Preface: Summit and Sierra Supercomputers

The Summit and Sierra Supercomputer Systems, deployed in 2018 at Oak Ridge (ORNL) and Lawrence Livermore (LLNL) Department of Energy (DOE) National Laboratories, provide a significant increase in computing capability relative to their predecessors and represent a major step in the path to Exascale computing. Developed by IBM under the DOE CORAL program, Summit and Sierra have been the #1 and #2 Top500 supercomputers since November 2018,[1] after being #1 and #3, respectively, on the prior list.

Supercomputers such as Summit and Sierra have unique design challenges in parallel computation, synchronous coordination, network bandwidth, communication patterns, and storage not typically found in traditional compute infrastructure deployments. While supercomputers and data centers are both large collections of interconnected compute nodes installed in a single location, they differ in the style of computing they perform. In particular, supercomputers are characterized for synchronous computation on "bare metal" resources, coordinated parallel work, message passing among parallel tasks, and intensive computation periods that can last many hours or even days so that job duration may be longer than the reliability window. In contrast, general-purpose data centers use virtualized images, interacting components driven by external events, and information exchange used to drive work, leading to massive short-lived computation pieces and software structures designed for failures. Although the requirements of these domains are different, there is a trend for incorporating supercomputer technology as the one in Summit and Sierra into hyperscale datacenters, reflecting the increasing widening spread of high-performance demands across a larger computing space.

Driven by the high efficiencies demanded by the required targets, the Summit and Sierra systems are based on a heterogenous compute approach, tying diverse compute engines together through a high bandwidth interconnect and state-of-the-art memory systems, key aspects toward meeting the high-performance application needs. Through OpenPOWER,[2] IBM and its partners created an environment for modularity, relying on strong IBM POWER9 central processing units (CPUs)[3] for the serial and limited parallel sections of applications, and powerful NVIDIA V100 graphics processing units (GPUs)[4] for massively parallel sections. Nodes are interconnected with Mellanox CX-5 InfiniBand[5] network interfaces, and storage is provided by IBM Elastic Storage Server (ESS)[6] servers.

[1]https://www.top500.org/lists/
[2]https://openpowerfoundation.org/
[3]https://www.ibm.com/it-infrastructure/power/power9/
[4]https://www.nvidia.com/en-us/data-center/volta-gpu-architecture/
[5]https://www.mellanox.com/products/
[6]https://www.ibm.com/us-en/marketplace/ibm-elastic-storage-server

Digital Object Identifier: 10.1147/JRD.2020.2976169

Additional infrastructure elements provide the support for the system management capabilities required for operation of such large-scale systems.

This special issue of the *IBM Journal of Research and Development* provides a description of salient hardware and software attributes of Summit and Sierra, as well as examples illustrating the use of these supercomputer systems in different contexts. Despite its comprehensive contents, this issue does not capture everything that these systems are capable of delivering nor everything that scientists and engineers are already doing with the systems to advance knowledge in their corresponding fields. Multiple other publications related to these systems are emerging throughout different technical venues, and undoubtedly there will be many more in the future as the systems are leveraged by more users on their applications.

The contents of this special issue are organized as follows: The first few articles address architecture and hardware aspects of the systems, followed by system software management and communications. Then, focus moves to summaries of early experiences collected by users at the two DOE laboratories during the development of the systems and during the period right after their deployment. The next articles are related to programmability and compiler technology, to complete the issue with articles describing the use of the systems on two application areas.

The first article, "The CORAL Supercomputer Systems" by Hanson, gives an overview of the most relevant aspects of the architecture of the system solutions deployed at ORNL and LLNL, fundamentally a data-centric architecture in which compute power is embedded where data resides, combining powerful CPUs with GPUs optimized for scientific computing and artificial intelligence workloads, plus the latest generation of interconnect technology, providing scientists with computing power to solve challenges in many research areas beyond what was previously possible.

The second article, "Redefining IBM POWER System Design for CORAL" by Roberts et al., emphasizes the innovations introduced in the compute nodes over the needs of nonsupercomputer use of similar technology. Specifically, the CORAL program stipulated compute nodes that scale to 200 petaflops with access to 2.5 PB of memory but should also be applicable to the commercial market for single-server applications, which led to the resultant AC922 POWER server used in Summit and Sierra; such a server is also being deployed on multiple other installations at different scales, leveraging air- and water-cooled versions, allowing use in a wide range of environments. The article also describes some novel design features that facilitate data movement and enable new coherent programming models.

The next article, "The High-Speed Networks of the Sierra and Summit Supercomputers" by Stunkel et al., addresses the attributes of the InfiniBand-based Fat-tree network topology that interconnects all compute, storage, administration, and management nodes into one linearly scalable network. The network uses Mellanox 100-Gb/s EDR InfiniBand ConnectX-5 adapters and Switch-IB2 switches, with compute-rack packaging and cooling contributions from IBM. The article details the hardware and software architecture and performance of the Summit and Sierra networks, which are based on the same underlying technology but differ in their overall architecture, and also describes a number of the enhancements engineered into this generation of InfiniBand.

The fourth article, "Building a High-Performance Resilient Scalable Storage Cluster for CORAL Using IBM ESS" by Islam and Shah, describes the design and implementation of the high-performance, scalable, and resilient storage subsystem that is essential for delivering and maintaining the consistent performance and high utilization expected from a modern supercomputer. This storage solution uses IBM Spectrum Scale and IBM ESS technologies. The storage solutions for the two systems are based on the same underlying storage technology but differ in the overall capacity. The larger of the two is composed of 77 ESS building blocks, each comprising a pair of high-performance IO Server nodes connected to four high-density storage enclosures; these ESS building blocks are interconnected via a redundant InfiniBand EDR network to form a storage cluster that provides a global namespace aggregating performance over 32,000 commodity disks. The article describes the innovations required to meet the performance, scale, manageability, and reliability goals, as well as the challenges faced when deploying a system of such unprecedented I/O capabilities.

The following article, "Summit and Sierra Supercomputer Cooling Solutions" by Tian et al., addresses the unique attributes of the water-cooled solution implemented for the CORAL supercomputers, providing a detailed view of the challenges and the innovative approach used to overcome those challenges. Achieving optimal efficiency requires water cooling of high-heat-density components, coupled with optimal water temperature and correct order of water preheating from air-cooled components. High-performance cold plates are used to directly water-cool all CPU and GPU processors with warm inlet water, and direct air-cooling for the approximately 15% of heat load generated by the lowest heat density components. In addition, the Summit system uses a rear-door heat exchanger to deliver zero net heat load to air.

The architecture and hardware section of this special issue ends with the article "Concurrent Installation and Acceptance of Summit and Sierra Supercomputers" by Liebsch, which gives a perspective on the methodology and process applied throughout the development and deployment of the supercomputers toward ensuring successful completion of their acceptance at the DOE laboratories. This topic is not widely described in the literature despite its crucial role in the deployment of large-scale high-performance computer systems, which typically includes an acceptance process to validate the system's specifications covering hardware, software, and delivered services. The article describes the efforts undertaken by IBM and its partners to accomplish early preparations and then concurrently deliver, stabilize, and accept the two fastest supercomputers in the world at the time of deployment.

The next set of articles addresses system management topics. The article "Cluster System Management" by Besaw et al. describes the CSM system administration tools co-designed with the DOE laboratories to provide the support necessary to effectively manage the Summit and Sierra supercomputers, tools that have been released as open source code. The CSM system administration tools provide a unified view of the large-scale clusters and the ability to examine and understand data collected from multiple sources throughout the system.

Cluster System Management operation is coordinated with the Job Step Manager (JSM) tool described in the next article, "Scalable, Fault-Tolerant Job Step Management for High Performance Systems" by Solt et al. JSM works in concert with CSM and also with the load sharing facility (LSF) scheduler to provide a scalable job launch infrastructure for complex workflows with multiple job steps within an allocation. This article explores the design decisions, implementation considerations, and performance optimizations of the JSM infrastructure to support the CORAL systems. JSM demonstrated launching over three-quarters of a million processes in under a minute while providing efficient services to communication libraries and tools over the management network. JSM relies on the Parallel Task Support library to provide a fault-tolerant, scalable communication medium among JSM daemons. A "resource set" concept gives JSM the opportunity to better organize process placement to optimize resource utilization.

The next article, "Communication Protocol Optimization for Enhanced GPU Performance" by Sharkawi and Chochia, addresses the topic of how to optimize performance for all combinations of data origin and destination in the heterogeneous CORAL architecture, including CPU memory and GPU memory, as well as in the same or a different node in the cluster given that different routing paths exist. The article describes the techniques used in Spectrum MPI PAMI layer to cache memory types and attributes in order to reduce the overhead associated with calling the CUDA API; in addition, the article details the different protocols used for the different memory types, namely Device Memory, Managed Memory, and Host Memory.

Having covered the salient aspects in the architecture of the CORAL systems, the following two articles summarize experiences gathered at the two DOE laboratories during the development of the systems and in the months following their deployment. Preparation for the arrival of the systems, in particular preparation for the transition to the heterogeneous CPU–GPU architecture, was accelerated by the establishment of a Center of Excellence (COE) at each of the laboratories, which were staffed with personnel from IBM and NVIDIA, in addition to DOE scientists. These COEs proved to be instrumental in enabling the prompt success in exercising the systems as soon as they became available at the corresponding sites, on applications of relevance to the laboratories.

The article "Sierra Center of Excellence: Lessons Learned" by Dahm et al. describes perspectives on the shift in direction for computational science at LLNL, and the preparation that took place prior to system deployment to address such shift. Over the five years prior to system deployment, the LLNL COE brought together people with specific areas of expertise in a concentrated effort to prepare the applications, system software, and tools for the Sierra supercomputer. This article shares the process applied for the LLNL COE, and also documents lessons learned during the collaboration with the hope that others will be able to learn from their success and intermediate setbacks.

In a similar context, the article "Pre-Exascale Accelerated Application Development: The ORNL Summit Experience" by Luo et al. summarizes the experiences and lessons learned at the ORNL COE with the goal of being able to provide guidance to address programming challenges such as scalability, performance portability, and software maintainability for heterogeneous high-performance computing (HPC) systems. Oak Ridge Leadership Computing Facility (OLCF) had already started the path towards compute heterogeneity with their prior system, thus Summit was the continuation of that trend albeit at a more demanding level of heterogeneity. The combination of newer technologies was recognized as pushing system performance to a new height, even breaking the exascale barrier by certain measures. Offloading work to multiple accelerators per compute node became a requirement for any application intending to effectively leverage the system, so a collection of applications from a wide spectrum of scientific domains was selected for early adoption on Summit, whose results are described in this article.

Continuing the thread of portability, the next article is "An Open-Source Solution to Performance Portability for Summit and Sierra Supercomputers" by Bercea et al., which describes the design, implementation, and delivery of a compliant OpenMP offloading implementation for the heterogeneous servers composing Summit and Sierra. This implementation was done in the mainline open-source Clang/LLVM compiler and OpenMP runtime projects, and the results illustrate performance using the GPUs that match handwritten CUDA code.

The next two articles delve into more detailed discussion of compiler related innovations. The article "Hybrid CPU/GPU Tasks Optimized for Concurrency in OpenMP" by Eichenberger et al. describes how exploiting device-level parallelism is key to achieving high-performance by reducing overheads typically associated with CPU and GPU task execution, done by hiding the complexity of exploiting hybrid intranode parallelism through using the OpenMP programming model abstraction. Complementing this discussion, the article "OpenMP 4.5 Compiler Optimization for GPU Offloading" by Tiotto et al. describes aspects of the OpenMP implementation in the IBM XL C/C++ and XL Fortran OpenMP compilers, which helps programmers achieve their performance objectives via novel compiler algorithms optimizations and runtime heuristics. In addition to performance enhancements, the article also presents an advanced diagnostic feature to aid in debugging OpenMP applications offloaded to the GPUs.

The following article, "Umpire: Application-Focused Management and Coordination of Complex Hierarchical Memory" by Beckingsale et al., presents an application-oriented library that provides support for querying memory resources, provisioning and allocating memory, and memory introspection, so that the memory hierarchy can be efficiently exploited by scientists without tying their applications to specific hardware or software.

Continuing the theme of enabling users to better leverage resources in the supercomputers, the article "Transformation of Application Enablement Tools on CORAL systems" by Maerean et al. discusses the efforts and challenges to meet the CORAL requirements on tools, detailing three such challenges that needed the most involvement. That description is complemented with a usage scenario showing how the tools help users migrate into the new heterogeneous computing environment and understand their application execution and data flow at scale.

The final two articles in this special issue address application examples from two different domains. The article "Porting a 3D Seismic Modeling Code (SW4) to CORAL Machines" describes an experience that uses a performance portability abstraction layer called RAJA, compares the performance of key kernels implemented using RAJA and CUDA, and estimates the performance penalty arising from the portability abstraction layer. This effort leads to a path for efficiently porting large code bases to heterogeneous compute architectures, while avoiding some of the issues arising when using a new architecture.

On a different domain, the final article, "Troubleshooting Deep-Learner Training Data Problems Using an Evolutionary Algorithm on Summit" by Coletti et al., addresses how a deep-learner (DL) model fidelity can be affected by malformed training and validation data, impeding model training progress. Thorough evaluation of such issues requires large compute capabilities as those available in Summit. The article illustrates the use of an evolutionary algorithm (EA) to troubleshoot this kind of DL problem, which evaluated thousands of DL configurations on Summit that yielded no overall improvement in DL performance. Addressing the issue, which suggested problems in the training and validation data, led to a diagnosis that prevented multiple weeks of unproductive efforts using the supercomputer and allowed a timely convergence on a viable solution.

The details regarding the technical characteristics of Summit and Sierra provided in this special issue by the teams that developed and deployed the systems, plus the experiences accumulated from those that participated in the earlier period through the Centers of Excellence at the DOE laboratories, convey a compelling message about the significant impact the CORAL systems are already having and will continue to have on the advancement of science, engineering, and machine learning. Novel results from applications in different fields, as well as the combinations of multiple applications into novel workflows, are leading the way into new achievements in digital discovery in many areas, whose impact will have long lasting effects.

The many participants in the development of the underlying technologies and the deployment of the resulting Summit and Sierra systems merit recognition for their contributions to the supercomputing field. The many scientists that have been and will continue to be the users of these systems merit recognition for the advances in knowledge they will achieve through their activities.

For us, guest editors of this special issue, it has been a personally rewarding experience working with the authors to deliver the articles composing the issue. We hope readers will enjoy and benefit from its contents while continuing to track the advances delivered through the use of the two most capable supercomputers in the world at this time.

Dr. Jaime H. Moreno
Distinguished Researcher, Senior Manager
IBM Thomas J. Watson Research Center

Dr. Yoonho Park
Principal Researcher, Senior Manager
IBM Thomas J. Watson Research Center
Guest Editors