

Software Defined Services (SDS) For High Performance Large Scale Science Data Streams Across 100 Gbps WANs

Joe Mambretti, Director, (j-mambretti@northwestern.edu)

International Center for Advanced Internet Research (www.icair.org)

Northwestern University

Director, Metropolitan Research and Education Network (www.mren.org)

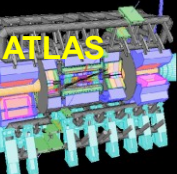
**Director, StarLight, PI StarLight SDX, Co-PI Chameleon, PI-iGENI, PI-
OMNINet (www.startap.net/starlight)**

**The International Conference for High Performance Computing,
Networking, Storage and Analysis (SC16)**

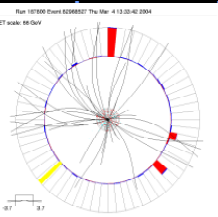
Salt Lake City, Utah

November 13-17, 2016





ALMA: Atacama Large Millimeter Array



DØ (DZ) www.d0.fnal.gov



IVOA International Virtual Observatory www.ivoa.net



Geological Drilling www.andrill.org



GEON: Geosciences Network www.geongrid.org



OSG www.opensciencegrid.org



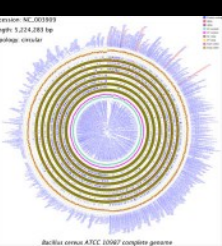
BIRN: Biomedical Informatics Research Network www.nbirn.net



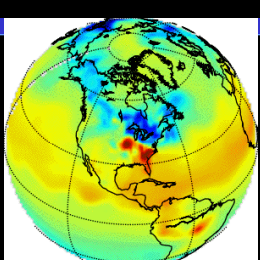
GLEON: Global Lake Ecological Observatory Network



Globus Alliance www.globus.org



CAMERA metagenomics camera.calit2.net



Carbon Tracker www.esrl.noaa.gov/gmd/ccgg/carbontrack



OOI-CI ci.oceanobservatories.org



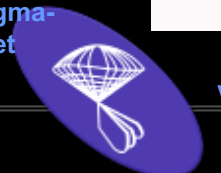
SKA www.skatelescope.org



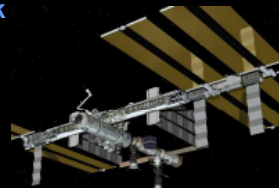
TeraGrid www.teragrid.org



XSEDE www.xsede.org



Sloan Digital Sky Survey www.sdss.org



ISS: International Space Station www.nasa.gov/station



LHCONE www.lhccone.net



CLASS www.class.noaa.gov



Compilation By Maxine Brown

STARLIGHTSM

Macro Network Science Themes

- **Transition From Legacy Networks To Networks That Take Full Advantage of IT Architecture and Technology**
- **Extremely Large Capacity (Multi-Tbps Streams)**
- **High Degrees of Communication Services Customization**
- **Highly Programmable Networks**
- **Network Facilities As Enabling Platforms for Any Type of Service**
- **Network Virtualization**
- **Highly Distributed Processes**



App1

App2

App3

App4

EP1

EP2

Ind1

Ind2

APIs Based On Messaging and Signaling Protocols

Network Programming Languages

Process Based Virtualization – Multi-Domain Federation –

Policies Cascading Through Architectural Components

Security Processes

Policy Processes

Orchestrator(s)

Policy Processes

Northbound Interface

Network OSs
SDN Control Systems

Network Hypervisors

Southbound Interface

State Machines

State Data Bases

Mon, Measurements
Real Time Analytics

Westbound Interfaces

Eastbound Interfaces

PhyR

PhyR

PhyR

PhyR

VirR

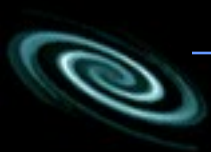
VirR

VirR

VirR

National Science Foundation's Global Environment for Network Innovations (GENI)

- **GENI Is Funded By The National Science Foundation's Directorate for Computer and Information Science and Engineering (CISE)**
- **GENI Is a Virtual Laboratory For Exploring Future Internets At Scale.**
- **GENI Is Similar To Instruments Used By Other Science Disciplines, e.g., Astronomers – Telescopes, HEP - Synchrotrons**
- **GENI Creates Major Opportunities To Understand, Innovate and Transform Global Networks and Their Interactions with Society.**
- **GENI Is Dynamic and Adaptive.**
- **GENI Opens Up New Areas of Research at the Frontiers of Network Science and Engineering, and Increases the Opportunity for Significant Socio-Economic Impact.**



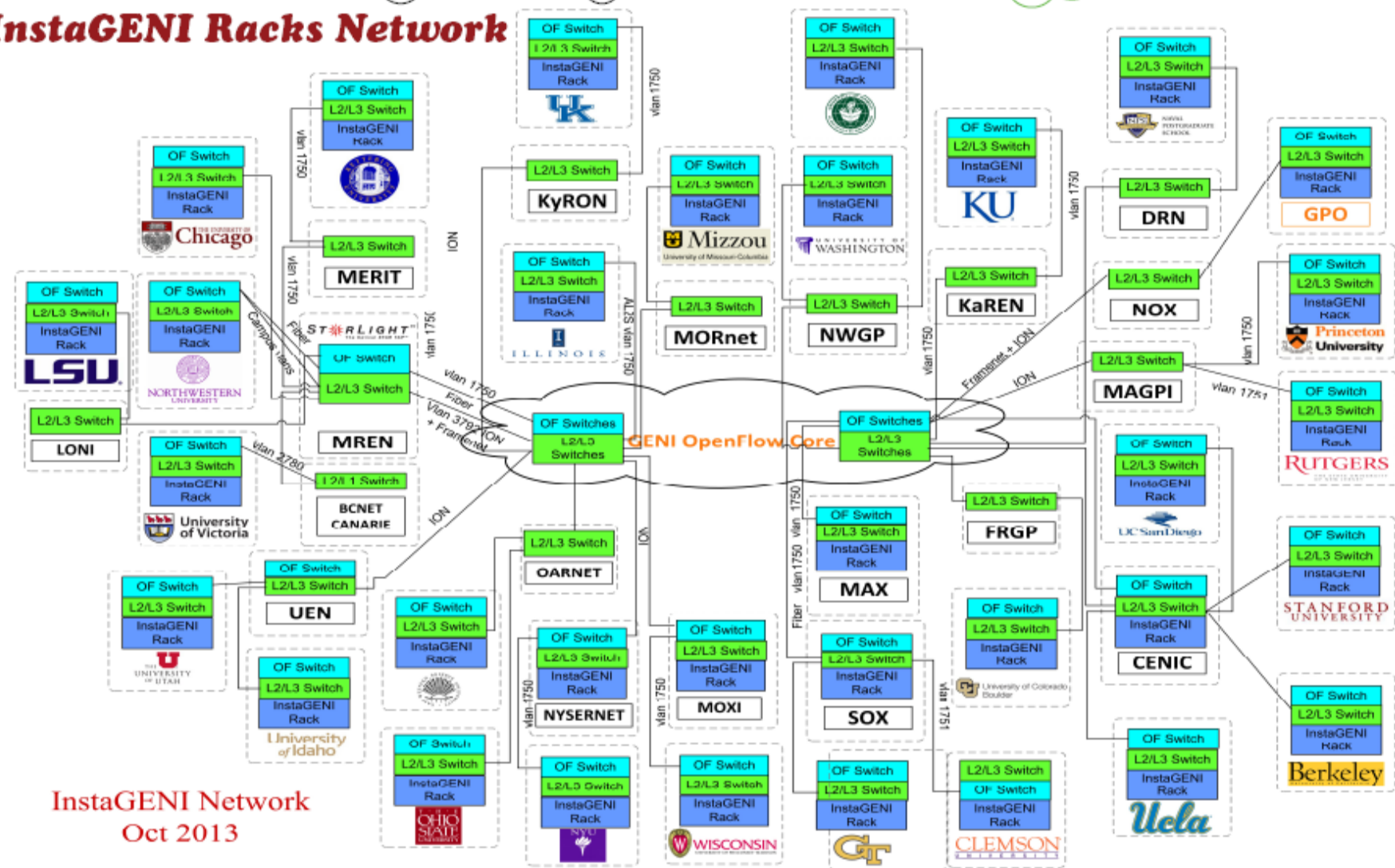
Future Cyberinfrastructure

- **Large Scale Highly Distributed Infrastructure That Can Support Multiple Empirical Research Testbeds At Scale**
- **Next Generation GENI, Edge Clouds, IOT, US Ignite, Platform for Advanced Wireless Research (PAWR) and Many Others**
- **Currently Being Planned – Will Be Designed, Implemented and Operated By Researchers for Researchers**



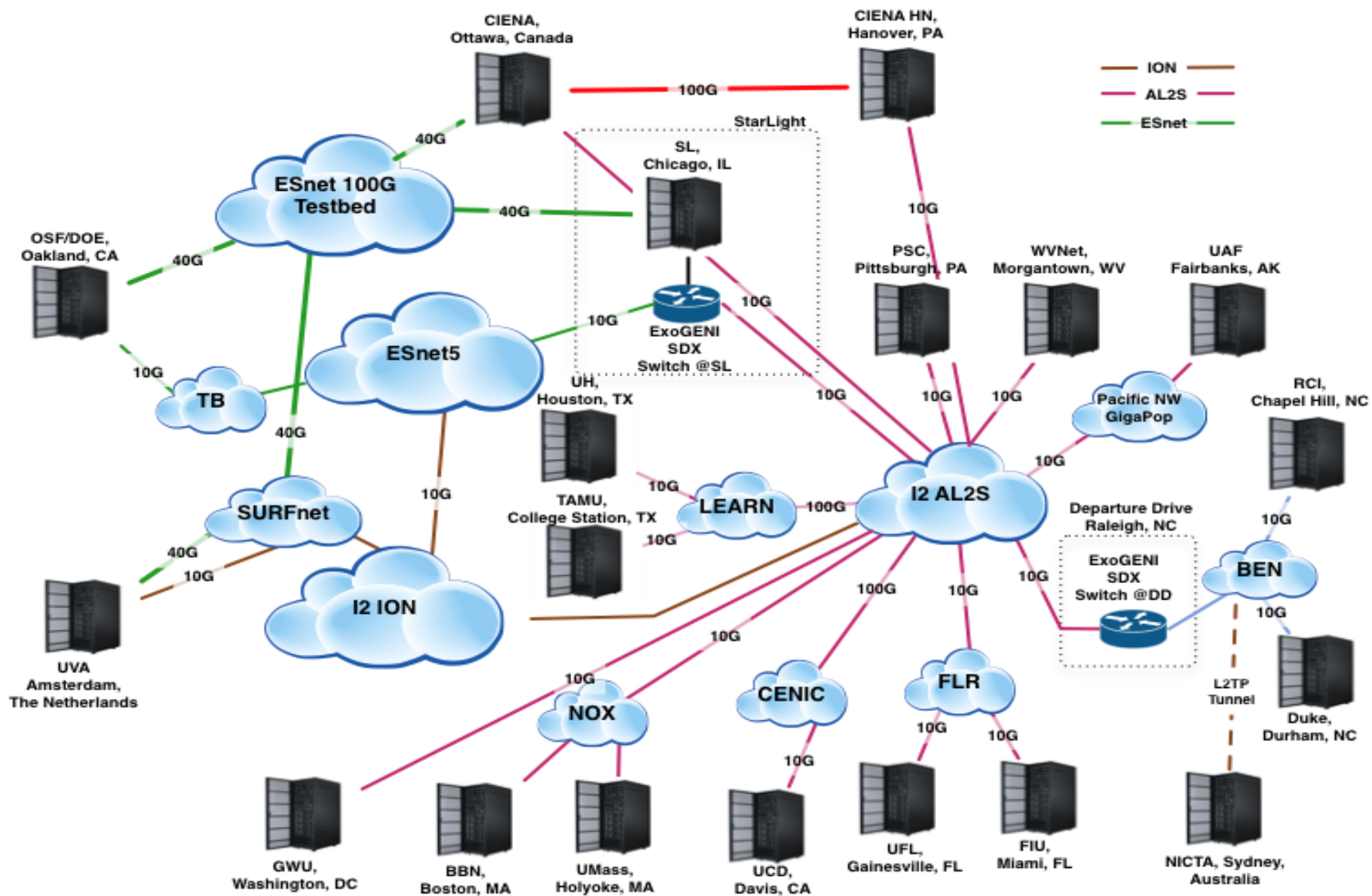
National Science Foundation Global Environment for Network innovations

InstaGENI Racks Network

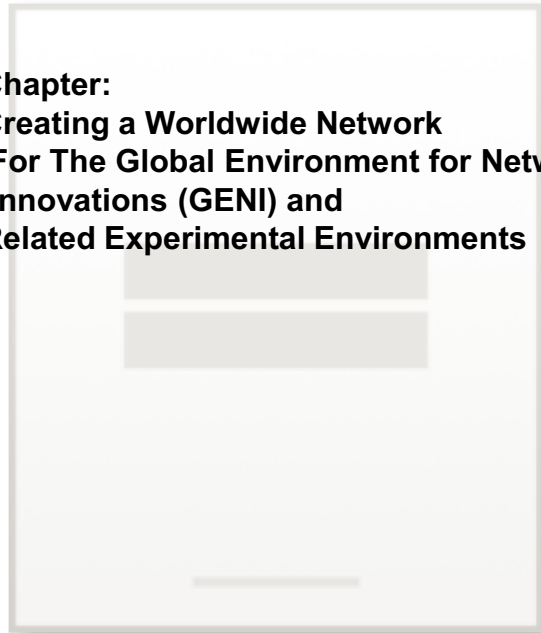


InstaGENI Network
Oct 2013

International 40G and 100 G ExoGENI Testbed



Chapter:
Creating a Worldwide Network
For The Global Environment for Network
Innovations (GENI) and
Related Experimental Environments



1st ed. 2016, XVIII, 655 p. 216 illus., 183
illus. in color.

 Printed book

R. McGeer, M. Berman, C. Elliott, R. Ricci (Eds.)

The GENI Book

- ▶ Provides a foundational overview of GENI's core architectural concepts
- ▶ Presents a detailed discussion of architecture and implementation
- ▶ Includes 24 chapters, divided into five sections, which outline GENI from precursors to architecture, development, applications, and then world federation
- ▶ Offers an extensive bibliography

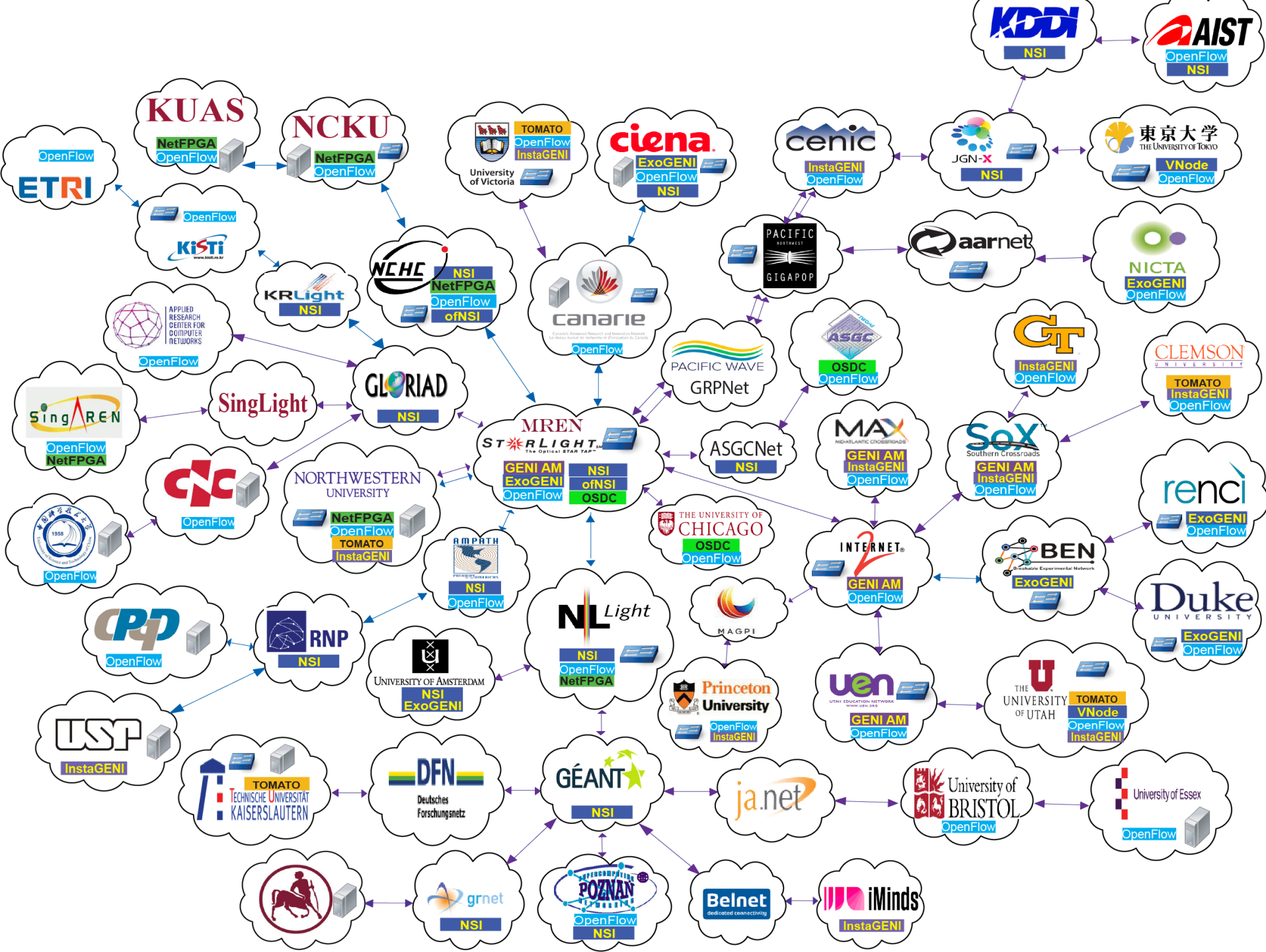
This book, edited by four of the leaders of the National Science Foundation's Global Environment and Network Innovations (GENI) project, gives the reader a tour of the history, architecture, future, and applications of GENI. Built over the past decade by hundreds of leading computer scientists and engineers, GENI is a nationwide network used daily by thousands of computer scientists to explore the next Cloud and Internet and the applications and services they enable, which will transform our communities and our lives. Since by design it runs on existing computing and networking equipment and over the standard commodity Internet, it is poised for explosive growth and transformational impact over the next five years.

iGENI: The International GENI

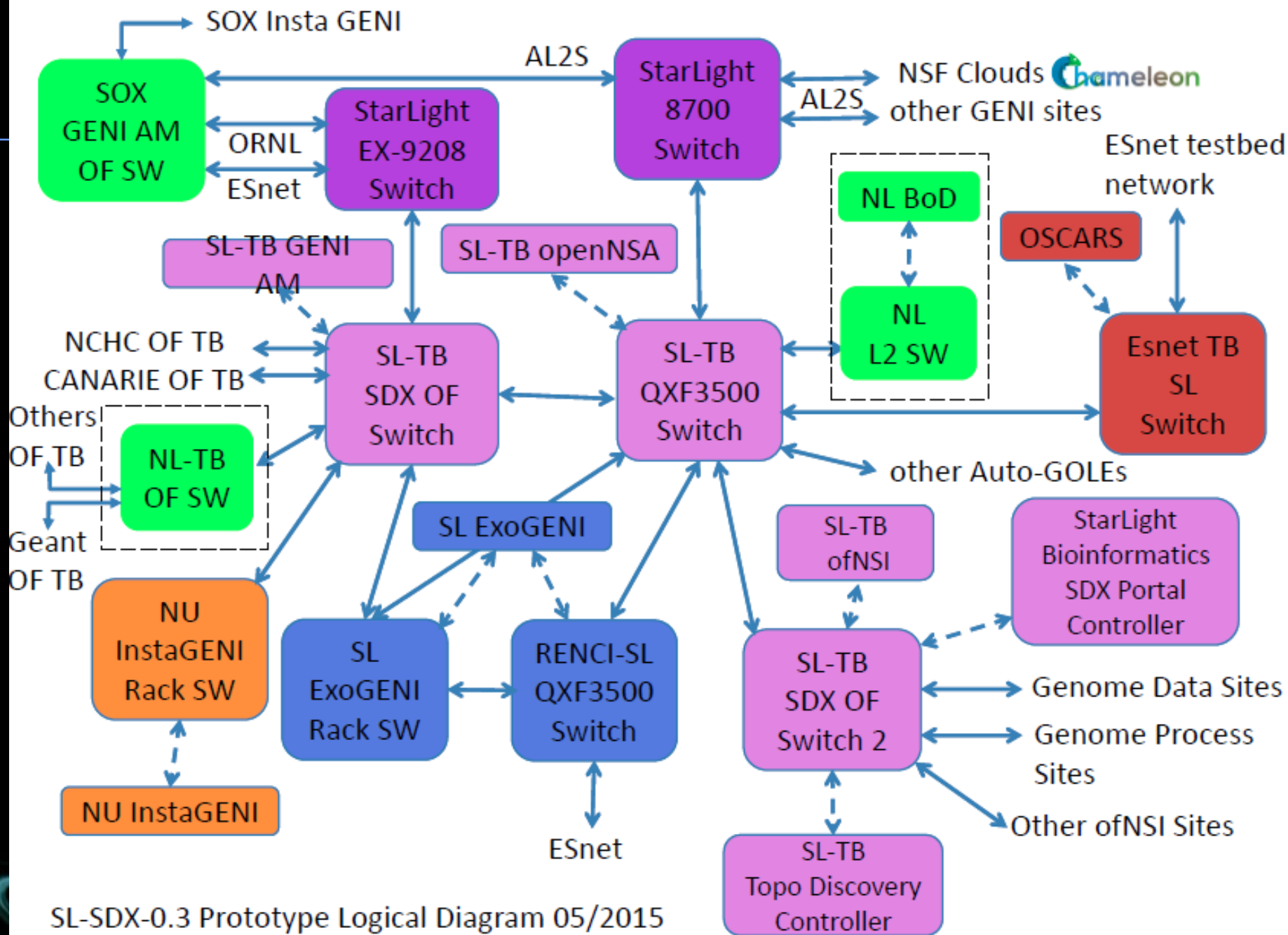
- The iGENI Initiative Will Design, Develop, Implement, and Operate a Major New National and International Distributed Infrastructure.
- iGENI Will Place the “G” in GENI Making GENI Truly Global.
- iGENI Will Be a Unique Distributed Infrastructure Supporting Research and Development for Next-Generation Network Communication Services and Technologies.
- This Infrastructure Will Be Integrated With Current and Planned GENI Resources, and Operated for Use by GENI Researchers Conducting Experiments that Involve Multiple Aggregates At Multiple Sites.
- iGENI Infrastructure Will Connect Its Resources With Current GENI National Backbone Transport Resources, With Current and Planned GENI Regional Transport Resources, and With International Research Networks and Projects,

•









SL-SDX-0.3 Prototype Logical Diagram 05/2015

SDX StarLight⇔NetherLight

STARLIGHT™
The Optical STAR TAP™

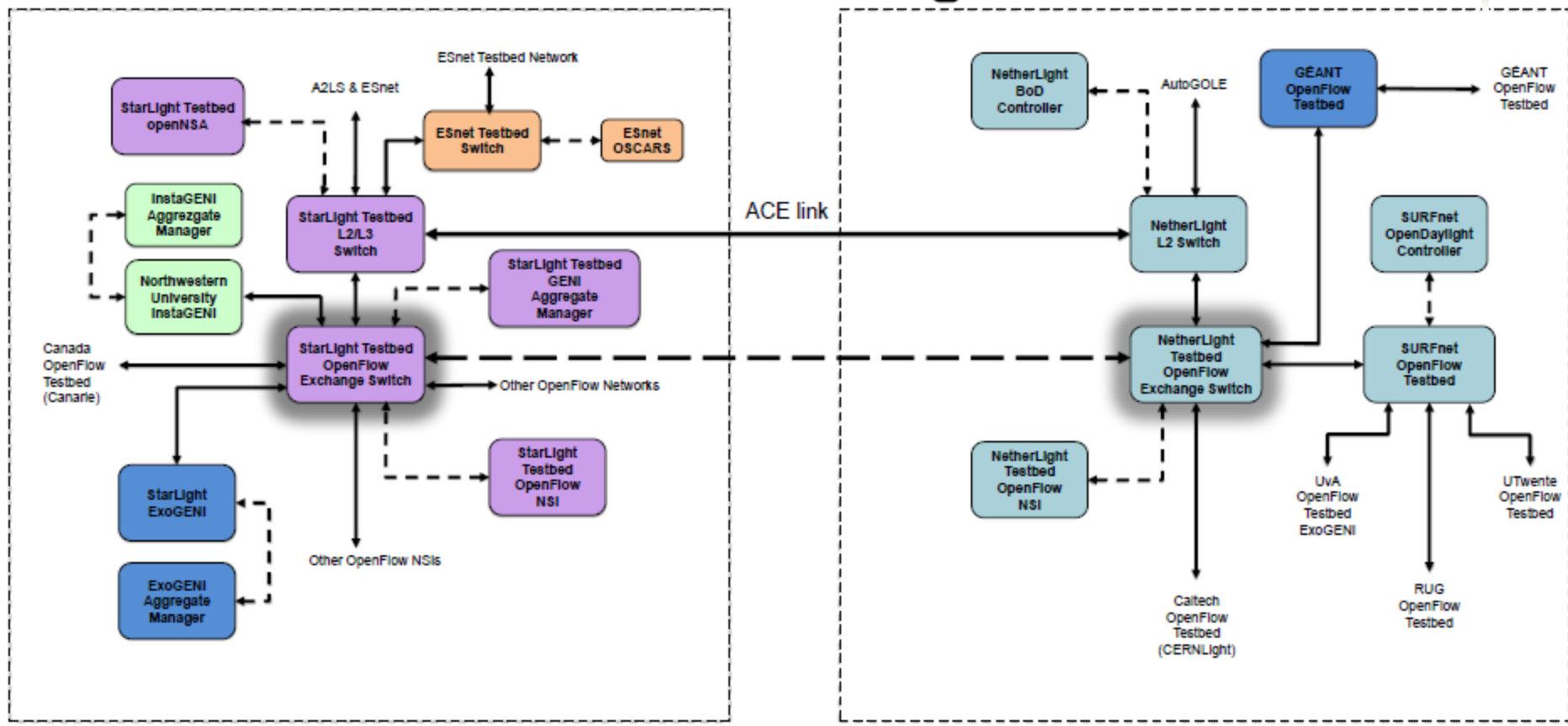
StarLight, Chicago

iCAIR

SURF NET

NetherLight, Amsterdam

NL Light



Ronald van der Pol, Joe Mambretti, Jim Chen, John Shillington

STARLIGHT™

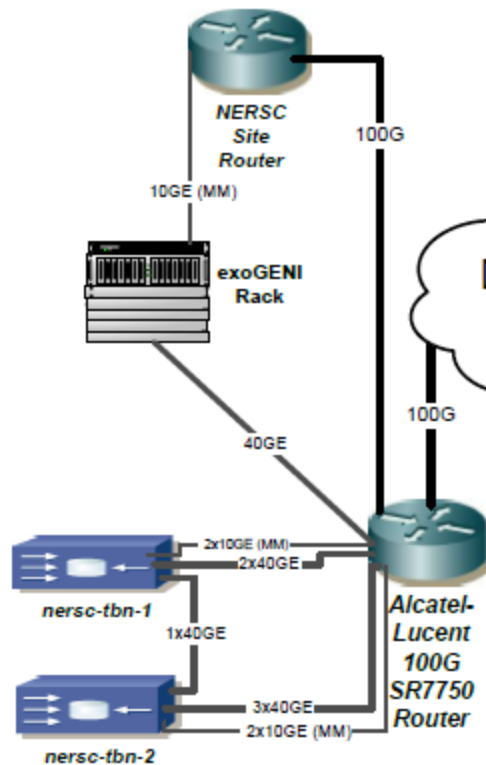
100G Component of Esnet SDN Testbed

NERSC

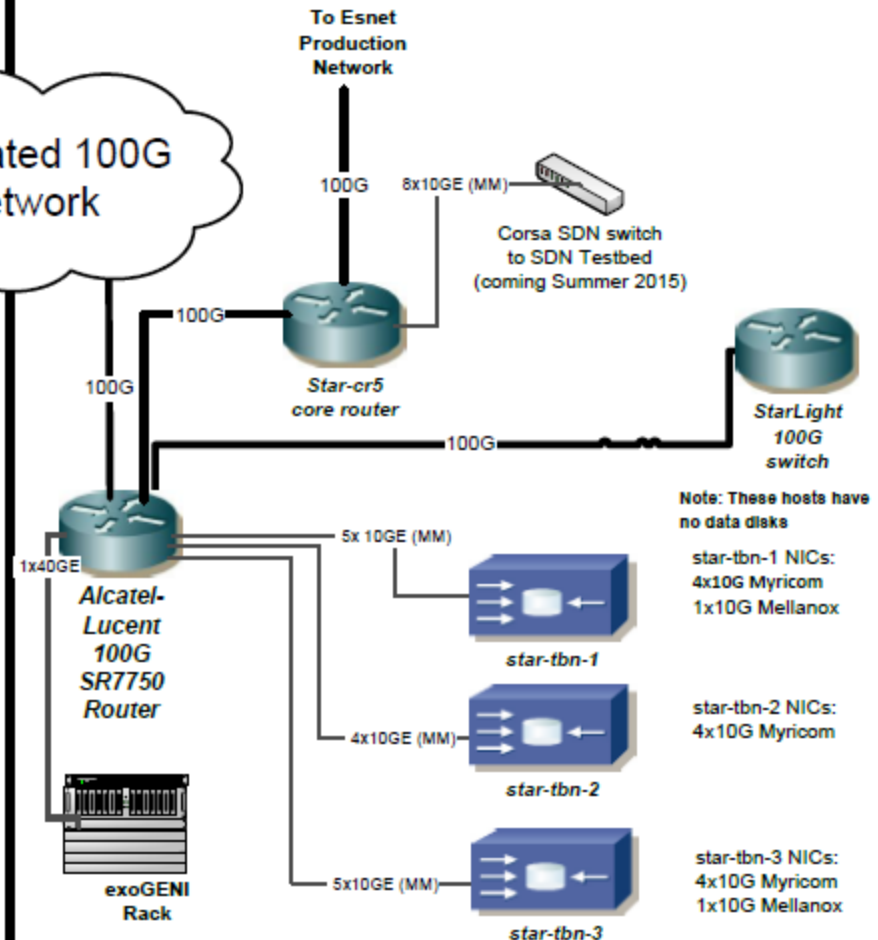
VLANS:
4012: All hosts
4020: Loop from NERSC to Chicago and back, all NERSC hosts

nersc-tbn-1 NICs:
2x40G Mellanox
1x40G Chelsio
2x10G Myricom
Disk: 24 HDDs

nersc-tbn-2 NICs:
4x40G Mellanox
1x40G Chelsio
2x10G Myricom
Disk: 24 SSDs



StarLight



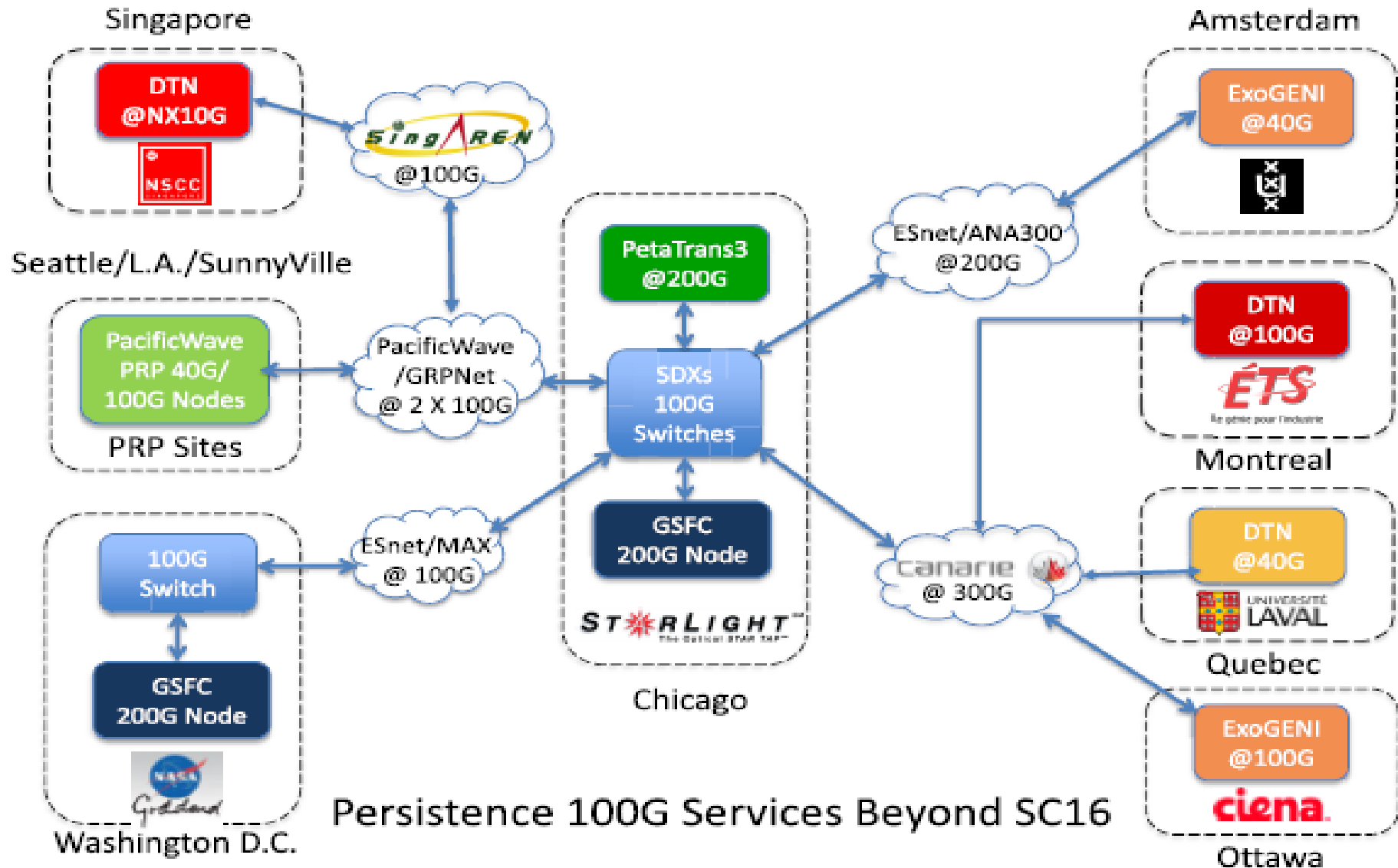
Note: These hosts have no data disks

star-tbn-1 NICs:
4x10G Myricom
1x10G Mellanox

star-tbn-2 NICs:
4x10G Myricom

star-tbn-3 NICs:
4x10G Myricom
1x10G Mellanox

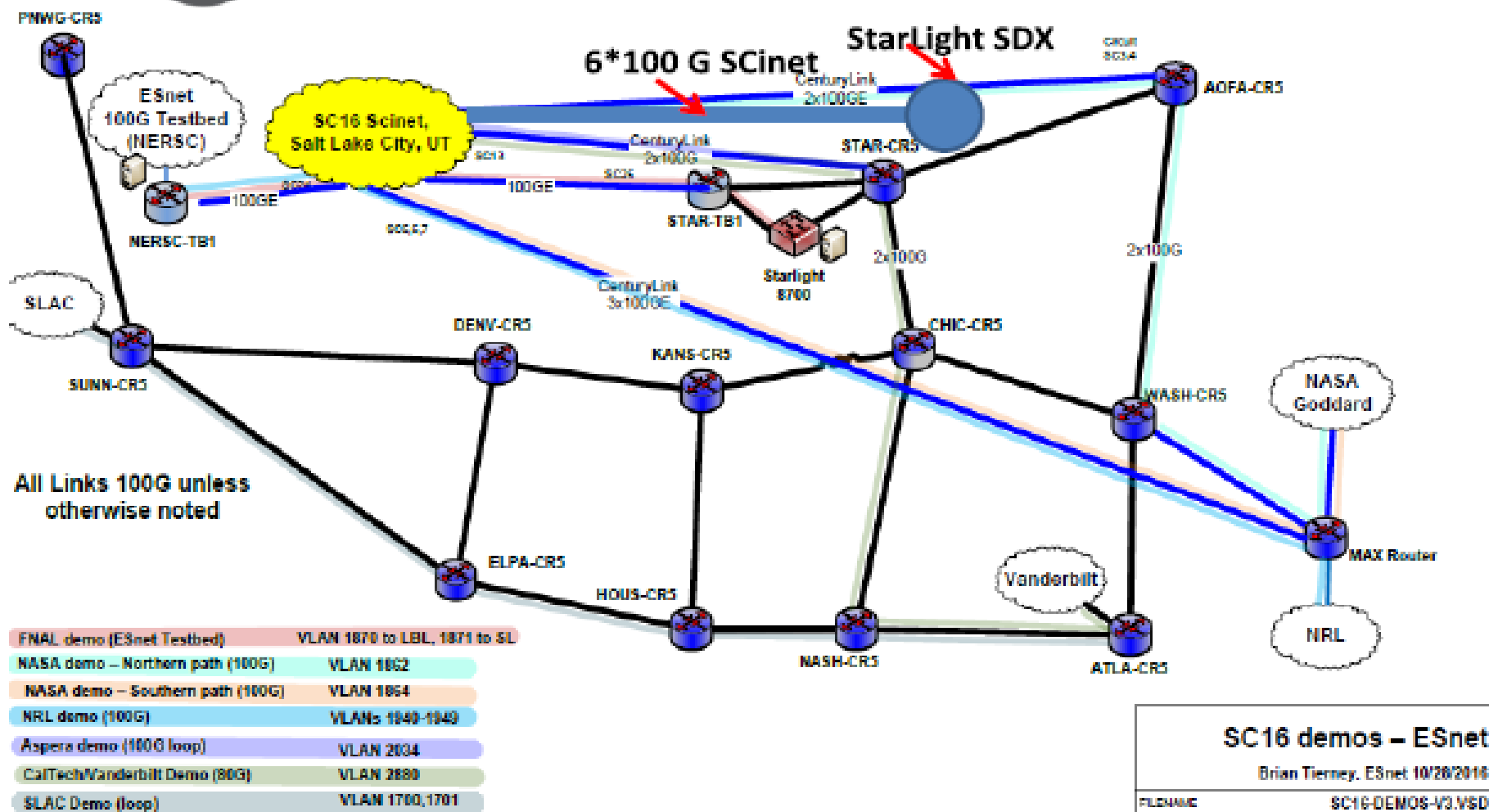
PetaTrans: Petascale Sciences Data Transfer





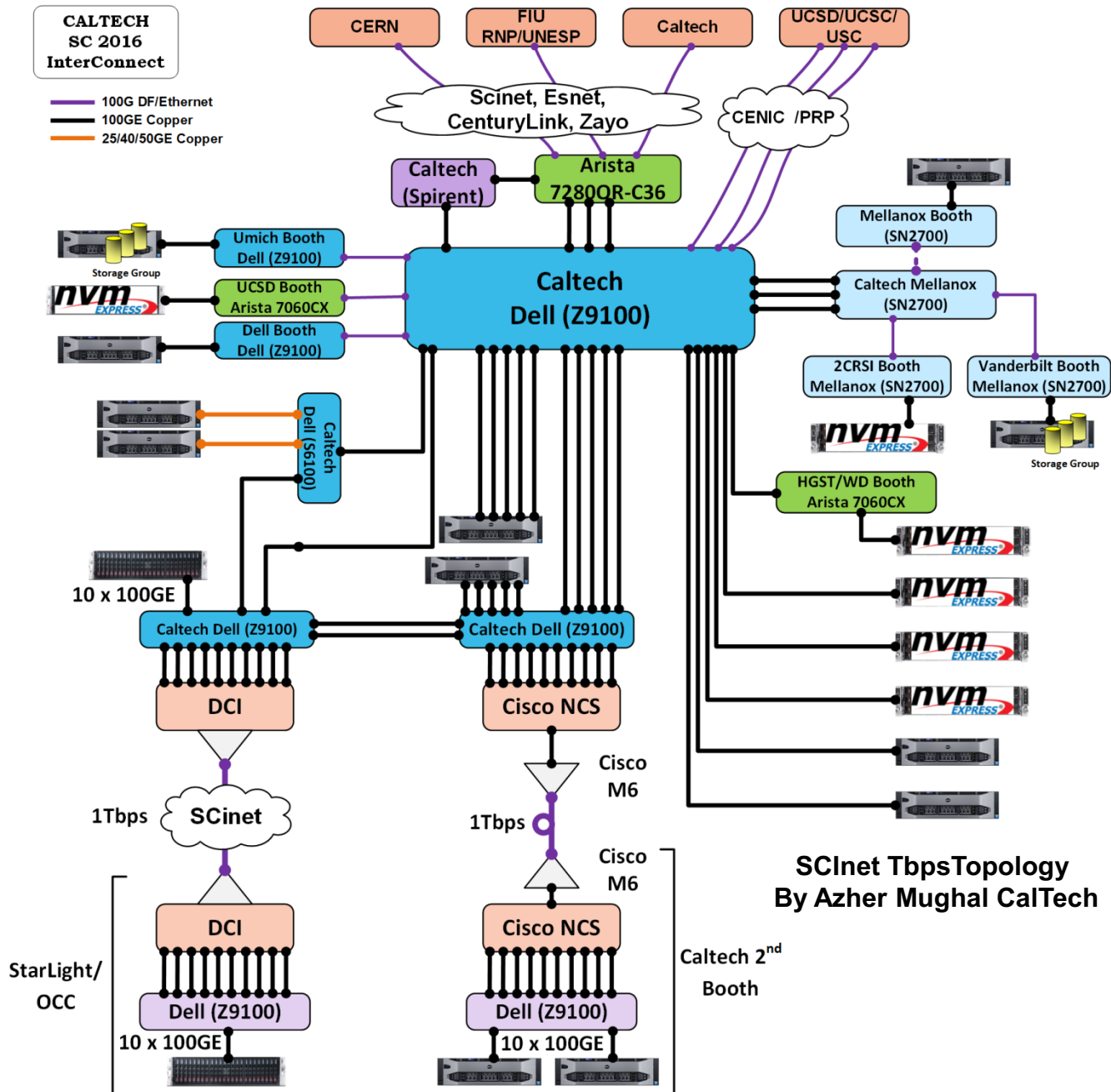
ESnet

ENERGY SCIENCES NETWORK




**CALTECH
SC 2016
InterConnect**

- 100G DF/Ethernet
- 100GE Copper
- 25/40/50GE Copper



0 1 2 3 4 5 6 7 8 9



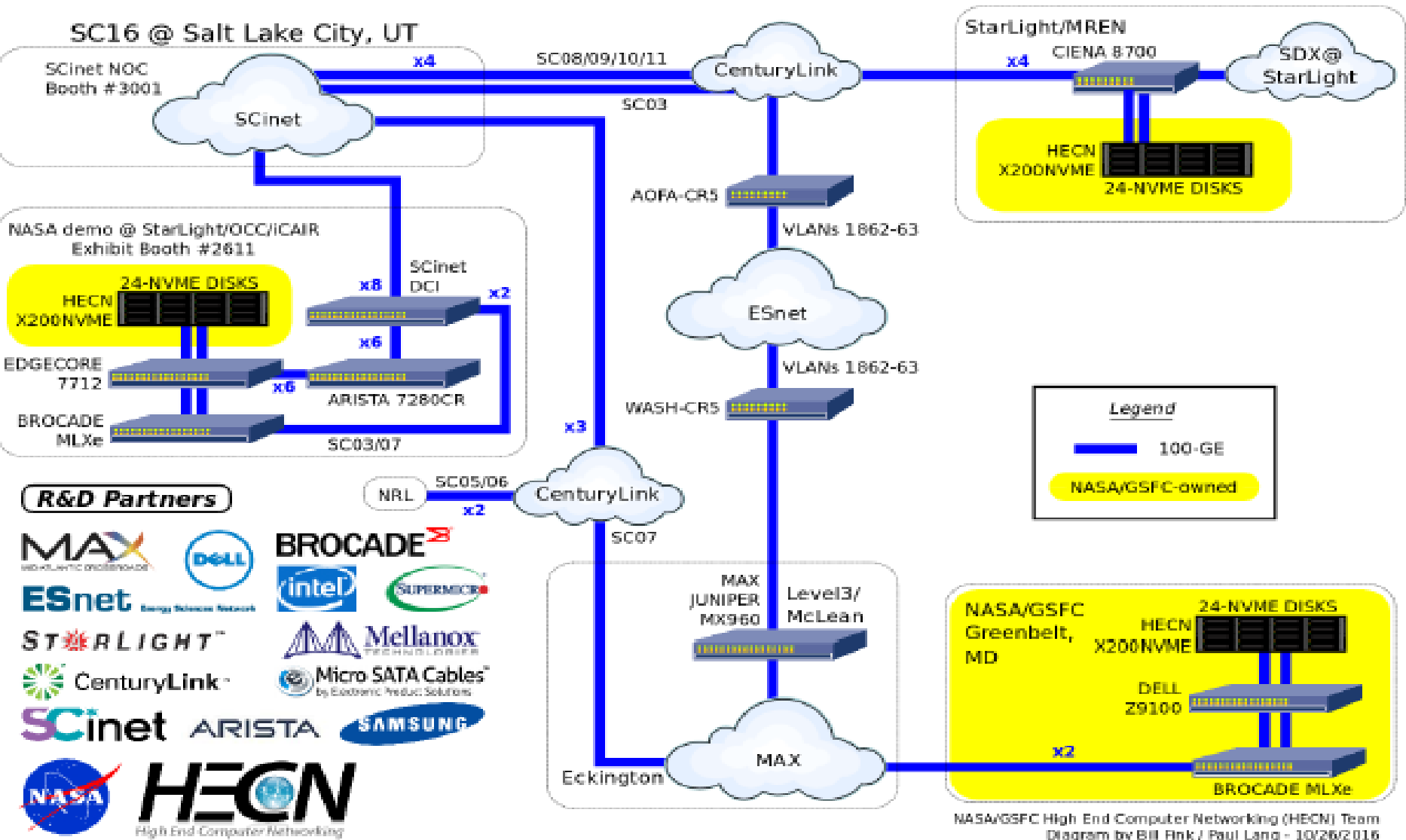
Salt Lake City, Utah **hpc**
matters.



SC16

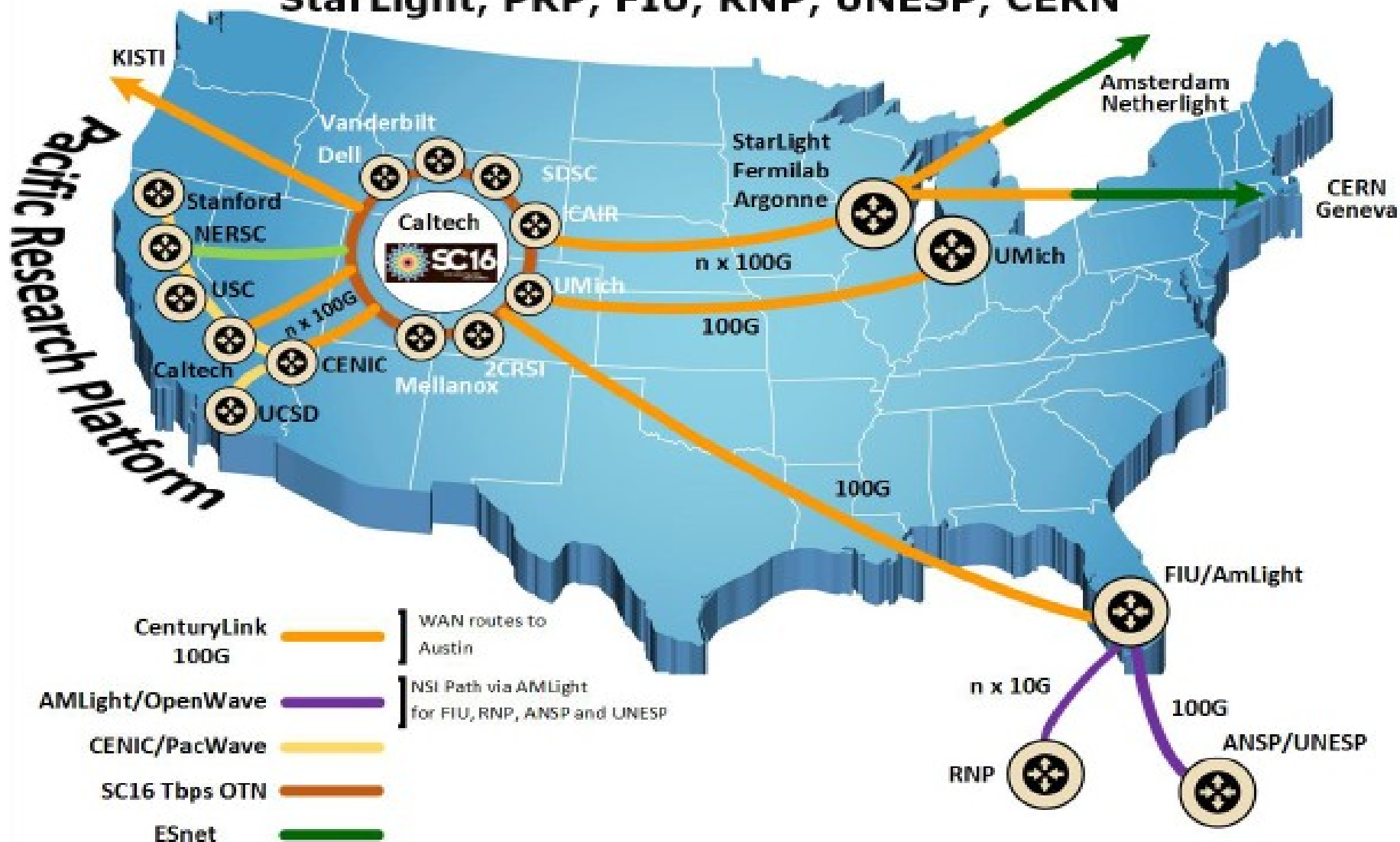
Demonstrations of 200 Gbps Disk-to-Disk WAN File Transfers using Parallelism across NVMe Drives

An SC16 Collaborative Initiative Among NASA and Several Partners



SC16 SDN-WAN Demonstration End-Points

Caltech, UM, Vanderbilt, UCSD, Dell, 2CRSI, KISTI, StarLight, PRP, FIU, RNP, UNESP, CERN







mdtmFTP @ 100Gbps Networks



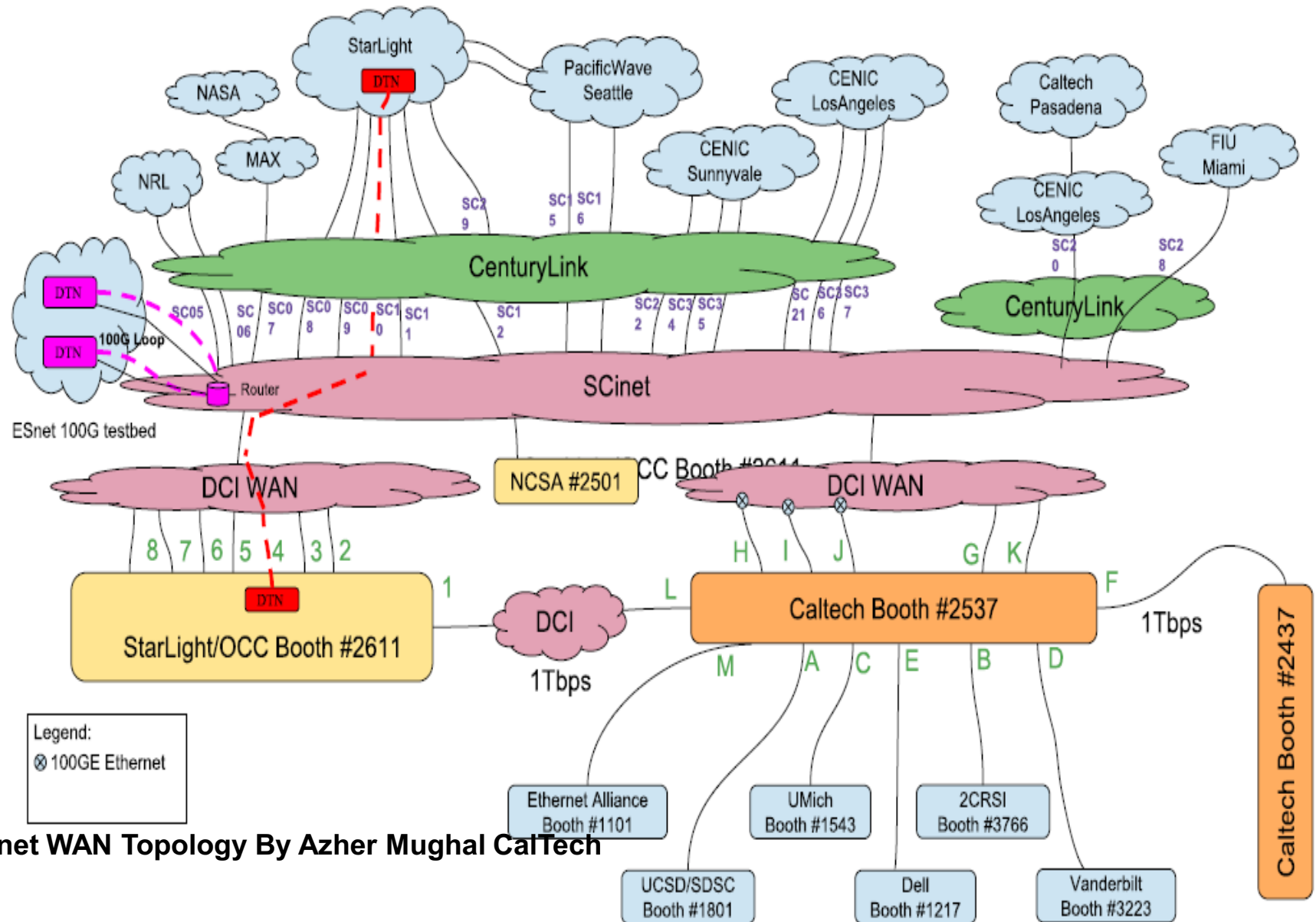
mdtmFTP: a high-performance data transfer tool

- Pipelined I/O-centric design to streamline data transfer
- Multicore-aware data transfer middleware (MDTM) optimizes use of underlying multicore system
- Extremely efficient in transferring of Lots Of Small Files (LOSF)



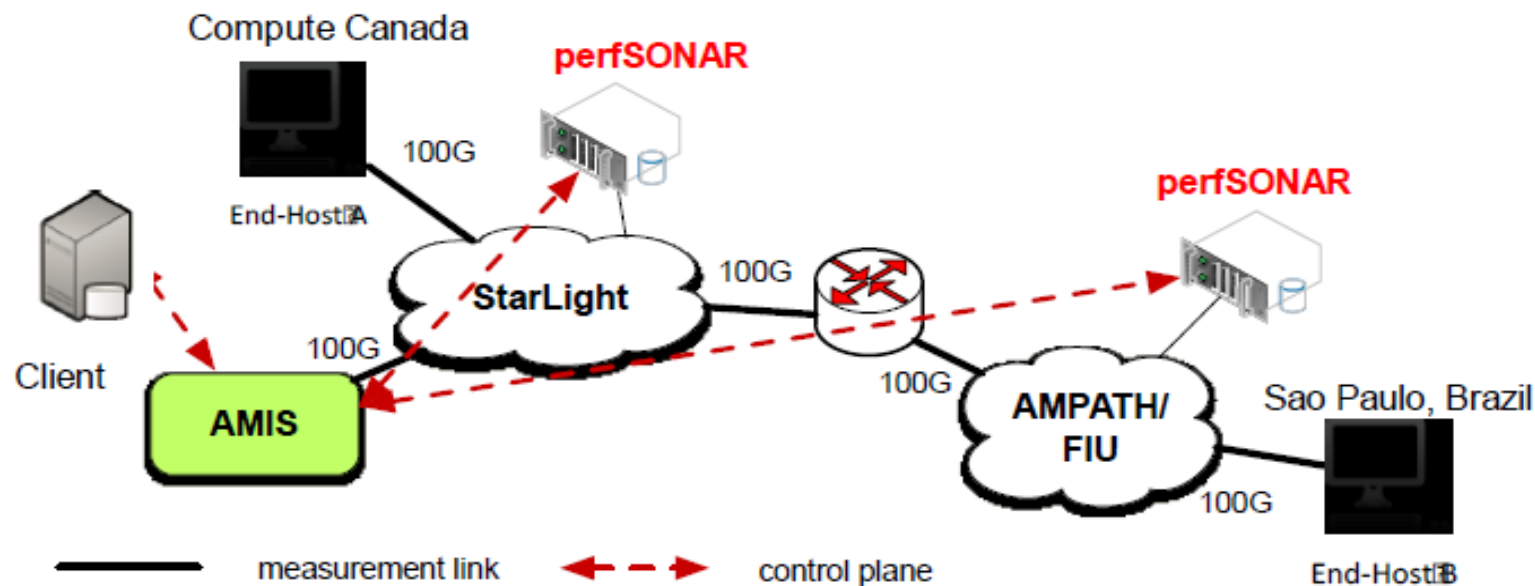
A DOE/SC/ASCR-sponsored research project

Software is available at: <http://mdtm.fnal.gov>



SCInet WAN Topology By Azher Mughal CalTech

Programmable Network Measurement of Data Intensive Flows on 100Gbps Networks



Demo1: Programmable Measurement with RESTful APIs

Demo2: Passive & Active Measurement (TCP window size)

Demo3: Passive & Active Measurement (TCP packet loss)

RNC AMIS Team: Yan Luo, PI, University of Massachusetts Lowell; Gabriel Ghinita, Co-PI, University of Massachusetts Boston; Cody Bumgardner, Co-PI, University of Kentucky; Michael McGarry, Co-PI, University of Texas El Paso. Contact: Yan_Luo@uml.edu

Collaborators: Jeo Mambretti, Jim Chen and Fei Yeh, StarLight/iCAIR/Northwestern University; Jeronimo Bezerra, AMPATH/Florida International University

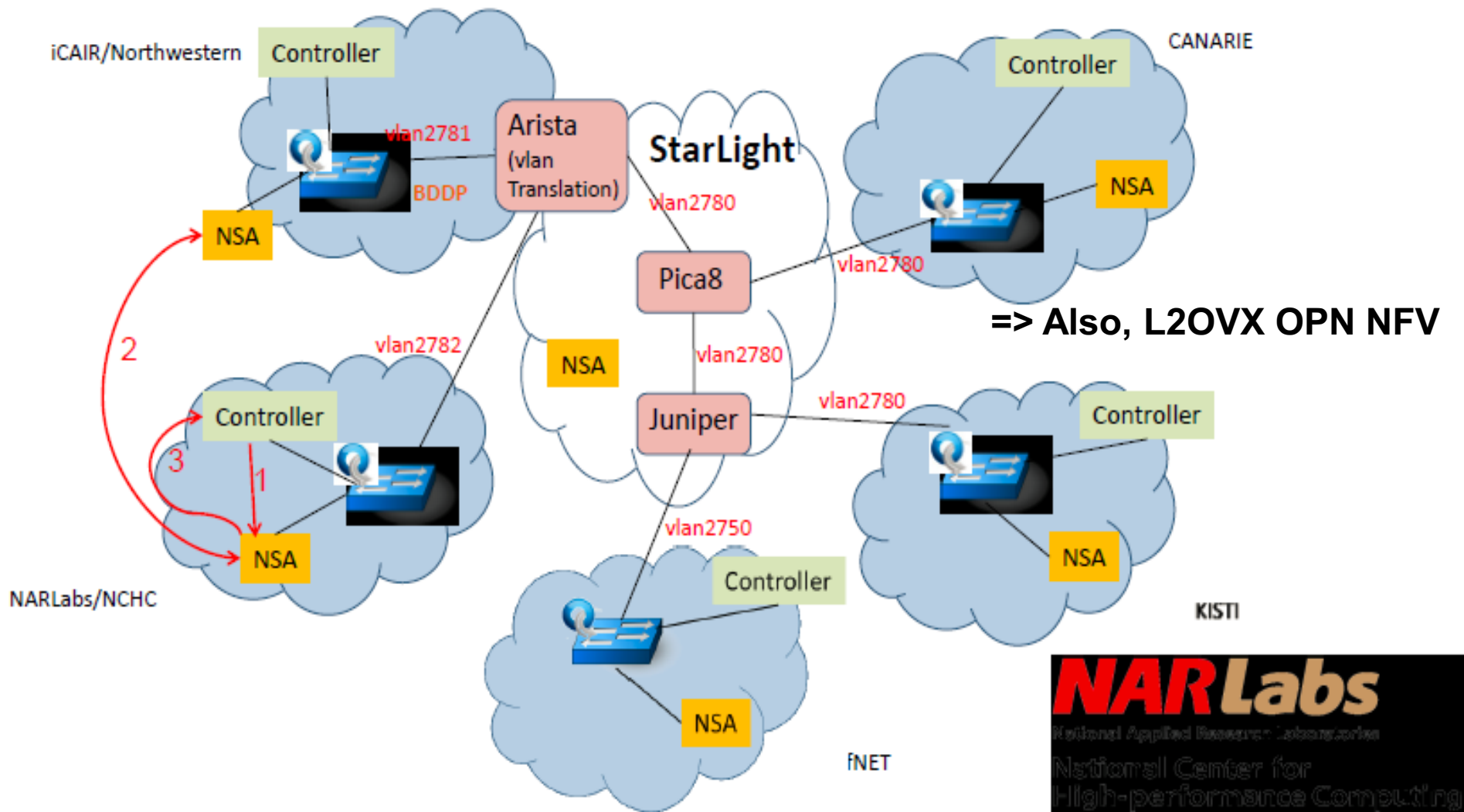


Northwestern
University



STARLIGHT

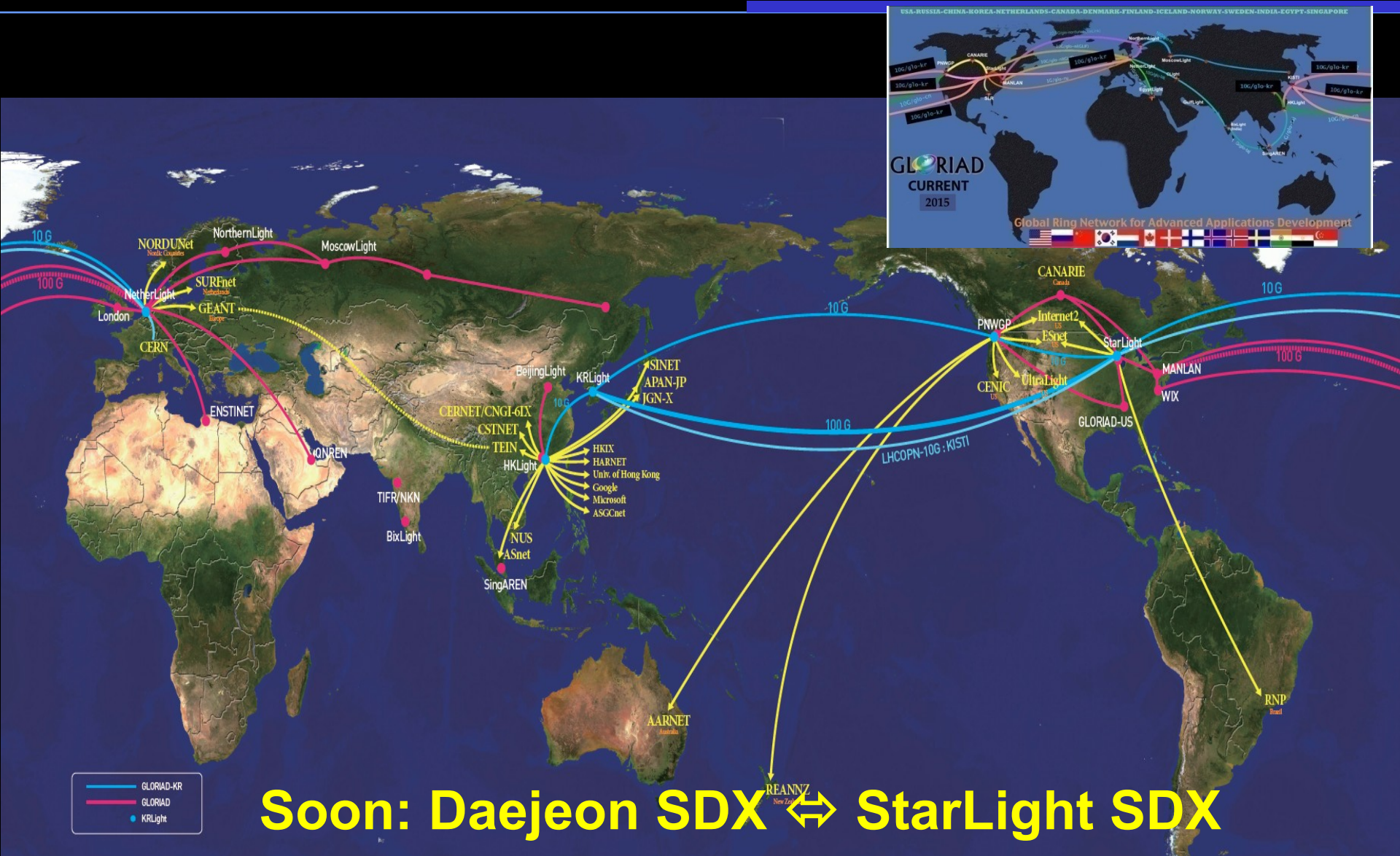
NSI-OpenFlow Hybrid Topology Exchange



Coming Soon => Taiwan SDX **STARLIGHTSM**

KREONet2 and GLORIAD-KR

KISTI Daejeon \Leftrightarrow 100 G \Leftrightarrow StarLight





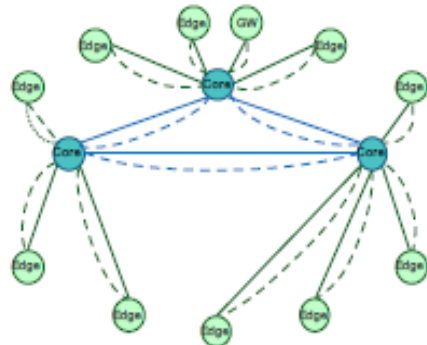
Nationwide 100 Gbps and Minimized Latency

- ◆ SINET5 will be a nationwide 100-Gbps backbone network using 100-Gigabit Ethernet technology and connect each pair of nodes with a minimized latency.

SINET4

- Star-like topology
- Resource-consuming secondary circuits

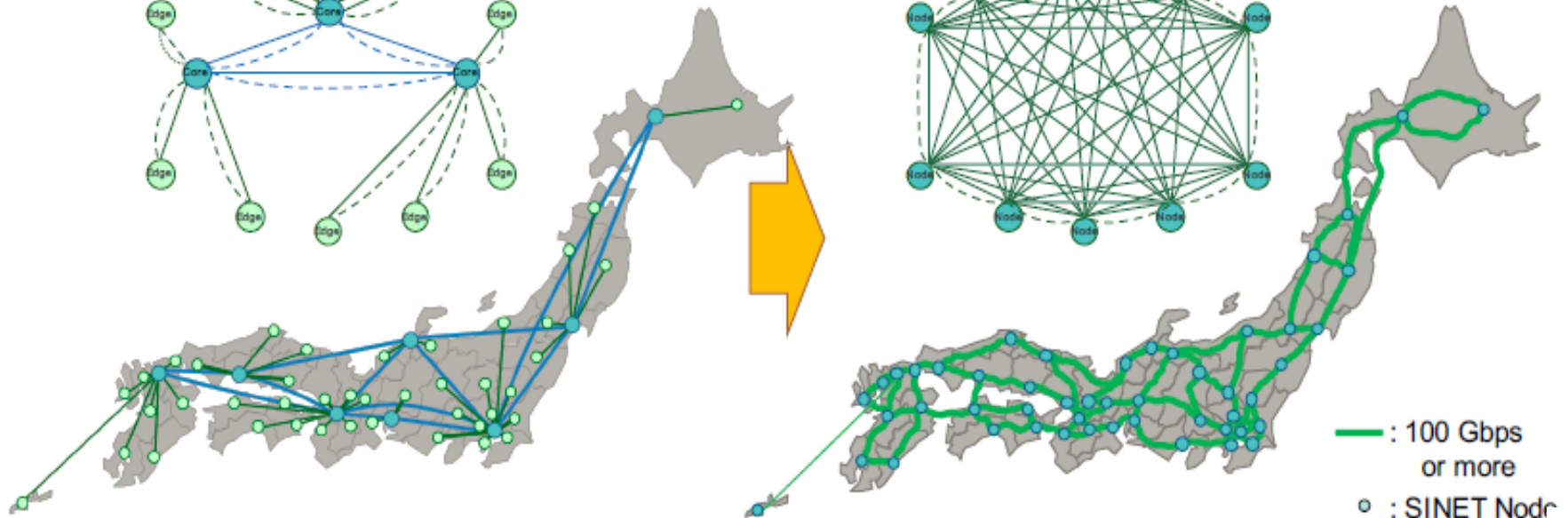
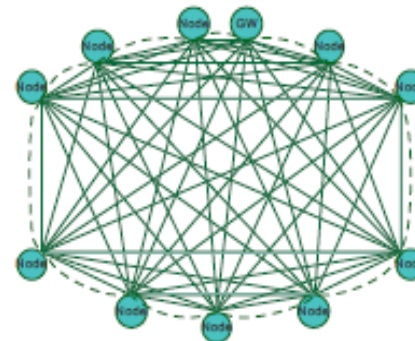
— : Leased Line (Primary Circuit)
- - : Leased Line (Secondary Circuit)



SINET5

- Fully-meshed topology with redundancy
- Non-resource-consuming secondary paths

— : MPLS-TP Path (Primary)
- - : MPLS-TP Path (Secondary)

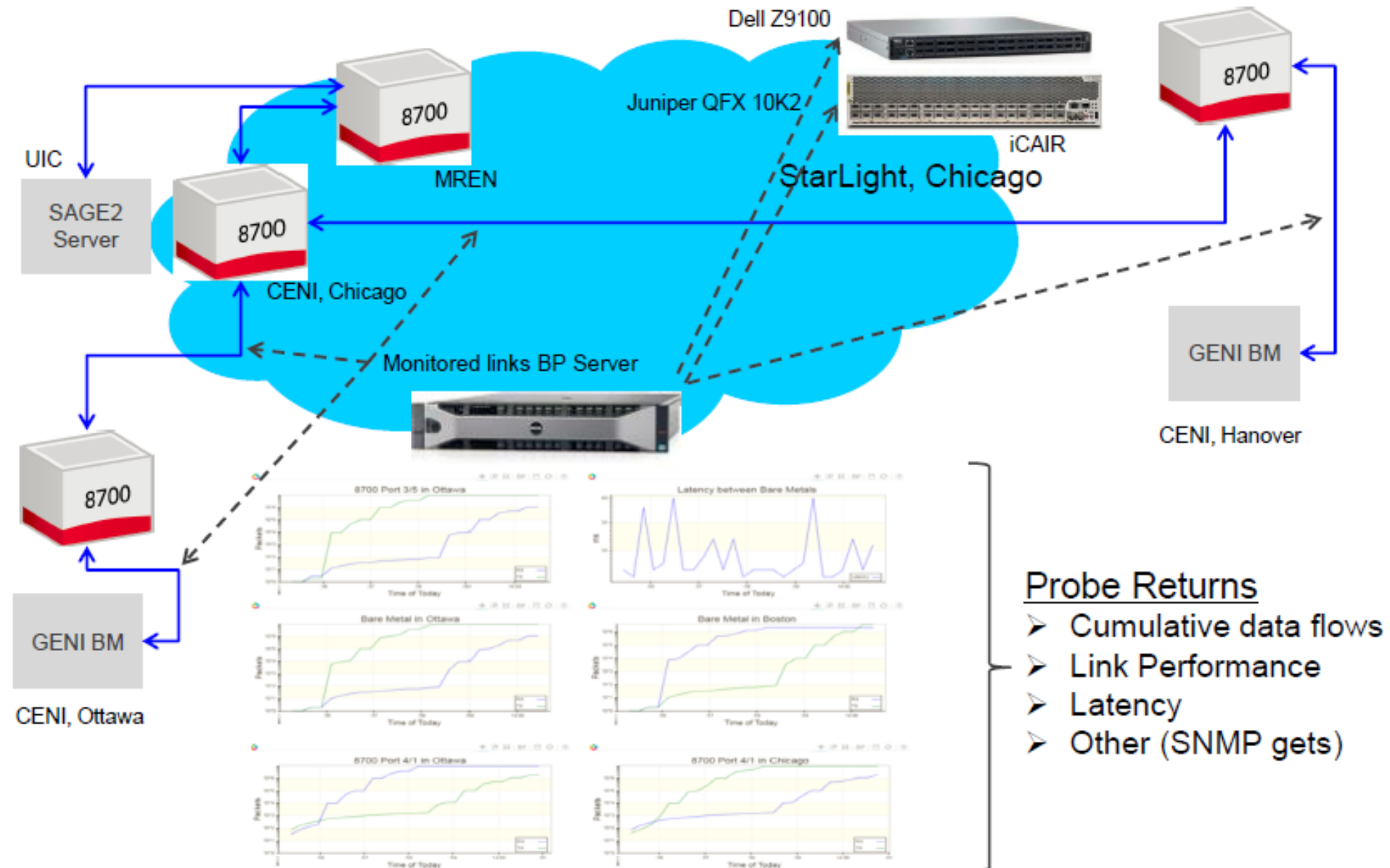


A*STAR Singapore

- **Singapore Supercomputing Center**
DTN ⇔ SingAREN ⇔ PacWavv ⇔ GRPnet ⇔
StarLight DTN ⇔ SC16
- **50-60 Gbps**



Network Diagram for Analytics Demonstration





Beyond Today's Internet Experiencing a Smart Future



Prototype SDX Bioinformatics Exchange: Demonstrating an Essential Use-Case for Personalized Medicine

Robert Grossman, Piers Nash, Allison
Heath, Renuka Arya
University of Chicago

Joe Mambretti, Jim Chen
Northwestern University



THE UNIVERSITY OF
CHICAGO
MEDICINE



NORTHWESTERN
UNIVERSITY

Genomic Data Commons Data Transfer

Data Commons Compute Status

[animate](#)

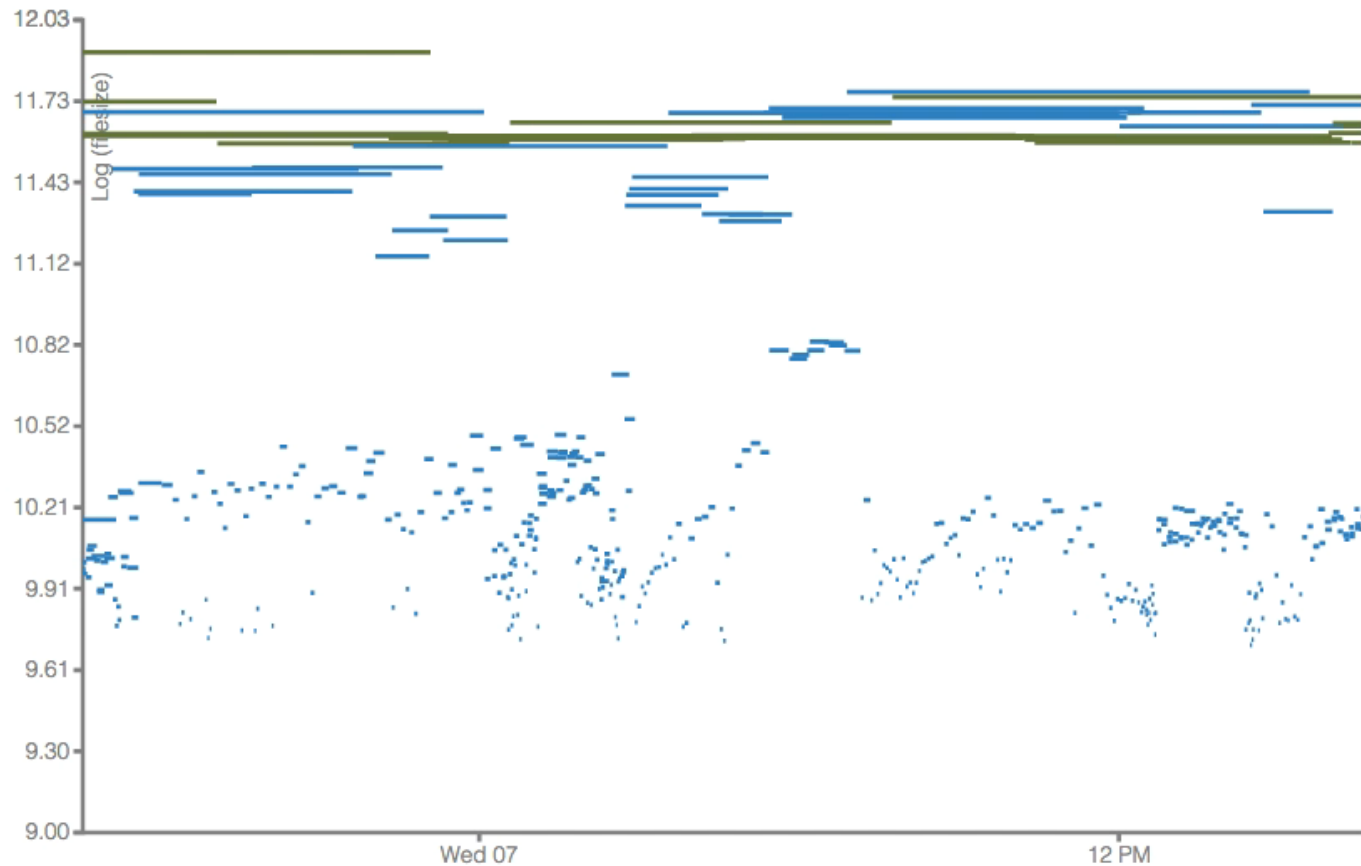
[stop](#)

■ ceph-TARGET

■ ceph-TCGA

■ cleversafe-TCGA

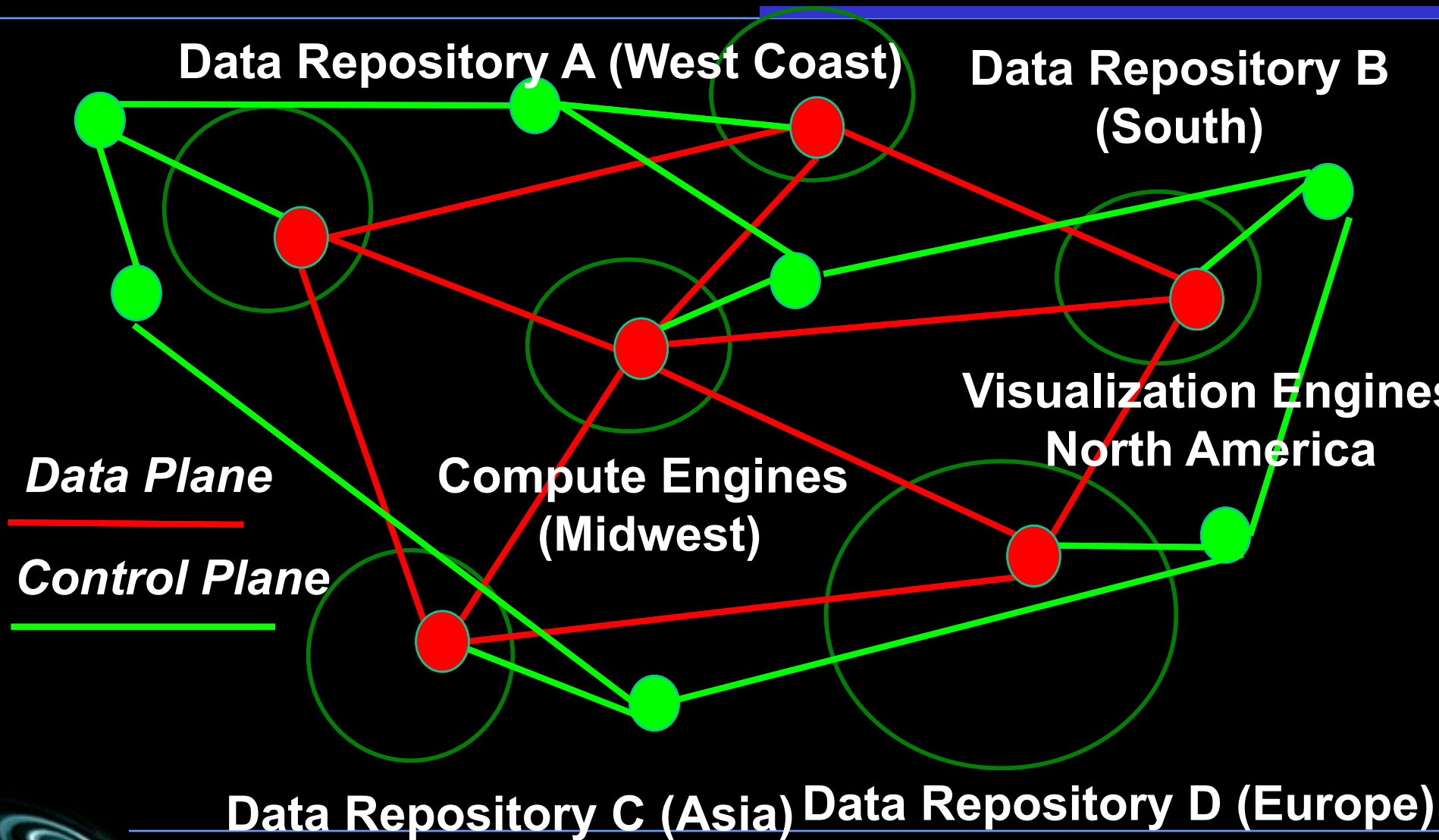
■ cleversafe-TARGET



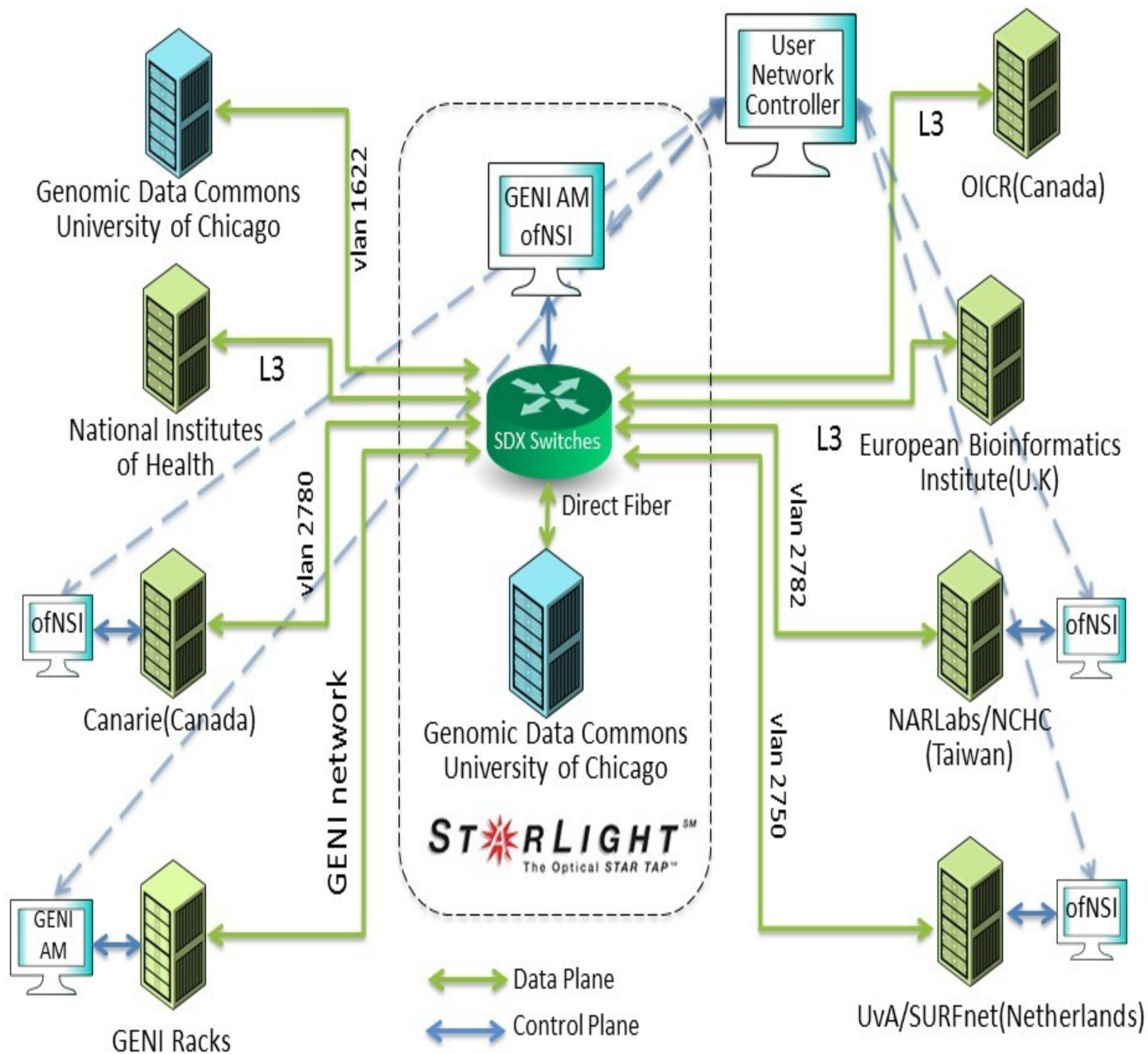


- STARLIGHTSM**

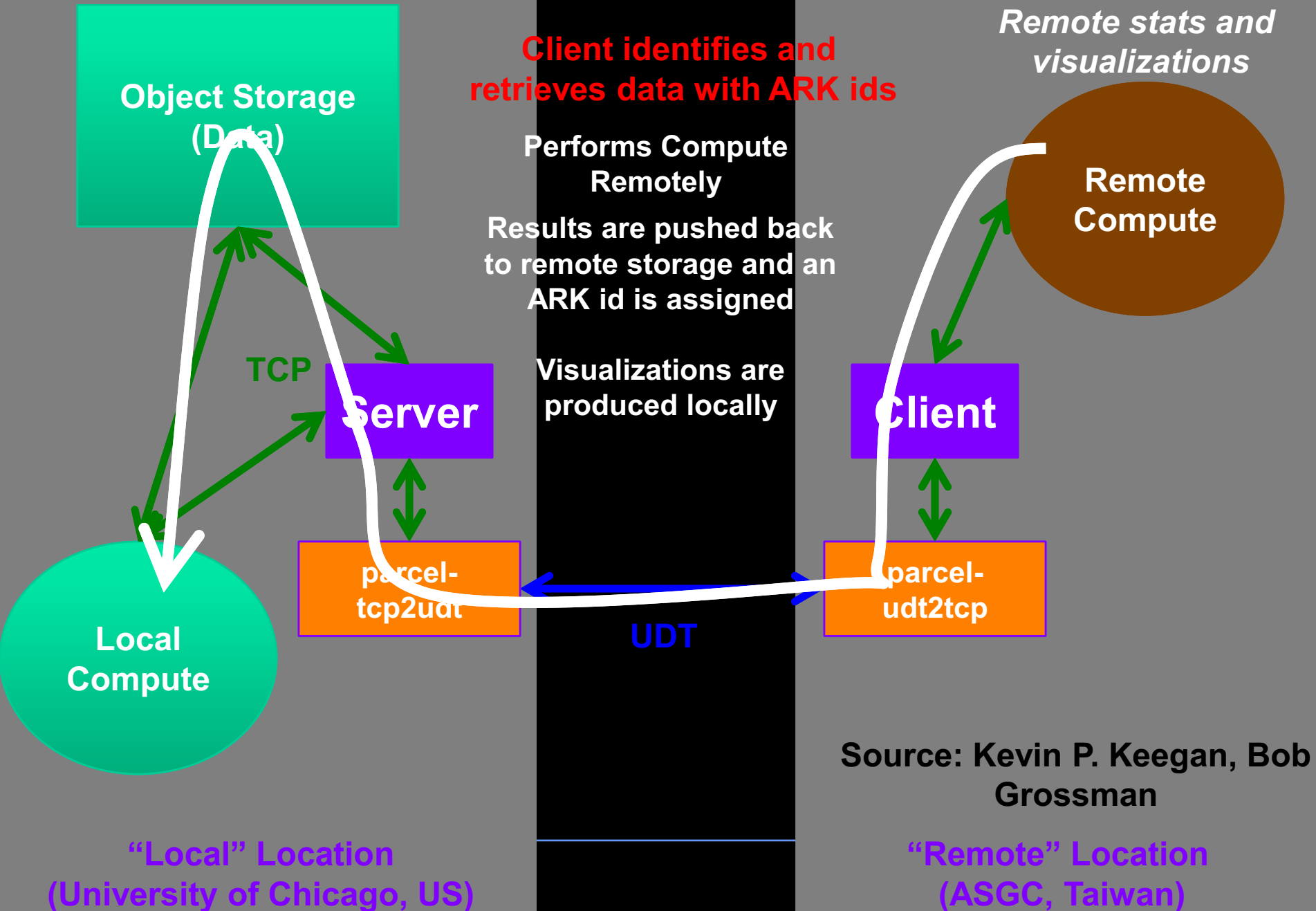
Biomedical Data Commons: Flow Orchestration: Control Plane + Data Plane



GEC22 Bioinformatics SDXs Demo Network



Parcel Based Collaboration



IRNC: RXP: StarLight SDX A Software Defined Networking Exchange for Global Science Research and Education

Joe Mambretti, Director, (j-mambretti@northwestern.edu)

International Center for Advanced Internet Research (www.icaair.org)

Northwestern University

Director, Metropolitan Research and Education Network (www.mren.org)

Co-Director, StarLight (www.startap.net/starlight)

PI IRNC: RXP: StarLight SDX

Co-PI Tom DeFanti, Research Scientist, (tdefanti@soe.ucsd.edu)

California Institute for Telecommunications and Information Technology (Calit2),

University of California, San Diego

Co-Director, StarLight

Co-PI Maxine Brown, Director, (maxine@uic.edu)

Electronic Visualization Laboratory, University of Illinois at Chicago

Co-Director, StarLight

Jim Chen, Associate Director, International Center for Advanced Internet

Research, Northwestern University

National Science Foundation

International Research Network Connections Program

Workshop

Chicago, Illinois

May 15, 2015





www.chameleoncloud.org

Another SDX Opportunity! An Experimental Testbed For Computer Science Research

CHAMELEON:

A LARGE-SCALE, RECONFIGURABLE EXPERIMENTAL ENVIRONMENT FOR CLOUD RESEARCH

Principal Investigator: Kate Keahey

Co-PIs: J. Mambretti, D.K. Panda, P. Rad, W. Smith, D. Stanzione

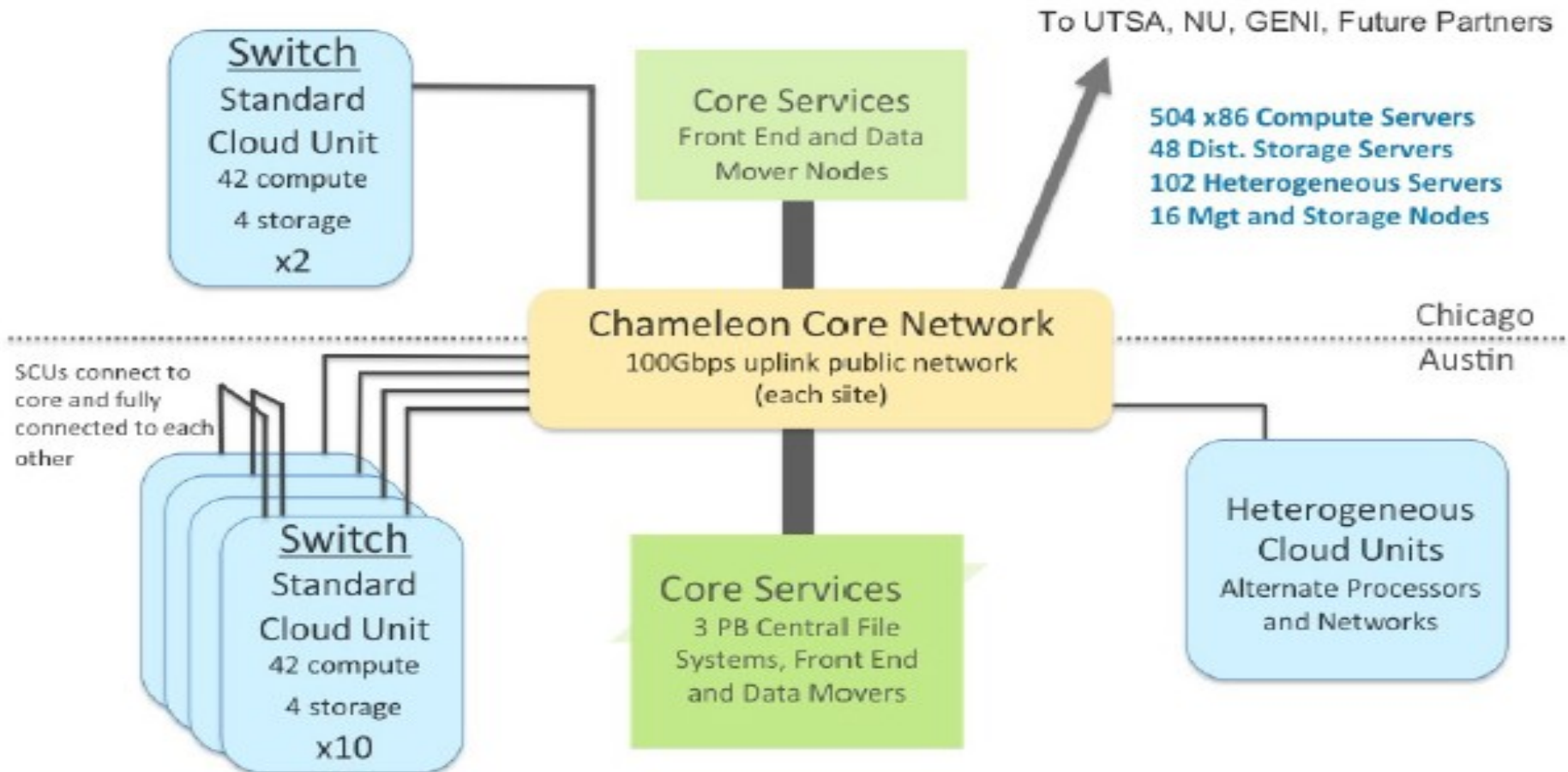
AUGUST 29, 2014

1

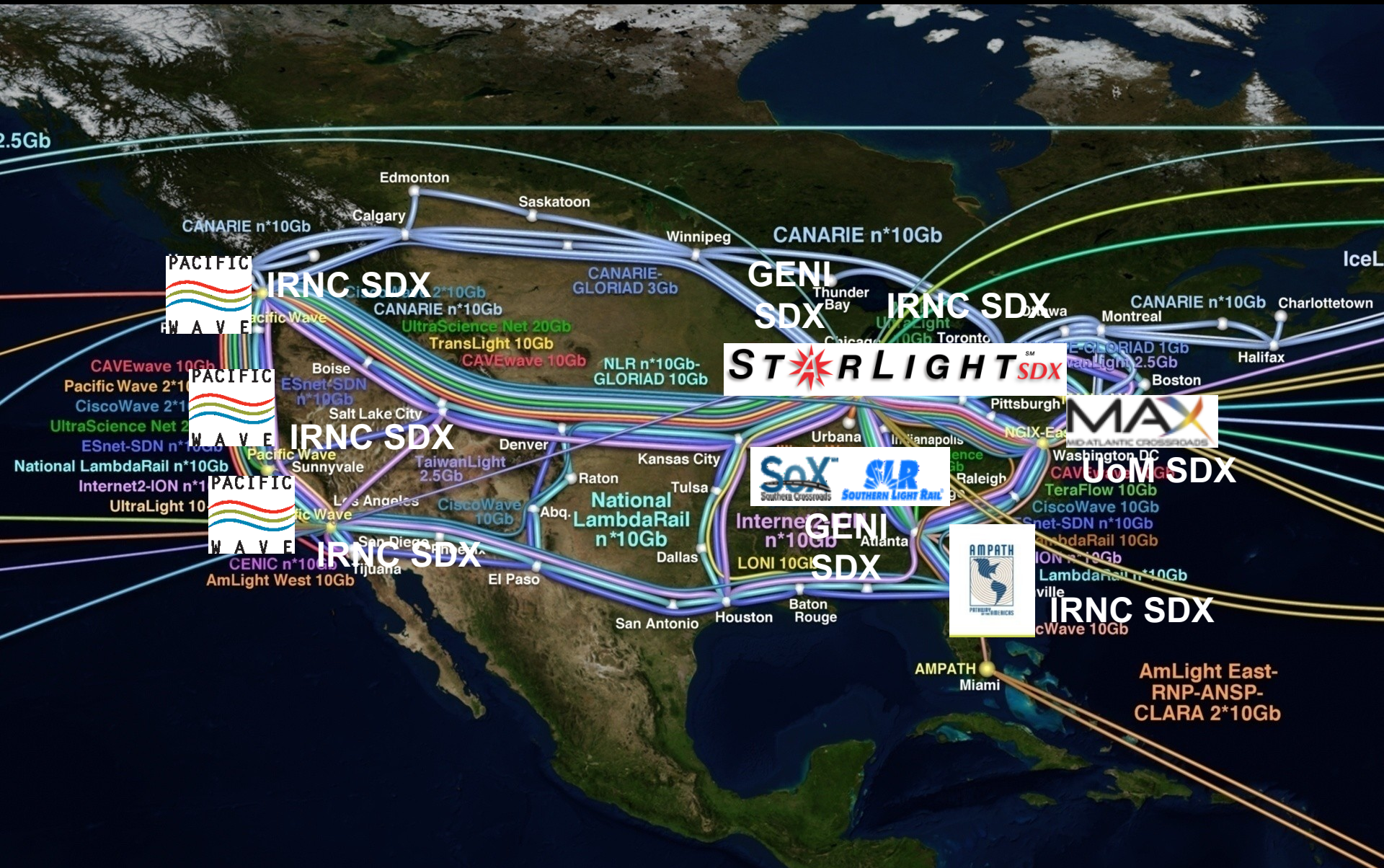


STARLIGHTSM

CHAMELEON ENVIRONMENT



Planned US SDX Interoperable Fabric



Next Step: Global Research Platform Building on CENIC/Pacific Wave and GLIF



INTERNATIONAL PEERING EXCHANGE



**Current
International
GRP Partners**



Global Research Platform

- A Emerging International Fabric
- A Specialized Globally Distributed Platform For Science Discovery and Innovation
- Based On State-Of-the-Art-Clouds
- Interconnected With Computational Grids, Supercomputing Centers, Specialized Instruments, et al
- Also, Based On World-Wide 100 Gbps Networks
- Leveraging Advanced Architectural Concepts, e.g., SDN/SDX/SDI – Science DMZs
- Ref: Demonstrations @ SC15, Austin Texas November 2015
- *New=> Global Research Platform 100 Gbps Network (GRPnet) On Private Optical Fiber Between PacificWave and StarLight via the PNWGP*

[**www.startap.net/starlight**](http://www.startap.net/starlight)

**Thanks to the NSF, DOE, DARPA
Universities, National Labs,
International Partners,
and Other Supporters**



STARLIGHTSM