Studying Latency and Throughput Constraints for Geo-Distributed Data in the National Science Data Fabric

Jakob Luettgau
Heberth Martinez
University of Tennessee
Knoxville, Tennessee, USA

Glenn Tarcea
University of Michigan
Ann Arbor, Michigan, USA

Giorgio Scorzelli
Valerio Pascucci
University of Utah
Salt Lake City, Utah, USA

Michela Taufer
University of Tennessee
Knoxville, Tennessee, USA

ABSTRACT

The National Science Data Fabric (NSDF) is our solution to the problem of addressing the data-sharing needs of the growing data science community. NSDF is designed to make sharing data across geographically distributed sites easier for users who lack technical expertise and infrastructure. By developing an easy-to-install software stack, we promote the FAIR data-sharing principles in NSDF while leveraging existing high-speed data transfer infrastructures such as Globus and XRootD. This work shows how we leverage latency and throughput information between geo-distributed NSDF sites with NSDF entry points to optimize the automatic coordination of data placement and transfer across the data fabric, which can further improve the efficiency of data sharing.

KEYWORDS

High-performance computing, Cloud computing, Data democratization, PerfSonar, XRootD

ACM Reference Format:

Jakob Luettgau, Heberth Martinez, Glenn Tarcea, Giorgio Scorzelli, Valerio Pascucci, and Michela Taufer. 2023. Studying Latency and Throughput Constraints for Geo-Distributed Data in the National Science Data Fabric. In *Proceedings of the 32nd International Symposium on High-Performance Parallel and Distributed Computing (HPDC '23), June 16–23, 2023, Orlando, FL, USA*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3588195.3595948

1 CHALLENGES AND PROPOSED SOLUTION

The need for frictionless integration of geographically distributed data becomes more important as research becomes increasingly distributed. Funding agencies are increasingly moving away from funding small-scale research clusters for individual projects and instead providing funding for storage and compute resource allocations on academic cloud providers and supercomputing sites. While this approach provides researchers access to more powerful computing resources, it also creates additional challenges for sharing data across different platforms and locations. While the National research and education networks (NRENs) enable researchers to exchange data seamlessly across institutions and domains, sharing data still poses significant challenges.

The National Science Data Fabric (NSDF) is a solution that provides communities with efficient data sharing and transfer capabilities while hiding the technical complexity of the process. NSDF achieves this by developing easy-to-install software stacks that utilize high-performance data transfer solutions such as Globus and XRootD. One of the key goals of NSDF is optimizing data placement through the automatic coordination of data placement and transfer in the data fabric. NSDF achieves this by leveraging the latency and throughput between sites with NSDF entry points. To evaluate the effectiveness of NSDF, we set up a testbed of geographically distributed entry points across eight locations in the United States. We monitor the testbed latency, throughput, and routing between entry points. The testbed results allow us to identify anomalous behaviors across the data fabric to be identified and used to inform users on how to set up NSDF services best. Our contributions are as follows: (a) We set up a testbed of geographically distributed entry points across eight locations in the United States, providing a realistic environment for testing and optimizing NSDF services [3–5] (b) We monitor latency, throughput, and routing between entry points over time, identifying areas of improvement and detecting anomalous behaviors across the data fabric. (c) We use the extracted knowledge to inform users on how best to set up NSDF services.

2 TESTBED

Different academic clouds and research institutions in the testbed reflect the reality of geographically distributed scientific collaborations. Our testbed includes eight NSDF entry points across different academic clouds and research institutions within the United States. The sites are deliberately chosen to be heterogeneous to provide a realistic environment for testing and optimizing NSDF services. All the sites have at least eight cores, 30 gigabytes of main memory, and 60 GiB of attached storage. This ensures that the sites can handle large data volumes and support the NSDF services effectively. The testbed includes sites provisioned through CloudLab, Chameleon Cloud, and Jetstream2. Five sites are provisioned through CloudLab,

HPDC '23, June 16–23, 2023, Orlando, FL, USA

ACM ISBN 979-8-4007-0155-9/23/06.
https://doi.org/10.1145/3588195.3595948

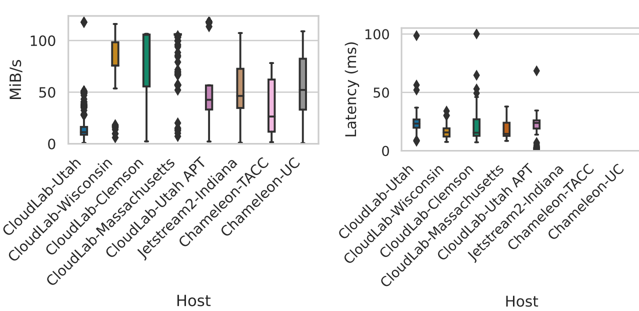
with two hosts at different locations in Utah (1G and 10G), one host each in Wisconsin (1G), Clemson (1G), and Massachusetts (1G). Two sites are provisioned through Chameleon Cloud at TACC (10G) in Texas and CHI (25G) in Illinois. One host is provisioned on Jetstream2 (10G) in Indiana.

3 PERFSONAR AND XROOTD BENCHMARKS

To evaluate the NSDF performance, we use two benchmarks: PerfSonar and XRootD. PerfSonar[2] is a service-oriented architecture for measuring and monitoring network performance. It provides tools for measuring network bandwidth, latency, and packet loss. We use PerfSonar to collect measurements of latency, throughput, and routes between the entry points over a period of two months in March and April of 2023. In total, the initiative collected more than 32,440 measurements, which provide valuable insights into the performance of the NSDF data fabric. XRootD [1] evaluates the performance of the NSDF data fabric in transferring different types of scientific data, including metadata and data-heavy project directories. We developed custom benchmarks for XRootD that mimic scientific use cases.

4 RESULTS

Our tests rely on existing infrastructures to connect the sites. We observe more than 210 regularly occurring network hops through which traffic is routed between the NSDF entry points. About half of the observed routes include Internet2 (93) or ESnet (13), which are fast backbone networks that quickly enable traffic to be routed across state boundaries. All platforms in the NSDF testbed supported containerized deployments. This means that researchers can use containerization to simplify the deployment and management of their scientific workflows. To this point, a latency and throughput measurements summary is plotted in Figure 1.



(a) Throughput (MiB/s) **(b) Latency (ms)**

Figure 1: Distribution of throughput and latency measurements for different hosts in the NSDF testbed.

We observe how instances on CloudLab allow us to run all tests using PerfSonar and their containerized XRootD clients and servers, suggesting that CloudLab provides a reliable infrastructure for scientific research. However, our tests also revealed that hosts on Jetstream2 and Chameleon Cloud, connected through networks utilizing Network Address Translation (NAT), only allow outgoing connections. This means there may be connection problems in one direction for 37.5% of the test locations that employ NAT. We also find that client components usually work well, even behind NAT. This suggests that researchers can still use the NSDF data fabric to transfer data even if some hosts are behind NAT.

Our tests on XRootD transfer service show the impact of two parameters, namely the number of jobs and the number of streams, on the throughput of transferring 1 GiB of data. In particular, we observe that when all the data is bundled into a single file, we achieve the best performance of up to 417 MiB/s when moving data from Wisconsin to Utah. However, when the data is spread across more but smaller files, we see a performance drop. For instance, we observe up to 71 MiB/s for ten files of 100 MiB each and up to 32 MiB/s for a thousand files of 1 MiB each. These findings suggest that the performance of the XRootD transfer service depends on the size and number of files being transferred, and the number of parallel streams and jobs used. Bundling data into larger files generally results in better performance, but there may be a trade-off between the size of the files and the ability to process them in parallel.

5 CONCLUSION

Our work advances the state-of-the-art in data transfer and management for scientific research. Our testbed and collected performance metrics are an important step toward improving the efficiency of data transfers in research sites across the United States. By allowing researchers to experience NSDF firsthand, we provide a platform for feedback that can help optimize distributed infrastructures. Future work includes extending the testbed with additional data points and monitoring it over longer periods. Comparing the performances of NSDF with other transfer tools, such as GridFTP, Globus, and S3 clients, can further help us identify each tool's strengths and weaknesses, allowing for further optimization.

ACKNOWLEDGMENTS

This research is supported by the National Science Foundation (NSF) #1841758, #2028923, #2103845, and #2138811; ACCESS TG-CIS210128; CloudLab PID-16202; Chameleon Cloud CHI-210923; and IBM Shared University Research Award.

REFERENCES

[1] Chuck Boeheim, Andy Hanushevsky, David Leith, Randy Melen, Richard Mount, Teela Pulliam, and Bill Weeks. 2007. Scalla: Scalable Cluster Architecture for Low Latency Access Using Xrootd and Olbd Servers. (2007).

[2] Andreas Hanemann, Jeff W. Boote, Eric L. Boyd, Jérôme Durand, Loukik Kudarimoti, Roman Łapacz, D. Martin Swany, Szymon Trocha, and Jason Zurawski. 2005. PerfSONAR: A Service Oriented Architecture for Multi-domain Network Monitoring. In *Service-Oriented Computing – ICSOC 2007*. Vol. 4749. Springer Berlin Heidelberg, Berlin, Heidelberg, 241–254. https://doi.org/10.1007/11596141_19

[3] Jakob Luettgau, Christine R Kirkpatrick, Giorgio Scorzelli, Valerio Pascucci, Glenn Tarcea, and Michela Taufer. 2022. NSDF-Catalog: Lightweight Indexing Service for Democratizing Data Delivering. In *IEEE ACM International Conference on Utility and Cloud Computing (UCC2022)*. Portland, Oregon.

[4] Jakob Luettgau, Paula Olaya, Naweiluo Zhou, Giorgio Scorzelli, Valerio Pascucci, and Michela Taufer. 2022. NSDF-Cloud: Enabling Ad-Hoc Compute Clusters Across Academic and Commercial Clouds. In *Proceedings of the 31st International Symposium on High-Performance Parallel and Distributed Computing (HPDC '22)*. Association for Computing Machinery, New York, NY, USA, 279–280. https://doi.org/10.1145/3502181.3533710

[5] Paula Olaya, Jakob Luettgau, Naweiluo Zhou, Jay Lofstead, Giorgio Scorzelli, Valerio Pascucci, and Michela Taufer. 2022. NSDF-FUSE: A Testbed for Studying Object Storage via FUSE File Systems. In *Proceedings of the 31st International Symposium on High-Performance Parallel and Distributed Computing (HPDC '22)*. Association for Computing Machinery, New York, NY, USA, 277–278. https://doi.org/10.1145/3502181.3533709