Evaluating ARM for Use in a Production Environment

Chris Hollowell <hollowec@bnl.gov>, Tom Smith <tsmith@bnl.gov>
Scientific Data and Computing Center (SDCC)

Throughput Computing 2023 - 7/12/2023

ARM Background

- RISC CPU architecture which was first introduced in the ‘80s
  - Arm (the company) does not sell CPUs, but licenses the ARM architecture and core design to other vendors

- Explosion in the use of ARM CPUs in the past decade, primarily due to popularity in mobile devices
  - 64-bit ARM (aarch64) CPUs introduced in 2012
  - After the launch of the M1 CPU in 2020, now the architecture in use in all of Apple’s new laptops/desktops

- Several ARM-based server CPUs available for the general market
  - Ampere Altra (Neoverse N1 core-based)
  - HPC-oriented CPUs
    - Nvidia Grace (Neoverse V2 core-based)
    - Fujitsu A64FX (Custom core-based)

- ARM CPUs have become increasingly available at commercial cloud providers
  - Amazon has made Graviton instances available since 2018
  - Google and Azure introduced Ampere-based ARM instances in 2022
    - Google reportedly also in the process of developing their own custom ARM CPUs

Arm's Neoverse Architecture

- Arm's Neoverse core architecture has facilitated/accelerated ARM adoption in server and HPC markets
  - Provides the building blocks needed to construct a server class CPU, reducing time/effort needed by CPU vendors - apparent in the number of commercial CPUs based on the core design
  - N series - server/datacenter applications
    - N1 and N2 currently available
  - V series - HPC CPUs
    - V1 and V2 available
    - Support for HBM (High Bandwidth Memory), additional or wider SVE (Scalable Vector Extension) units
  - E series - for edge computing use cases
    - Focus on power efficiency

- Ampere switched to a custom core-design for the Ampere One CPU

Potential ARM Advantages

- Energy efficiency
  - ARM's energy efficiency has been one of the primary drivers of adoption in mobile devices
  - Does it hold true for server applications?
    - A study performed by E. Simili, et al. at University of Glasgow presented at CHEP 2023 comparing HS23 score and power usage for an ARM Altra Q80 vs AMD 7643 shows promising results

Plot from CHEP2023's "ARMing up for HEP" - E. Simli, et al.

- Purchase cost
  - Retail pricing for Ampere's CPUs seem considerably lower than similarly spec'd x86 counterparts
  - Additional $/HS23 studies needed to confirm

LHC Software ARM64 Readiness

- The LHC experiments have been working on porting their code to ARM over the past several years, and significant progress has been made in this area
  - Physics validations may not be complete in some cases

- The HEPscore23 benchmark fully supports ARM
  - Includes ARM-compiled workloads for all LHC experiments, and Belle-II:
    - atlas-gen_sherpa-ma-bmk
    - atlas-reco_mt-ma-bmk
    - cms-gen-sim-run3-ma-bmk
    - cms-reco-run3-ma-bmk
    - lhcb-sim-run3-ma-bmk
    - belle2-gen-sim-reco-ma-bmk
    - alice-digi-reco-core-run3-ma-bmk
  - The same sub-benchmark set used for x86_64

- ATLAS simulation and reconstruction software fully functional and validated on ARM
  - The ATLAS Experiment Software On ARM - J. Elmsheuser, et al., CHEP 2023
  - Planned evaluation/validation of additional production workflows

OSG Software ARM64 Readiness

- RHEL/Rocky/Alma 8/9 all support aarch64
- OSG-related packages currently only provided for x86_64 in https://repo.opensciencegrid.org

Index of /osg/3.6/el9/release

Name	Last modified	Size	Description
Parent Directory		-	
source/	2023-07-07 17:08	-	
x86_64/	2023-07-07 17:09	-	

- HTCondor 10.x builds for aarch64 are available in https://research.cs.wisc.edu/htcondor/repo/
- Apptainer aarch64 packages are available for EL8/9 in EPEL
- CVMFS aarch64 packages are available: https://cvmrepo.web.cern.ch/cvmrepo/yum/cvmfs/EL/

ARM OSG Queue at BNL/SDCC

- ***ARM is increasing in popularity, and it appears it will become a viable platform for production LHC workloads***
  - Initial studies suggest the platform may offer significant energy efficiency and purchase cost benefits to sites

- Therefore, as part of an ATLAS R&D milestone, we plan to evaluate the deployment of an OSG ARM CPU-based queue/site at the US ATLAS T1, and will work with PanDA developers to enable the utilization of these resources
  - Purchased two Ampere Altra-based ARM nodes for the queue, and will provision with Alma 9
  - Will evaluate/debug the deployment of OSG software (HTCondor-CE, HTCondor, etc.) on the systems
    - It may be necessary to build some of the aarch64 packages ourselves
  - Plan to test if cross-platform job submission from CEs (i.e. CE x86_64, aarch64 batch node) is possible/functional
  - Will conduct performance (HEPscore23) and energy studies

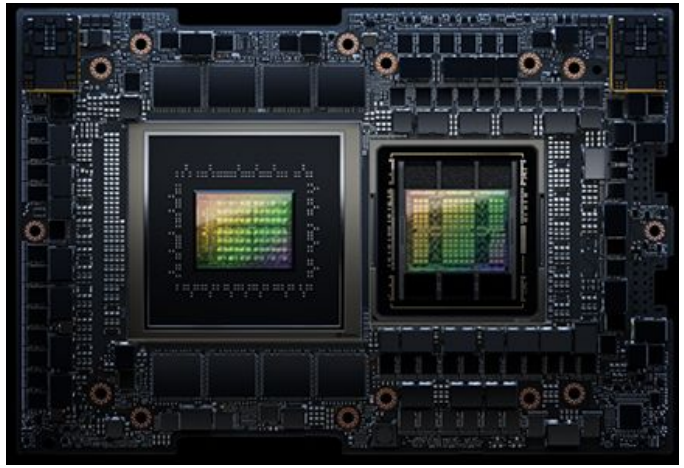
ARM Evaluation System Specs

- Two Supermicro MegaDC ARS-110M-NR systems with different configurations were purchased:
  - One with an Ampere Altra with Q64-2.6 CPU:
    64 physical cores @ 2.6 GHz
    - 256 GB RAM (8x32 GB DDR4-3200 DIMMs)
    - 4x2 TB NVMes
  - One with an Ampere Altra max with M128-2.8 CPU:
    128 physical cores @ 2.8 GHz
    - 512 GB RAM (8x64 GB DDR4-3200 DIMMs)
    - 6x2 TB NVMes

  - *Note: Ampere Altra CPUs do not support SMT/hyperthreading, and have 8 memory channels*
  - Systems will be configured to utilize one 10 Gbps NIC
  - Expect the nodes to arrive in August/September
  - Interested in evaluating both lower-end and higher-end CPU models
    - The Q64-based host will be used as a CE if there are issues with cross-architecture job submission

Nvidia Grace

- Nvidia will begin shipping Grace ARM superchips later this year
  - Available in CPU/GPU and CPU/CPU combinations
    - ARMv9, Neoverse V2 core-based
    - CPUs or CPU/GPUs interconnected with NVLink
    - Provide on-chip High Bandwidth Memory (HBM)
      - Up to 1 TB/s memory access
  - GPU versions shipping with Hopper H100
- Looking into the possibility of purchasing one or more Grace-Grace based systems for evaluation
  - These could also eventually be included in the test queue
  - May evaluate Grace-Hopper later if there is interest from our HPC community



Nvidia Grace-Hopper

Nvidia Grace-Grace

Conclusions

- ARM is becoming an increasingly available/popular x86 alternative for server/HPC use in datacenters and cloud environments
  - Neoverse core designs have accelerated the process
  - A number of vendors are now offering server/HPC ARM CPUs

- It appears ARM will become a viable platform for production LHC experiment workflows in the near future
  - ATLAS has already validated their simulation and reconstruction software on this architecture
  - Initial studies suggest there may be cost and energy efficiency benefits to using this platform

- The US ATLAS T1 at BNL will be evaluating the use of ARM in our computing environment, and plan to create a test ARM OSG queue
  - Will consist of two recently purchased Ampere Altra-based hosts, which are expected to be delivered in August/September