

Resilient Distributed Processing and Reconfigurable Networks

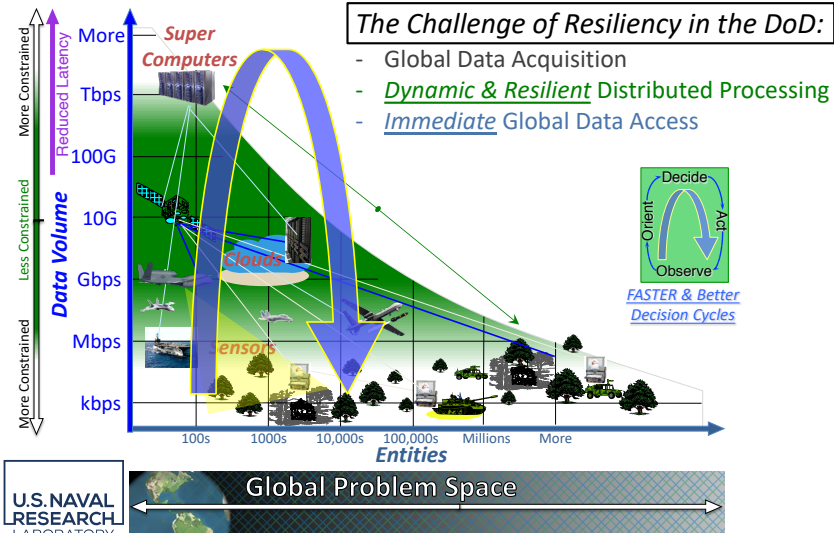
SC23 Demonstration
Naval Research Laboratory
Center for Computational Science
November 12-17, 2023



SC23
Denver, CO | i am hpc.

Basil Decina (basil.decina@nrl.navy.mil)
Linden Mercer (linden.mercer.ctr@nrl.navy.mil)
Dardo Kleiner (dardo.kleiner@nrl.navy.mil)
Larry O'Ferrall (larry.oferrall@nrl.navy.mil)
Louis Berger (louis.berger.ctr@nrl.navy.mil)
Richard Elliott (richard.elliott.ctr@nrl.navy.mil)

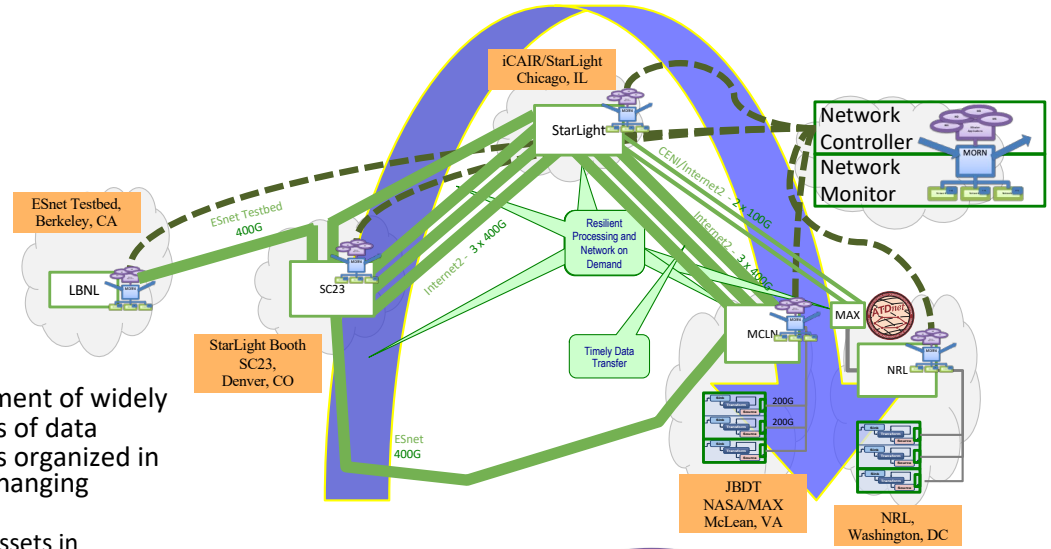
DISTRIBUTION STATEMENT A. Approved for public release.
This material is based upon work supported by the
Department of Defense, US Naval Research Laboratory.



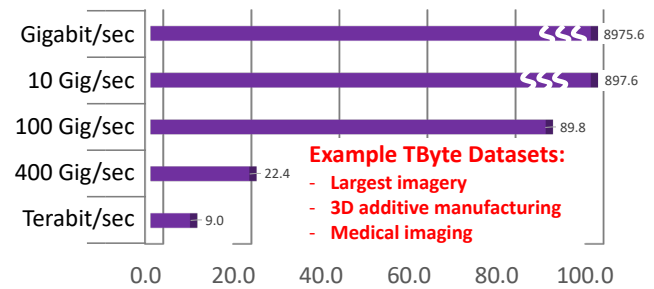
Naval Research Laboratory
Center for Computational Science
SC23 Demonstrations

For More Information Contact:
Basil Decina, basil.decina@nrl.navy.mil
Linden Mercer, linden.mercer.ctr@nrl.navy.mil

- NRL aims to demonstrate:
- Dynamic arrangement and re-arrangement of widely distributed processing of large volumes of data across compute and network resources organized in response to resource availability and changing application demands
 - From SC23 floor to compute and storage assets in Washington, DC; McLean, VA; Chicago, IL; and Berkeley, CA

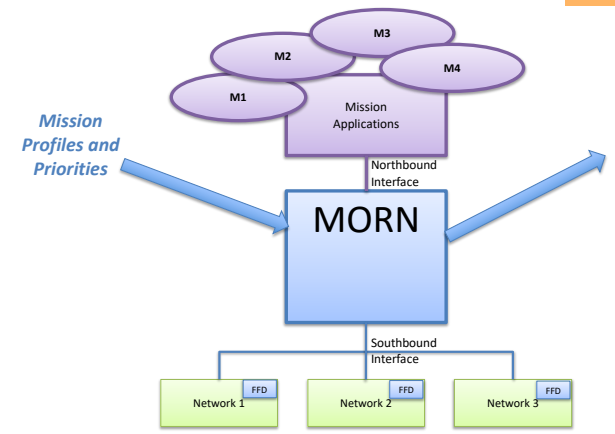


Terabyte Data Movement



The number of seconds to transfer One Terabyte

1. Specific goals: Network deployment, monitoring, reporting, and redeployment.
2. Tbps RDMA performance over global distance for timely Terabyte bulk data transfers (goal << 1 min Tbyte transfer on N by 400G network).
3. Dynamic shifting of processing and network resources from one location/path/system to another (in response to demand and availability).

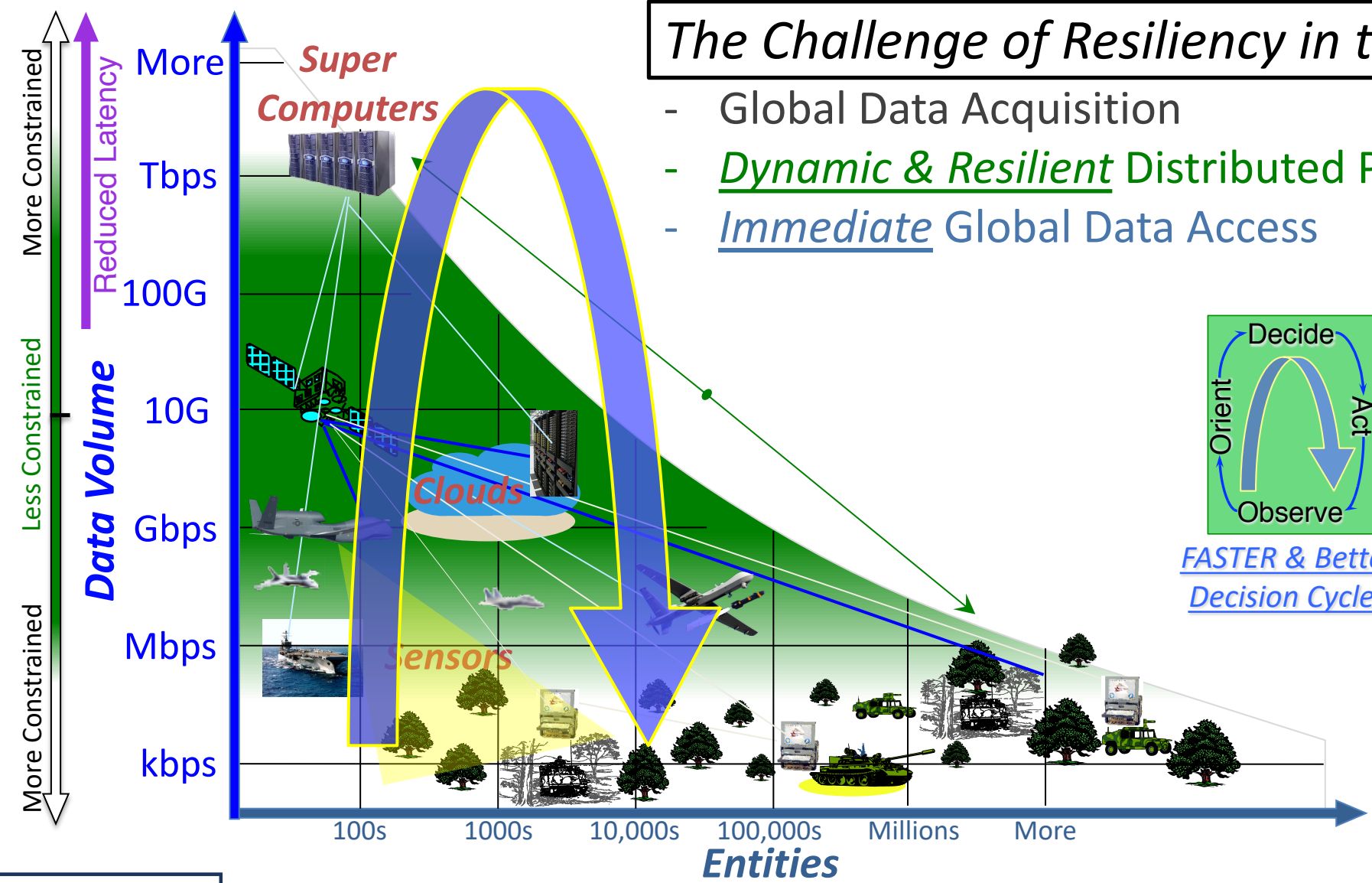


Interconnected and interlocking problems demand a resilient distributed DoD, high performance, low latency data sharing infrastructure



The Challenge of Resiliency in the DoD:

- Global Data Acquisition
- Dynamic & Resilient Distributed Processing
- Immediate Global Data Access



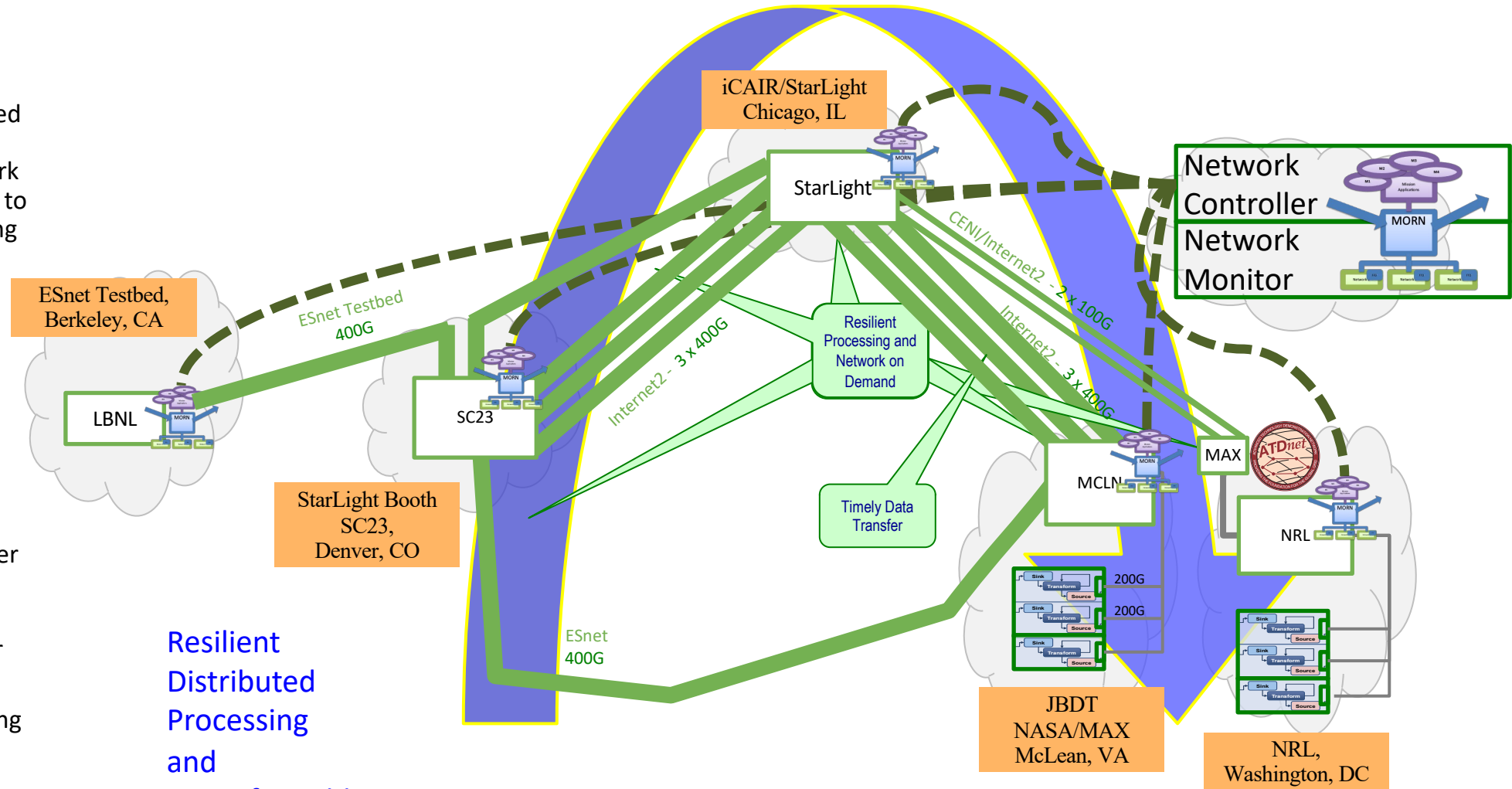
**U.S. NAVAL
RESEARCH
LABORATORY**

Global Problem Space

NRL aims to demonstrate:

- Dynamic arrangement and re-arrangement of widely distributed processing of large volumes of data across compute and network resources organized in response to resource availability and changing application demands
- From SC23 floor to compute and storage assets in Washington, DC; McLean, VA; Chicago, IL; and Berkeley, CA

1. Specific goals: Network deployment, monitoring, reporting, and redeployment.
2. Tbps RDMA performance over global distance for timely Terabyte bulk data transfers (goal << 1 min Tbyte transfer on N by 400G network).
3. Dynamic shifting of processing and network resources from one location/path/system to another (in response to demand and availability).



Resilient Distributed Processing and Reconfigurable Networks

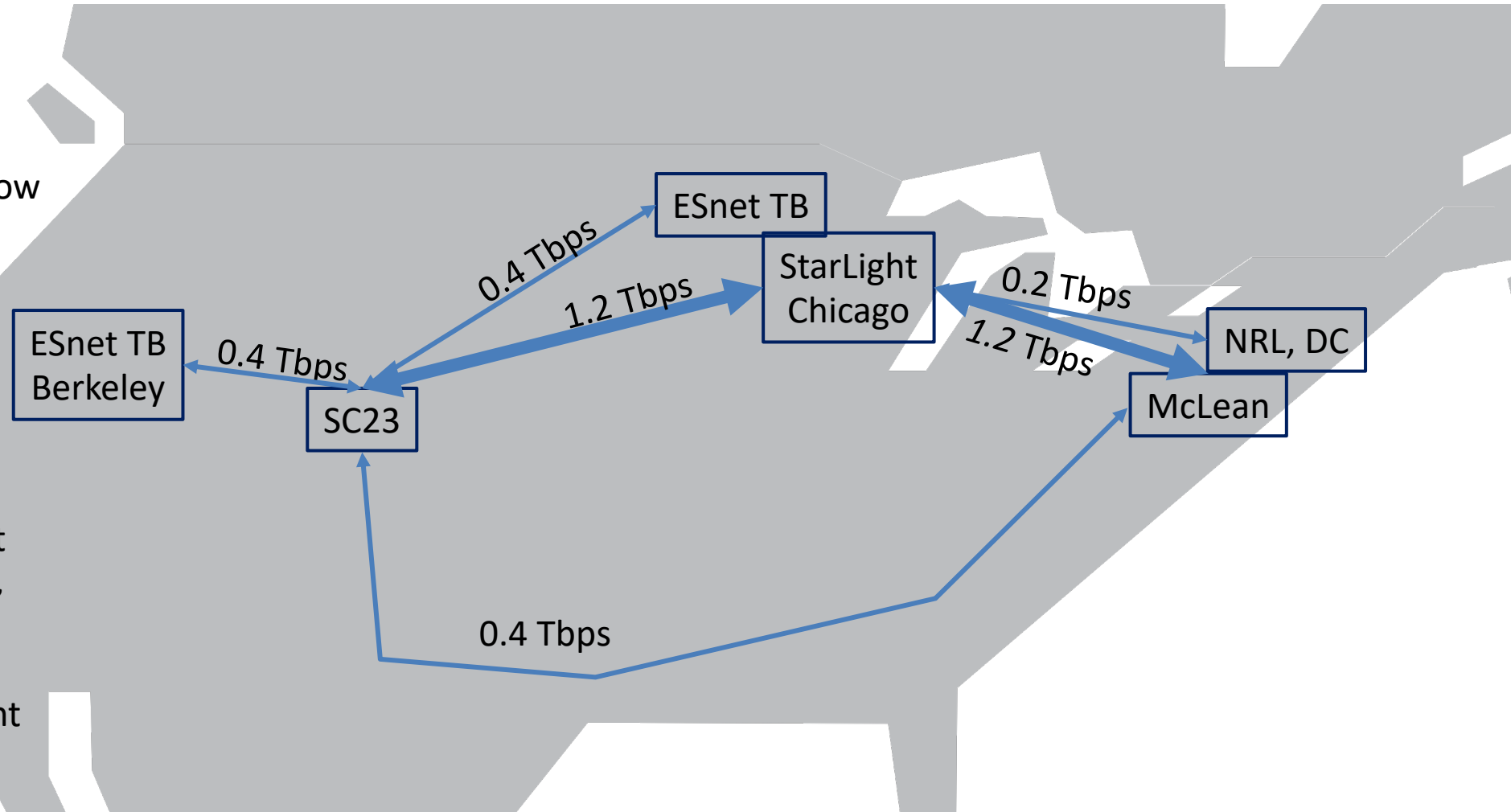
Naval Research Laboratory Center for Computational Science SC23 Demonstration

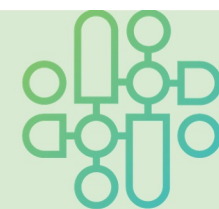


"Interconnected and interlocking problems" demand a high performance dynamic distributed data centric infrastructure

JBDT SC23 WAN (NW iCAIR, NRL, NASA)

- StarLight at NW, Chicago, IL
- Lumen in McLean, VA
- SC23 in Denver, CO
- Bandwidth of >1 Tbps to allow Tbps application flows (multiple – time shared)
- Multiple paths between locations to allow more concurrent activities and network control/restoration test/demos
- 400G systems established at StarLight, McLean, Berkeley, and Dallas by iCAIR, NRL, NASA
- Multiple vendors to augment and expand





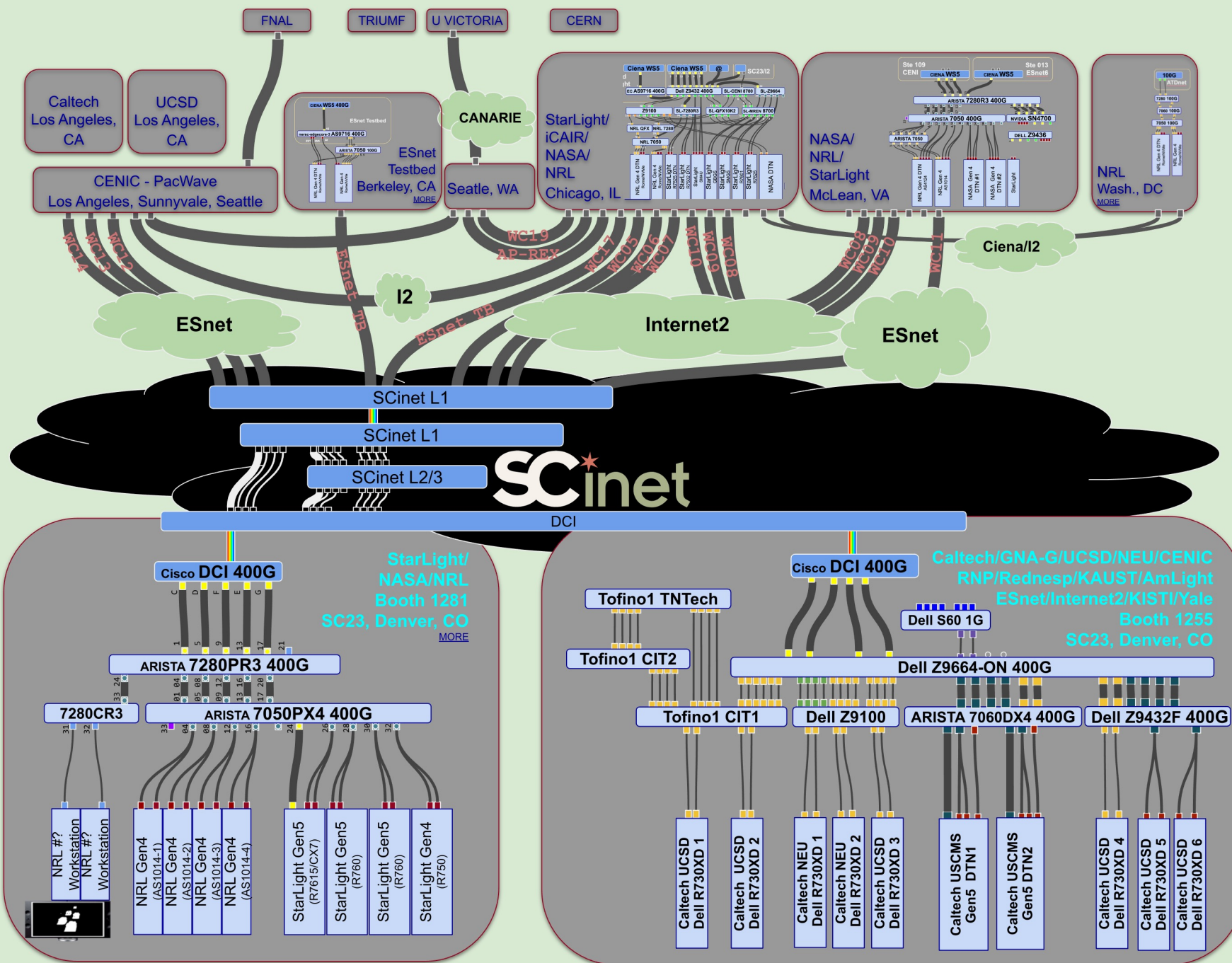
SC23
Denver, CO | i am hpc.

JOINT BIG DATA TESTBED

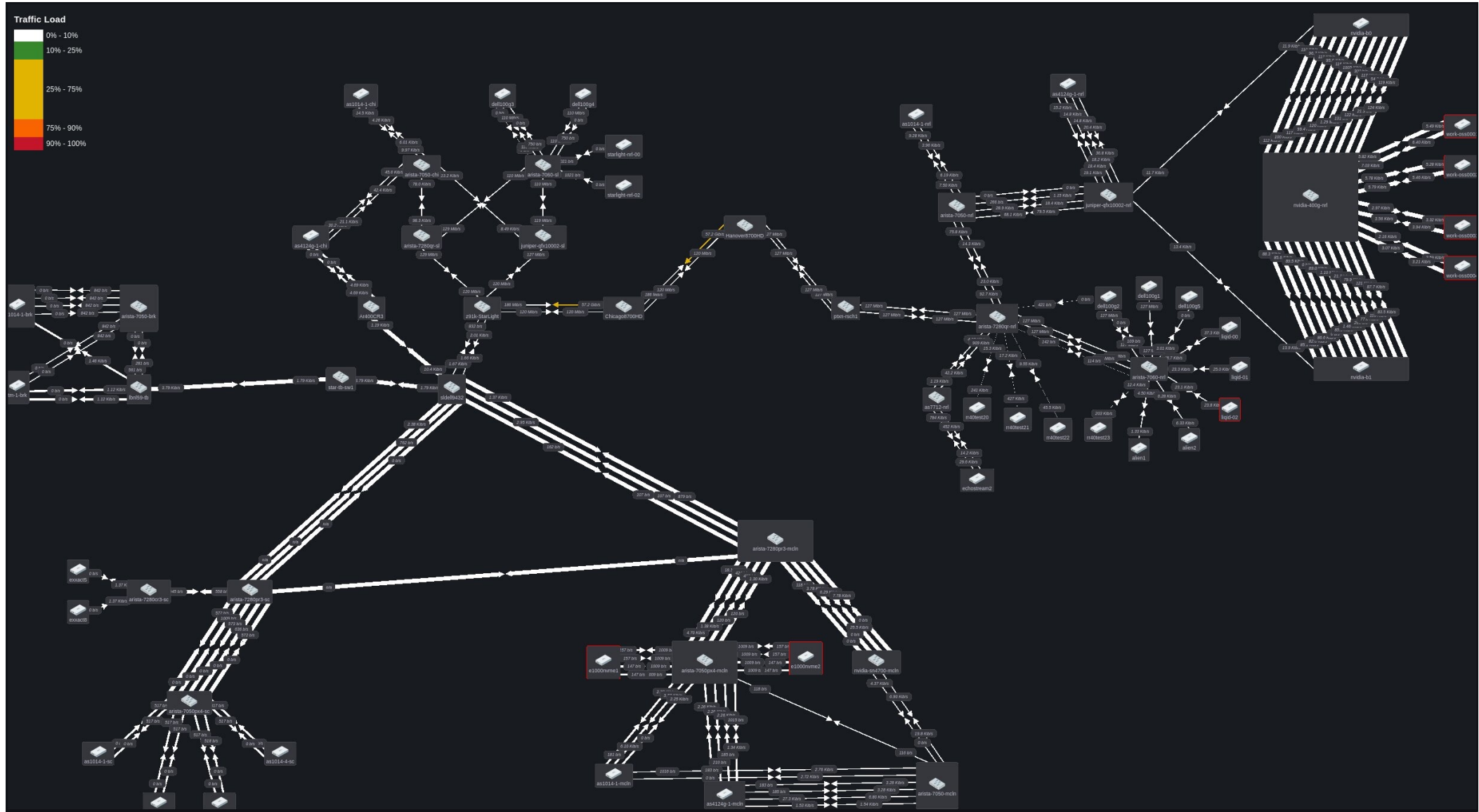
- 400G - LR4
- 400G - FR4
- 400G - DAC
- 200G - SR4 or DAC
- 100G - CWDM4
- 100G - LR4
- 100G - SR4
- 100G - DAC
- 40G - SR4
- 40G - DAC
- 10G
- 1G

10/16/2023

Latest Version at:
<https://tinyurl.com/SC23-JBDT>
To request changes, please leave a comment



Live Testbed Map



Abstract

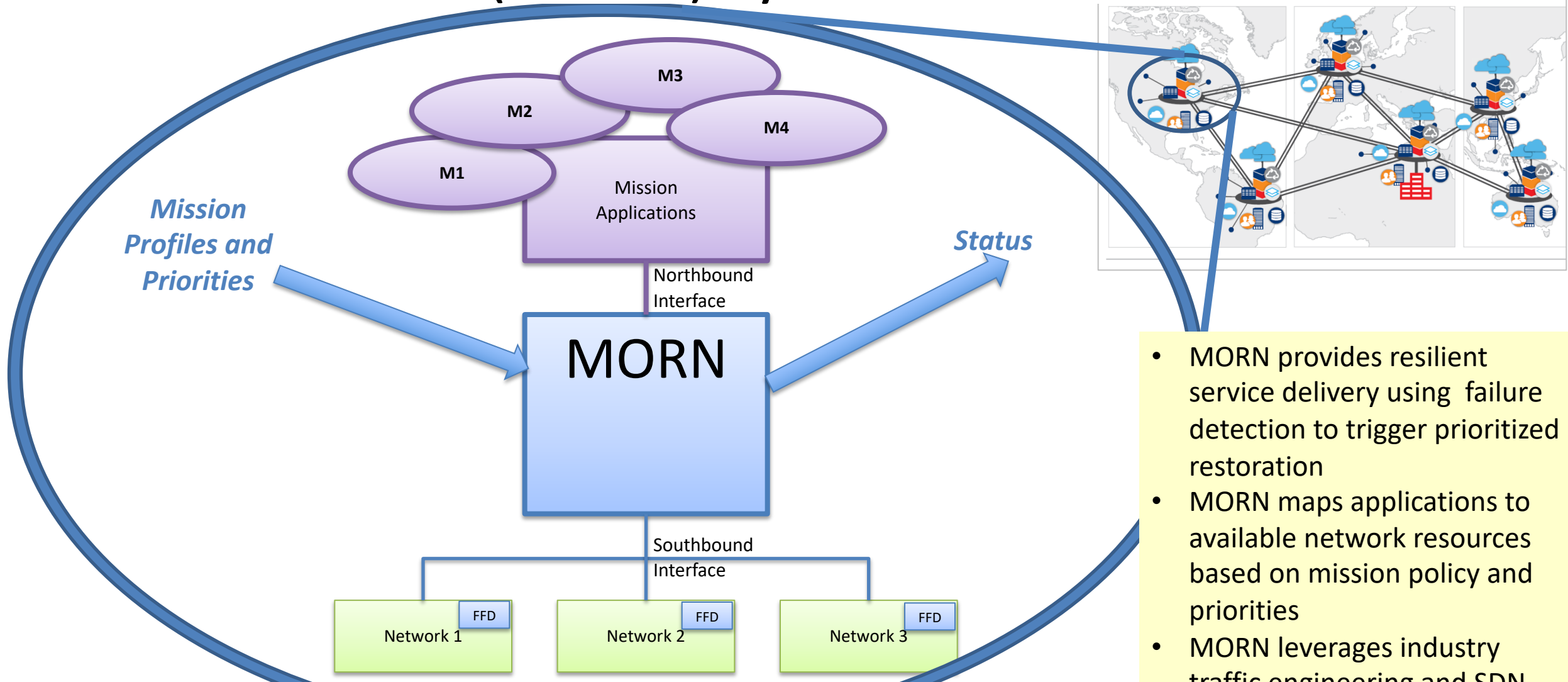
This demonstration will build on our previous NRE SC demonstrations. We aim to show dynamic arrangement and re-arrangement of widely distributed processing of large volumes of data across a set of compute and network resources organized in response to resource availability and changing application demands. We also aim to explore performance limitations and enablers for high volume bulk data transfers. A software controlled network will be assembled using a number of switches and multiple SCinet 400G/100G connections from DC and Chicago to Denver. We plan to show rapid automated deployment and redeployment, real-time monitoring and QOS management application data flows with very different network demands. Technologies we intend to leverage include SDN, RDMA, RoCE, NVMe, GPU acceleration and others.

Goals

Similar to previous efforts but SC23 focus is on dynamic network monitoring and control, and 400G enabled data movement.

1. Network deployment, monitoring, reporting, and redeployment.
2. Tbps RDMA performance over global distance for timely Terabyte bulk data transfers (goal \ll 1 min Tbyte transfer on N by 400G network).
3. Dynamic shifting of processing and network resources from one location/path/system to another (in response to demand and availability).

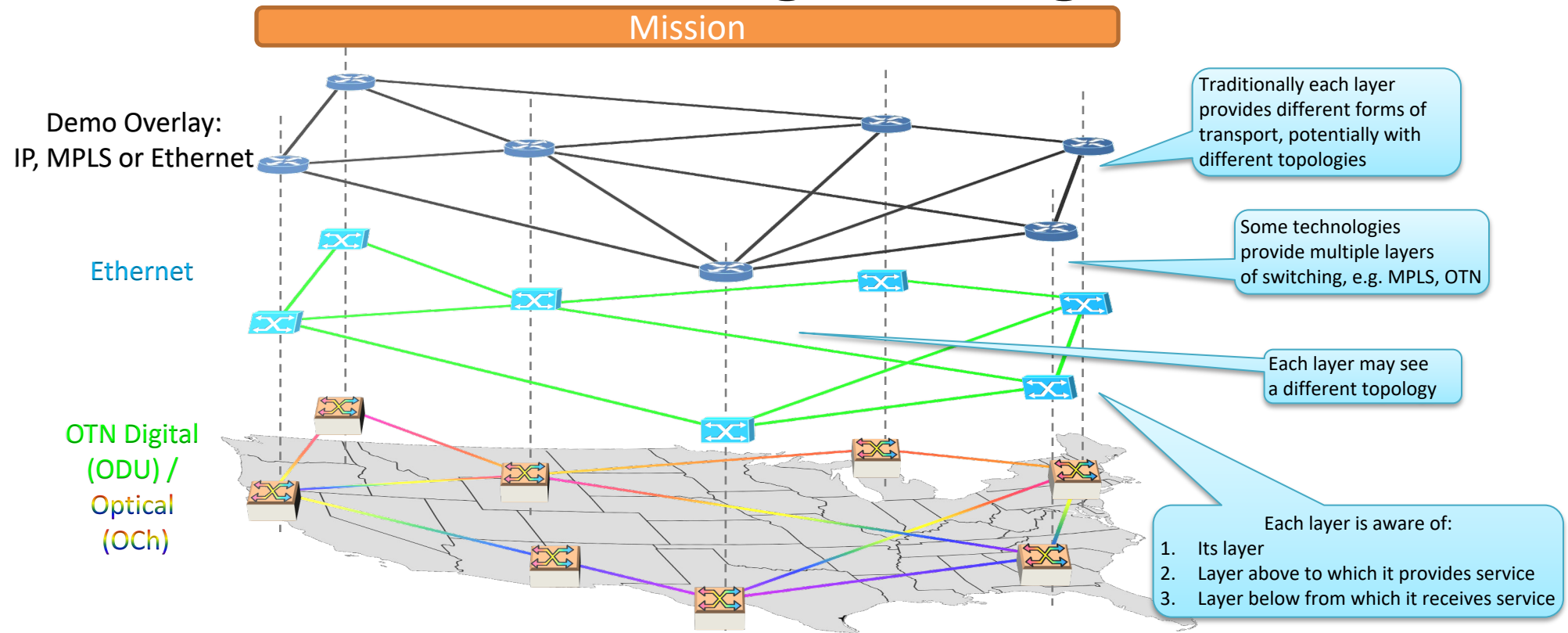
Mission Oriented Reconfigurable Networking (MORN) System View



- MORN provides resilient service delivery using failure detection to trigger prioritized restoration
- MORN maps applications to available network resources based on mission policy and priorities
- MORN leverages industry traffic engineering and SDN

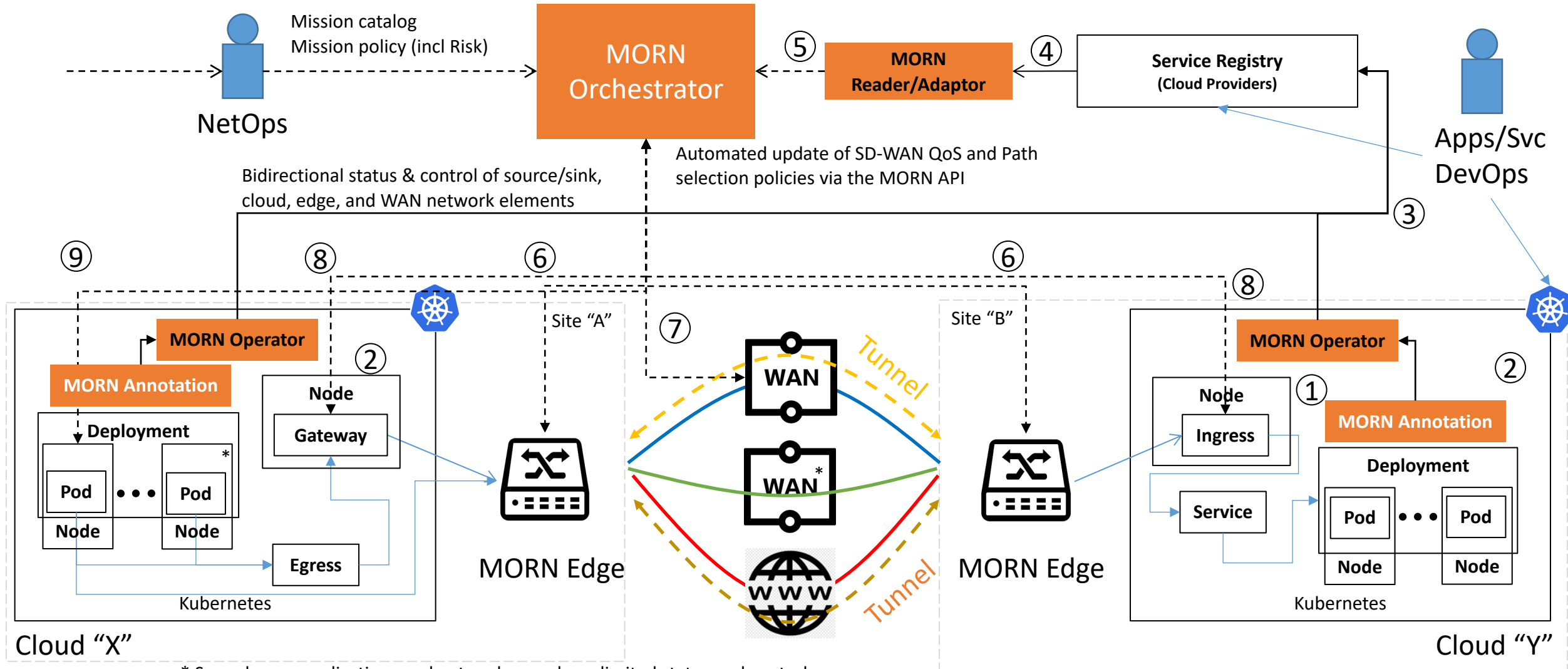
MORN is our approach to optimizing the allocation of resources to meet mission requirements

MORN Builds on Classic Network Layering and Traffic Engineering



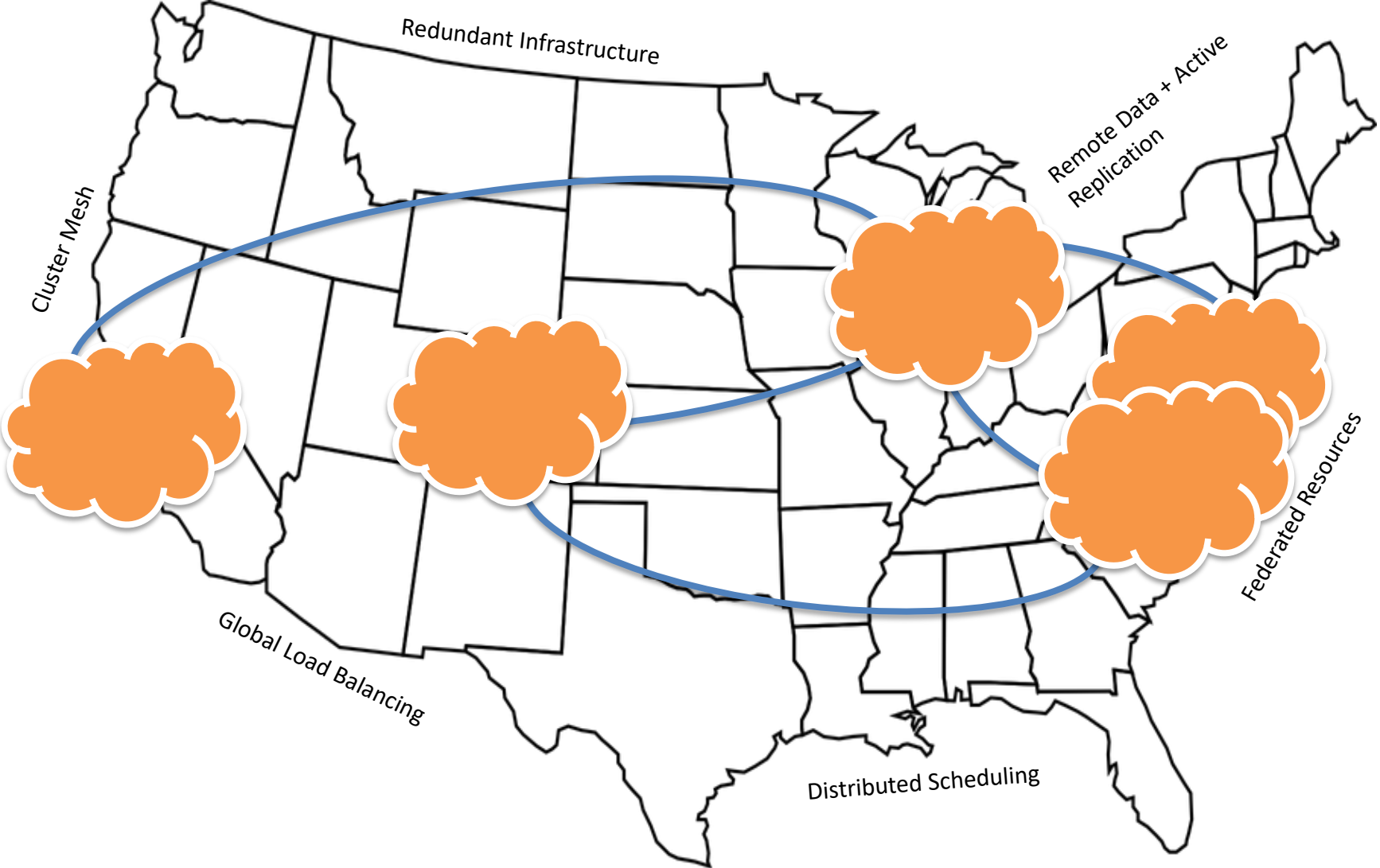
- Network services are provided per layer
 - And may differ in traffic identification and service delivered
- MORN network overlay provides a tailored service interconnection service
 - Mission services prioritized and based on policy
 - Sensor-based failure reporting and restoration

Integrated Orchestration of Cloud & Network

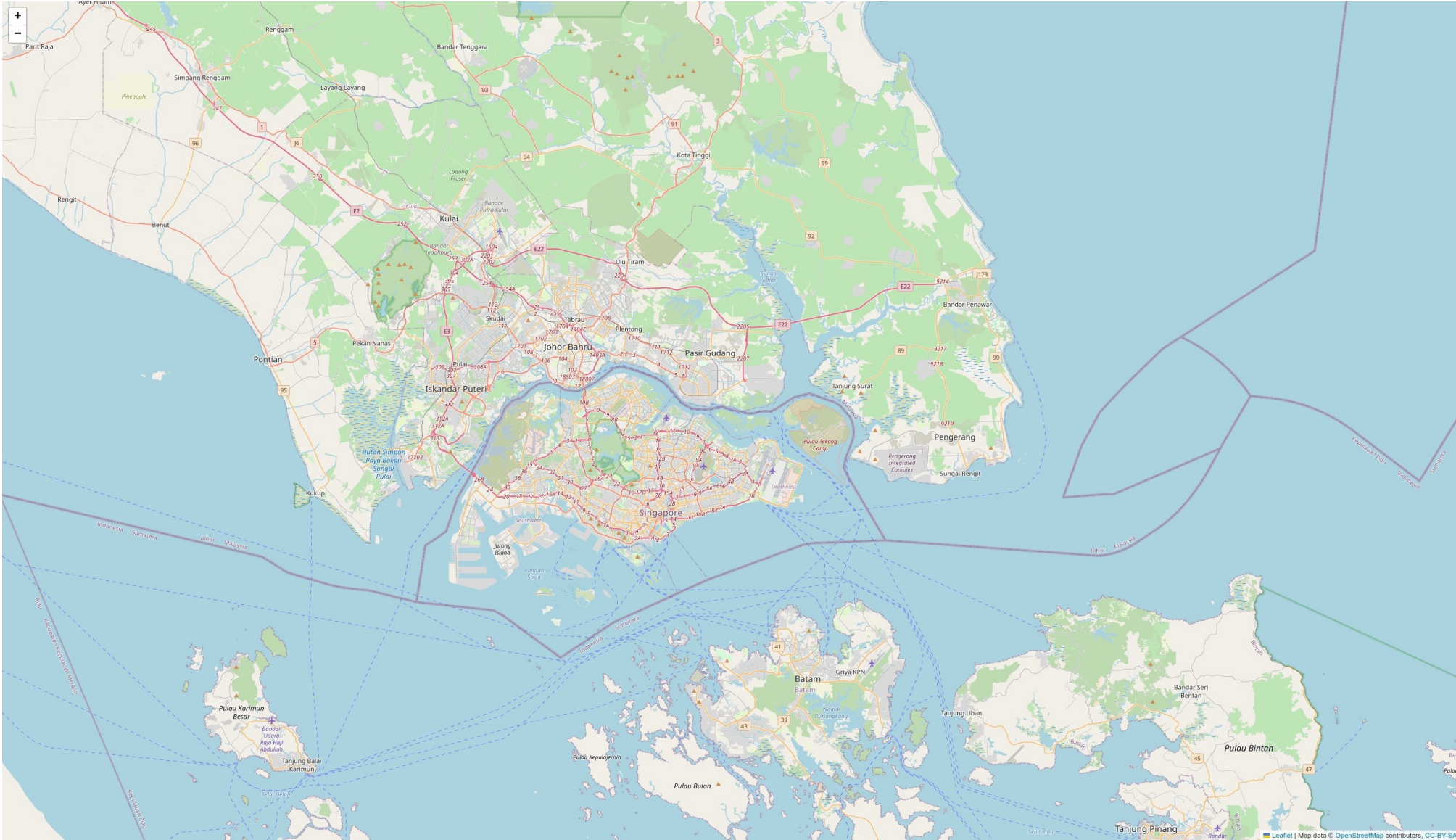


* Some legacy applications and networks may have limited status and control

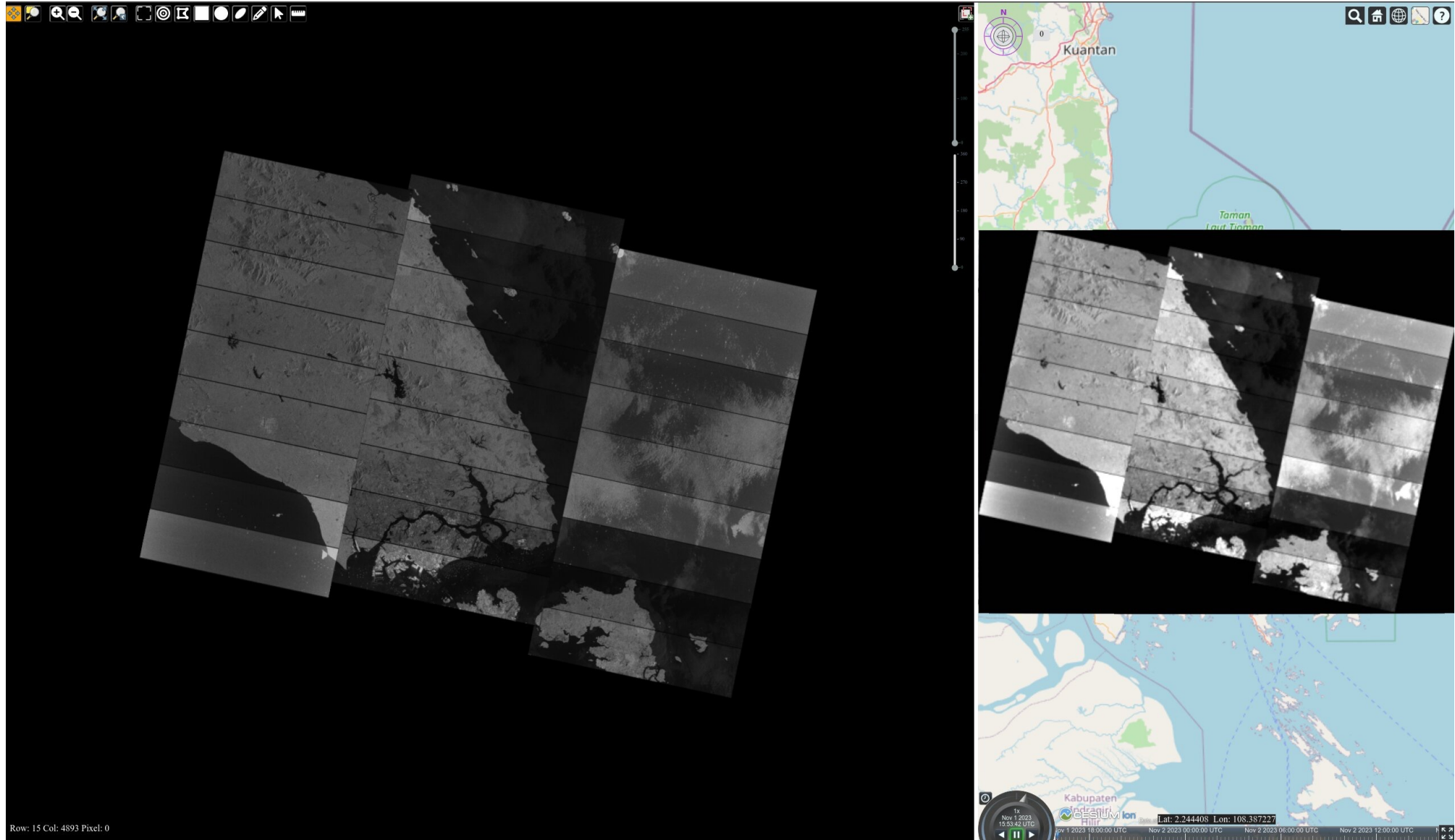
Resilient Distributed On-Demand GIS/HPC



Base Map

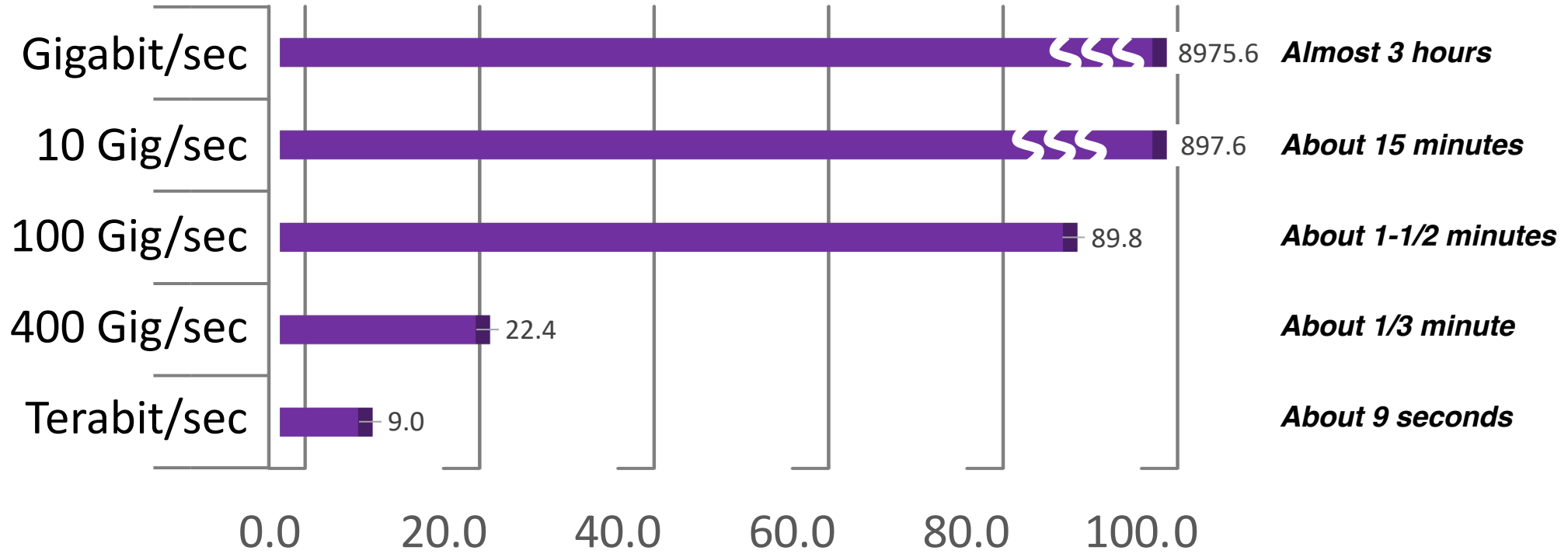


SAR Processing Viewer



What is a Terabyte* and How Fast Can You Move It ?

Network Speed



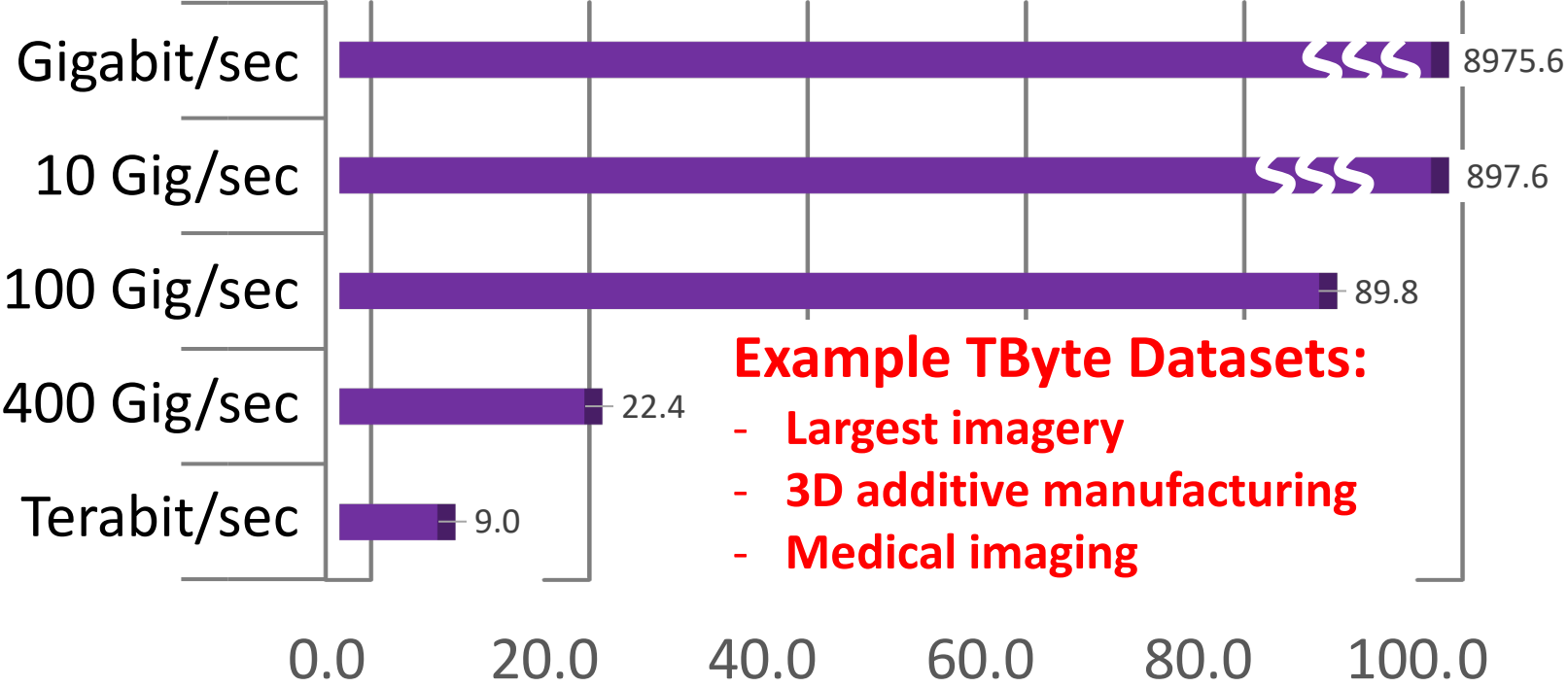
The number of seconds to transfer One Terabyte

Example TByte Datasets:

- Largest imagery
- 3D additive manufacturing
- Medical imaging

* a Terabyte is 8 Terabits (storage to bandwidth conversion)
 * ~3 hours of high-quality compressed 4K UHD video (H.265) is about 100 Gigabytes

Terabyte Data Movement



Example TByte Datasets:

- Largest imagery
- 3D additive manufacturing
- Medical imaging

The number of seconds to transfer One Terabyte

"Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program."

The Networking and Information Technology Research and Development
(NITRD) Program

Mailing Address: NCO/NITRD, 2415 Eisenhower Avenue, Alexandria, VA 22314

Physical Address: 490 L'Enfant Plaza SW, Suite 8001, Washington, DC 20024, USA Tel: 202-459-9674,
Fax: 202-459-9673, Email: nco@nitrd.gov, Website: <https://www.nitrd.gov>

