CSCE 735 Minor Project (Spring 2024) [UIN: 734003886]

Supercomputers play a crucial role in advancing computational capabilities across scientific research, engineering, and data analysis. Known for their high FLOPS processing speeds, extensive memory and storage capacities, and advanced parallel processing architectures, they drive innovation in weather forecasting, climate modeling, drug discovery, astrophysics, and artificial intelligence. Comparing leading supercomputers like Frontier and Fugaku, currently ranked no. 1 and no. 4 in the TOP500, helps determine their suitability for specific applications. This paper explores various aspects of both, including architecture, technical specifications, use cases, and performance, offering insights into optimizing cutting-edge capabilities.

Frontier is the world's first exascale supercomputer built in collaboration with Oak Ridge National Laboratory, HPE Cray and AMD. The machine was built at a cost of US $600 million. It began deployment in 2021 and reached full capability in 2022. It clocked 1.194 Exaflops in May 2022, making it the world's fastest supercomputer as measured in the June 2022 edition of the TOP500 list [7]. Frontier utilizes AMD EPYC 64C 2GHz processors and is based on the latest HPE Cray EX235a architecture. The system has a total of 8,699,904 combined CPU and GPU cores. Additionally, Frontier has a power efficiency rating of 52.59 GFlops/watt and relies on HPE's Slingshot 11 network for data transfer [1]. These specifications make Frontier a highly efficient and powerful tool used for scientific research and engineering including materials science, nuclear physics, astrophysics, and machine learning.

Frontier uses 9,472 AMD EPYC 7453s "Trento" 64 core 2 GHz CPUs (606,208 cores) and 37,888 Radeon Instinct MI250X GPUs [2]. A high-level diagram of a single Frontier Node is given in Figure 1. They can perform double precision operations at the same speed as single precision. "Trento" is an optimized 3rd Gen EPYC CPU, which is based on the Zen 3 microarchitecture. It occupies 74 19-inch (48 cm) rack cabinets. Each cabinet hosts 64 high density compute blades, each consisting of 2 nodes. Blades are interconnected by HPE Slingshot 64-port switches that provides 12.8 terabits/second of bandwidth. Groups of blades are linked in a dragonfly topology with at most three hops between any two nodes. Cabling is either optical or copper, customized to minimize cable length. Total cabling runs 145 km (90 mi). Frontier is liquid-cooled, allowing 5x the density of air-cooled architectures. Each node consists of one CPU, 4 GPUs and 4 terabytes of flash memory. Each GPU has 128 GB of RAM soldered onto it. Frontier has coherent interconnects between CPUs and GPUs, allowing GPU memory to be accessed coherently by code running on the EPYC CPUs [8]. Frontier consumes 21 megawatts (MW), compared to its predecessor Summit's 13 MW.

The CPUs developed by AMD have two unique features - AMD Infinity Guard and AMD Infinity Architecture [9]. AMD Infinity Guard is responsible for ensuring data security, while Infinity Architecture increases scalability and performance. Furthermore, Frontier uses HPE's Cray OS software, which is designed to run scalable and highly complex applications. The supercomputer consumes 22,703 kW of power and has a Linpack benchmark of 1,194 PFlop/s and a theoretical peak of 1,679.82 PFlop/s.

Moreover, the HPE Cray Slingshot interconnect, utilized in the Frontier supercomputer, features a dragonfly network topology as shown in Figure 2, offering high bandwidth and low latency communication between nodes [10]. With aggregate bandwidths in the terabits per second range, it ensures efficient data exchange for large-scale parallel computations. Adaptive routing algorithms optimize communication performance, while fault tolerance mechanisms enhance reliability. Integrated with the HPE Cray software stack, the Slingshot interconnect enables seamless system management and programming, contributing to Frontier's exceptional performance and scalability.

Frontier hosts a diverse range of scientific, data analysis, and computational software. It includes simulation tools like ANSYS Fluent and LAMMPS, alongside data analysis software such as MATLAB, Python (with libraries like NumPy and pandas), and R. High-performance computing libraries like MPI and CUDA support parallel tasks, while machine learning frameworks like TensorFlow and PyTorch aid AI research. It also accommodates bioinformatics tools like BLAST and GROMACS, quantum computing simulators like Qiskit, and computational chemistry software like Gaussian and NWChem. Additionally, it supports Earth system models like CESM and WRF, CFD software like OpenFOAM, and financial modeling tools. This software suite empowers researchers to tackle complex challenges on Frontier Supercomputer.

The Frontier Supercomputer utilizes advanced parallelization techniques across distributed and shared memory architectures. Task parallelism is achieved through MPI for concurrent execution across nodes, and OpenMP enables parallel execution within shared memory nodes. It also employs data parallelism with SIMD instructions and GPU acceleration via CUDA and OpenACC [4]. Hybrid parallelism combines distributed and shared memory approaches, while specialized libraries streamline parallel I/O. Dynamic parallelism dynamically spawns new threads or processes as needed. These techniques enable Frontier to efficiently handle diverse computational challenges, solidifying its status as a pioneering solution for scientific research and engineering.

In contrast, Fugaku, located at the RIKEN Center for Computational Science in Kobe, Japan, is a petascale supercomputer jointly developed by Fujitsu and RIKEN, completed in March 2021. It is dedicated to exploring vast amounts of data for diverse applications, including drug discovery, weather forecasting, and disaster prevention. Fugaku has also been instrumental in supporting research related to the COVID-19 pandemic. With a performance exceeding 442 petaflops, Fugaku utilizes Fujitsu's A64FX CPU, based on the Arm architecture [3]. The supercomputer comprises 158,976 nodes and incorporates the second generation of High Bandwidth Memory, operating at 1024 GB/s. Designed to conform to the ARMv8.2-A SVE 512-bit architecture, it supports 7,630,848 cores. Fugaku's power consumption was optimized to 26248.36 kW, boasting an Rmax of 442.01 PFlop/s and an Rpeak of 537.21 PFlop/s.

The A64FX processor features four memory core groups, each with thirteen cores (twelve for computation and one assistant core) as shown in Figure 3. It includes Level 2 cache, HBM2 memory, and a ring bus interface connecting TofuD Interconnect and PCIE Interface. It utilizes 32 Gb of HBM2 memory with a bandwidth of 1 Tb/s. Storage includes Light Weight Layered IO Accelerator for job execution, Lustre-based file system for shared areas, and Cloud Storage services. Branch prediction employs piecewise detection of simple loop structures. Instruction decoding and resource management allocate multiple virtual Store/Fetch ports. A gather-load instruction writes to registers after reading multiple element data from memory. The level 2 cache consists of local and global pipelines, managing cache coherence via TAG Directory (TAGD) in the global pipeline.

The four Core Memory Groups (CMG), Interrupt Controller, and Tofu/PCIE Controller are connected via two ring buses with six ring stops as shown in Figure 4. The A64FX processor links adjacent CMGs using CMG interconnect paths, each offering over 115Gb/s throughput.

TofuD connects Input/Output devices, mutually linking CPUs and PCIs. It features twenty 28 Gbps lanes with a 6.8 Gbps bandwidth, connecting up to 10 CPUs. PCIe includes sixteen 8 Gbps lanes with a 16 GBps bandwidth. TofuD utilizes a six-dimensional mesh/torus topology, scalable to over 100,000 nodes, with full-duplex links peaking at 10 Gbps. Each node connects to an InterConnect Controller (ICC) chip, housing four Tofu interfaces and a router. TofuD features six Tofu Network Interfaces (TNI), interconnecting with 10 CPUs via a network router. An overview of the TofuD interconnect system is shown in Figure 5.

Fugaku's software includes Fujitsu's in-house MPI, based on OpenMPI, and RIKEN-MPICH, based on MPICH. These two interfaces serve different purposes within the supercomputer. OpenMPI focuses on usage and network conduits, while MPICH is designed as a high-quality implementation for niche cases. In addition to hardware, Fugaku employs a multi-kernel operating system that includes Red Hat Enterprise Linux 8 (RHEL) and McKernel, both running on the Interface for Heterogeneous Kernels (IHK) [5]. McKernel is dedicated to running high-performance simulations, while RHEL is used for other services (e.g., POSIX-related). Both kernels contribute to high-power computing and greatly reduce machine noise. The $1 billion computer has been used to examine the spread of COVID-19 droplets from the mouth and potential virus treatments. Fugaku has also been utilized for advancements in solar and fuel cells.

Fugaku has been utilized in various applications, including innovative drug discovery infrastructure to accelerate precision medicine, human-scale whole brain simulation with connectome analysis, and addressing the heart failure pandemic by integrating a multi-scale heart simulator with clinical data. It has also contributed significantly to combating COVID-19 by predicting virus droplet infection indoors, analyzing pandemic phenomena through simulation, and conducting host genetic analysis for severe cases. The supercomputer utilizes Fortran, C++, GNU, OpenMP, and Java compilers, with programming primarily handled by Xcapable MP and FPDPS. Key math libraries include BLAS, LAPACK, and ScaLAPACK.

The Fugaku supercomputer supports multiple forms of parallelism, such as task parallelism through MPI and thread-level parallelism with OpenMP, leveraging the capabilities of the Arm instruction set utilized by Fugaku's processor architecture. The Arm SIMD instruction set, Scalable Vector Extension (SVE), with its 512-bit width SIMD arithmetic units, facilitates data-level parallelism, enabling simultaneous processing of multiple data elements [6]. Additionally, the use of a Lustre file system in the second layer of storage supports task-level parallelism, complementing Fugaku's support for task parallelism. Fugaku's support for various forms of parallelism, coupled with the capabilities of the Arm instruction set and storage system, enables efficient utilization of computational resources and accelerates a wide range of applications.

In conclusion, Frontier and Fugaku are advanced high-performance computing systems tailored for scientific data processing and analysis. Both excel in fundamental capabilities such as FLOPS and parallel processing, enabling efficient handling of large datasets. However, they diverge in performance metrics: Frontier leads with 1.194 exaflops, while Fugaku boasts superior power efficiency and larger memory capacity. Understanding these differences is crucial for selecting the most suitable system for specific scientific computing needs, highlighting the importance of ongoing evaluation in advancing scientific research.

Comparison between Frontier and Fugaku supercomputers:

Parameter	Frontier	Fugaku
Manufacturer	HPE	Fujitsu
Processor	AMD Optimized 3rd Gen EPYC	Fujitsu A64FX
Operating System	HPE Cray OS	Linux (RedHat 8) and McKernel
Node	1 AMD EPYC CPU 4 AMD Radeon Instinct GPUs	Fujitsu A64FX 48C Armv8.2-A SVE 512 bit
No. of Nodes	9,472	158,976
Total Cores	8,699,904	7,630,848
Total System Memory	9.2 PB	4.85 PB
Frequency	2 GHz	2.2 GHz
Interconnects	Slingshot-11	Tofu Interconnect D
Power Consumption	22,703.0 kW	26,248.36 kW
Theoretical Peak (Rpeak)	1,679.82 PFlops/s	537.21 PFlop/s
Linpack Performance (Rmax)	1,194 PFlops/s	442.01 PFlops/s
HPCG Benchmark	14,054.00	16,004.50
Cost	$600 Million	$1.2 Billion

References:

[1] "TOP500 List | November 2023" TOP500.
https://www.top500.org/lists/top500/2023/11/

[2] "Frontier Supercomputer" Wikipedia.
https://en.wikipedia.org/wiki/Frontier_(supercomputer)

[3] "Fugaku Supercomputer" Wikipedia.
https://en.wikipedia.org/wiki/Fugaku_(supercomputer)

[4] "Frontier Specifications" Oak Ridge Leadership Computing Facility,
https://www.olcf.ornl.gov/wp-content/uploads/2019/05/frontier_specsheet.pdf

[5] "Fugaku Specifications" Fujitsu Global.
https://www.fujitsu.com/global/about/innovation/fugaku/specifications/

[6] Matsuoka, Satoshi. "Overview of Fugaku and its Future Perspectives".
https://science.osti.gov/-/media/ascr/ascac/pdf/meetings/202107/ASCAC_meeting_202107-Fugaku.pdf?la=en&hash=778587C4FA8139C1165B3DBDD31073A034819CBA

[7] Butler, Georgia. "Frontier remains world's most powerful supercomputer on Top500 list". Data Center Dynamics, 14 November 2023.
https://www.datacenterdynamics.com/en/news/frontier-holds-onto-number-one-spot-on-top500-list/

[8] J. Dongarra, A. Geist. "Report on the Oak Ridge National Laboratory's Frontier System". University of Tennessee, 30 May 2022.

https://icl.utk.edu/files/publications/2022/icl-utk-1570-2022.pdf

[9] A. Geist. "ORNL's Frontier Exascale Computer". U.S. Department of Energy, 27 August 2019.
https://smc.ornl.gov/wp-content/uploads/2019/09/Geist-presentation-2019.pdf

[10] S. Atchley. "Frontier's Architecture". U.S. Department of Energy, 12 July 2022.
https://olcf.ornl.gov/wp-content/uploads/Frontiers-Architecture-Frontier-Training-Series-final.pdf

Figures:

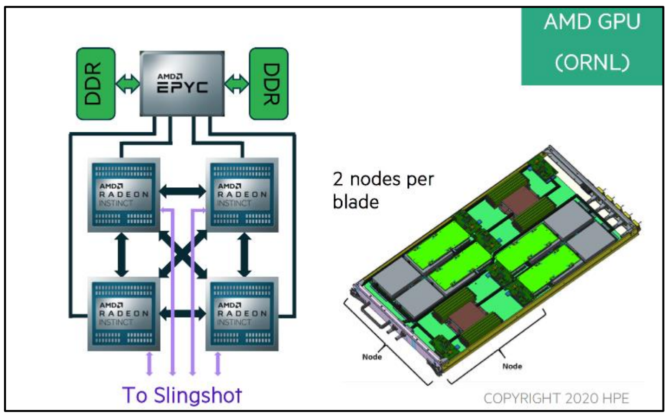

Figure 1. Frontier Node containing 1- AMD EPYC CPU & 4 - AMD Instinct MI250X GPUs

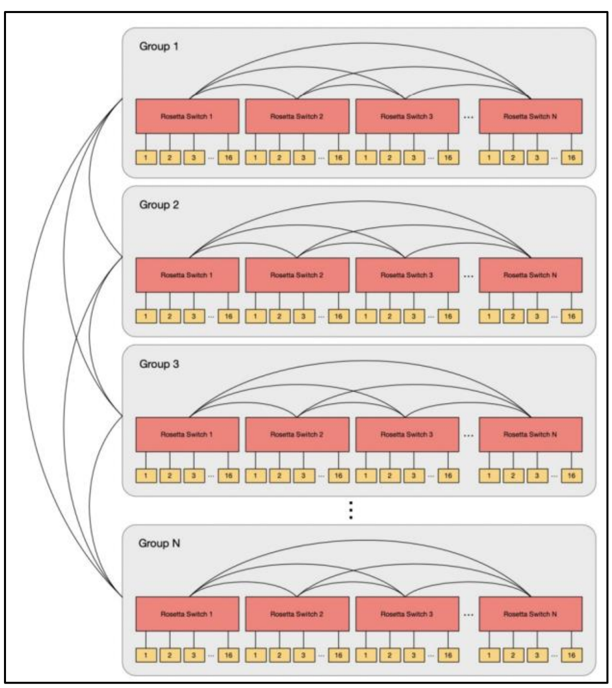

Figure 2. Dragonfly Topology

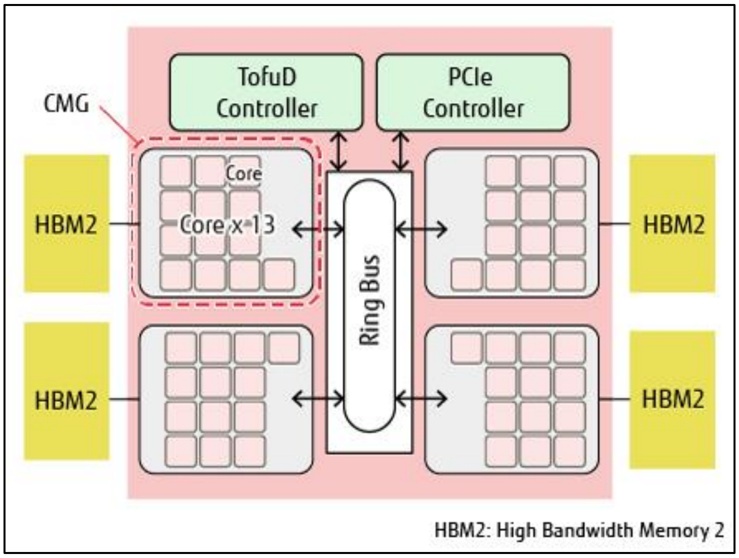

Figure 3. A64FX Processor Overview

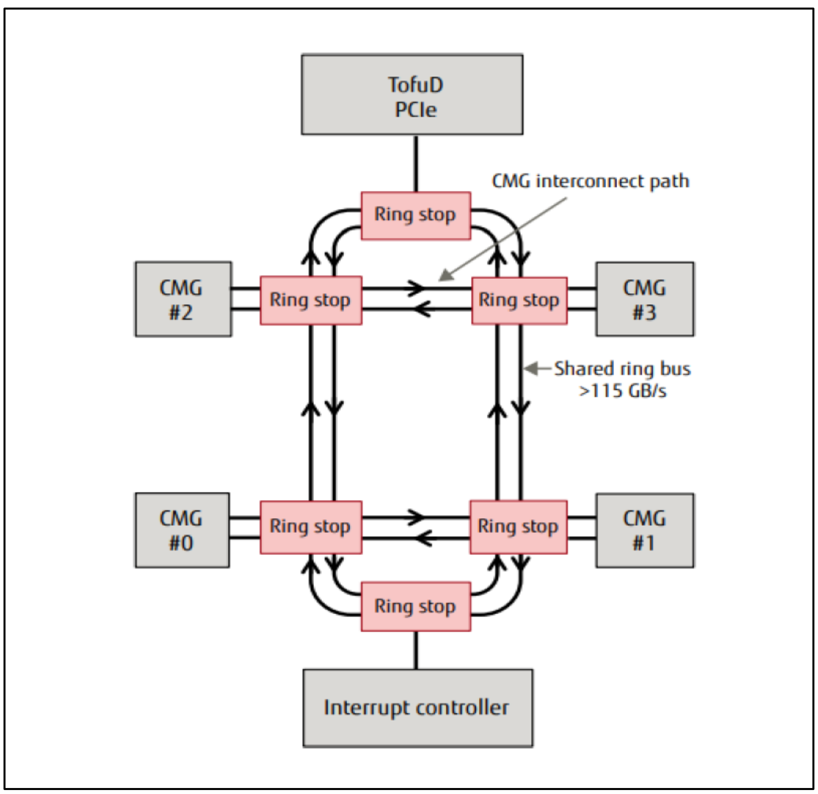

Figure 4. Core Memory Groups, Interrupt Controller, TofuD PCIe in ring configuration

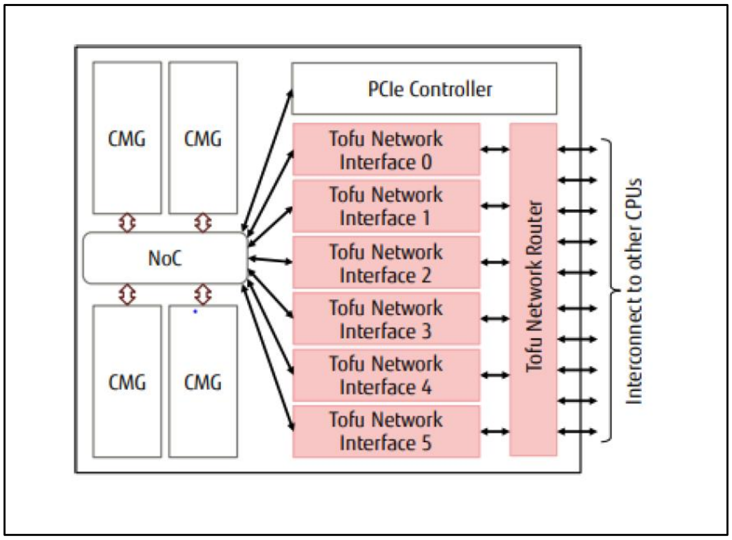

Figure 5. TofuD Interconnect Overview