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

Jakob Luettgau[1], Heberth Martinez[1], Giorgio Scorzelli[2],
Glenn Tarcia[3], Valerio Pascucci[2], Michela Taufer[1]
[1]University of Tennessee Knoxville, [2]University of Utah, [3]University of Michigan,

Motivation

- Sharing and transferring research data across geographically distributed sites is increasingly important
- Setting up infrastructure to support data movements often comes with unexpected challenges especially for science teams without strong technical expertise
- We are designing the National Science Data Fabric (NSDF) to resolve and automate common tuning parameters to allow optimizing data placement, throughput and latency

Lessons Learned from Observations

- Our testbed setup includes 8 hosts deployed across t**hree different academic cloud providers** (Chameleon Cloud, CloudLab, and Jetstream2) as well as a physical node at the University of Michigan (UMich).
  - **To connect these 8 hosts,** we rely on significant amounts of existing infrastructure: In particular, **we observe more than 210 regularly occurring network hops through which traffic is routed**.
  - About half of the **observed routes include Internet2 (93) or ESnet (13) and are thus using the fast backbone networks** to route across state boundaries.
- Depending on how the resources are provisioned and how network interfaces are configured, **setting up infrastructure can be a large burden to science teams**:
  - While **all platforms allowed us to use containerized deployments**, only instances on Cloudlab could be automatically configured to always connect NSDF services, the PerfSonar testpoints, as well as our containerized XRootD client and server.
  - Hosts connected through networks utilizing **Network Address Translation (NAT)** such as Jetstream2, Chameleon Cloud, and our hardware testbed at UMich, **can pose challenges and problems that may be hard to troubleshoot especially for science teams that may not enjoy a lot of technical support.**

Network Monitoring with PerfSonar

We use PerfSonar as a containerized sidecar in our NSDF Entry Points to monitor health and point-to-point performance metrics such as latency and throughput. Depending on how the IP addresses are exposed and if Network Address Translation (NAT) is used different metrics can be obtained without manual configuration. In particular, hosts behind NAT can usually serve as clients (e.g., consuming tasks) but often not as server:

THROUGHPUT

SOURCE \ DESTINATION	CloudLab Utah	CloudLab Wisconsin	CloudLab Clemson	CloudLab Massachusetts	CloudLab APT Utah	Jetstream Indiana	Chameleon Cloud TACC	Chameleon Cloud UC
CloudLab Utah	NA	ok	ok	ok	ok	fails	fails	fails
CloudLab Wisconsin	ok	NA	ok	ok	ok	fails	fails	fails
CloudLab Clemson	ok	ok	NA	ok	ok	fails	fails	fails
CloudLab Massachusetts	ok	ok	ok	NA	ok	fails	fails	fails
CloudLab APT Utah	ok	ok	ok	ok	NA	fails	fails	fails
Jetstream2 Indiana	ok	ok	ok	ok	ok	NA	fails	fails
Chameleon Cloud - TACC	ok	ok	ok	ok	ok	fails	NA	fails
Chameleon Cloud - UC	ok	ok	ok	ok	ok	fails	fails	NA

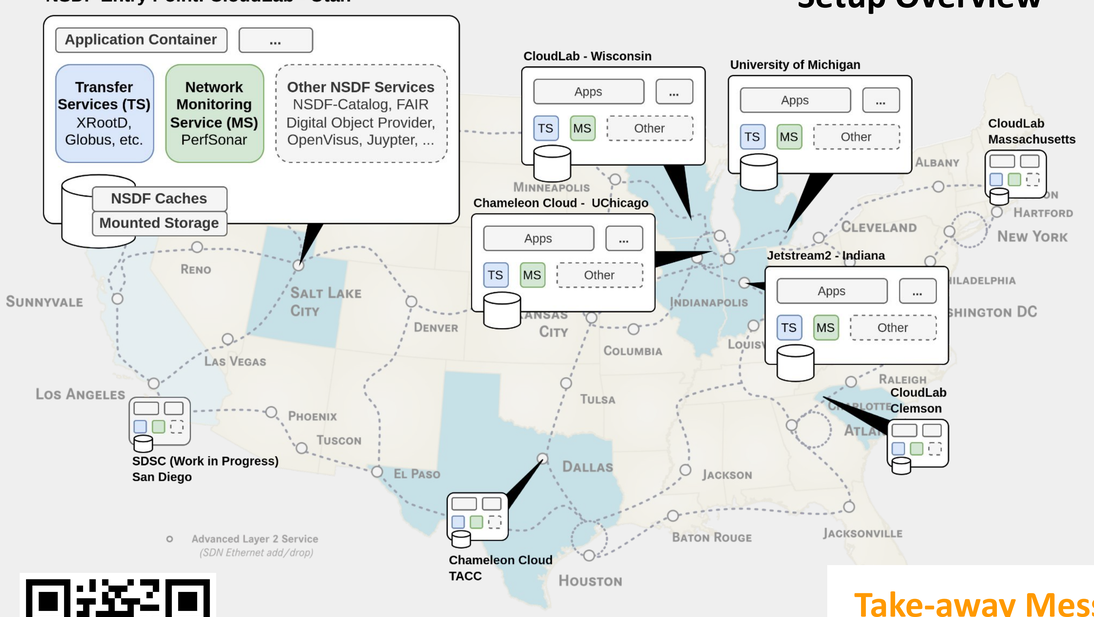
LATENCY

SOURCE \ DESTINATION	CloudLab Utah	CloudLab Wisconsin	CloudLab Clemson	CloudLab Massachusetts	CloudLab APT Utah	Jetstream Indiana	Chameleon Cloud TACC	Chameleon Cloud UC
CloudLab Utah	NA	ok	ok	ok	fails	fails	fails	fails
CloudLab Wisconsin	ok	NA	ok	ok	fails	fails	fails	fails
CloudLab Clemson	ok	ok	NA	ok	fails	fails	fails	fails
CloudLab Massachusetts	ok	ok	ok	NA	fails	fails	fails	fails
CloudLab APT Utah	ok	ok	ok	ok	NA	fails	fails	fails
Jetstream2 Indiana	fails	fails	fails	fails	fails	NA	fails	fails
Chameleon Cloud - TACC	fails	fails	fails	fails	fails	fails	NA	fails
Chameleon Cloud - UC	fails	fails	fails	fails	fails	fails	fails	NA

TRACE & Round Trip Time (RTT)

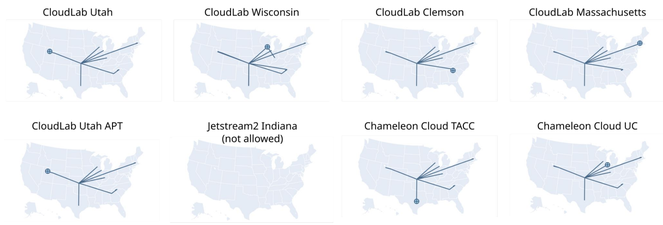
SOURCE \ DESTINATION	CloudLab Utah	CloudLab Wisconsin	CloudLab Clemson	CloudLab Massachusetts	CloudLab APT Utah	Jetstream Indiana	Chameleon Cloud TACC	Chameleon Cloud UC
CloudLab Utah	NA	ok	ok	ok	ok	ok	ok	ok
CloudLab Wisconsin	ok	NA	ok	ok	ok	ok	ok	ok
CloudLab Clemson	ok	ok	NA	ok	ok	ok	ok	ok
CloudLab Massachusetts	ok	ok	ok	NA	ok	ok	ok	ok
CloudLab APT Utah	ok	ok	ok	ok	NA	ok	ok	ok
Jetstream2 Indiana	not allowed	not allowed	not allowed	not allowed	not allowed	NA	not allowed	not allowed
Chameleon Cloud - TACC	ok	ok	ok	ok	ok	ok	NA	ok
Chameleon Cloud - UC	ok	ok	ok	ok	ok	ok	ok	NA

Setup Overview

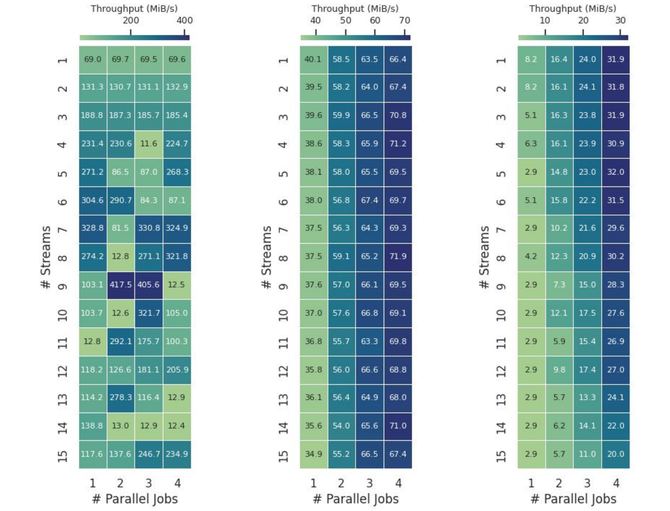


Routing and Performance Analysis

Over a period of one month we collected 39.528 measurements of different metrics such as latency, throughput and routes taken between the different endpoints in our testbed:



In addition, to continuously taking measurements to understand performance variability over time we developed a special benchmark suite with differently structured project directories. The use cases are designed to allow inquiry on how different transfer services are parallelizing transfers. The following plots summarize parameter sweeps that allow NSDF to determine XRootD transfer parameters for different project structures (1 GiB, split into 1, 100, 1000 files):



Find out more about NSDF on: nationalsciencedatafabric.org

Take-away Message

Using continuous performance monitoring and a co-designed suite of benchmarks, NSDF can help science teams improve transfer performance matching their workflows data as well as team and resource locations.

Acknowledgments:
This research was supported by the National Science Foundation (NSF) under grant numbers #1841758, #2028923, #2103845, and #2138811; the Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS) program, under allocation TG-CIS210128; Chameleon Cloud under allocation CHI-210923; and IBM through a Shared University Research Award.