$ Xeon 9242 has 512-bit vector unit
  - Lower overall performance (25%)
  - Around 10% better performance per joule
  - Highly variable compiler performance
    - ARM compiler is **worse** than GCC!
  - Similar scalability performance to x86

[Man+20]

Existing Benchmarks

- One report focusing on Ampere Altra processors
- Mostly focused on GPU accelerated workloads
- CPU Hydrodynamics workload (SPH-EXA2 Sedov-Taylor blast wave explosion)
  - ThunderX2 ARM (2S/64C/256T): 6.7 iter/min (0.10/core)
  - Ampere Altra ARM (1S/80C): 9.9 iter/min (0.12/core)
  - EPYC 7762 x86 (2S/128C/256T): **31.9 iter/min** (0.25/core)

[Elw+22]

Burning Questions

- Is ARM falling behind as the preliminary results suggest?
- What about power consumption?
- Is this down to optimization?

Outline

Software Evaluation

- Portability evaluation targeted
  - Optimizations are great
    - ... but not if you have to change your entire workflow
    - ... and not if it takes forever to compile
  - Preliminary testing suggests high degree of application portability
  - Benchmark data also suggest that default optimizations may be lacking
  - Some programs use hand-tuned assembly / intrinsics for critical paths
- Limitations are self-imposed on portability
  - Compile on intended application toolchain (e.g. gcc/clang + cmake) without significant hand-tuning or knowledge
    - Just adding `-O2` or `-march=` is probably fine
    - Switching compilers + compiling subdependencies from source, probably not

Software Tuning

- Optimizing Linpack configuration for high performance
- Use ARM-optimized libraries where **reasonably** possible

Scalability Testing

- Due to limited nodes, only scaling to all cores on one node
  - 1 core → 1 core group → 1 CPU → N CPUs
  - Ampere Altra: 1C → 80C
  - Kunpeng 920: 1C → 32C → 64C → 128C
  - Xeon 9242: 1C → 24C → 48C → 96C

Current Concerns

- Limited software availability on NVIDIA ARM nodes
- Mismatched BLAS, OpenMPI, etc. on different clusters
- Older kernel version on GWDG compute cluster
  - Missing optimizations?
  - Benchmark fairness?
- Getting reliable power consumption data is still an issue

Summary

- ARM is still an interesting HPC architecture
- Preliminary benchmarks showed poor performance, reason unknown
- Some issues getting access to the cluster
- Some time investment needed to get software running well
- Further evaluation will be performed
  - Optimized x86 vs Optimized ARM
  - Scalability Benchmarks
  - Power Consumption

References

[Elw+22] Wael Elwasif et al. *Application Experiences on a GPU-Accelerated Arm-based HPC Testbed*. 2022. DOI: 10.48550/ARXIV.2209.09731. URL: https://arxiv.org/abs/2209.09731.

[Man+20] Filippo Mantovani et al. “Performance and energy consumption of HPC workloads on a cluster based on Arm ThunderX2 CPU”. In: *Future Generation Computer Systems* 112 (2020), pp. 800–818. ISSN: 0167-739X. DOI: https://doi.org/10.1016/j.future.2020.06.033. URL: https://www.sciencedirect.com/science/article/pii/S0167739X19309781.