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)

Middleware And Grid Interagency Coordination (MAGIC) – NITRD
Washington DC
February 1, 2017

GriPhyN
Data Intensive Science

NEPTUNE
A Fiber-Optic Telescope to Inner Space

NEES

GENOMES to LIFE
BIOLOGICAL SOLUTIONS FOR ENERGY CHALLENGES
U.S. DEPARTMENT OF ENERGY

NG Digital Sky Survey

NEON
National Ecological Observatory Network

ATLAS

ALMA: Atacama Large Millimeter Array

ANDRILL: Antarctic Geological Drilling
www.andrill.org

BIRN: Biomedical Informatics Research Network
www.nbirn.net

CAMERA metagenomics
camera.calit2.net

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

CineGrid
www.cinegrid.org

LHCONE
www.lhcone.net

GEON: Geosciences Network
www.geongrid.org

GLEON: Global Lake Ecological Observatory Network

OOI OCEAN OBSERVATORIES INITIATIVE
CYBERINFRASTRUCTURE
Providing a link between ocean research and discovery

OOI-CI
ci.oceanobservatories.org

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

Comprehensive Large-Array Stewardship System
www.class.noaa.gov

DØ (DZero)
www-d0.fnal.gov

Two neutron stars orbit each other

LIGO
www.ligo.org

WLCG
Worldwide LHC Computing Grid

WLCG
lcg.web.cern.ch/LCG/public/

Pacific Rim Applications and Grid Middleware Assembly
www.pragma-grid.net

TeraGrid
www.teragrid.org

IVOA: International Virtual Observatory
www.ivoa.net

Open Science Grid

OSG
www.opensciencegrid.org

the globus alliance

Globus Alliance
www.globus.org

SKA
SQUARE KILOMETRE ARRAY

SKA
www.skatelescope.org

Sloan Digital Sky Survey
www.sdss.org

XSEDE
www.xsede.org

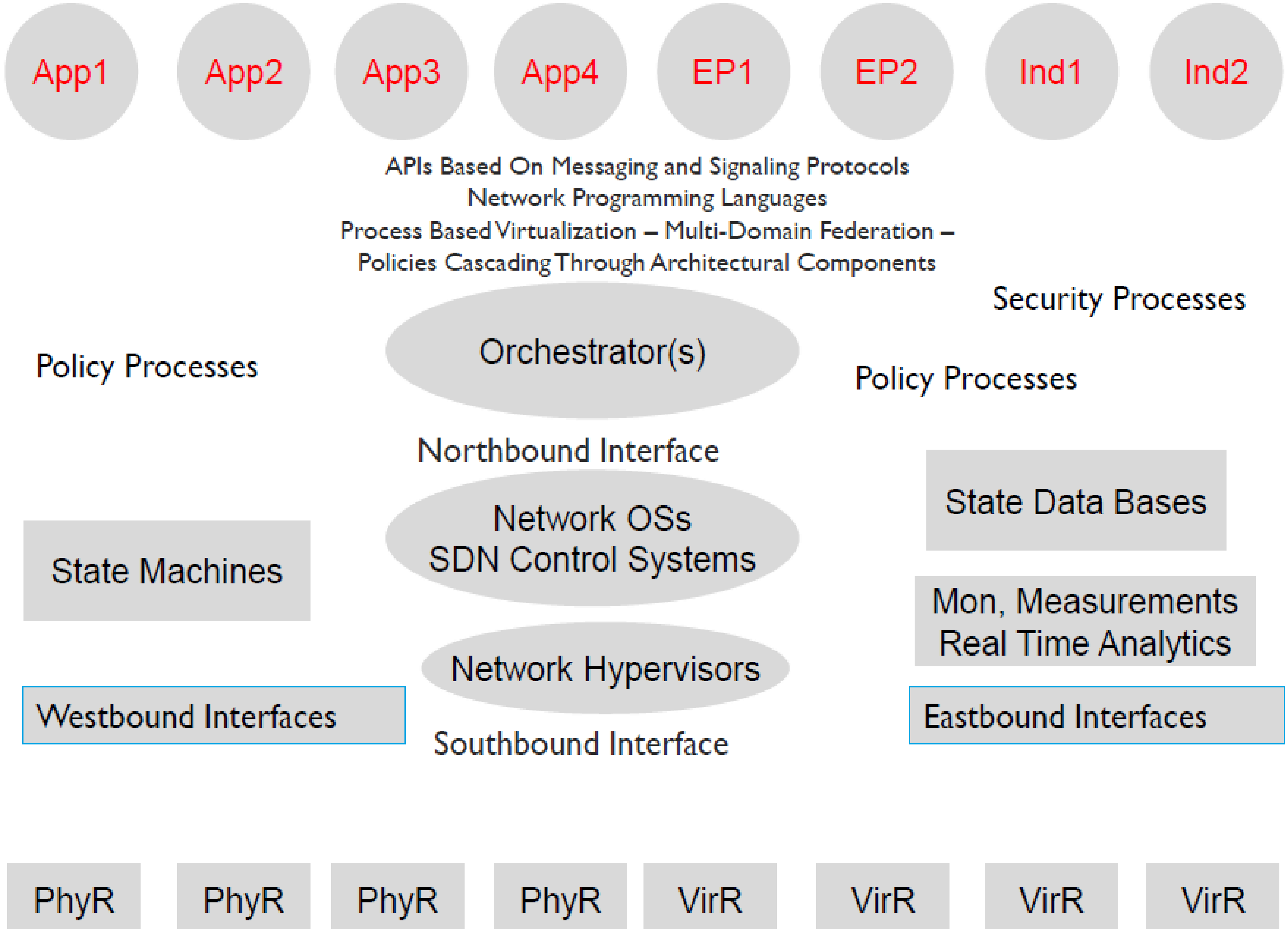
LIGO

Compilation By Maxine Brown

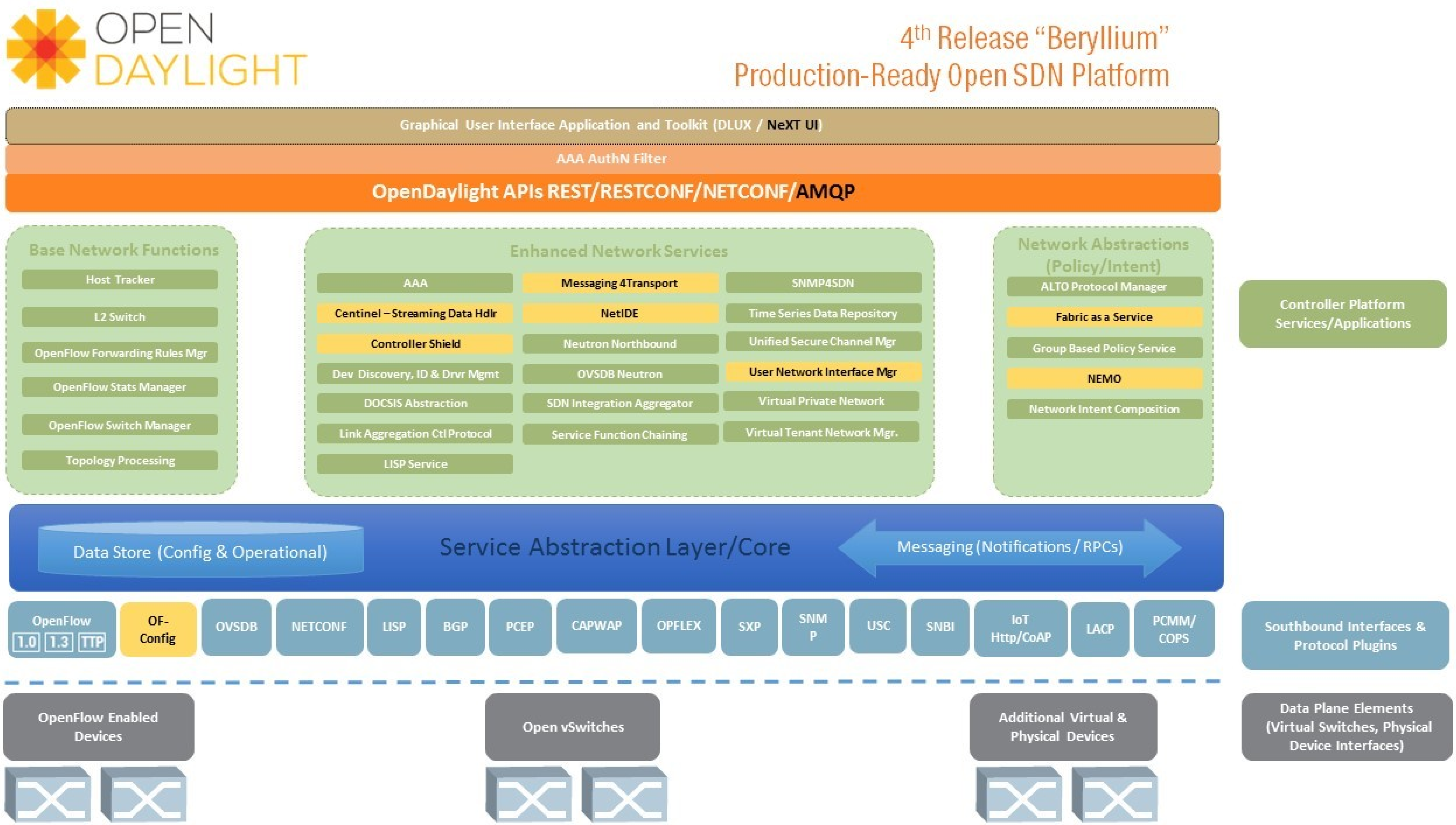
STARLIGHT℠

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
State Data Bases
Network OSs
SDN Control Systems
State Machines
Mon, Measurements
Real Time Analytics
Network Hypervisors
Westbound Interfaces
Eastbound Interfaces
Southbound Interface
PhyR
PhyR
PhyR
PhyR
VirR
VirR
VirR
VirR

Opendaylight 4th Release: Baryllium



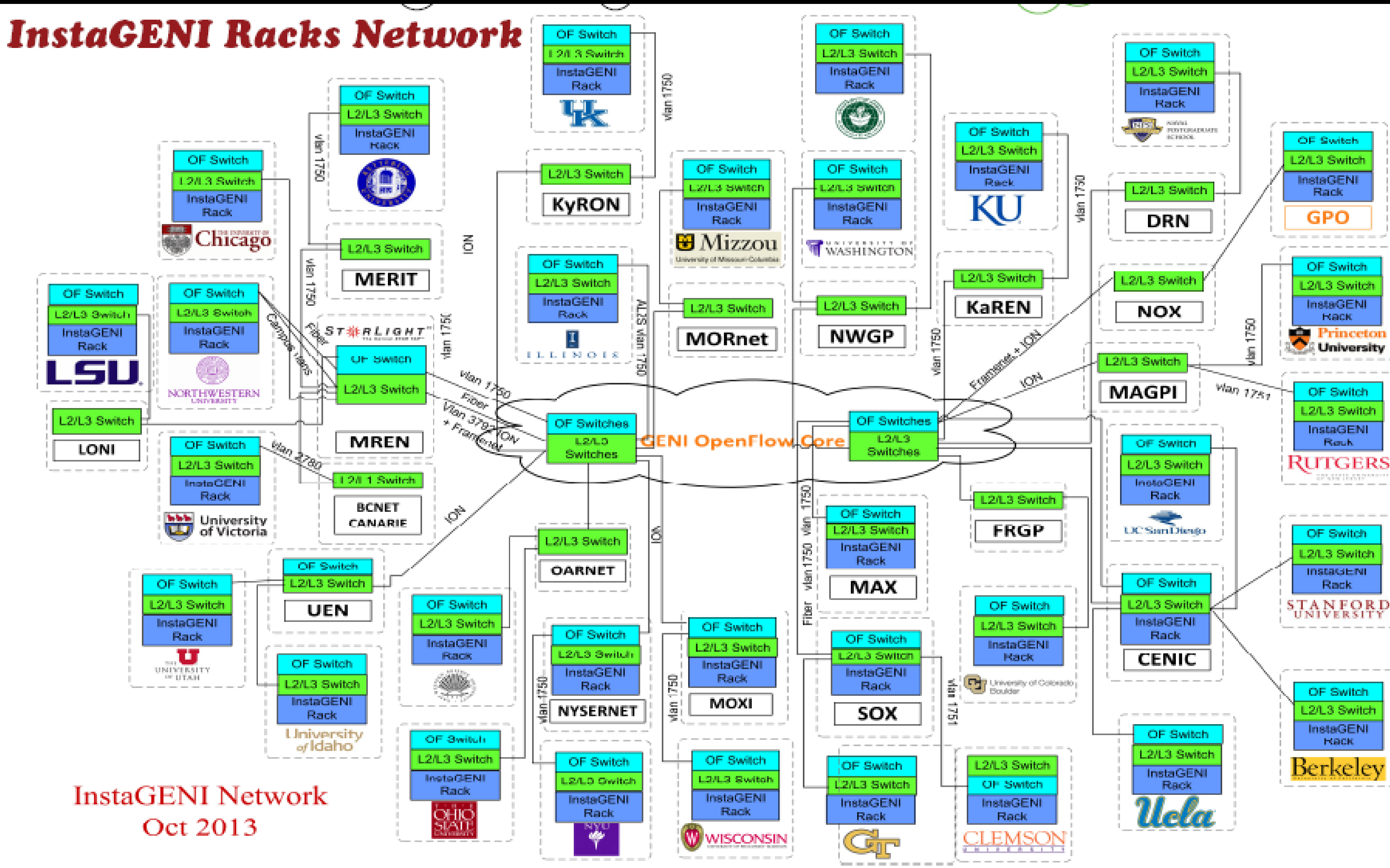
National Science Foundation's Global Environment for Network Innovations (GENI)

- 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

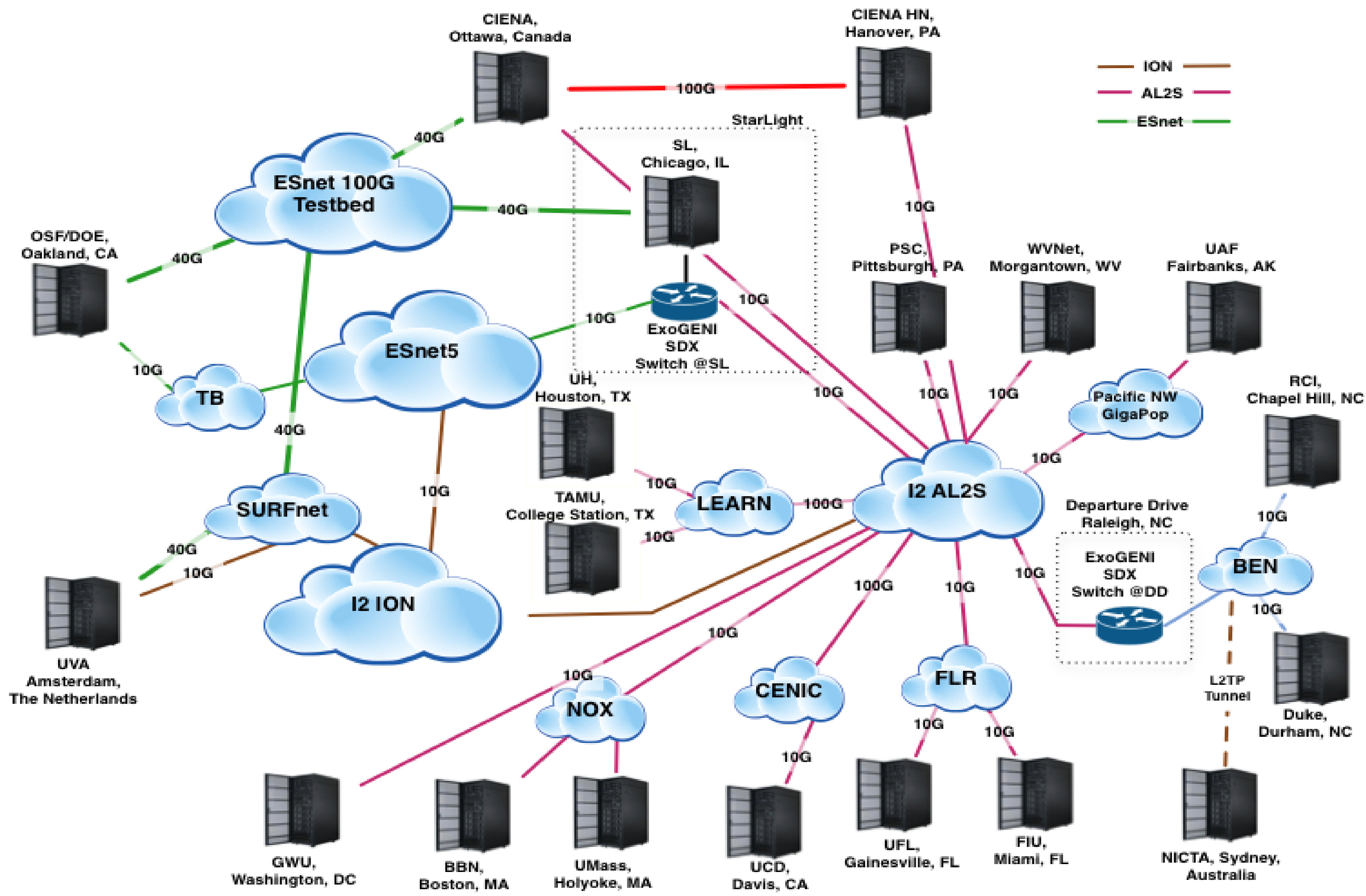
- Large Scale Highly Distributed Infrastructure That Can Support Multiple Empirical Research Testbeds At Scale
- Next Generation GENI, Edge Clouds, IOT, US Ignite, Platform for Advanced Wireless Research (PAWR) and Many Others
- Currently Being Planned – Will Be Designed, Implemented and Operated By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

National Science Foundation Global Environment for Network innovations



STARLIGHT SDX

International 40G and 100 G ExoGENI Testbed



Springer

springer.com

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.)

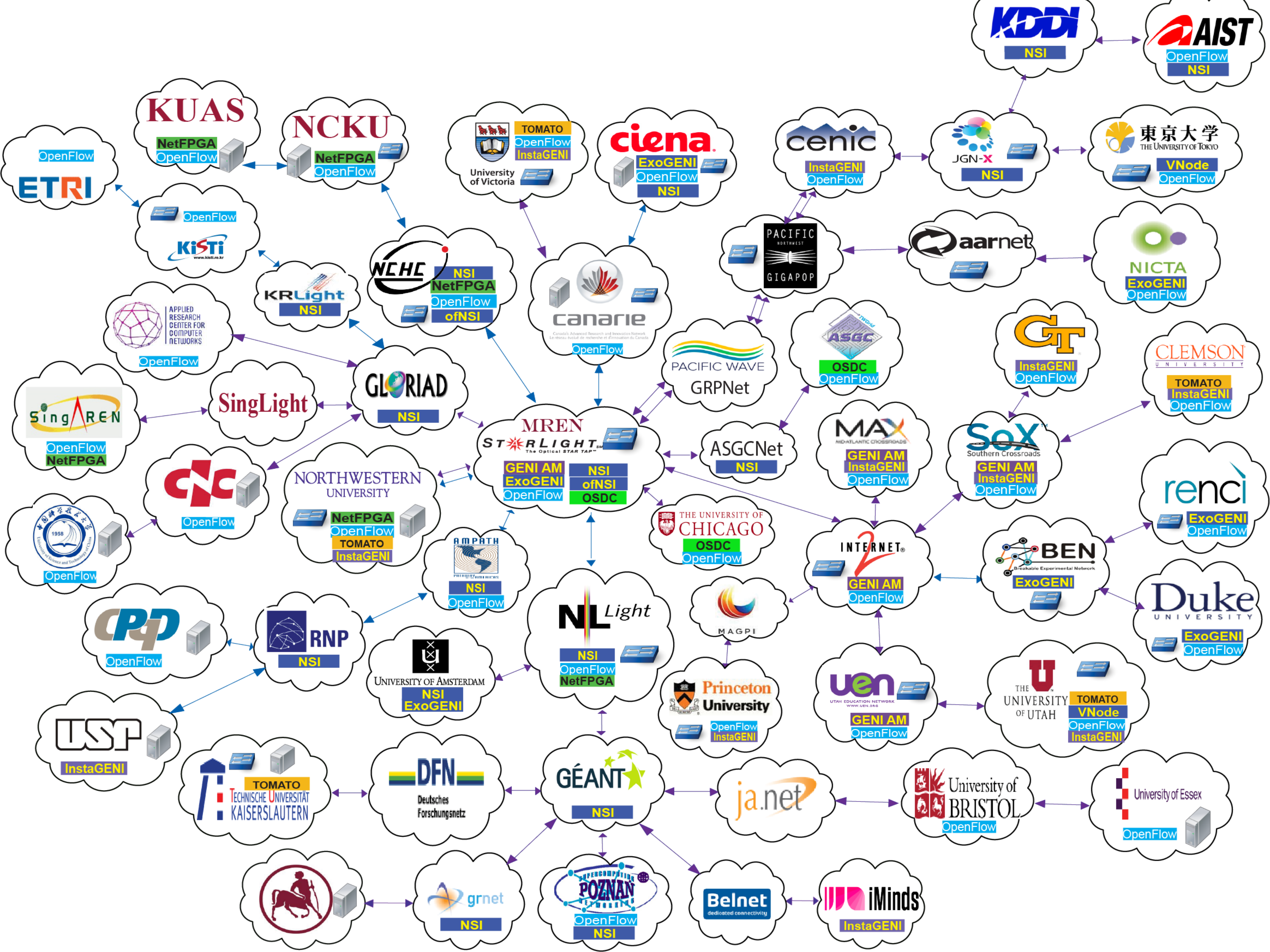
The GENI Book

- 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

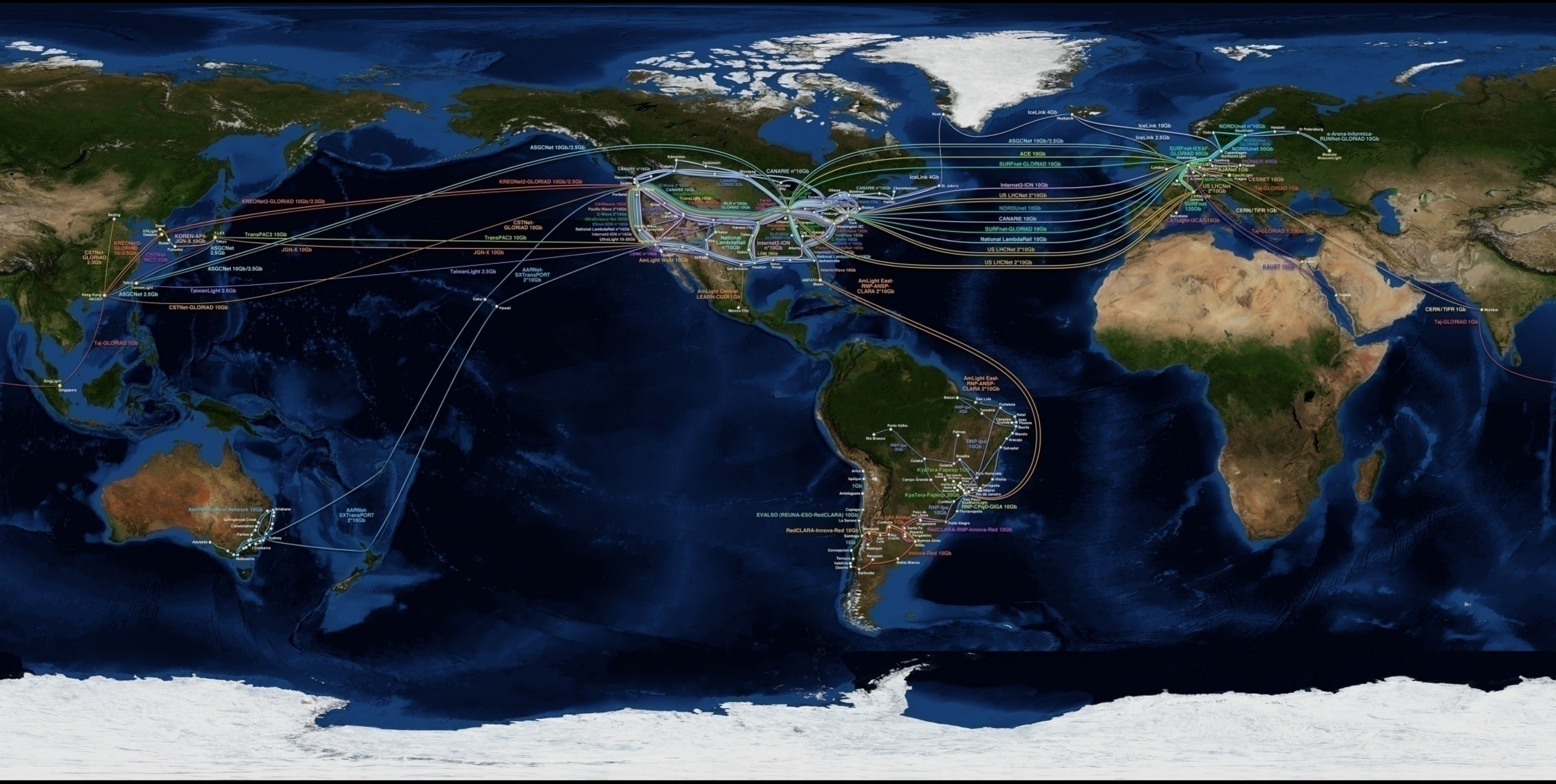
This book, edited by four of the leaders of the National Science Foundation's Global Environment and Network Innovations (GENI) project, gives the reader a tour of the history, architecture, future, and applications of GENI. Built over the past decade by hundreds of leading computer scientists and engineers, 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,
-


KDDI
NSI
AIST
OpenFlow
NSI
KUAS
NetFPGA
OpenFlow
NCKU
NetFPGA
OpenFlow
University of Victoria
TOMATO
OpenFlow
InstaGENI
ciena
ExoGENI
OpenFlow
NSI
cenic
InstaGENI
OpenFlow
JGN-X
NSI
東京大学
THE UNIVERSITY OF TOKYO
VNode
OpenFlow
OpenFlow
ETRI
OpenFlow
KISTI
KRLight
NSI
NCHC
NSI
NetFPGA
OpenFlow
ofNSI
canarie
OpenFlow
PACIFIC NORTHWEST GIGAPOP
aarnet
NICTA
ExoGENI
OpenFlow
APPLIED RESEARCH CENTER FOR COMPUTER NETWORKS
OpenFlow
PACIFIC WAVE
GRPNet
ASGC
OSDC
OpenFlow
GT
InstaGENI
OpenFlow
CLEMSON UNIVERSITY
TOMATO
InstaGENI
OpenFlow
SingAREN
OpenFlow
NetFPGA
SingLight
GLORIAD
NSI
MREN
STARLIGHT
The Optical STAR TAP
GENI AM
ExoGENI
OpenFlow
NSI
ofNSI
OSDC
ASGCNet
NSI
MAX
MID-ATLANTIC CROSSROADS
GENI AM
InstaGENI
OpenFlow
SoX
Southern Crossroads
GENI AM
InstaGENI
OpenFlow
renci
ExoGENI
OpenFlow
CNC
OpenFlow
NORTHWESTERN UNIVERSITY
NetFPGA
OpenFlow
TOMATO
InstaGENI
1958
OpenFlow
THE UNIVERSITY OF CHICAGO
OSDC
OpenFlow
INTERNET2
GENI AM
OpenFlow
BEN
Breakable Experimental Network
ExoGENI
AMPATH
NSI
OpenFlow
Duke UNIVERSITY
ExoGENI
OpenFlow
CPqD
OpenFlow
RNP
NSI
NetherLight
NSI
OpenFlow
NetFPGA
MAGPI
UNIVERSITY OF AMSTERDAM
NSI
ExoGENI
Princeton University
OpenFlow
InstaGENI
uen
UTAH EDUCATION NETWORK
GENI AM
OpenFlow
THE UNIVERSITY OF UTAH
TOMATO
VNode
OpenFlow
InstaGENI
USP
InstaGENI
TOMATO
TECHNISCHE UNIVERSITÄT KAISERSLAUTERN
DFN
Deutsches Forschungsnetz
GÉANT
NSI
ja.net
University of BRISTOL
OpenFlow
University of Essex
OpenFlow
grnet
NSI
POZNAN
OpenFlow
NSI
Belnet
dedicated connectivity
iMinds
InstaGENI

The Global Lambda Integrated Facility: a Global Programmable Resource

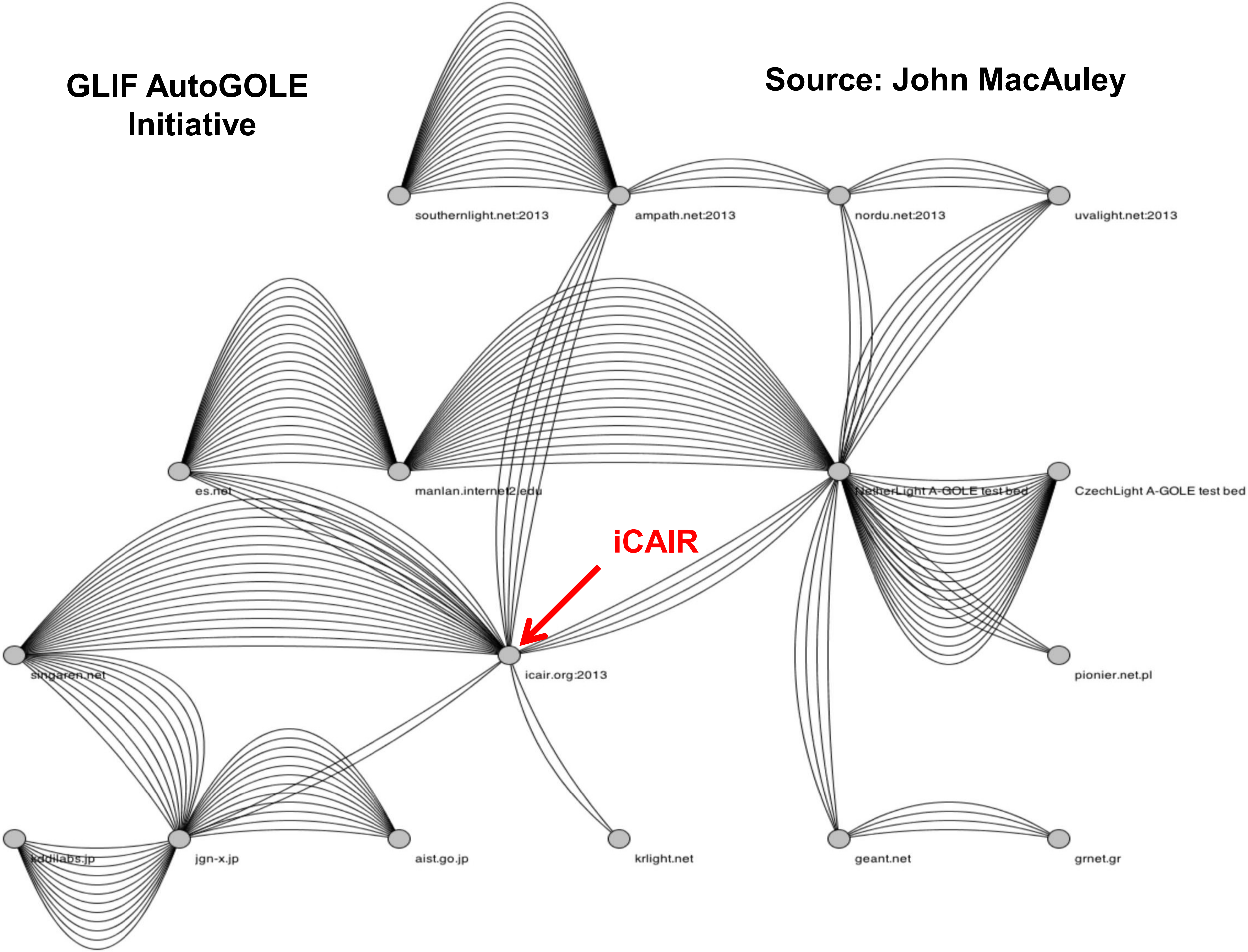


iCAIR

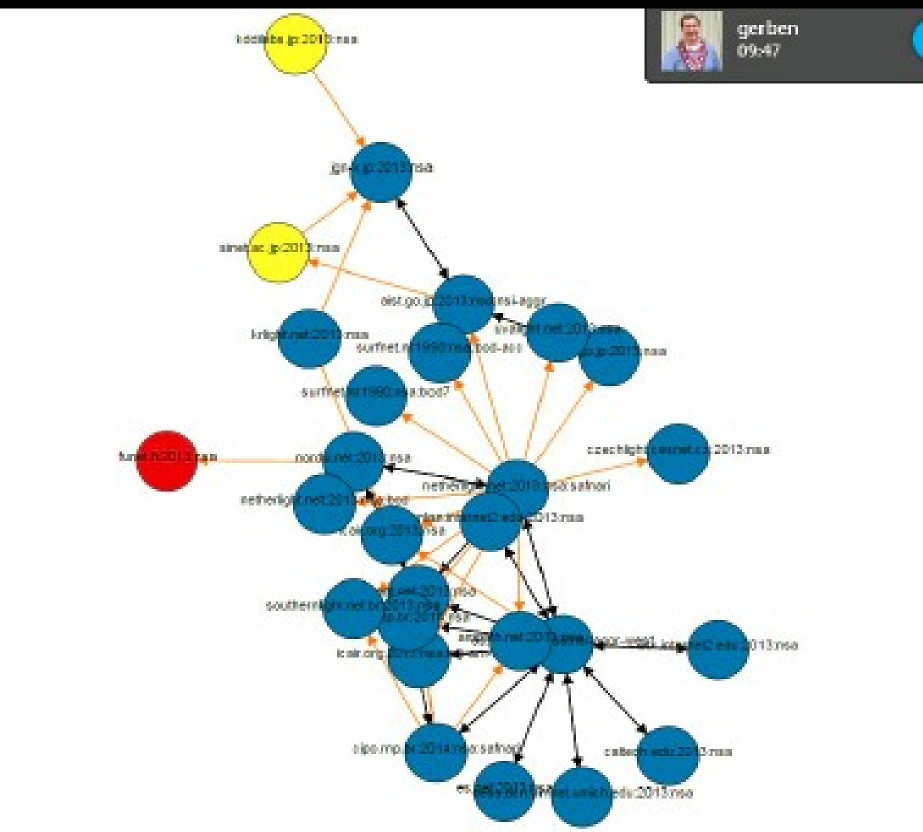
STARLIGHT SDX

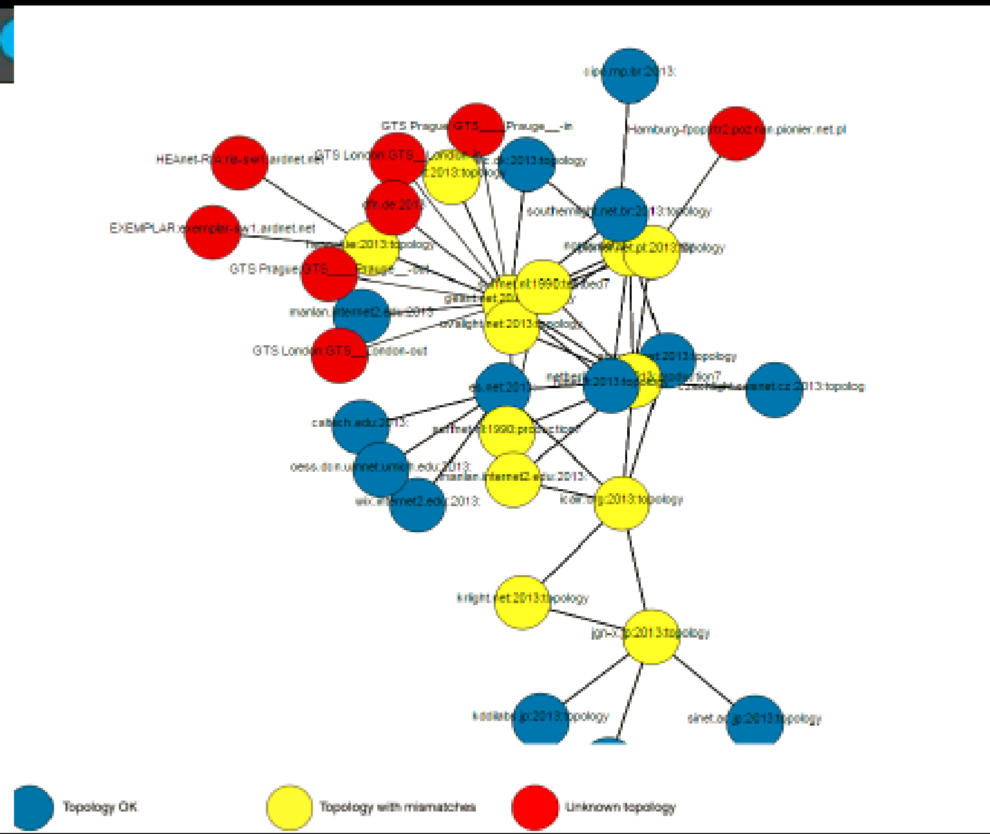
GLIF AutoGOLE Initiative

Source: John MacAuley

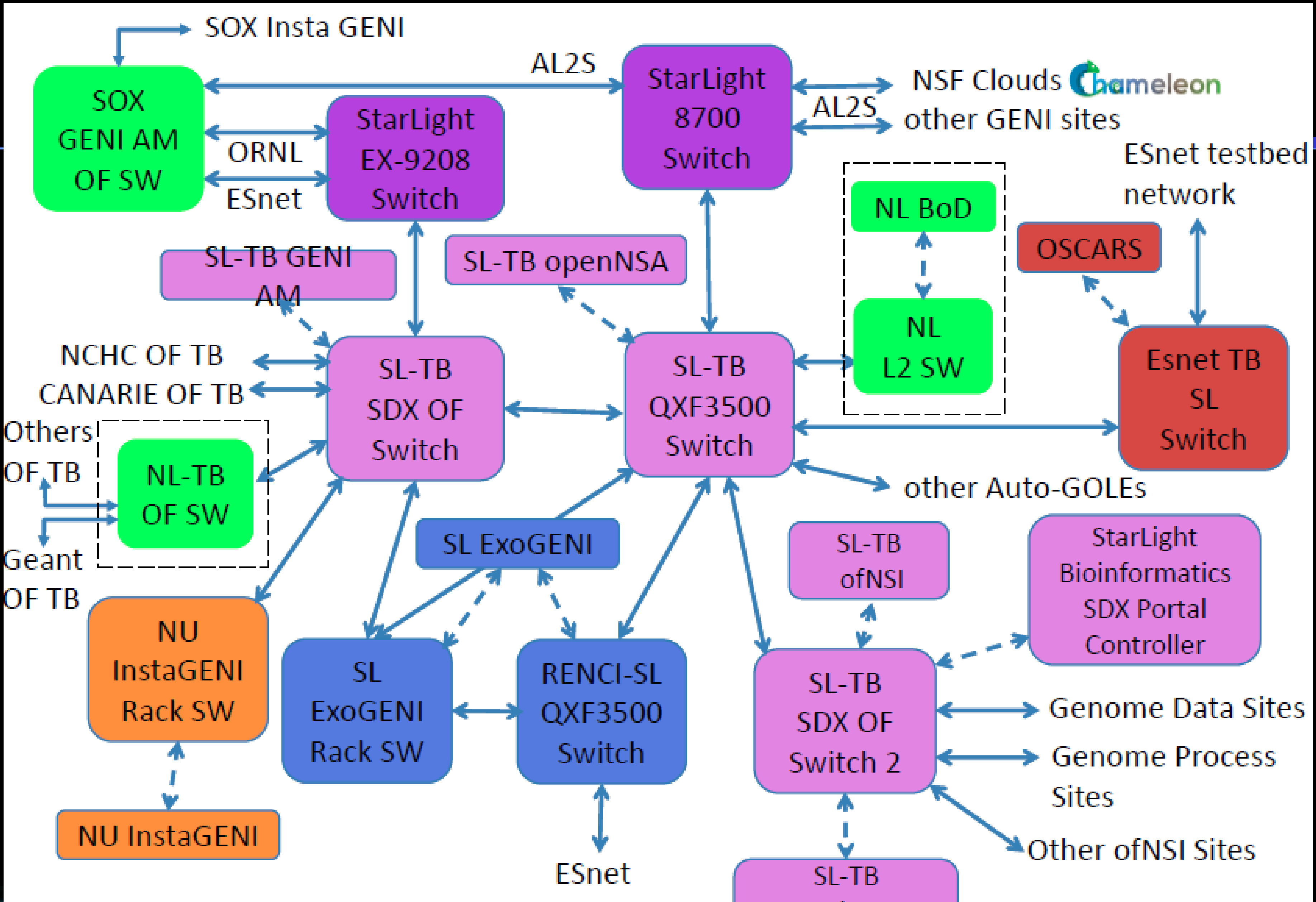


AutoGOLE Dashboard



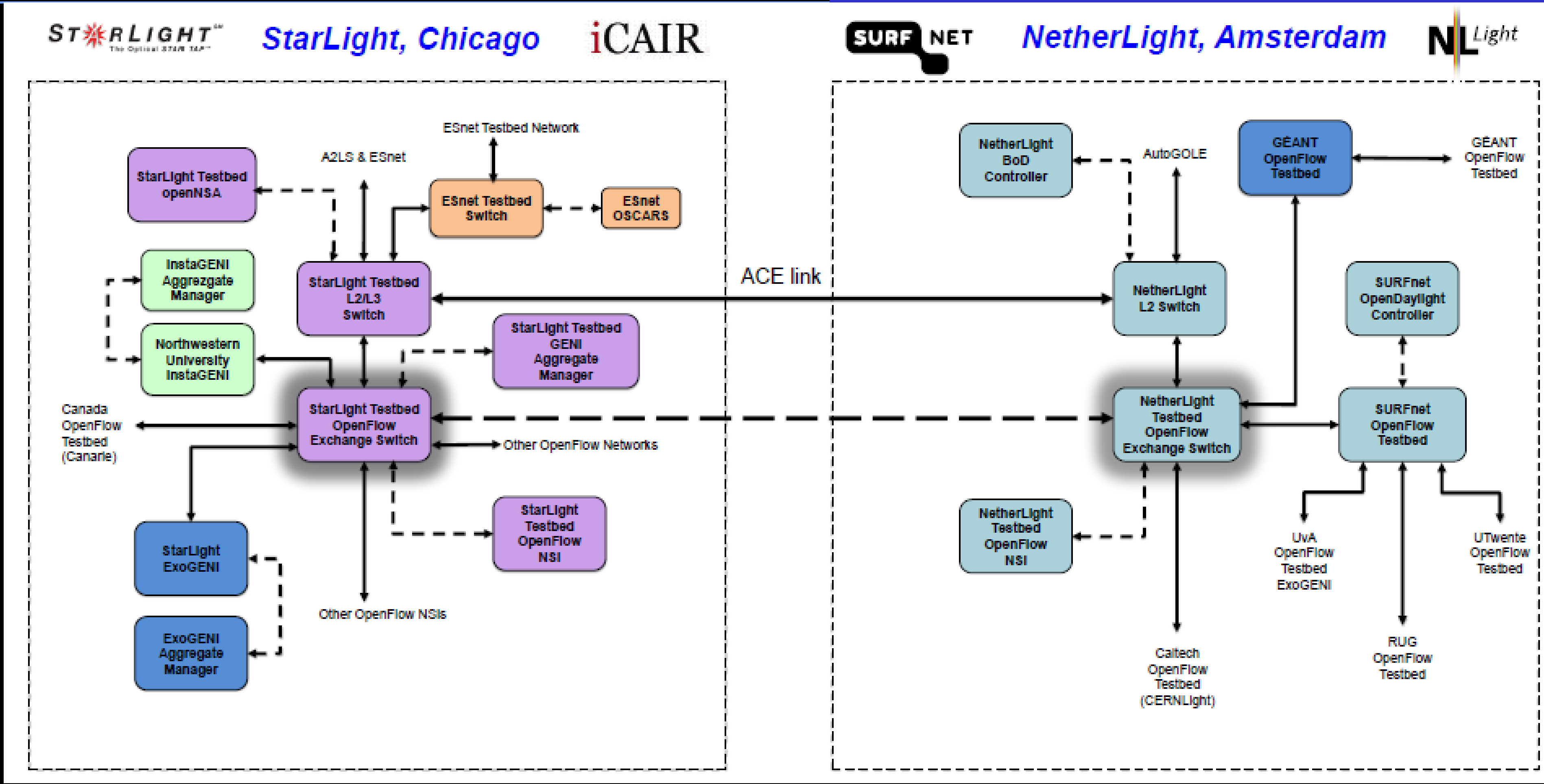


Control Plane **Data Plane**



SL-SDX-0.3 Prototype Logical Diagram 05/2015

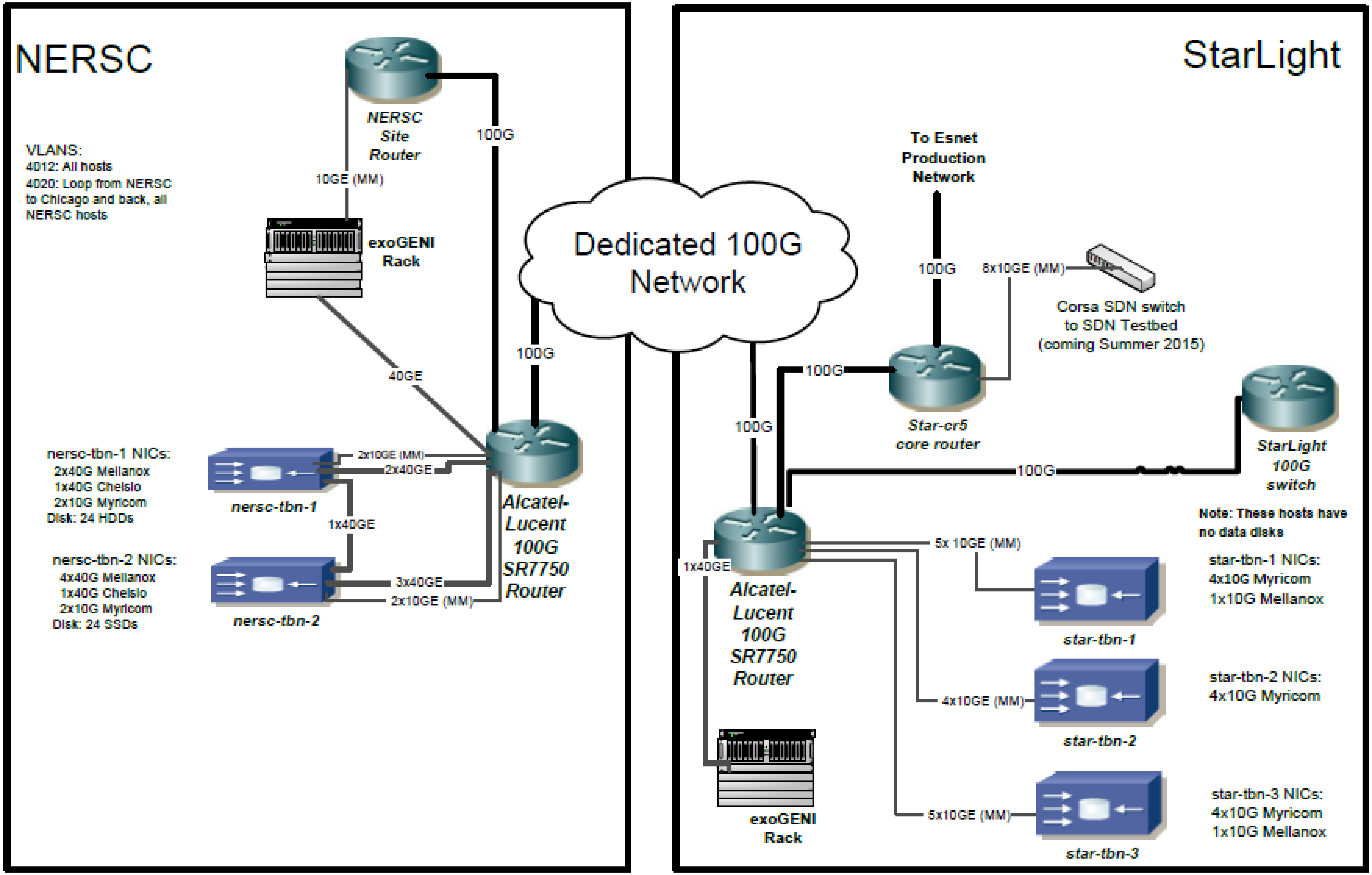
SDX StarLight⇔NetherLight



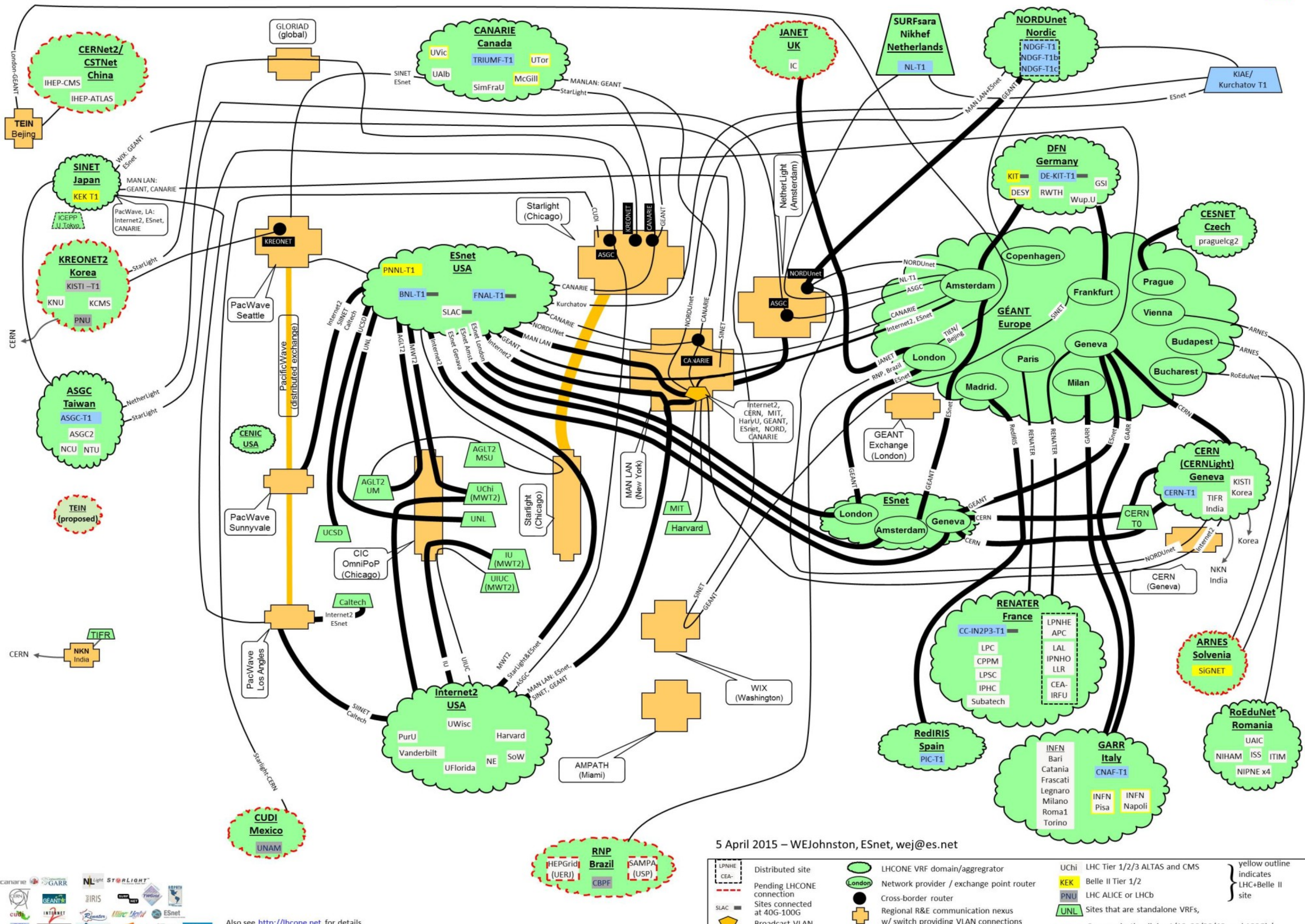
Ronald van der Pol, Joe Mambretti, Jim Chen, John Shillington

STARLIGHT℠

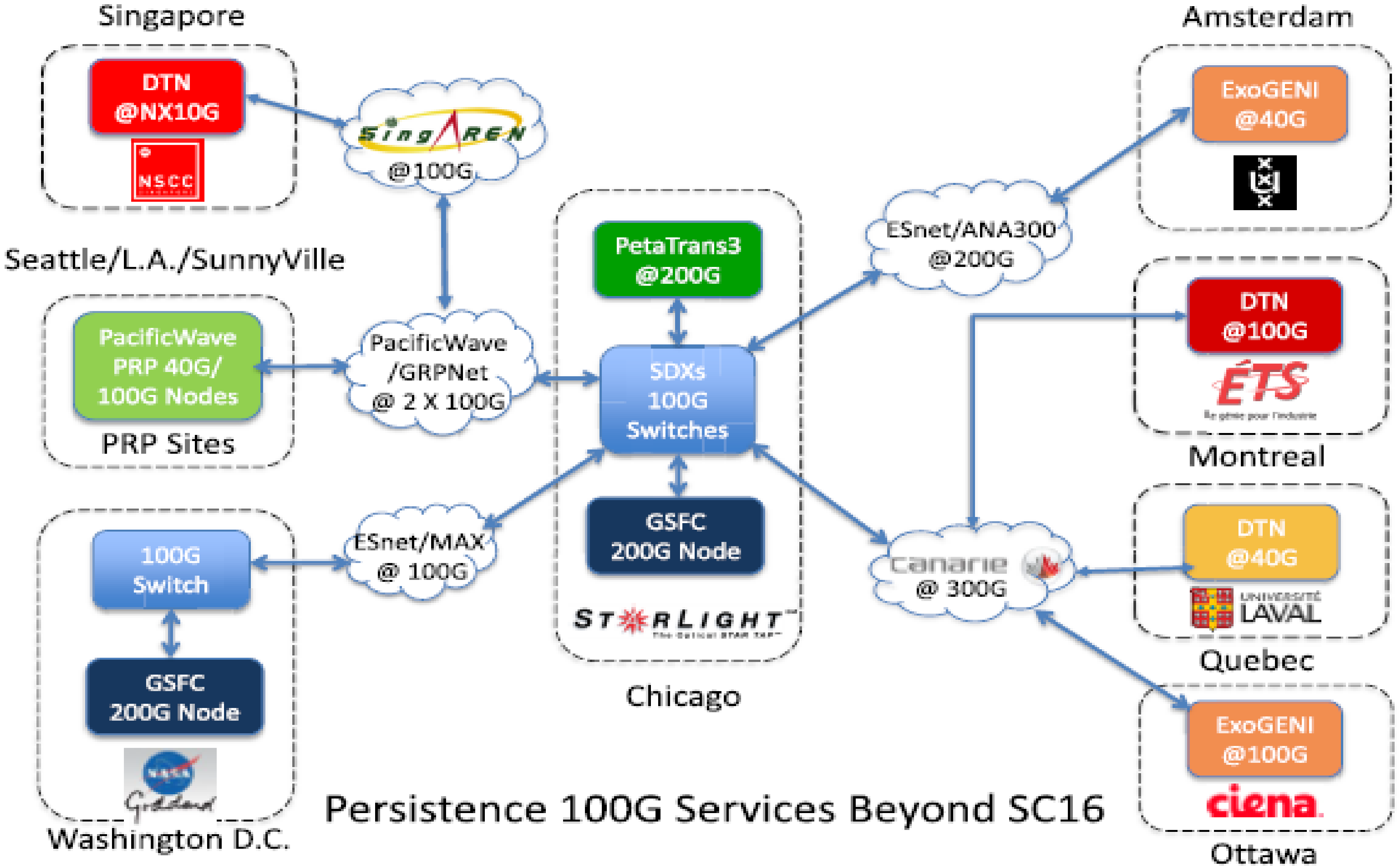
100G Component of Esnet SDN Testbed



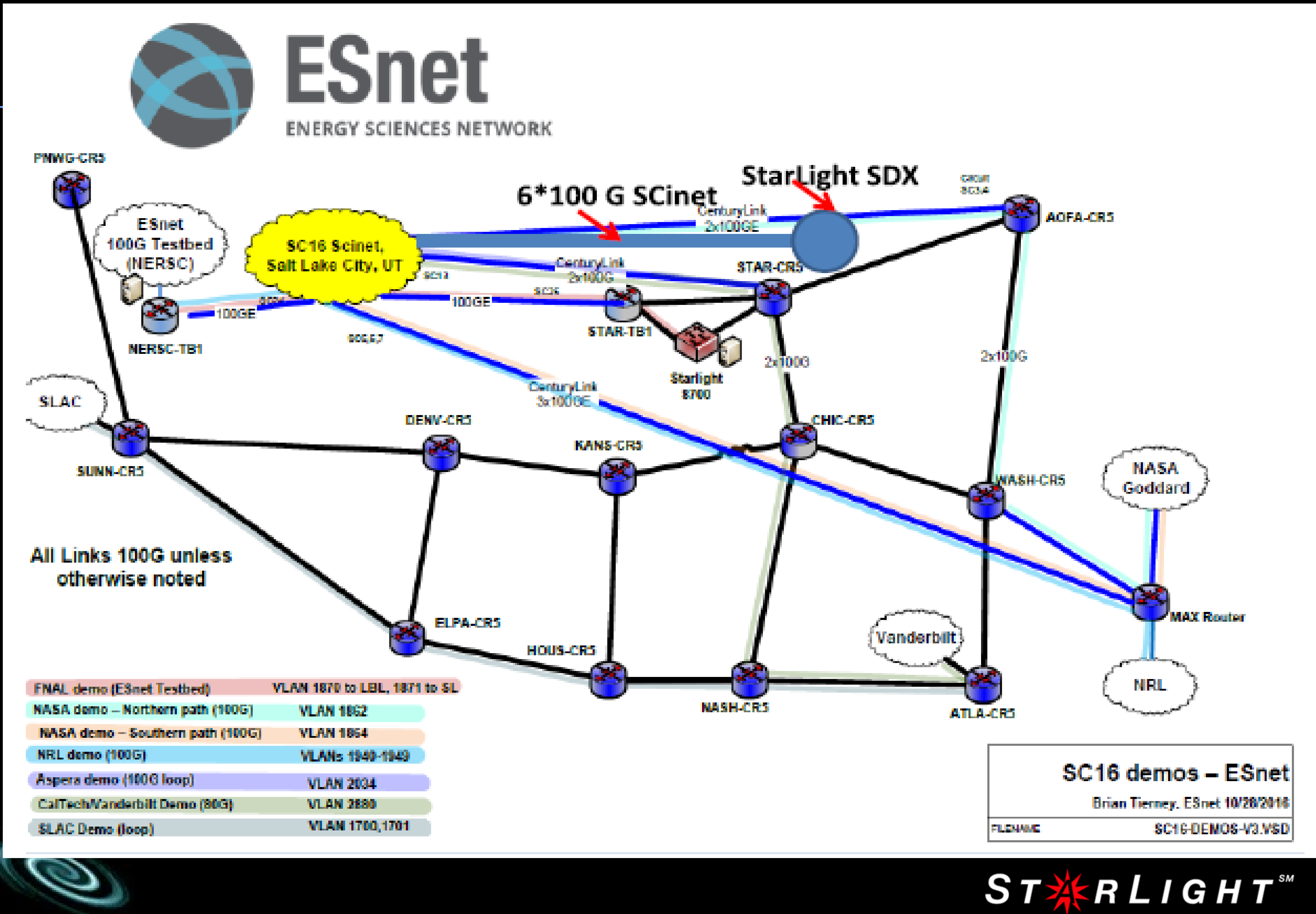
LHCONE: A global infrastructure for the High Energy Physics (LHC and Belle II) data management



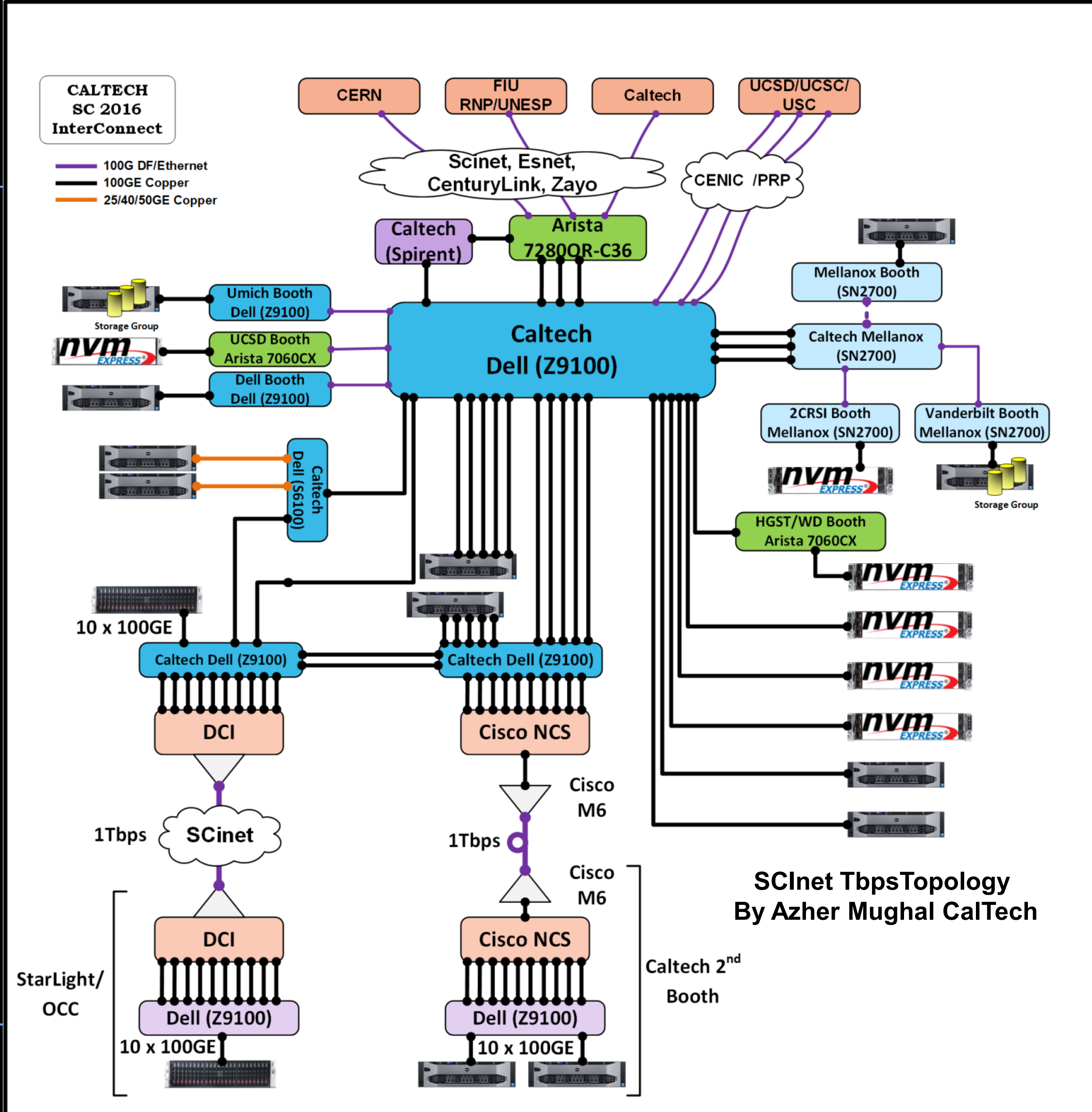
PetaTrans: Petascale Sciences Data Transfer



Persistence 100G Services Beyond SC16

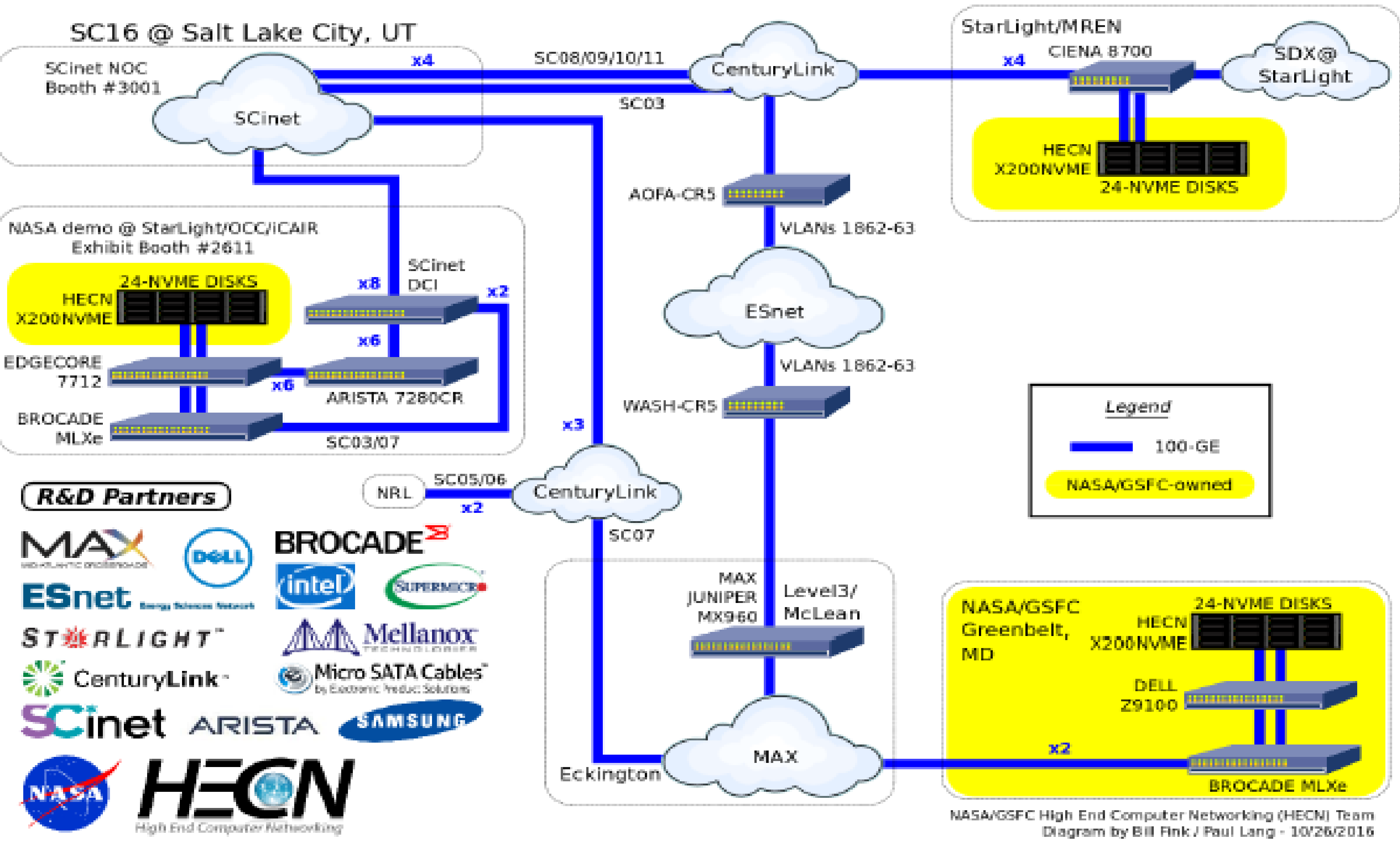


ESnet
ENERGY SCIENCES NETWORK
StarLight SDX
6*100 G SCinet
SC16 Scinet, Salt Lake City, UT
ESnet 100G Testbed (NERSC)
PNWG-CR5
AOFA-CR5
STAR-CR5
STAR-TB1
NERSC-TB1
Starlight 8700
CenturyLink 2x100GE
CenturyLink 2x100G
CenturyLink 3x100GE
100GE
2x100G
SLAC
DENV-CR5
KANS-CR5
CHIC-CR5
SUNN-CR5
WASH-CR5
NASA Goddard
MAX Router
ELPA-CR5
HOUS-CR5
NASH-CR5
ATLA-CR5
Vanderbilt
NRL
All Links 100G unless otherwise noted
FNAL demo (ESnet Testbed) VLAN 1870 to LBL, 1871 to SL
NASA demo – Northern path (100G) VLAN 1862
NASA demo – Southern path (100G) VLAN 1864
NRL demo (100G) VLANs 1940-1949
Aspera demo (100G loop) VLAN 2034
CalTech/Vanderbilt Demo (80G) VLAN 2880
SLAC Demo (loop) VLAN 1700,1701
SC16 demos – ESnet
Brian Tierney, ESnet 10/28/2016
SC16-DEMOS-V3.VSD
STARLIGHT


CALTECH
SC 2016
InterConnect
100G DF/Ethernet
100GE Copper
25/40/50GE Copper
CERN
FIU
RNP/UNESP
Caltech
UCSD/UCSC/
USC
Scinet, Esnet,
CenturyLink, Zayo
CENIC /PRP
Caltech
(Spirent)
Arista
7280QR-C36
Mellanox Booth
(SN2700)
Umich Booth
Dell (Z9100)
Storage Group
UCSD Booth
Arista 7060CX
Dell Booth
Dell (Z9100)
Caltech
Dell (Z9100)
Caltech Mellanox
(SN2700)
2CRSI Booth
Mellanox (SN2700)
Vanderbilt Booth
Mellanox (SN2700)
Storage Group
Caltech
Dell (S6100)
HGST/WD Booth
Arista 7060CX
10 x 100GE
Caltech Dell (Z9100)
Caltech Dell (Z9100)
DCI
Cisco NCS
Cisco
M6
1Tbps
SCinet
1Tbps
Cisco
M6
DCI
Cisco NCS
StarLight/
OCC
Dell (Z9100)
Dell (Z9100)
10 x 100GE
10 x 100GE
Caltech 2nd
Booth
SCInet TbpsTopology
By Azher Mughal CalTech
IGHT SM

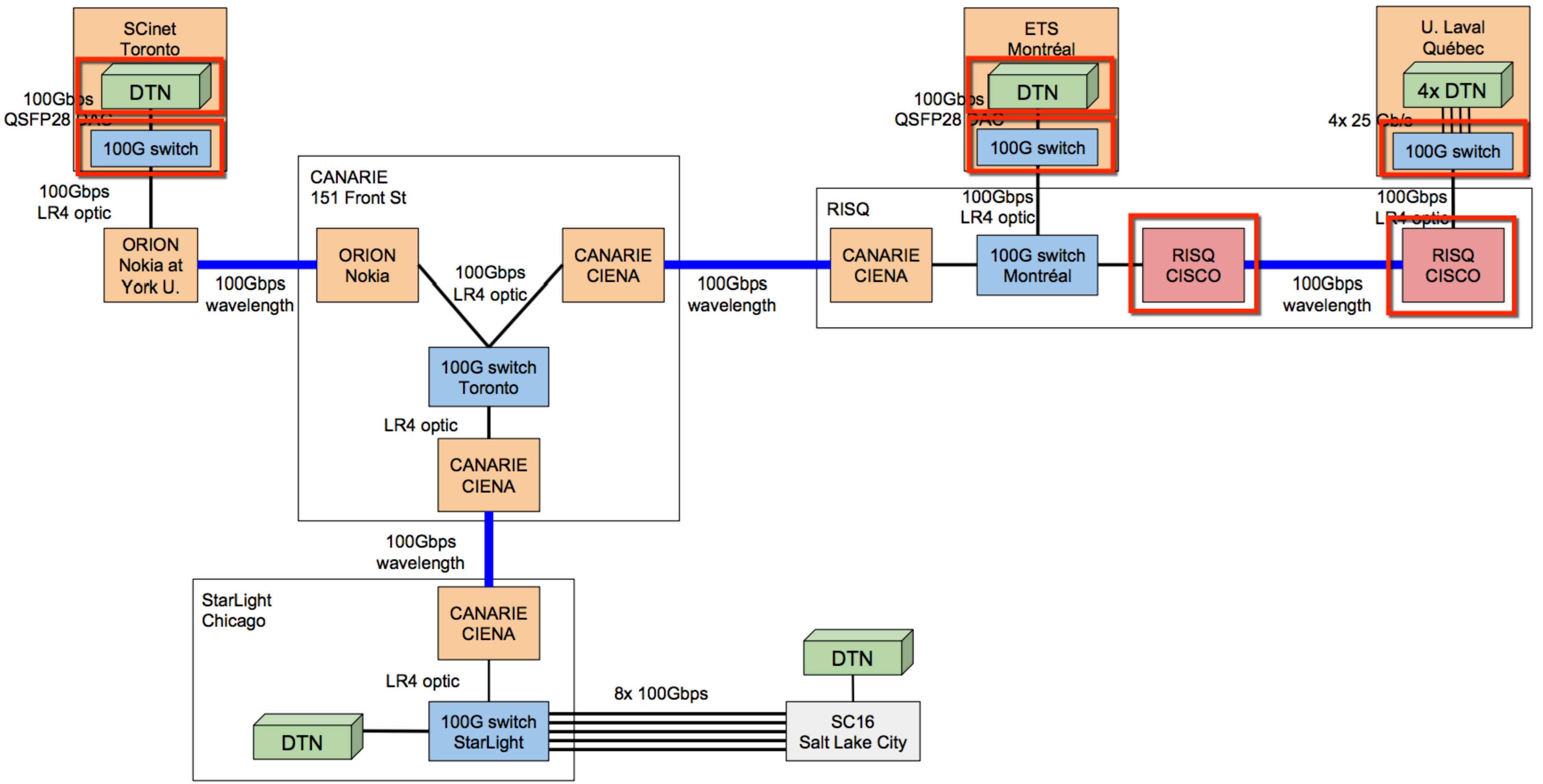
SC16 Demonstrations of 200 Gbps Disk-to-Disk WAN File Transfers using Parallelism across NVMe Drives

An SC16 Collaborative Initiative Among NASA and Several Partners

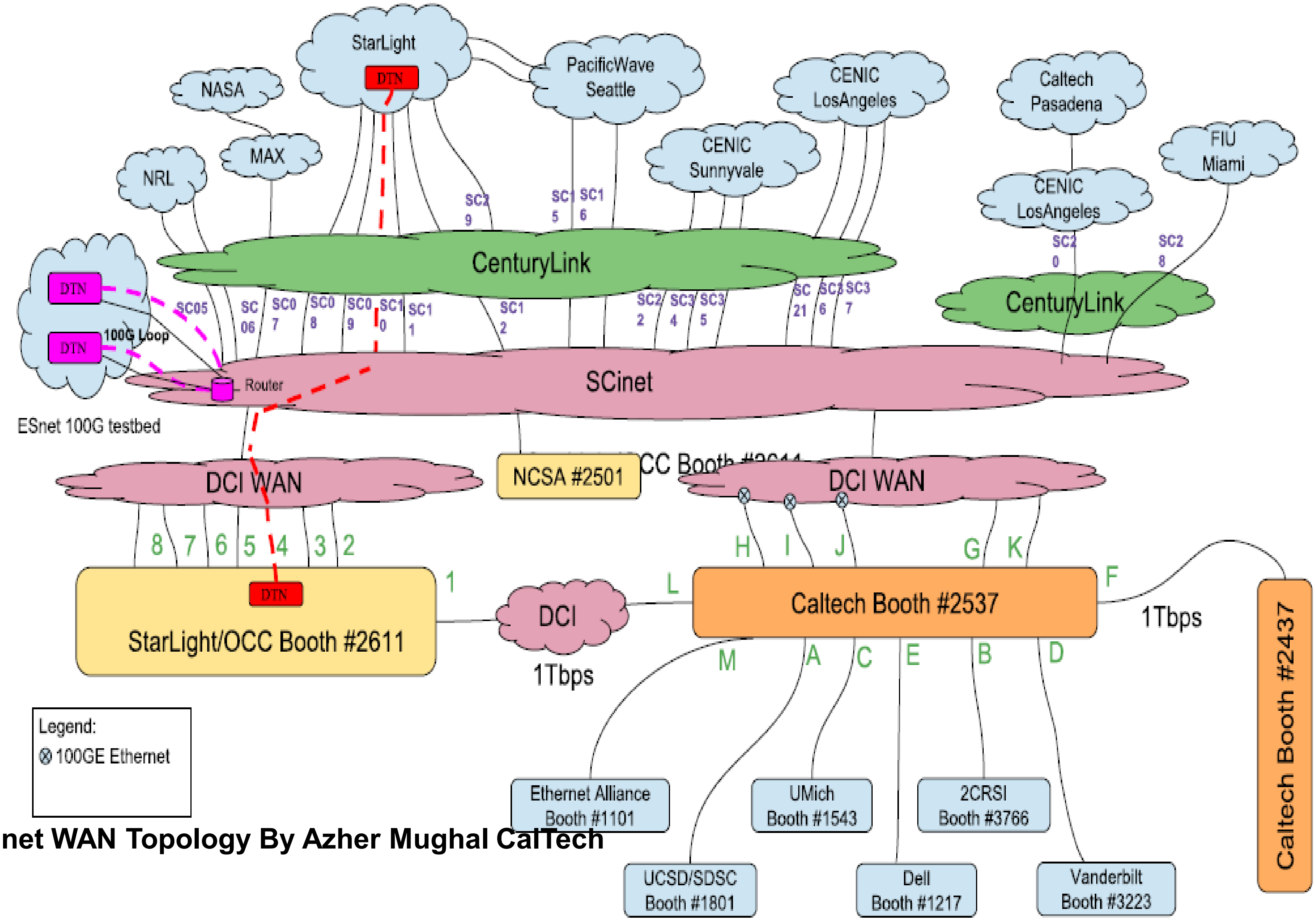


SC16 SDN-WAN Demonstration End-Points
Caltech, UM, Vanderbilt, UCSD, Dell, 2CRSI, KISTI, StarLight, PRP, FIU, RNP, UNESP, CERN

DTN Flows@100 Gbps=>Compute Canada⇔CANARIE⇔StarLight<+>SC16

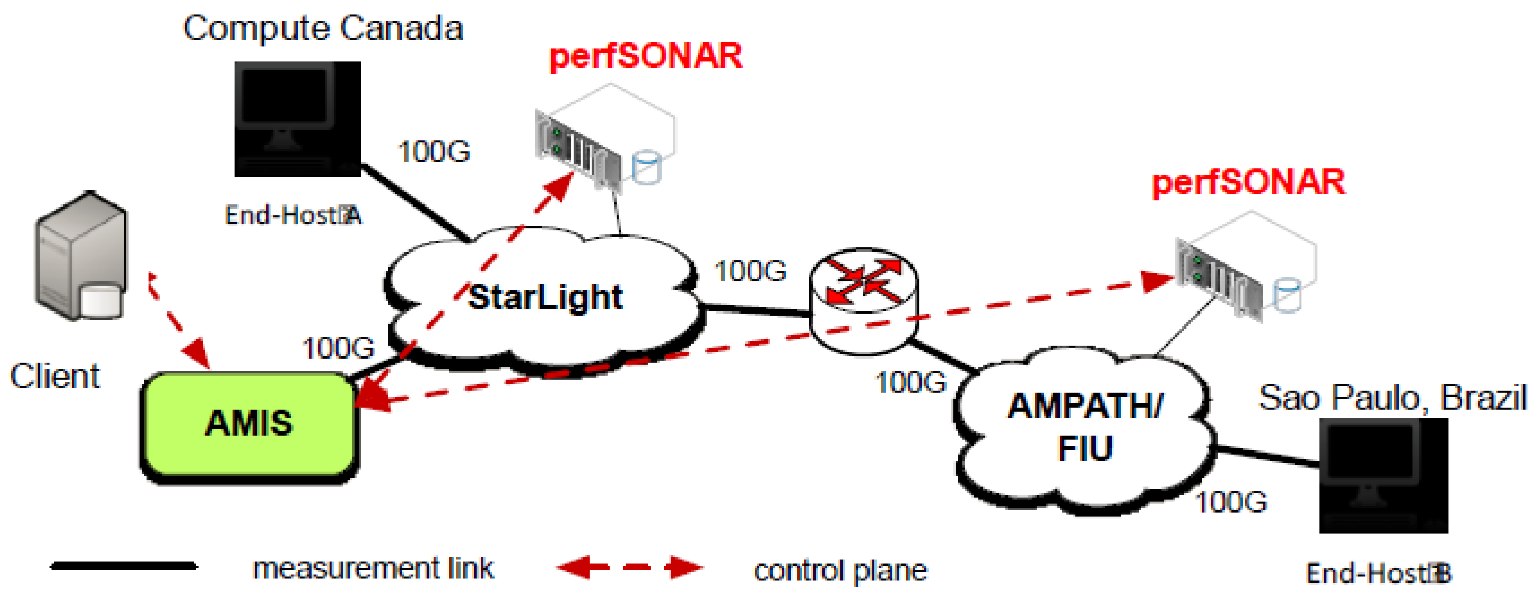


mdtmFTP @ 100Gbps Networks



SCInet WAN Topology By Azher Mughal CalTech

Programmable Network Measurement of Data Intensive Flows on 100Gbps Networks



Demo1: Programmable Measurement with RESTful APIs

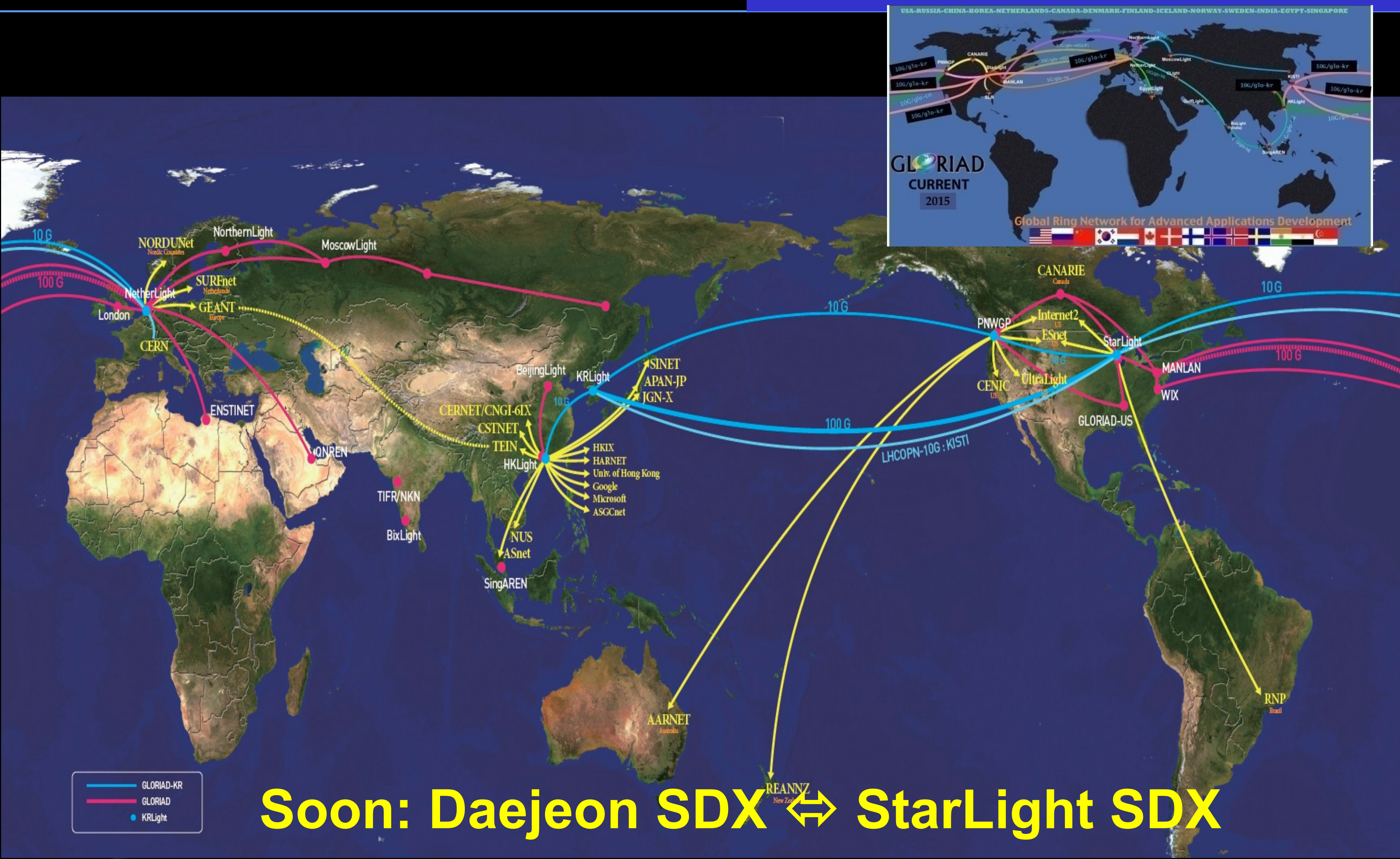
Demo2: Passive & Active Measurement (TCP window size)

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

KREONet2 and GLORIAD-KR
KISTI Daejeon ⇔ 100 G ⇔ StarLight



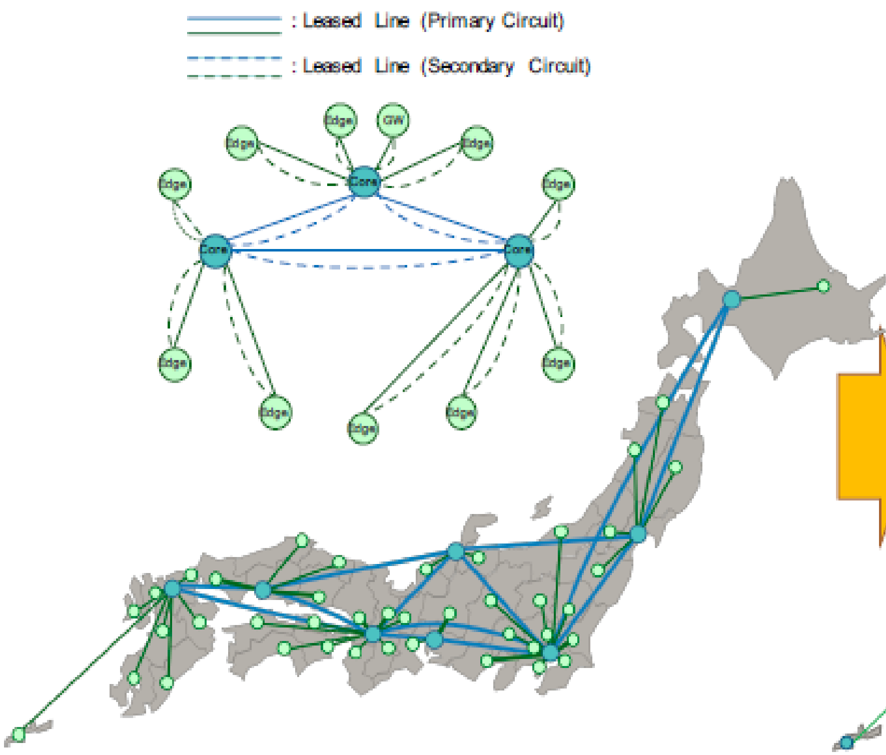
Soon: Daejeon SDX ⇔ StarLight SDX

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.

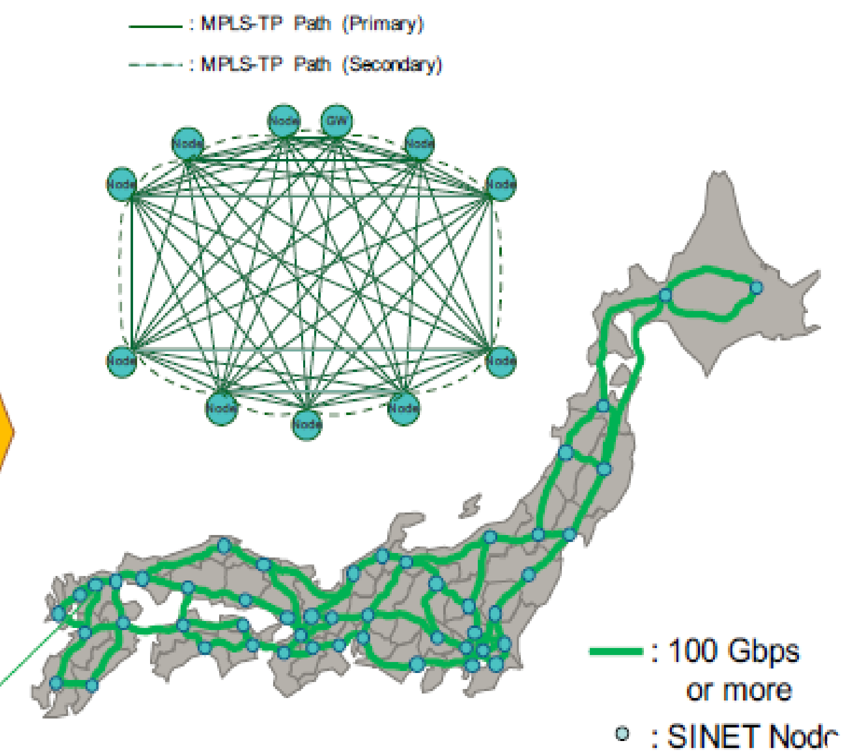
SINET4

- Star-like topology
- Resource-consuming secondary circuits



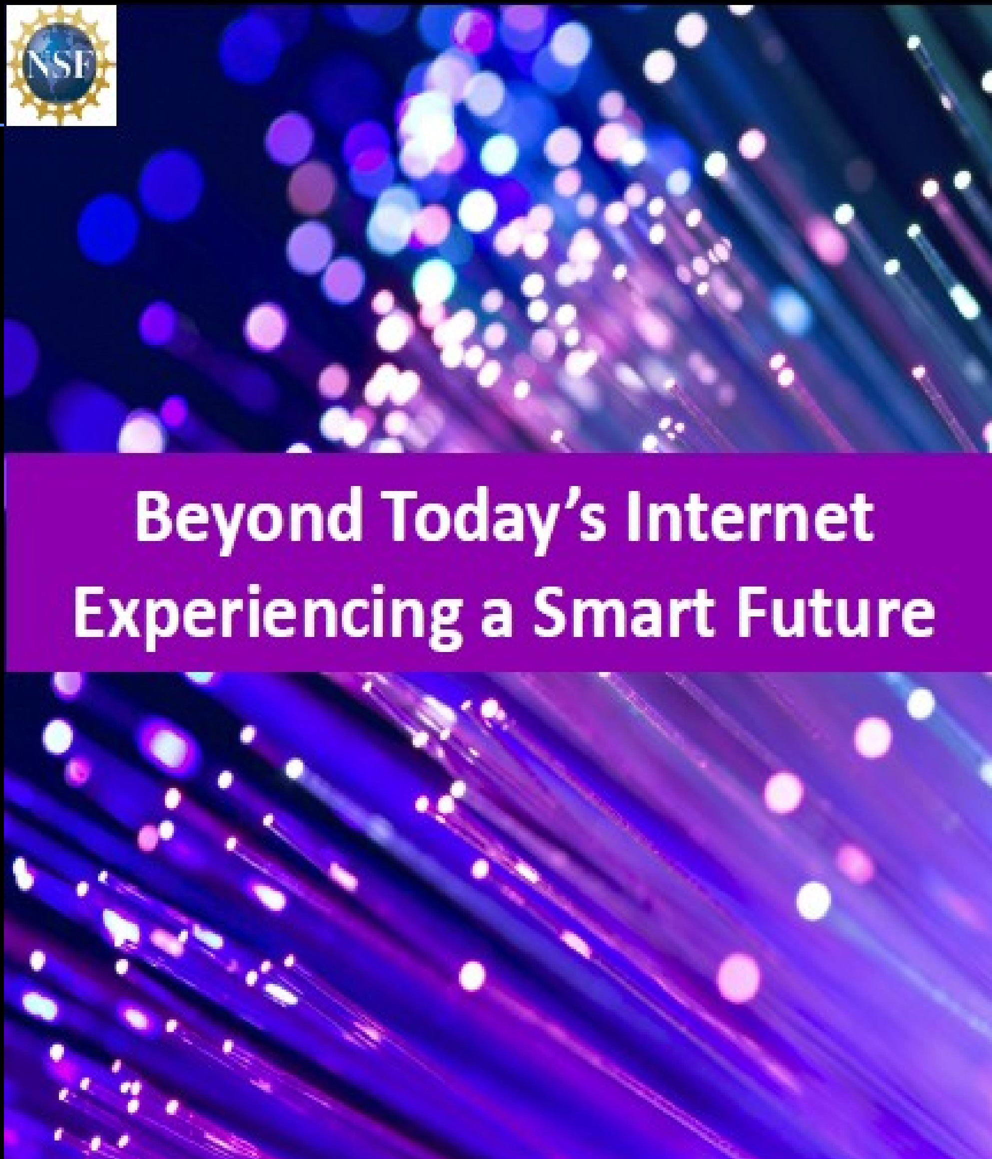
SINET5

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



A*STAR Singapore

- Singapore Supercomputing Center DTN⬄SingAREN⬄PacWavw⬄GRPnet⬄ StarLight DTN⬄SC16
- 50-60 Gbps

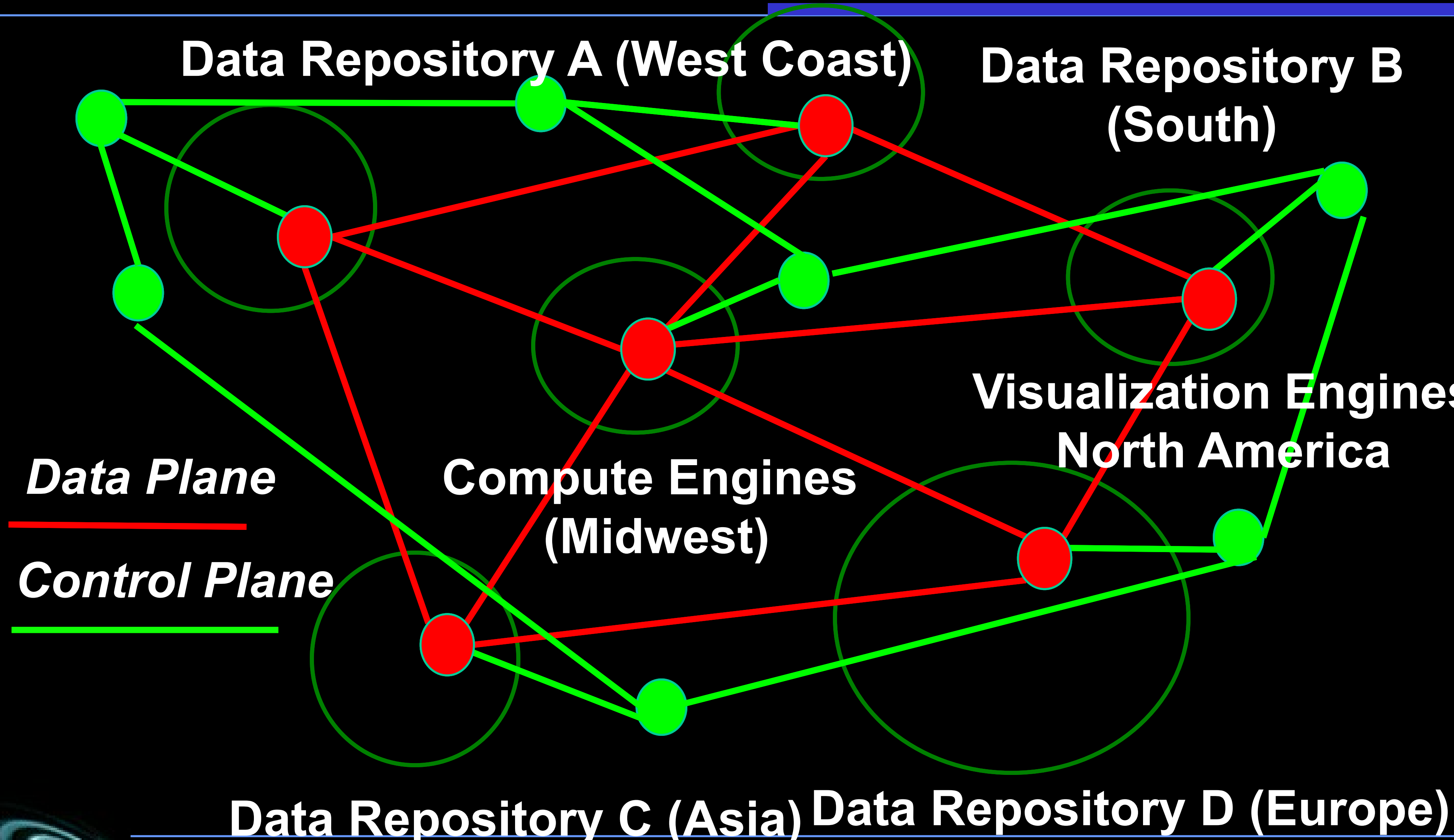


Prototype SDX Bioinformatics Exchange: Demonstrating an Essential Use-Case for Personalized Medicine

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

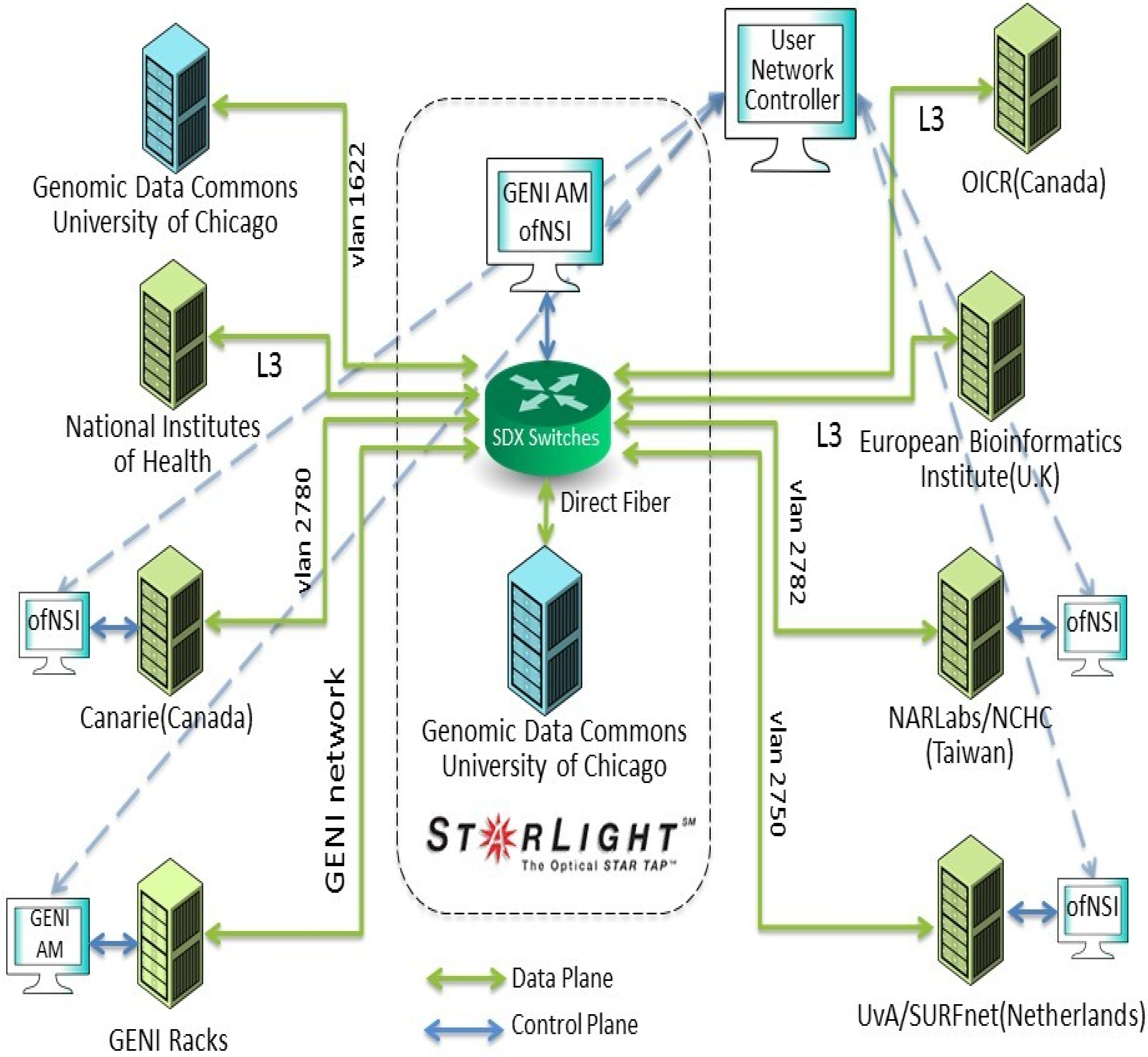
Joe Mambretti, Jim Chen
Northwestern University

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



STARLIGHT℠

GEC22 Bioinformatics SDXs Demo Network



An Experimental Testbed For Computer Science Research

www.chameleoncloud.org

CHAMELEON: A LARGE-SCALE, RECONFIGURABLE EXPERIMENTAL ENVIRONMENT FOR CLOUD RESEARCH

Principal Investigator: Kate Keahey

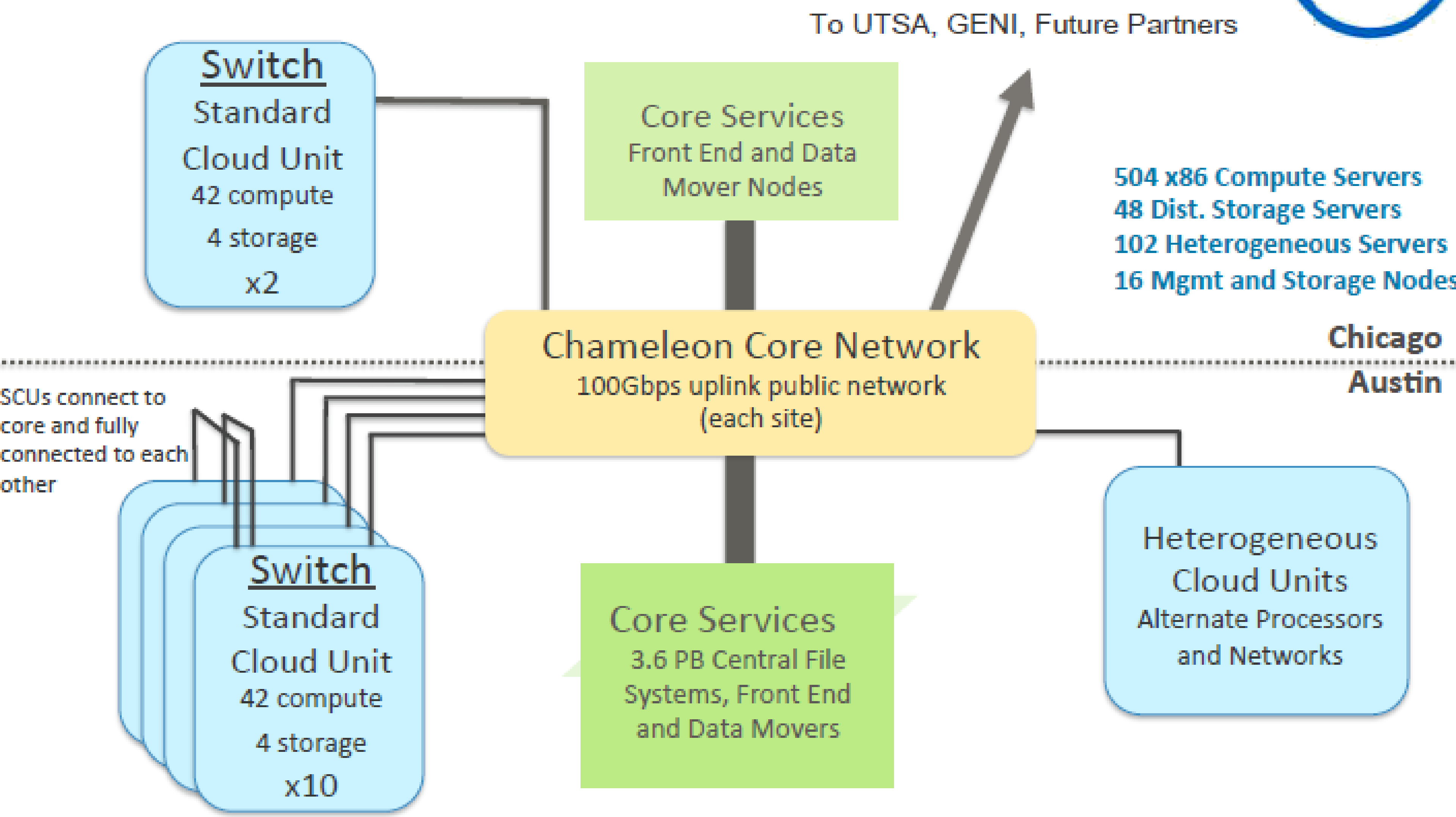
Co-PIs: J. Mambretti, D.K. Panda, P. Rad, W. Smith, D. Stanzione

Another SDX Opportunity and Model For "Tenant" Networking – Experimenters As Tenants

AUGUST 29, 2014

THE UNIVERSITY OF CHICAGO · TACC · NORTHWESTERN UNIVERSITY · The Ohio State University · UTSA · NSF

CHAMELEON HARDWARE



IMPLEMENTING THE EXPERIMENTAL WORKFLOW

discover resources → provision resources → configure and interact → monitor

discover resources

- Fine-grained
- Complete
- Up-to-date
- Versioned
- Verifiable

provision resources

- Advance reservations & on-demand
- Fine-grained allocations
- Isolation

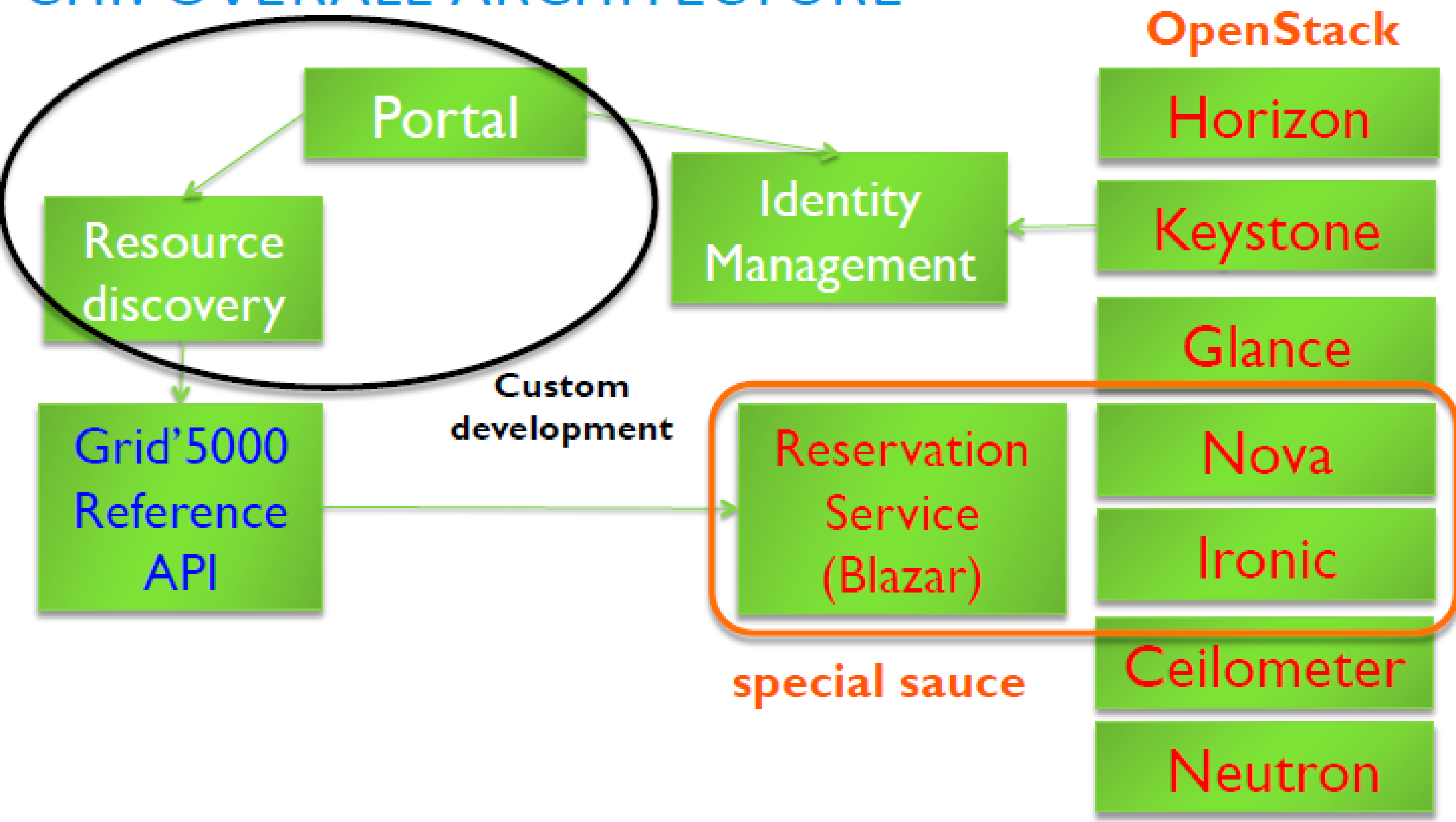
configure and interact

- Bare metal
- Deeply reconfigurable
- Map multiple appliances to a lease
- Snapshotting
- Complex Appliances

monitor

- Hardware metrics
- Fine-grained information
- Aggregate and archive

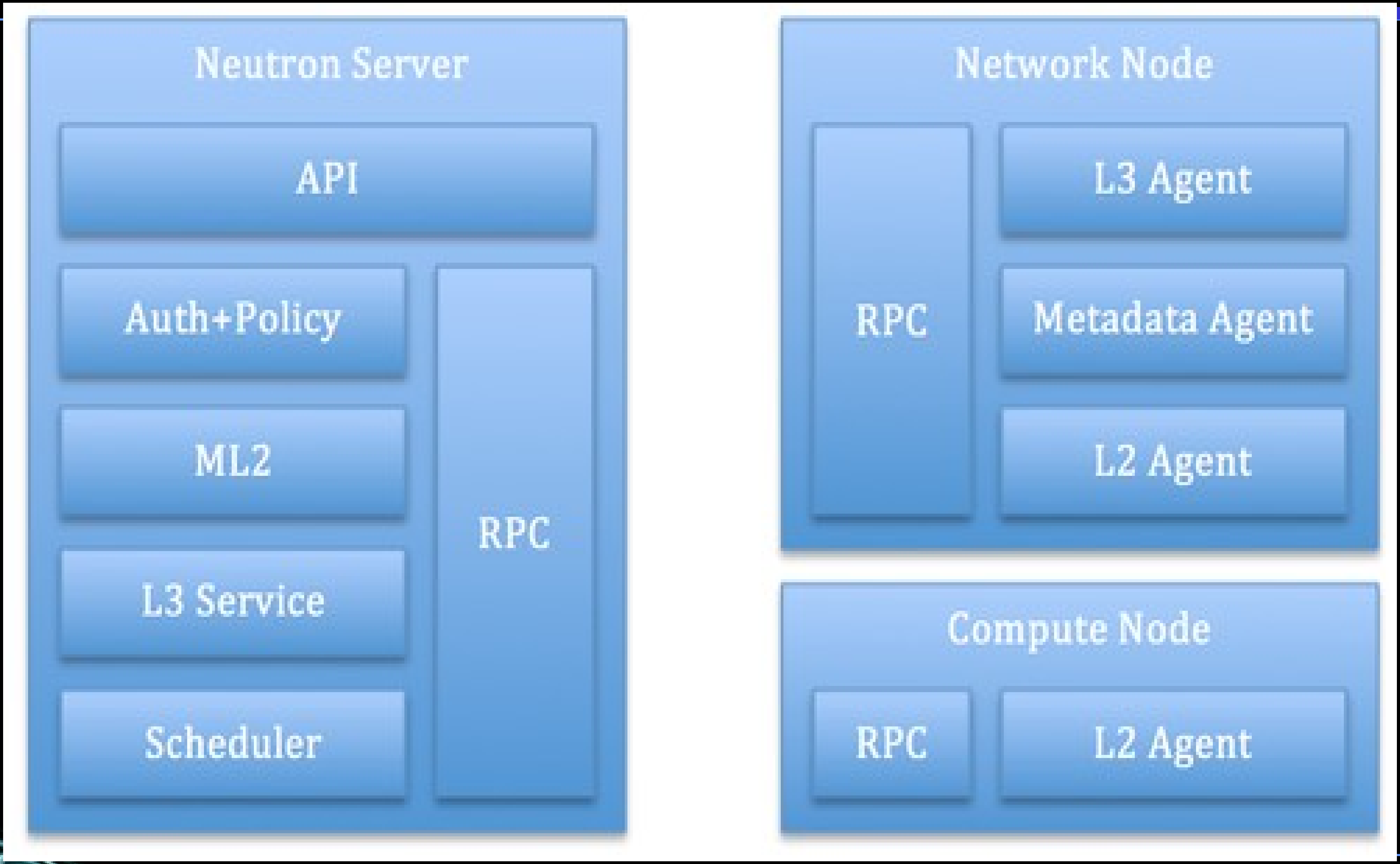
CHI: OVERALL ARCHITECTURE



Chameleon and OpenStack Net API

- The Chameleon Network Control API Is the OpenStack Network API -- Neutron, A Cloud Networking Controller And a Networking-as-a-Service (NaaS) Platform.
- The Implementation Includes Chameleon Enhancements, e.g., Pre-Set Network Configurations that Experimenters Can Select, Implement, and Toggle Among.
- Options The Control Plane Can Support Include Standard Routing Services, Pre-configured VLANs, Extensible VXLANs, Generic Routing Encapsulation (GRE), and OpenFlow with Hybrid Networking.
- Because of the Address Limitations of VLANs (~4k), the VXLAN Encapsulation Protocol Architecture Is Important To Multi-Tenant Cloud Providers, Especially For SDN, Because It Extends the Address Space To 16 million With a 24-bit Segment ID, Enabling Provisioning of Large Numbers of Overlay Networks On Shared Infrastructure.

Open Stack SW Architecture For Neutron Reference Platform



Neutron

- Neutron Provides APIs (Via Dashboard Web-Based GUI), interfaces to Devices (e.g., Routers, Switches, virtual routers, virtual switches, and SDN Controllers), And Policy Based Control Software Components.
- Neutron Enables Networks To Be Created and Managed Within IaaS Platforms, e.g., L2 Paths, Routed L3 Paths, IP address Management Processes, and Gateways Through Which It is Possible To Interconnect With External Networks.
- Minimal Neutron Implementation includes At Least One Controller Node (w/ At Least 1 Network Interface On The Management Plane), One Network Node, and One Compute Node.

Options For Bare Metal Integration With Tenant Networks

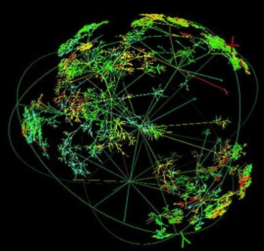
- For Some Services and Applications, VMs Alone Do Not Provide All Capabilities Required, Especially Services Requiring Exceptionally High Performance.
- OpenStack Supports Bare Metal Implementations Through the Ironic Environment, Which Can Be Used To Establish, Configure, and Use Bare Metal Nodes.
- This Option Requires a Network That Connects to a Boot Image Server That Can Support a PXE Boot Function Over The Network and Other Services, From Neutron, to Configure Bare Metal Nodes.
- The Ironic Process Creates a Flat Network, a Single VLAN, Among Bare Metal Nodes.
- Chameleon Is Building On This By Developing Options For Network Isolation For Bare Metal Nodes.

Extending Tenant Networks Through Federation

- Tenant Networks Are Not Only Implemented In Cloud Facilities, And Among Cloud Facilities
- They Must Be Extended To Multiple External Sites – Across Multiple Domains – Across The US and Internationally
- Creating Global Private Networks Enhanced Through Federation
- E.g., GENI Has Been Federated With the Future Internet Research and Experimentation (FIRE), With the Smart Applications On Virtual Environments (SAVI) Testbed, with NSFCloud Testbeds, and the International GENI.

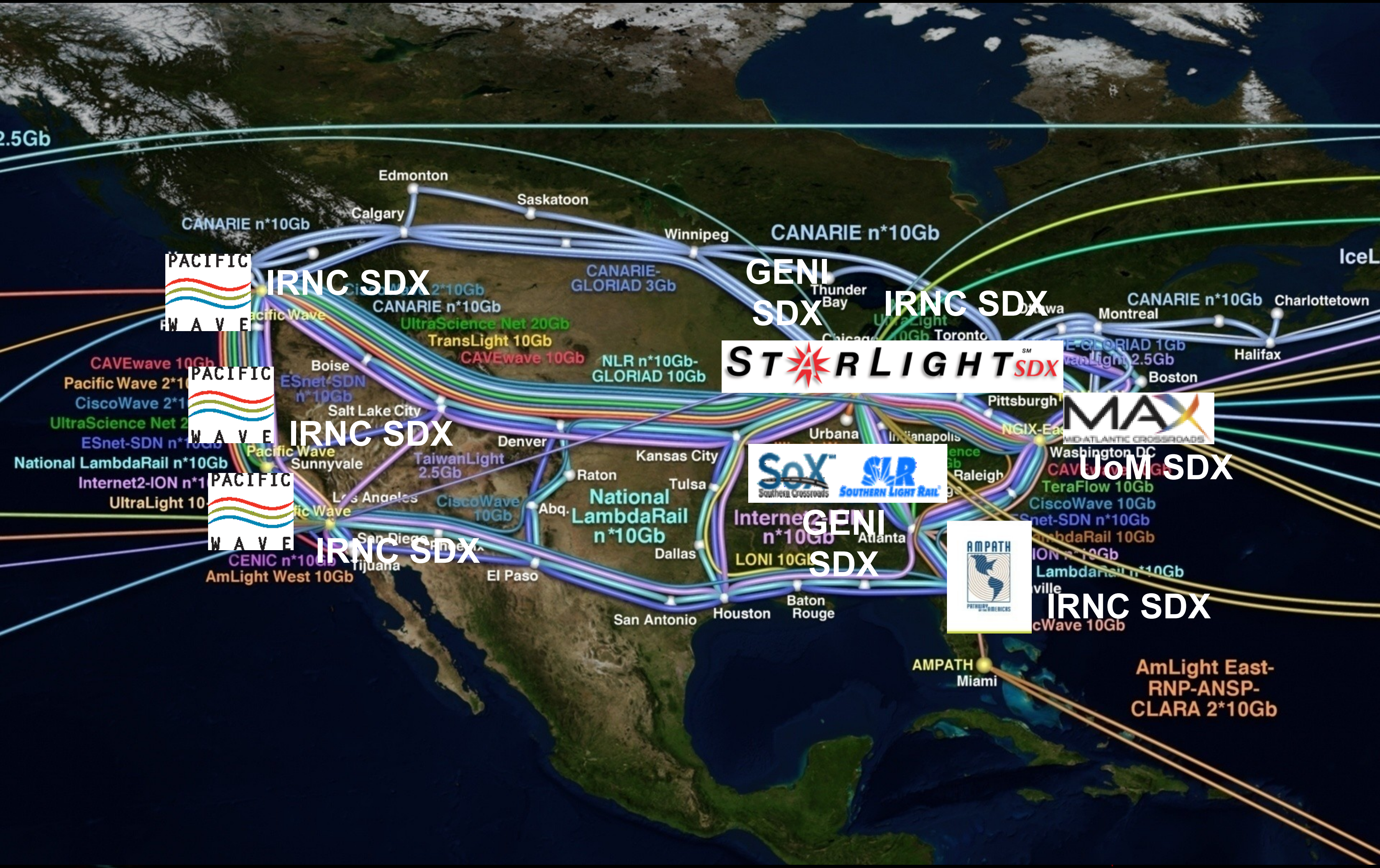
IRNC: RXP: StarLight SDX A Software Defined Networking Exchange for Global Science Research and Education

Joe Mambretti, Director, (j-mambretti@northwestern.edu)
International Center for Advanced Internet Research (www.icair.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

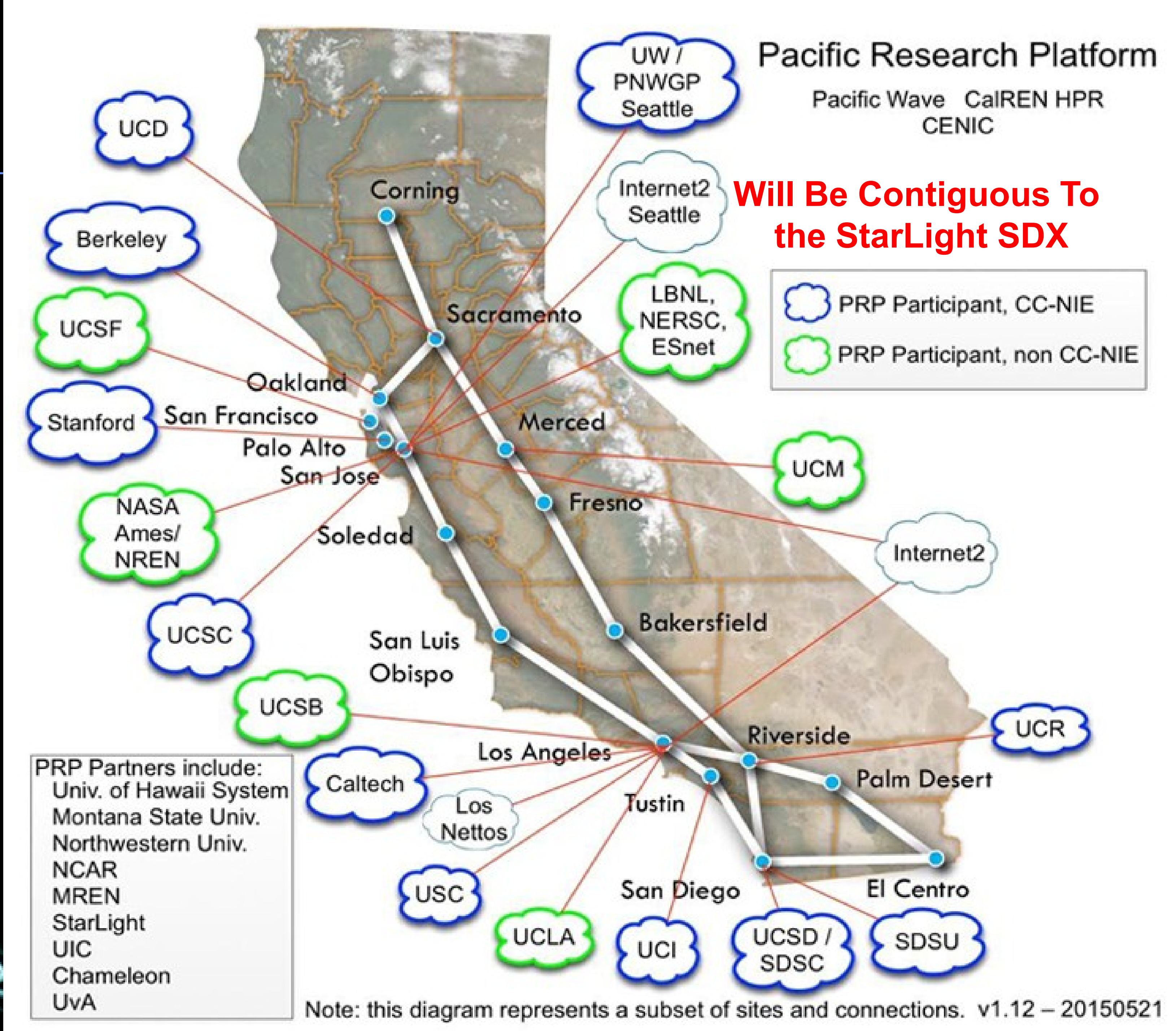


National Science Foundation
International Research Network Connections Program Workshop
Chicago, Illinois
May 15, 2015

Planned US SDX Interoperable Fabric



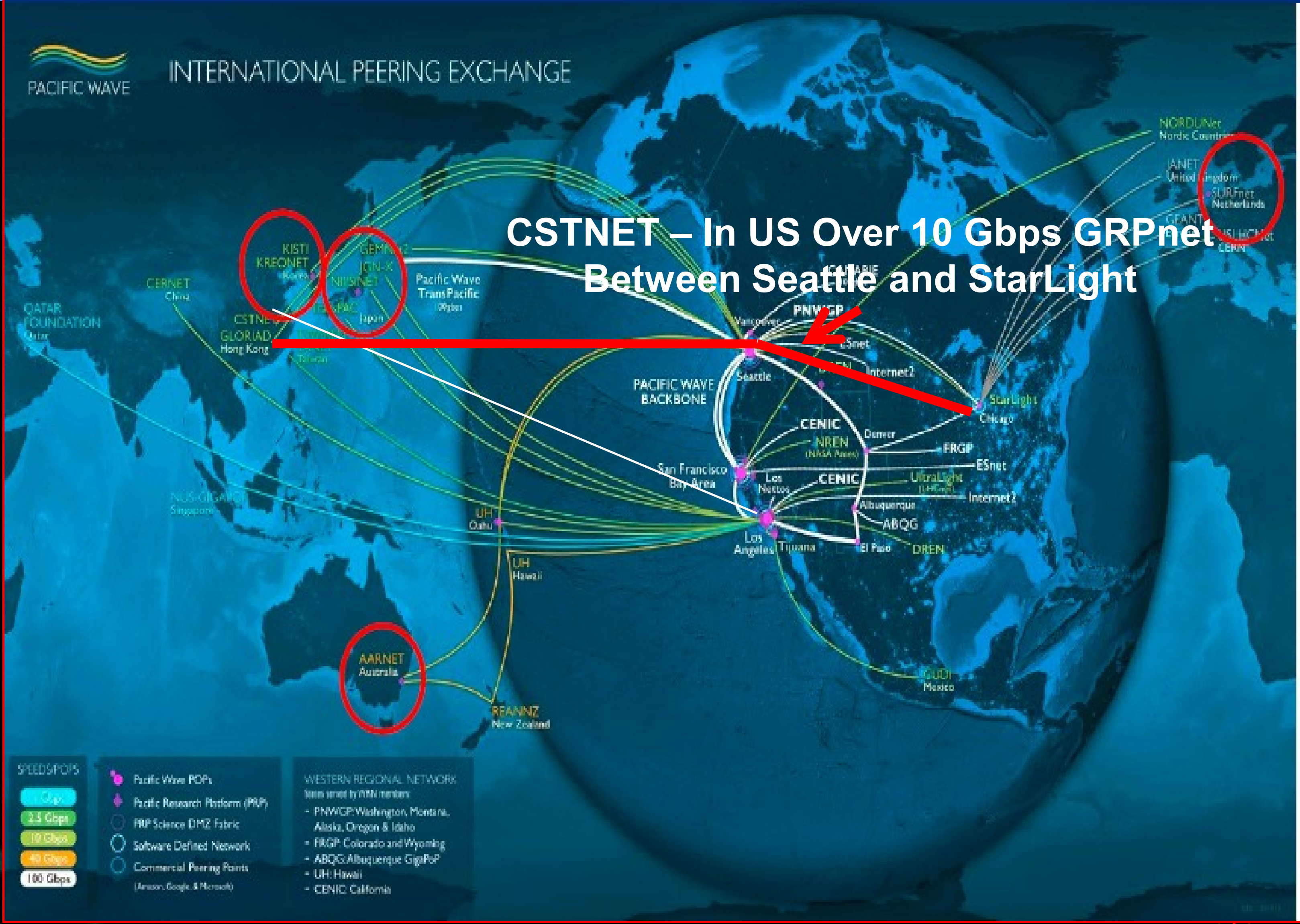
Will Be Contiguous To the StarLight SDX



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

Current International GRP Partners

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



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: 1st Demonstrations @ SC15, Austin Texas November 2015
- Subsequent Demonstrations @ SC16 Salt Lake City Utah, November 2016
- ***New=> Global Research Platform 100 Gbps Network (GRPnet) On Private Optical Fiber Between PacificWave and StarLight via the PNWGP***

STARLIGHT℠

www.startap.net/starlight