Bridges-2: A Platform for Rapidly-Evolving and Data Intensive Research

Shawn T. Brown
Pittsburgh Supercomputing Center,
Carnegie Mellon University/
University of Pittsburgh
Pittsburgh, PA, USA
stbrown@psc.edu

Paola Buitrago
Pittsburgh Supercomputing Center,
Carnegie Mellon University/
University of Pittsburgh
Pittsburgh, PA, USA
paola@psc.edu

Edward Hanna
Pittsburgh Supercomputing Center,
Carnegie Mellon University/
University of Pittsburgh
Pittsburgh, PA, USA
ehanna@psc.edu

Sergiu Sanielevici
Pittsburgh Supercomputing Center,
Carnegie Mellon University/
University of Pittsburgh
Pittsburgh, PA, USA
sergiu@psc.edu

Robin Scibek
Pittsburgh Supercomputing Center,
Carnegie Mellon University/
University of Pittsburgh
Pittsburgh, PA, USA
flaus@psc.edu

Nicholas A. Nystrom
Peptilogics, Inc.
Pittsburgh, PA, USA
nicholas.nystrom@peptilogics.com

ABSTRACT

Today's landscape of computational science is evolving rapidly, with a need for new, flexible, and responsive supercomputing platforms for addressing the growing areas of artificial intelligence (AI), data analytics (DA) and convergent collaborative research. To support this community, we designed and deployed the *Bridges-2* platform. Building on our highly successful *Bridges* supercomputer, which was a high-performance computing resource supporting new communities and complex workflows, *Bridges-2* supports traditional and nontraditional research communities and applications; integrates new technologies for converged, scalable high-performance computing (HPC), AI, and data analytics; prioritizes researcher productivity and ease of use; and provides an extensible architecture for interoperation with complementary data intensive projects, campuses, and clouds. In this report, we describe *Bridges-2*'s hardware and configuration, user environments, and systems support and present the results of the successful Early User Program.

CCS CONCEPTS

• **Computer systems organization** → **Distributed architectures**; • **Software and its engineering** → *Software implementation planning*; *System administration.*

KEYWORDS

high performance computing, data analytics, artificial intelligence

ACM Reference Format:

Shawn T. Brown, Paola Buitrago, Edward Hanna, Sergiu Sanielevici, Robin Scibek, and Nicholas A. Nystrom. 2021. Bridges-2: A Platform for Rapidly-Evolving and Data Intensive Research. In *Practice and Experience in Advanced Research Computing (PEARC '21), July 18–22, 2021, Boston, MA, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3437359.3465593

PEARC '21, July 18–22, 2021, Boston, MA, USA

ACM ISBN 978-1-4503-8292-2/21/07.
https://doi.org/10.1145/3437359.3465593

1 INTRODUCTION

Today's landscape of computational science is evolving rapidly, requiring new, flexible, responsive supercomputing platforms to address the growing areas of artificial intelligence (AI), data analytics (DA) and convergent collaborative research. Scientists', researchers', and engineers' research has greatly changed over the last decade. They are deluged with vast amounts of data as it streams in from instruments (e.g., telescopes, microscopes, weather instruments, particle accelerators, and genome sequencers), simulations (e.g., astrophysics, materials science, and biophysics), the Internet (e.g., news and social networks), and societal, historical, and cultural collections (e.g., libraries). With much of the data unstructured and highly complex, no longer can data preparation and processing be done manually by human experts. Emergent technologies and methodologies have evolved in the area of data analytics that can leverage AI and machine learning (ML) algorithms and high-performance computing (HPC) to help researchers make sense of the data and facilitate discovery. Increasingly, several disciplines such as the physical and social sciences, engineering, computer science, and humanities are leveraging similar data analytic techniques to gain insight into massive amounts of data outputs created by vast numbers of simulations runs to fully explore experimental scopes never before achievable through manual analysis techniques. Additionally, these communities are increasingly adopting principles of open science, where data and results are openly shared with others, and vitally need platforms that allow them to store and efficiently access open data close to powerful computational resources.

Bridges-2 is the latest supercomputing platform deployed at the Pittsburgh Supercomputing Center (PSC) designed to support the rapidly evolving AI/ML, data analytics, and convergent collaborative research community[3]. Deployed in collaboration with Hewlett Packard Enterprise (HPE), *Bridges-2* builds on our previous, highly successful platform *Bridges* [4, 9], which was designed to support novel research communities and complex workflows.

Bridges-2 provides heterogeneous hardware resources to support multiple types of computing for complex workflows; unified data access across all processing elements; a novel network topology and hierarchical filesystem optimized for data-intensive capacity computing; a flexible user environment to support multiple modes of computing including interactive computing, Jupyter notebooks and R through web interfaces, batch processing, science gateways; containerization for ease deployment and cloud interactivity; and a dedicated and highly experienced support team providing help from basic questions to research consulting. To support the ever growing AI/ML community, *Bridges-2* offers a large number of NVIDIA Volta GPUs with advanced NVLink connectivity, high internode bandwidth, and a robust AI software environment.

In this report, we describe *Bridges-2*'s hardware and systems configuration, user environments, and the results of our successful Early User Program. *Bridges-2* officially began operations in March 2021 and was made available to the national research community at no cost through the NSF XSEDE program [11].

2 HARDWARE DESCRIPTION

2.1 Compute Nodes

Three tiers of computing nodes varying by processor and amount of shared memory are available to *Bridges-2* users for various use cases. All nodes in Bridges-2 include 4-8 TB of NVMe SSD capacity for fast data staging.

2.1.1 Regular Memory (RM) and Large Memory (LM) Nodes. Regular Memory (RM) and Large Memory (LM) nodes represent the main component of capacity computing in *Bridges-2*, and are the most general purpose compute nodes available in the platform. They are provided to support most traditional and non-traditional use-cases including high-throughput computing such as parameter sweeping, optimization, and simulations. Each RM and LM node has 2 64 core, 2.25/3.4 GHz AMD EPYC 7742 processors. The difference between the RM and LM Nodes is 256 GB and 512 GB of 2933 MHz DDR4 coherent shared memory RAM, respectively. Each of these nodes are housed in an HPE Apollo 2000 Gen11 server. *Bridges-2* has 488 RM nodes and 16 LM Nodes providing a large amount of capacity and flexibility to users.

2.1.2 Extreme Memory (EM) Nodes. Extreme Memory (EM) nodes provide a very large shared memory capability, ideal for applications in genomics, graph analytics, and large-scale data analysis. Each EM node contains 4 24-core Intel Xeon 8260M "Cascade Lake" 2.4/3.9 GHz processors and has 4 TB of 2933 MHz DDR4 coherent shared memory RAM housed in an HPE DL560 Gen10 server. *Bridges-2* contains 4 EM Nodes.

2.1.3 GPU Nodes. Graphical Processing Unit (GPU) nodes are provided to give users a large capacity of GPUs with features to enable scalable analysis over multiple processing units, for large-scale accelerated AI/ML and simulations. Each GPU Node contains 8 NVIDIA Volta 100 GPUs with 640 tensor cores (5,120 CUDA cores), 32GB of HBM2 GPU memory, and all eight of the GPUs are connected within a node via 300GB/s NVLink interconnect. Each node contains 2 Intel Xeon 6248 "Cascade Lake" 20 core 2.5/3.9GHz processors and 512GB of of 2933 MHz DDR4 coherent shared memory RAM housed in an HPE Apollo 6500 Gen10 server. Additionally, the 24 GPU Nodes in *Bridges-2* are connected via Mellanox InfiniBand HDR-200 providing direct network access to the GPUs, bypassing the CPU's, providing a extremely powerful set of capabilities for scalable AI/ML algorithms.

2.2 Hierarchical Data Storage

Fast, reliable storage is critical for supporting rapidly-evolving AI/ML and data analytics. To meet this need, *Bridges-2* includes a hierarchical HPE ClusterStor filesystem consisting of three levels: fast SSD, capacity, and archival storage. The *Fast SSD Storage* level is available to facilitate low-latency, high-bandwidth data access, especially for data sets that require high IOPs (i.e. many small I/O requests). It is implemented as a 200 TB NVMe SSD-based Lustre filesystem called *Jet*. With 30 InfiniBand HDR-200 connections, *Jet* has over 100GB/s read/write bandwidth and delivers over 9 million IOPs. The *Capacity Storage* level for large, high-speed storage capacity is provided by our *Ocean* spinning disk Lustre filesystem, which is optimized to store large datasets with relatively few small files. *Ocean* has 15 PB of usable capacity with performance of over 150 GB/s read/write bandwidth. For providing long-term, infrequently accessed storage capacity, *Ocean* also includes an HPE StoreEver MSL6480 tape library. In addition to these three levels, each compute node on *Bridges-2* includes local NVMe SSDs of 4-8 TB for very fast node-specific access. While *Jet* and local NVMe storage are available for file transfer via explicit scripting, *Ocean's Capacity* and *Archival* storage share the same name space and are presented transparently to the user through HPE's Data Management Framework (DMF), which automatically chooses the best storage based on policy and usage patterns. DMF also provides the capabilities for enhanced project file tracking, job scheduler integration and data-intensive workflow automation, which are features to be deployed in the future.

2.3 Flexible Service Nodes

For flexible deployment of various research services such as science gateways, virtual machines, databases, and other web-based services users may need, *Bridges-2* has 32 *Service Nodes* that can be purposed for individual needs of research projects. Each *Service Node* has 2 64 core, 2.25/3.4 GHz AMD EPYC 7742 processors and 256 GB of DDR4-2933 shared coherent RAM housed in an HPE DL395 Gen10 server. These nodes can be configured on request for a number of different use cases and provide the flexibility to support highly complex research platforms. Currently, a small number of these nodes serve various management tasks for *Bridges-2* such as providing login nodes, Slurm services, Open OnDemand, and general containerization services.

2.4 Interconnect

A key aspect of *Bridges-2* is the fact that all of the various heterogeneous compute components of the platform are connected through a single, high-speed interconnect for maximum accessibility and inseparability between them. We chose Mellanox InfiniBand (IB) HDR-200 for its high-performance, low-latency, and hardware optimizations. A fat tree Clos topology with a modest 2.3:1 oversubscription allows typical workloads to perform nearly at full bandwidth. With this interconnect, users have powerful and rapid access to both compute and data throughout the *Bridges-2*system to enable research workflows that can flexibly take advantage of the whole of the system.

3 ADVANCED SYSTEM SUPPORT

PSC's Advanced Systems team supports the on-going operations of the *Bridges-2* platform, performing system administrative tasks and maintenance. There are a number of individual features that have been implemented that are worth highlighting.

3.1 Provisioning

It is crucial in providing a flexible, highly-configurable environment, that *Bridges-2* has a highly manageable means for configuring nodes, installing operating systems and needed software, and tracking provenance of configurations. Based on the lessons learned from operating *Bridges*, we chose the Warewulf framework[12] for building and hosting the netbooted images, preparing hard drives, and setting up the basic operating systems for all of the nodes and developed a custom wrapper, called Lyncanthrope, that defines default and server group specific provisioning parameters in a YAML file. To define, maintain, and provision the system software configuration after booting a node, Puppet is used for tasks such as setting specific drivers, mounting filesystems, and installing packages. Puppet modules are created to define profiles for this provisioning and are version-controlled through Gitlab, all giving us the flexibility of bringing up custom-defined, domain-specific and cluster-generic nodes very quickly.

3.2 Monitoring

For system monitoring and reporting, we chose Prometheus, an open source metrics and alerting software and InfluxDB[7], a time series database, both integrated with Grafana[6] observabilty dashboards. Third party services, such as StatusCake[10], are used to monitor and test our monitoring, login, data transfer nodes and other components outside PSC's firewalls. Servers, file systems and services are monitored to verify that *Bridges-2* is operating as expected and to assure the best computing experience and availability for our user-community.

3.3 User Application and Software Installation

Providing a number of standard, optimized software packages is important o support our broad computational research community. We provide and install this software via three primary channels. First, our experienced Scientific User Support team has 100s of person years of experience in working with multiple domain-specific scientific tools, and have experience testing methods for installing and using these packages. The other two methods take advantage of community driven work to provide such software. Docker and Singularity container recipes are provided for many software packages (e.g. bioinformatics community) that can be immediately pulled and used on the system. And finally, we have chosen to install many packages through the Spack framework[5] which provides community-developed compilation recipes and automatic LMOD [8] integration. Through all of these methods, at the writing of this paper, we are providing over 120 software packages for users.

4 USER ENVIRONMENT

Bridges-2 supports many familiar and convenient methods for users to interact with the platform, spanning the spectrum of research to achieve broad applicability to diverse communities. The user environment of *Bridges-2* supports both traditional advanced research computing users as well as new, more interactive modes for the rapidly evolving community.

4.1 Traditional Batch Computing Environment

For users comfortable with HPC and using command-line interfaces, *Bridges-2* provides direct access to command line tools and the scheduler. Users can log in via *ssh* to a head node and leverage a robust development environment, including tools available through standard Linux distributions. Users can write scripts, customize their environments, interact with the batch queuing system, and easily access their file storage. PSC makes over 120 application packages available for users to conduct research, which are available through the LMOD environment management system. *Bridges-2* like its predecessor, must support an extremely flexible environment when it comes to scheduling computational work on its heterogeneous resources. The Slurm batch scheduling system is used for this purpose and provides many useful functionalities for helping users complete their work. Each different type of resource on *Bridges-2* is provided as a "partition" for which users can target their workload appropriately when submitting, with flexible memory, CPU, and GPU needs specified. Additionally, given the fact that many of our compute resources have a large number of cores available per node, each of these resources can be requested in our "shared" queues which allow users to specify portions of node available resources to maximize their allocations and workloads. Customized, sophisticated, and modular scheduling policies can be implemented over the various resources to tune the throughput of jobs and equitably schedule users workloads. The Slurm system can be accessed through the command-line interface or via other computing modalities described below.

4.2 Interactive Computing Environment

Emergent modes of interacting with supercomputing resources have evolved, including Jupyter Notebooks, MATLAB, RStudio, and other user-focused models. While command-line interfaces are possible for many of these tools, users have become more productive using these tools interactively and have a need for large resources to be provided interactively to facilitate big data analytics and other types of analysis. To meet this need, *Bridges-2* provides the Open OnDemand platform, where users can request resources via a web-based interface, request the interactive platforms mentioned above, and manage their data resources. Through this mechanism, users can run interactive computing on all available compute and data types available in *Bridges-2*.

4.3 Web Services Environment

As data has become more plentiful and more complex, there is a growing requirement for computing platforms to help researchers execute complex analytics, as addressed through the development of web-based platforms such as research portals and science gateways. Our *Flexible Service Nodes* are available to researchers and platform developers to be able to deploy complex web-based platforms for direct use on *Bridges-2*, with these nodes being able to host web-services, host databases, and provide access to resources on *Bridges-2*. Our team is also working towards providing access to *Bridges-2* via a generic web-service oriented architecture so that platform developers can utilize the resources on the machine without the need to literally host their platform there.

4.4 Getting Help

There are several ways for users to get support for working on the *Bridges-2* platform. The *Bridges-2 Users Guide* [2] provides a comprehensive set of documentation on the various aspects of using the resource. Additionally, PSC support staff stands ready to help users in their various support needs. We provide a general email address, help@psc.edu, where users can send any question. The email address is automatically routed to our ticket tracking system where it is appropriately routed. Basic, front-line questions are answered rapidly by our support staff, and when necessary, our advanced scientific support consultants work with users on issues that require in-depth support.

5 EARLY USER PROGRAM

Beginning in November, 2020 through January, 2021, PSC in collaboration with HPE deployed and tested the *Bridges-2* platform to ensure all components were fully functional and reliable. In January, we started a 30-day Early User Program (EUP) where we openly invited researchers from around the country to come perform research at no cost to show the utility of the platform to the research community. We broadly announced the EUP in October 2020. XSEDE researchers already awarded Bridges-2 allocations were invited to participate in the EUP by providing software and dataset requirements, and their agreement to provide us feedback. A special form was created by XSEDE for other groups to request participation in the EUP, where they outlined their proposed research, specified their software and dataset needs and request installation support. We explained that usage would be recorded but not charged during the EUP. On the first day of the EUP, we conducted an introductory webinar for the accepted participants, and recorded it for the benefit of those unable to attend [1].

This process resulted in 187 projects participating in the EUP, representing 158 institutions in 33 US states, and 61 fields of science. During the 30 day period of the EUP, we measured 85% overall system utilization, allowing 253 distinct users to run a total of 158,127 jobs. Science progress was shared by over 80 projects with some highlights including:

- Scientists at the Catholic University of America created all-atom simulations of the Bacteriophage T4, herpes simplex and SARS-CoV-2 viruses,which will aid in therapeutic development.
- A group of scientists from The Water Institute on the Gulf, the Coastal Protection and Restoration Authority of Louisiana, the RAND Corporation and Purdue University used *Bridges-2* to help validate the 2023 Louisiana Coastal Master Plan for preserving and mitigating the state's coastline in the face of climate change.
- At the University of Oklahoma, researchers used WRF simulations in combination with random forest ML techniques to improve the probabilistic forecasting of severe weather with a 1-day lead time.

Participants were asked to complete a survey, and we received 149 responses. On a scale of 1 (lowest) to 5 (highest), the stability of *Bridges-2* was rated 4.59; the usability 4.43; the helpfulness of user support 4.67; and the helpfulness of the documentation 4.41. The total average rating was 4.53.

6 SUMMARY

Bridges-2 is a uniquely capable platform for rapidly evolving and data-intensive research, serving a broad spectrum of science, engineering, and other applications. *Bridges-2* is available at no cost for open research through the NSF XSEDE program. PSC's experienced user support and advanced technologies personnel amplify *Bridges-2*'s inherent capabilities by working closely with users, with the project's success measured by the scientific impact it enables.

ACKNOWLEDGMENTS

Bridges-2 is supported by the National Science Foundation grant OAC-1928147. The authors would like to thank HPE for their ongoing help and support. The authors would also like to thank the entire PSC team for all of their many contributions.

REFERENCES

[1] Bridges-2 EUP Seminar 2021. *Bridges-2 Early User Program Seminar.* https://www.psc.edu/wp-content/uploads/2021/01/Bridges-Early-User-Workshop.pdf
[2] Bridges-2 User Guide 2021. *Bridges-2 User Guide.* https://www.psc.edu/resources/bridges-2/user-guide-2/
[3] Paola A. Buitrago and Nicholas A. Nystrom. 2021. Neocortex and Bridges-2: A High Performance AI+HPC Ecosystem for Science, Discovery, and Societal Good. In *High Performance Computing*, Sergio Nesmachnow, Harold Castro, and Andrei Tchernykh (Eds.). Springer International Publishing, Cham, 205–219.
[4] Paola A. Buitrago, Nicholas A. Nystrom, Rajarsi Gupta, and Joel Saltz. 2020. Delivering Scalable Deep Learning to Research with Bridges-AI. In *High Performance Computing: 6th Latin American Conference, CARLA 2019: Turrialba, Costa Rica, September 25–27, 2019: Revised Selected Papers (Communications in Computer and Information Science, Vol. 1087)*, Juan Luis Crespo-Mariño and Esteban Meneses-Rojas (Eds.). Springer International Publishing, Switzerland, 200–214. https://doi.org/10.1007/978-3-030-41005-6_14
[5] T. Gamblin, M. LeGendre, M. R. Collette, G. L. Lee, A. Moody, B. R. de Supinski, and S. Futral. 2015. The Spack package manager: bringing order to HPC software chaos. In *SC15: International Conference for High-Performance Computing, Networking, Storage and Analysis.* IEEE Computer Society, Los Alamitos, CA, USA, 1–12. https://doi.org/10.1145/2807591.2807623
[6] Grafana 2021. Grafana website. https://grafana.com/.
[7] InfluxDB 2021. InfluxDB product website. https://www.influxdata.com/products/influxdb/.
[8] Robert McLay, Karl W. Schulz, William L. Barth, and Tommy Minyard. 2011. Best Practices for the Deployment and Management of Production HPC Clusters. In *State of the Practice Reports* (Seattle, Washington) *(SC '11)*. Association for Computing Machinery, New York, NY, USA, Article 9, 11 pages. https://doi.org/10.1145/2063348.2063360
[9] Nicholas A Nystrom, Paola A Buitrago, and Philip D Blood. 2019. Bridges: Converging HPC, AI, and Big Data for Enabling Discovery. In *Contemporary High Performance Computing: From Petascale toward Exascale, Volume Three*, Jeffrey S. Vetter (Ed.). CRC Press, Boca Raton, FL.
[10] Statuscake 2021. Statuscake website. https://www.statuscake.com.
[11] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, and N. Wilkins-Diehr. 2014. XSEDE: Accelerating Scientific Discovery. *Computing in Science Engineering* 16, 5 (2014), 62–74. https://doi.org/10.1109/MCSE.2014.80
[12] Warewulf3 2021. Warewulf3 website. https://warewulf.lbl.gov/.