Next Generation Clouds, The Chameleon Cloud Testbed, and Software Defined Networking (SDN)

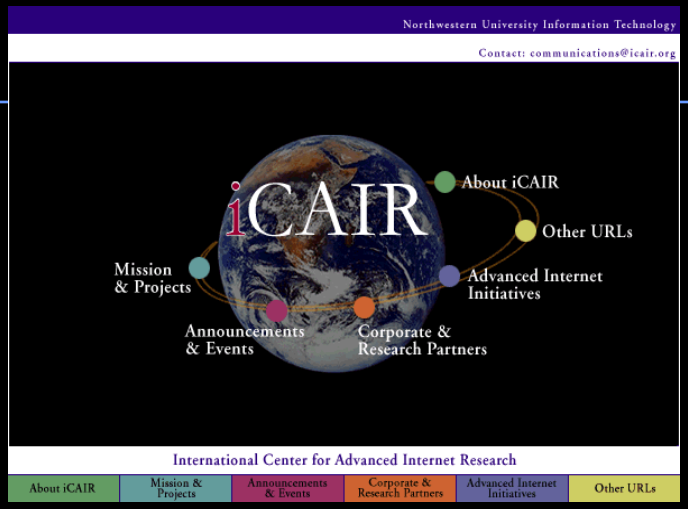
Joe Mambretti, Director, (j-mambretti@northwestern.edu)
International Center for Advanced Internet Research (www.icair.org)
Northwestern University
Director, Metropolitan Research and Education Network (www.mren.org)
Co-Director, StarLight, PI- StarLight SDX, PI-iGENI, PI-OMNINet, Co-PI Chameleon, (www.startap.net/starlight)

International Conference on Cloud Computing
Research and Innovation
Singapore
October 28-29, 2015

iCAIR

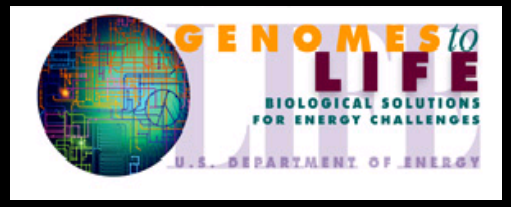
STARLIGHT℠

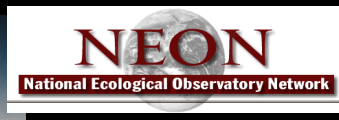
Introduction to iCAIR:

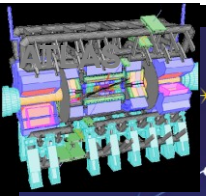


Accelerating Leading Edge Innovation and Enhanced Global Communications through Advanced Internet Technologies, in Partnership with the Global Community

- **Creation and Early Implementation of Advanced Networking Technologies - The Next Generation Internet All Optical Networks, Terascale Networks, Networks for Petascale Science**
- **Advanced Applications, Middleware, Large-Scale Infrastructure, NG Optical Networks and Testbeds, Public Policy Studies and Forums Related to NG Networks**
- **Three Major Areas of Activity: a) Basic Research b) Design and Implementation of Prototypes c) Operations of Specialized Communication Facilities (e.g., StarLight)**



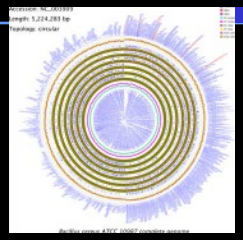




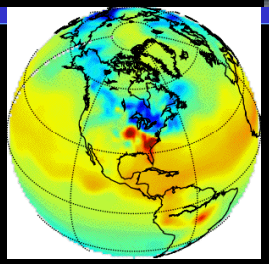
ALMA: Atacama Large Millimeter Array

ANDRILL: Antarctic Geological Drilling
www.andrill.org

BIRN: Biomedical Informatics Research Network
www.nbirn.net



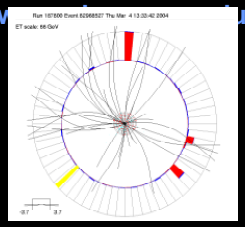
CAMERA metagenomics
camera.calit2.net



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

CineGrid
www.cinegrid.org

LHCONE
www.lhcone.net



DØ (DZero)
www-d0.fnal.gov

GEON: Geosciences Network
www.geongrid.org

GLEON: Global Lake Ecological Observatory Network

OOI-CI
ci.oceanobservatories.org

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

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

LIGO
www.ligo.org

IVOA: International Virtual Observatory
www.ivoa.net

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

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

TeraGrid
www.teragrid.org

OSG
www.opensciencegrid.org

Globus Alliance
www.globus.org

SKA
www.skatelescope.org

Sloan Digital Sky Survey
www.sdss.org

XSEDE
www.xsede.org

Compilation By Maxine Brown

STARLIGHT SM

Petascale Computational Science

For Decades, Computational Science Has Driven Network Innovation

Today –

Petascale Computational Science

BLUE WATERS

SUSTAINED PETASCALE COMPUTING

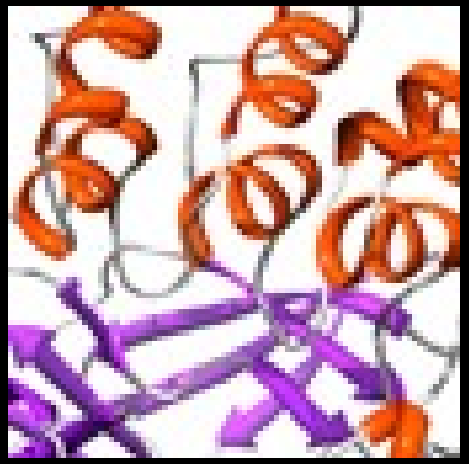
National Center for Supercomputing Applications, UIUC

STARLIGHT℠

TeraGrid=> XSEDE

Concept of Private Network As Backplane To
Distributed Computational Environment Continues In Next Iteration

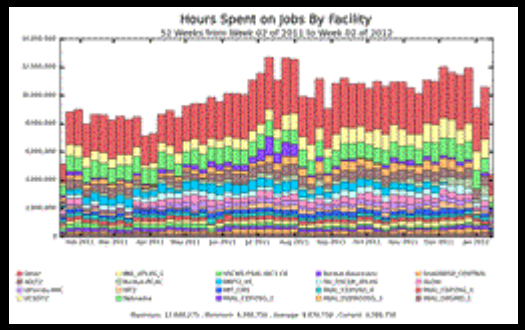
Open Science Grid: Selected Investigations



DNA Modeling

Gravity Wave Modeling

Nutrino Studies

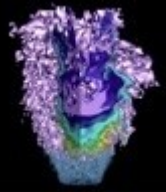


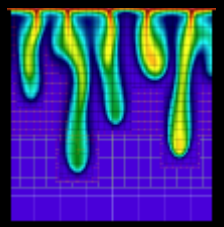
Usage

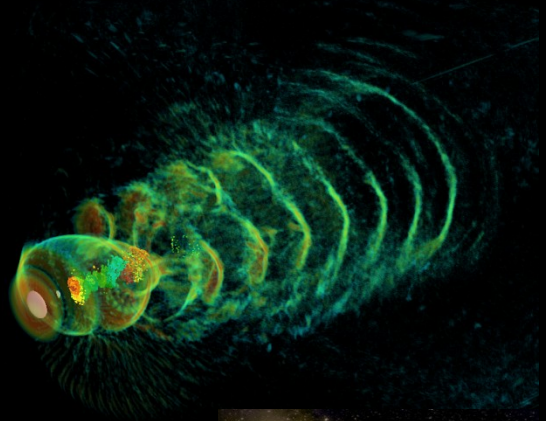
Open Science Grid

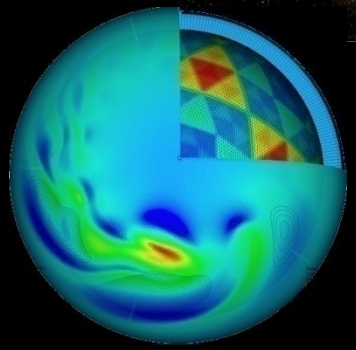
This Distributed Facility
Supports Many Sciences

HPC Cloud Computing









DOE Magellan Initiative: Testbed To Explore Cloud Computing For Science

STARLIGHT℠

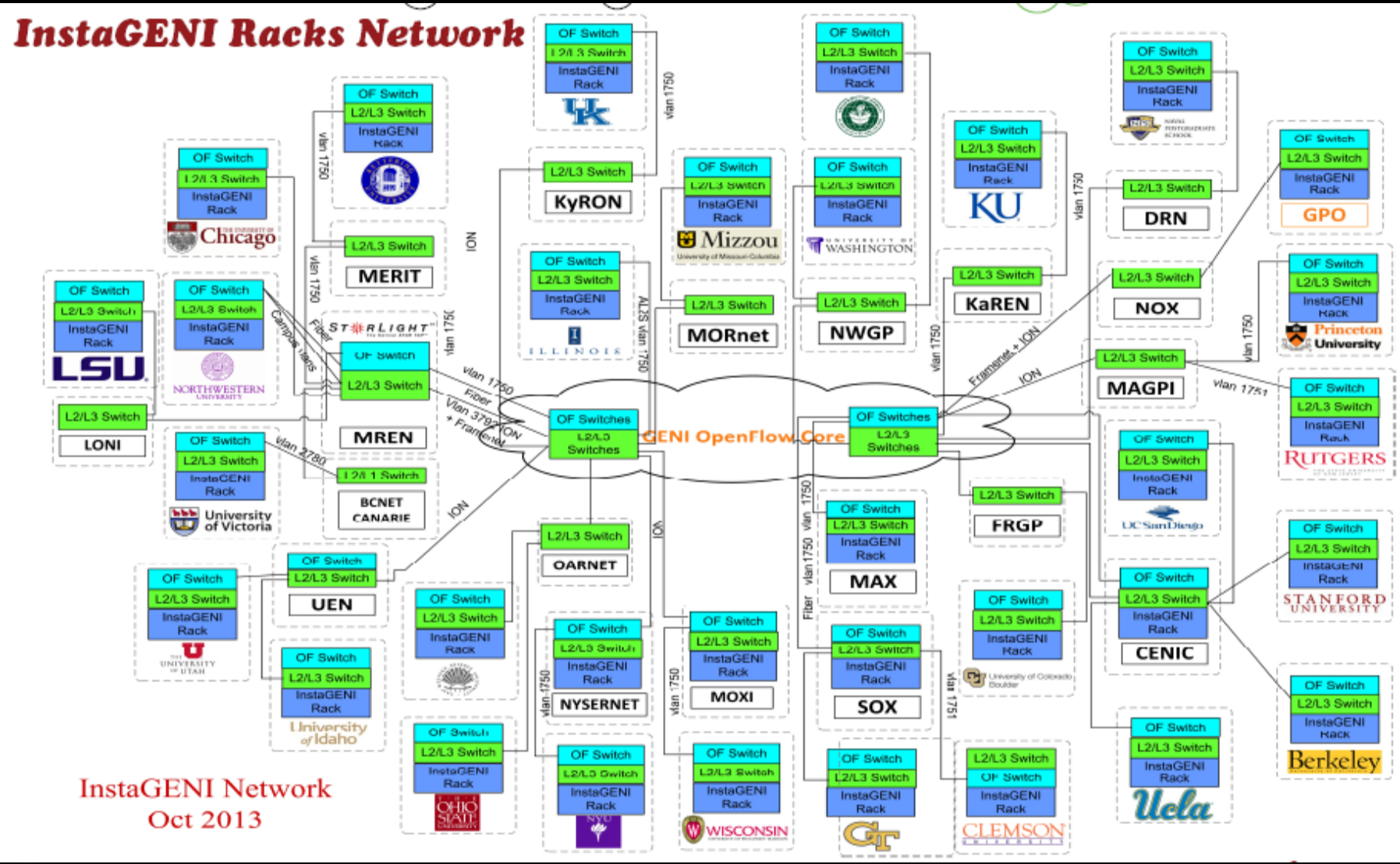
Multiple HPC Cloud Computing Testbeds Specifically Designed for Science Research

At Scale Experimentation
Integerated With High Performance Networks

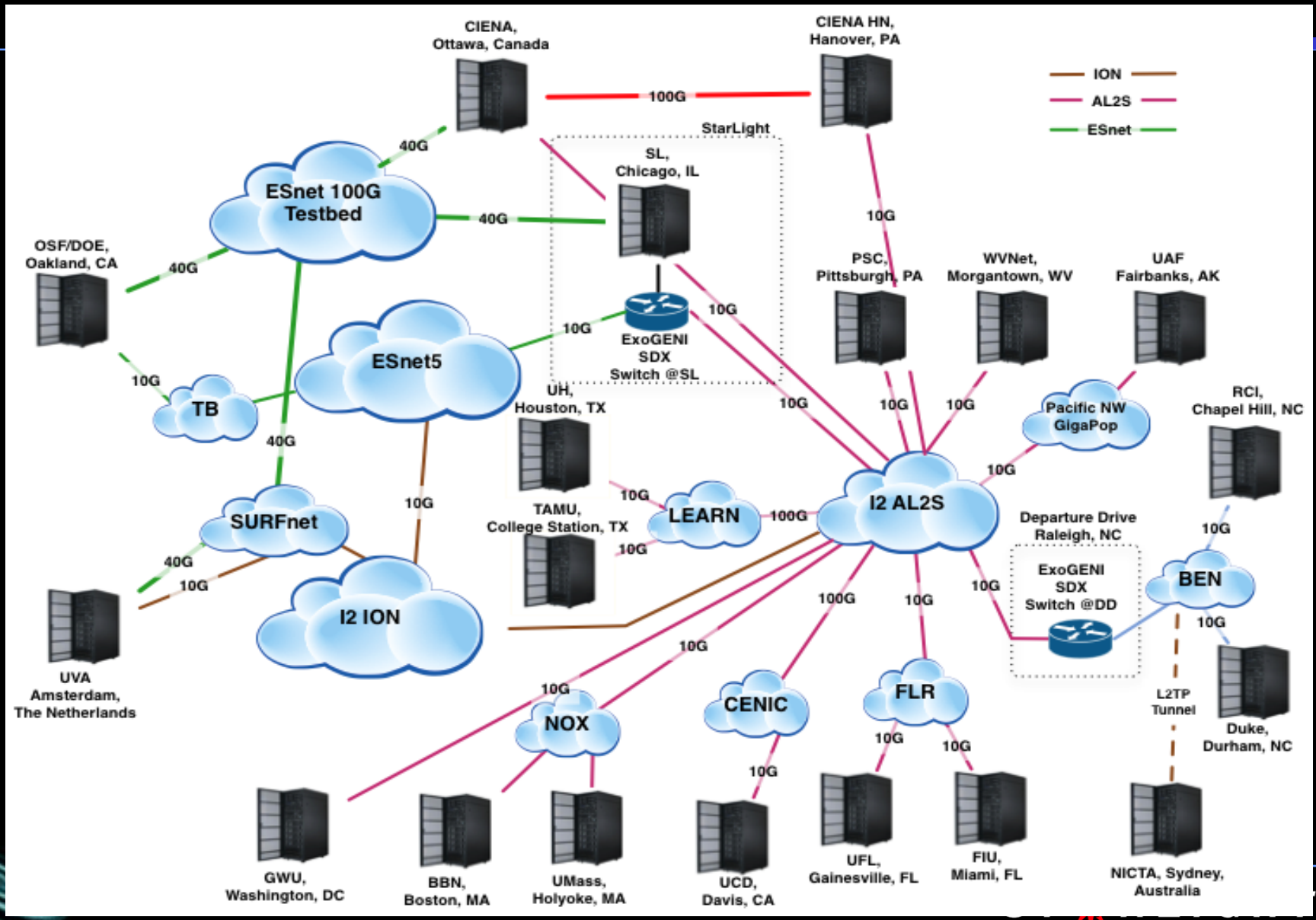
Global Environment for Network Innovations (GENI)

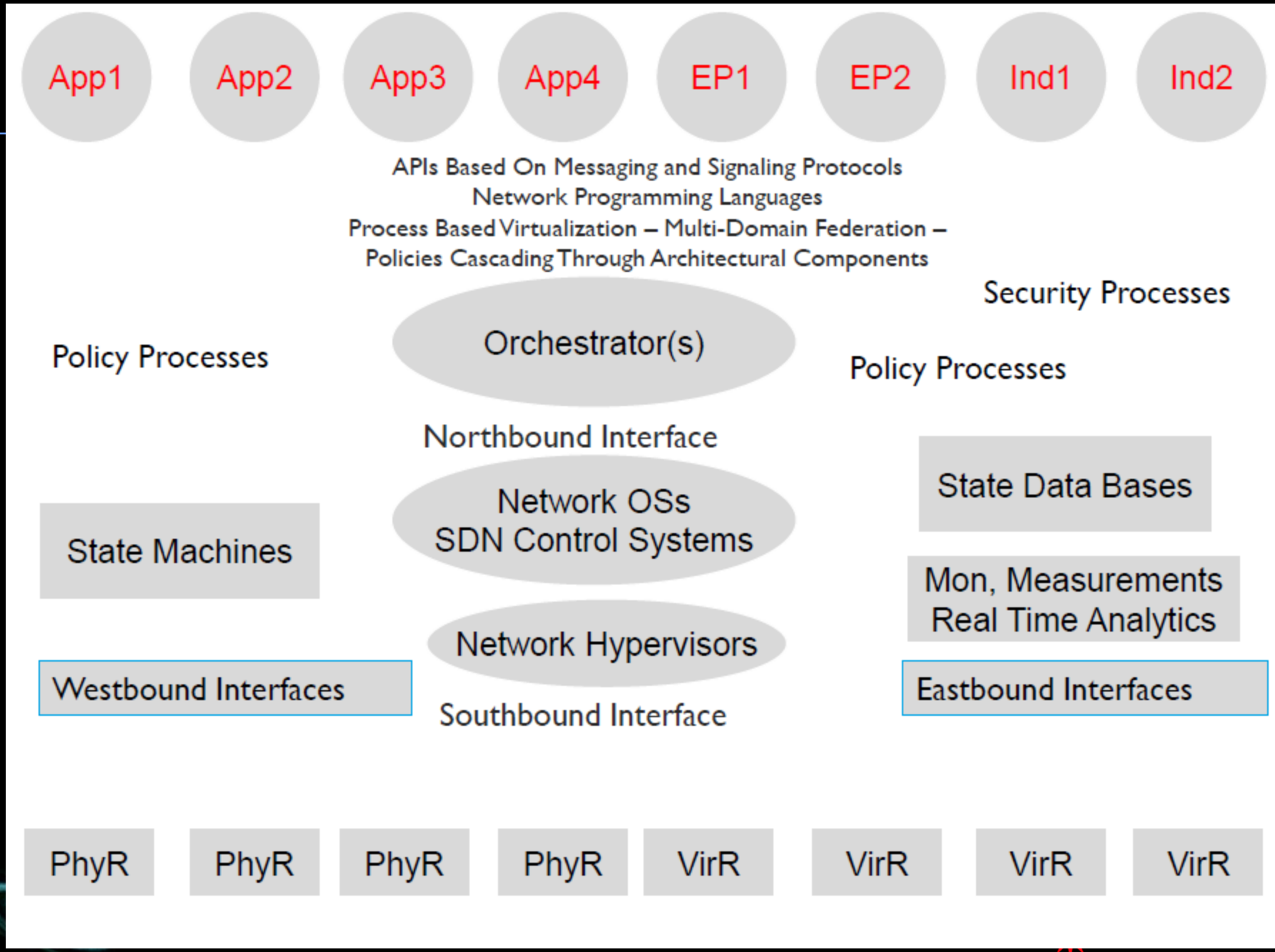
- GENI
  - Supports At-Scale Experimentation on Shared, Heterogeneous, Highly Instrumented Infrastructure
  - Enables Deep Programmability Throughout the Network,
  - Promotes innovations in Network Science, Security, Technologies, Services and Applications
  - Provides Collaborative and Exploratory Environments for Academia, Industry and the Public to Catalyze Groundbreaking Discoveries and Innovation.

Each Rack Is On a Different University Campus – Each Hosts a Cloud – A Distributed Sliceable Environment



International 40G and 100 G ExoGENI Testbed




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

An Experimental Testbed For Computer Science Research

www.chameleoncloud.org

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

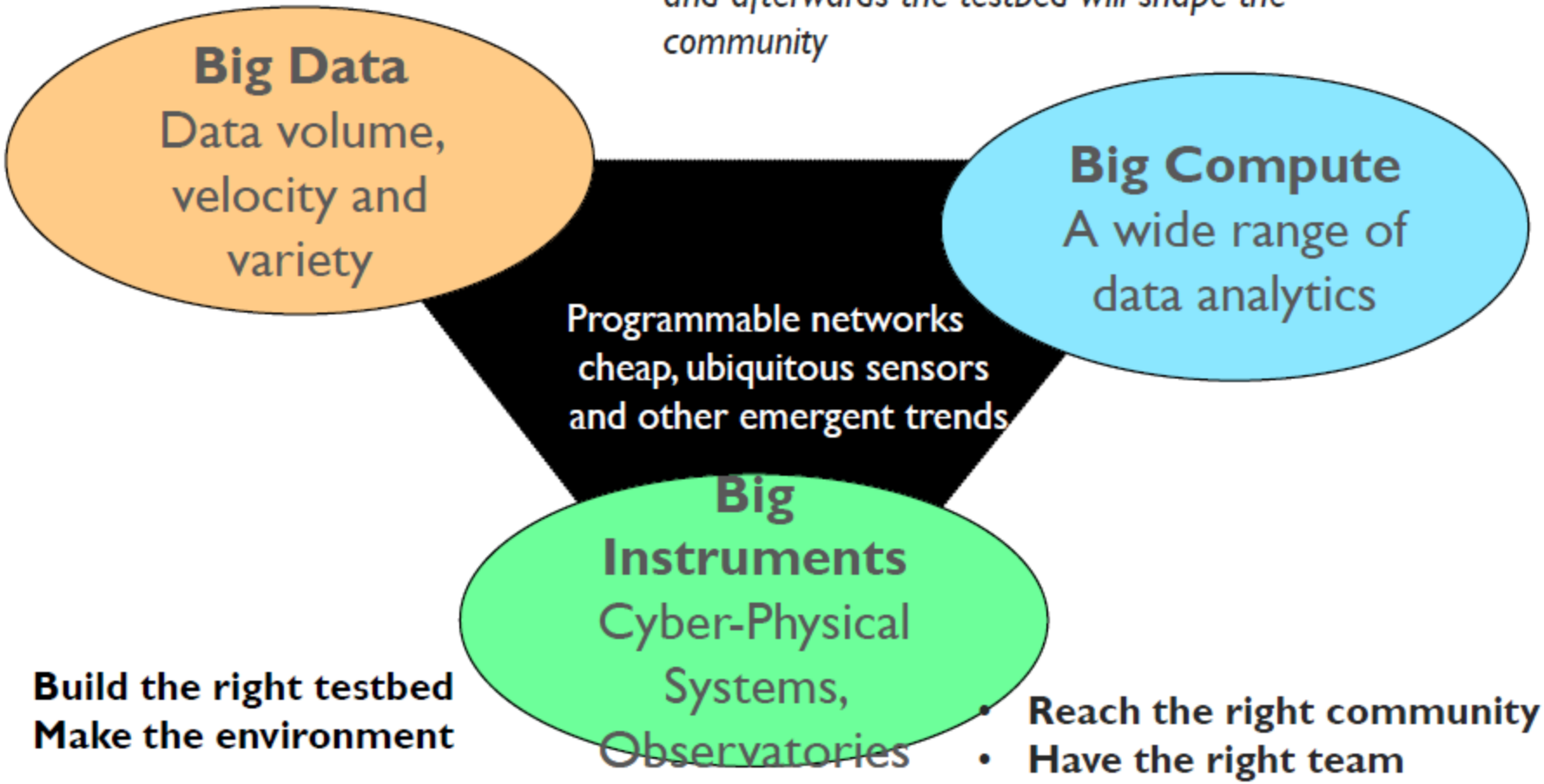
Principal Investigator: Kate Keahey

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

THE UNIVERSITY OF CHICAGO · TACC · NORTHWESTERN UNIVERSITY · THE OHIO STATE UNIVERSITY · UTSA · NSF · STARLIGHT℠

TESTBED TO SUPPORT THE COMMUNITY'S RESEARCH CHALLENGES

The community builds the testbed, and afterwards the testbed will shape the community



- **Build the right testbed**
- **Make the environment**

- **Reach the right community**
- **Have the right team**

CHAMELEON: A POWERFUL AND FLEXIBLE EXPERIMENTAL INSTRUMENT

- Large-scale
  - Targeting Big Data, Big Compute, Big Instrument research
  - Over 650 nodes, 5 PB disk, 100G network
- Reconfigurable
  - Bare metal reconfiguration, single instrument, graduated approach for ease-of-use
- Connected
  - Workload and Trace Archive, partners with production clouds
- Complementary
  - Complementing GENI, Comet, Wrangler, XSEDE
  - Partnering with GENI, Grid'5000, OCC, et al (FIRE?)
- Sustainable
  - Strong industry connections

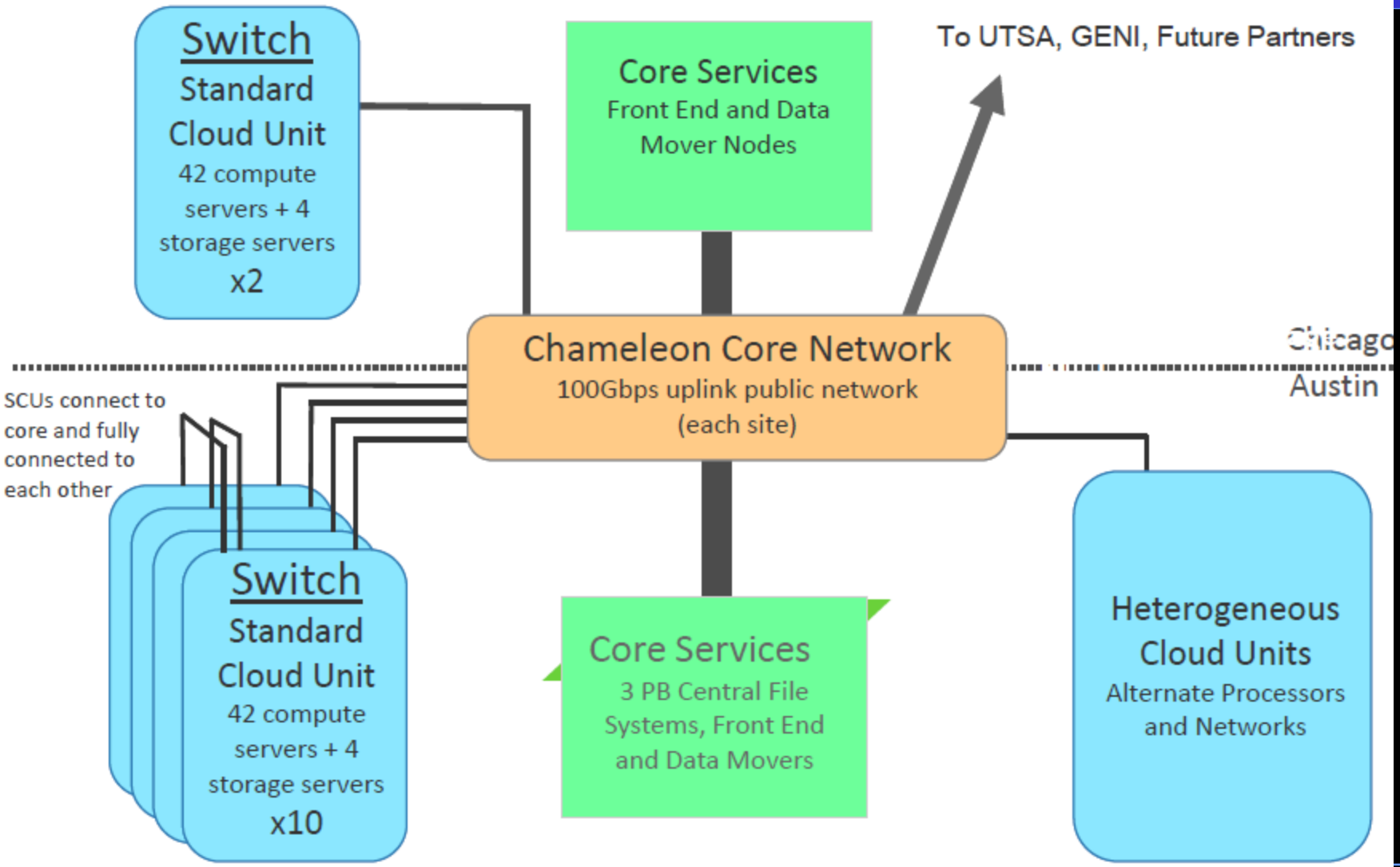
RESEARCH COMMUNITIES AND CAPABILITIES

Users	New models, algorithms, platforms, auto-scaling HA, etc., Application and educational uses	
	Persistent, reliable, shared cloud	
Core Researchers and Users	Repeatable experiments in new models, algorithms, platforms, auto-scaling, HA, etc.	
	Isolated partition, pre-configured images reconfiguration	
Core Researchers	Virtualization technology (SR-IOV, accelerators, etc.) Infrastructure-level resource management	
	Isolated partition, full bare metal reconfiguration	

SUPPORTED APPLICATIONS

- CPS
  - Offloading, muti-criteria trade-off analysis (response time vs cost), auto-scaling, high availability, etc.
- Machine learning, data mining
  - Mix of Big Compute and Big Data simulations and models, design of novel data processing frameworks
- System Software/Virtualization
  - Hypervisors optimizing a range of qualities, SR-IOV, virtualizing accelerators, etc.
- Networking
  - Programmable networks & QoS, refinement and effects of SR-IOV, large dataflows, end-to-end QoS

ARCHITECTURE



SYSTEM SOFTWARE: CORE CAPABILITIES

Users

Persistent clouds
OpenStack

Pre-configured Image Catalog
VM images

Core Researchers and Users

Pre-configured Image Catalog
Bare metal images

Core Researchers

Provisioning
LosF (TACC)
OpenFlow (GENI)

Scheduling
OAR2 (Grid'5000)

Orchestration
Cloudinit.d

CHAMELEON SERVICES AND FEATURES

- User Services
  - Allocation management through reservations, automatic image deployment
  - Dedicated Web portal for reservations, docs, stats, etc.
- Experiment Support
  - Trace and Workload Archive
  - Experiment enhancement (e.g., load generators)
- Additional Features
  - Reconfigurable, connected instrument
  - Development-focused approach
  - Ease-of- use: one stop shopping for experimental needs
  - Distinct from off-the-shelf cloud services
  - Code ownership and collaboration

NETWORKING CAPABILITIES

- Expose SDN, OpenFlow, etc. to users
  - Isolation
  - Hybrid Network Capabilities
  - Programmable Topologies
  - Integration With Other Resources Within and External to the Testbed
- Pushing 100G Networks To Their Limit
  - Using 100G + SDN Optimally
  - Chameleon appliances and services allow experimenters a highly granulated view into -- and control -- over traffic flows
- Integration/Federation with GENI
- Common Policy Context

Federation Among Multiple International Testbeds (Federation-as-a-Service)

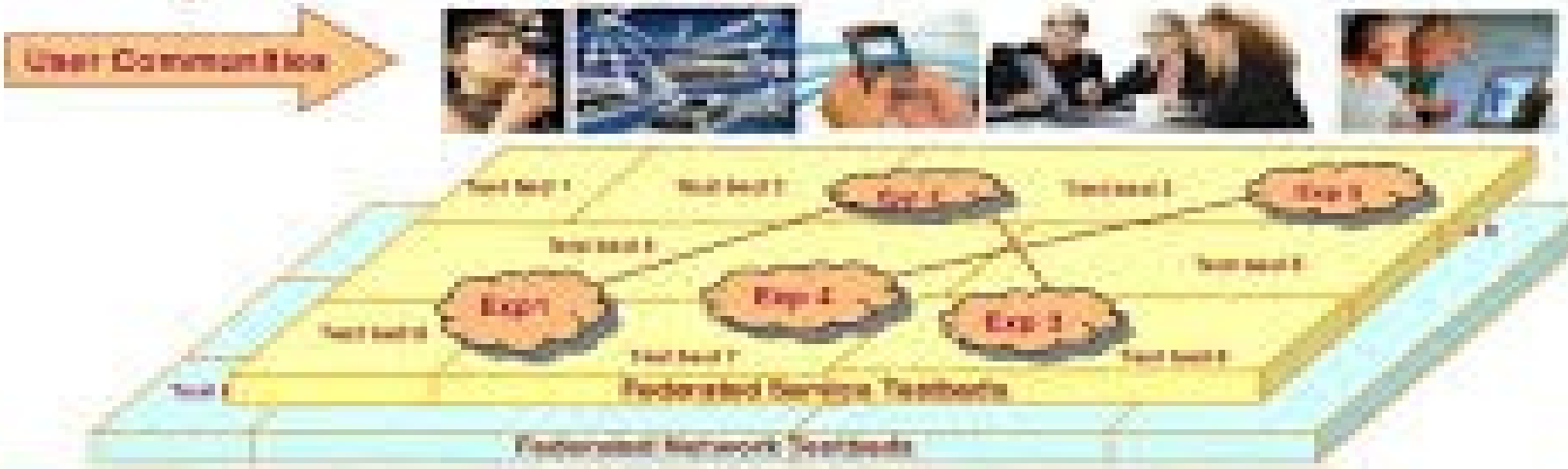
- Pair-Wise Approaches Among Multiple Is A Challenge
- Centralized Approaches Do Not Scale
- Decentralized Processes Will Scale
- SDXs Are Key Resources For Federation
- Requirements
  - Multi-Domain
  - Multi-Resource
  - Discovery
  - Integration
  - Policy Based Resource Utilization
  - Flexibility
  - Selectivity (e.g. A⇔N⇔T ⇔NOT X ⇔ Conditionally Z)

Future Internet Research and Experimentation

What is FIRE?

FIRE Research

FIRE Experimental Facility



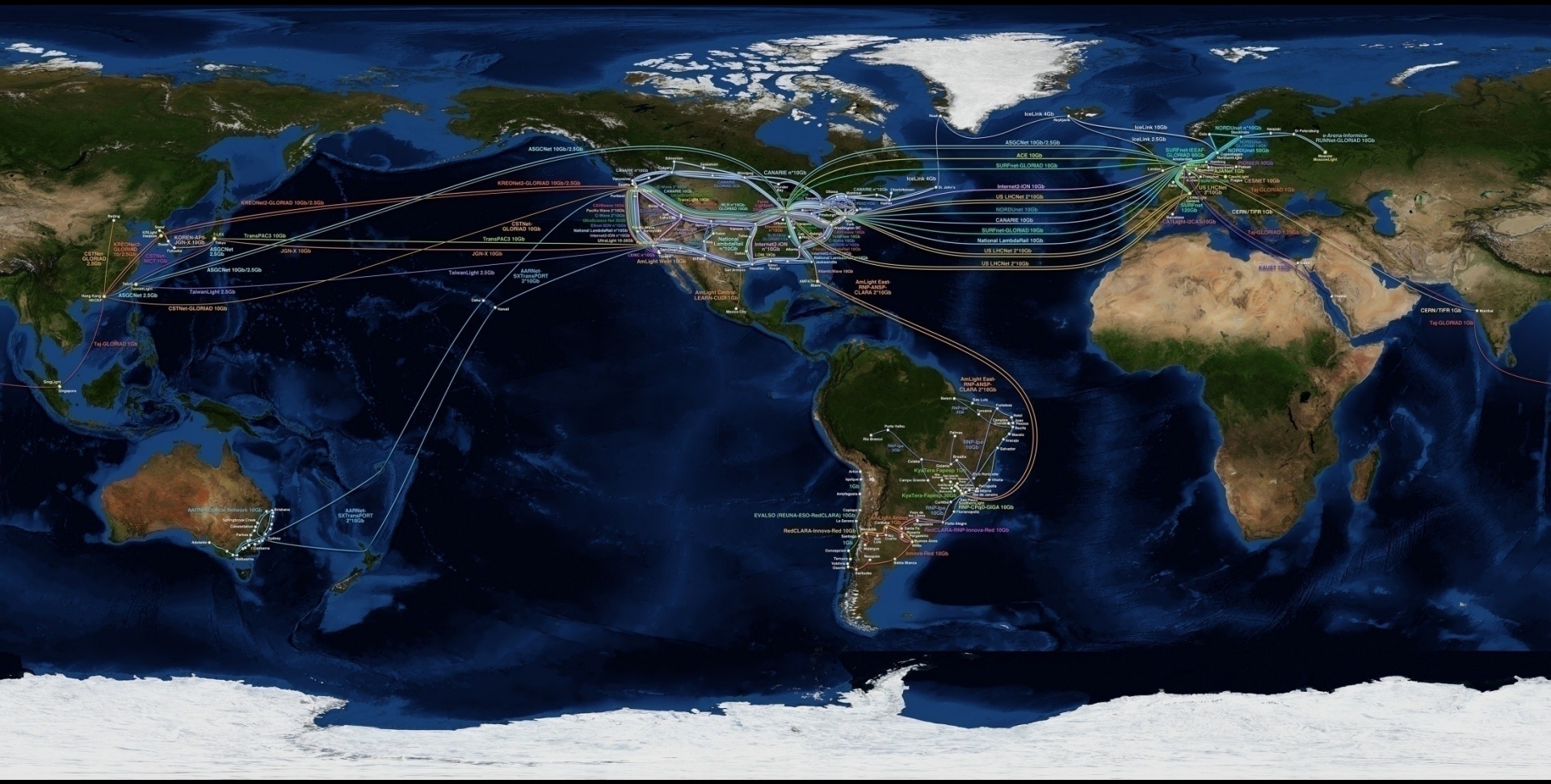
FIRE - Future Internet Research and Experimentation

FIRE⇔FIRE Federation Project⇔ GENI⇔Chameleon

iGENI: The International GENI

- The iGENI Initiative Designed, Developed, Implemented, and Operated a Major National and International Distributed Infrastructure.
- iGENI Placed the "G" in GENI Making GENI Truly Global.
- iGENI Is a Unique Distributed Infrastructure Supporting Research and Development for Next-Generation Network Communication Services and Technologies.
- This Infrastructure Has Been 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 Has Connected Its Resources With Current GENI National Backbone Transport Resources, With Current and Planned GENI Regional Transport Resources, and With International Research Networks and Projects
-

GLIF is a consortium of institutions, organizations, consortia and country National Research & Education Networks who voluntarily share optical networking resources and expertise to develop the *Global LambdaGrid* for the advancement of scientific collaboration and discovery – a *Federation!.*



StarLight International/National Communications Exchange Facility– "By Researchers For Researchers"

StarLight Is an Innovation Platform For Advanced Communications Services Architecture and Technologies, Including Experimental Testbeds Optimized For High-Performance Data Intensive Applications

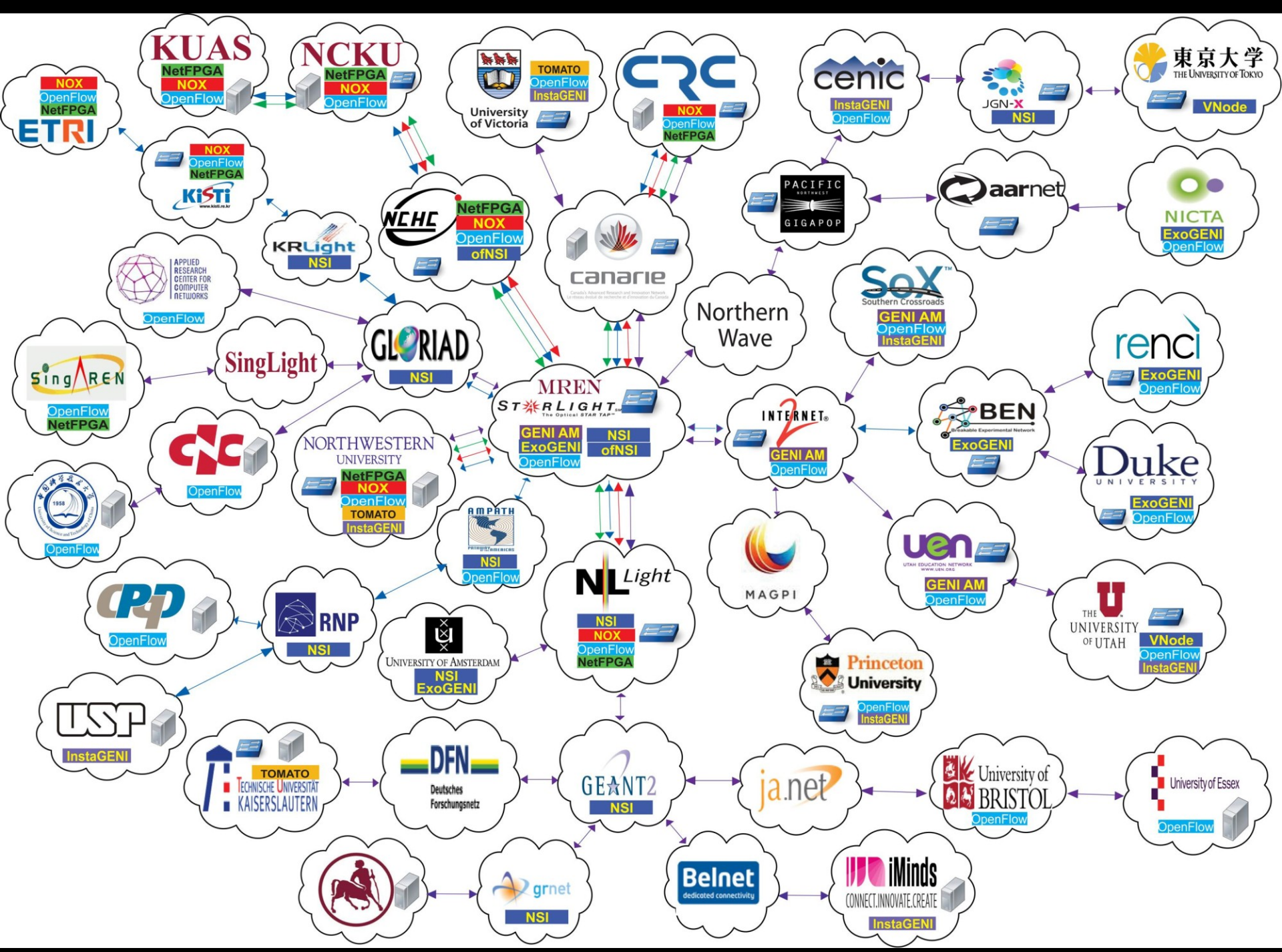
Multiple 10GE+100 Gbps Over Optics – World's "Largest" 10G/100G Exchange

First of a Kind Enabling Interoperability At L1, L2, L3

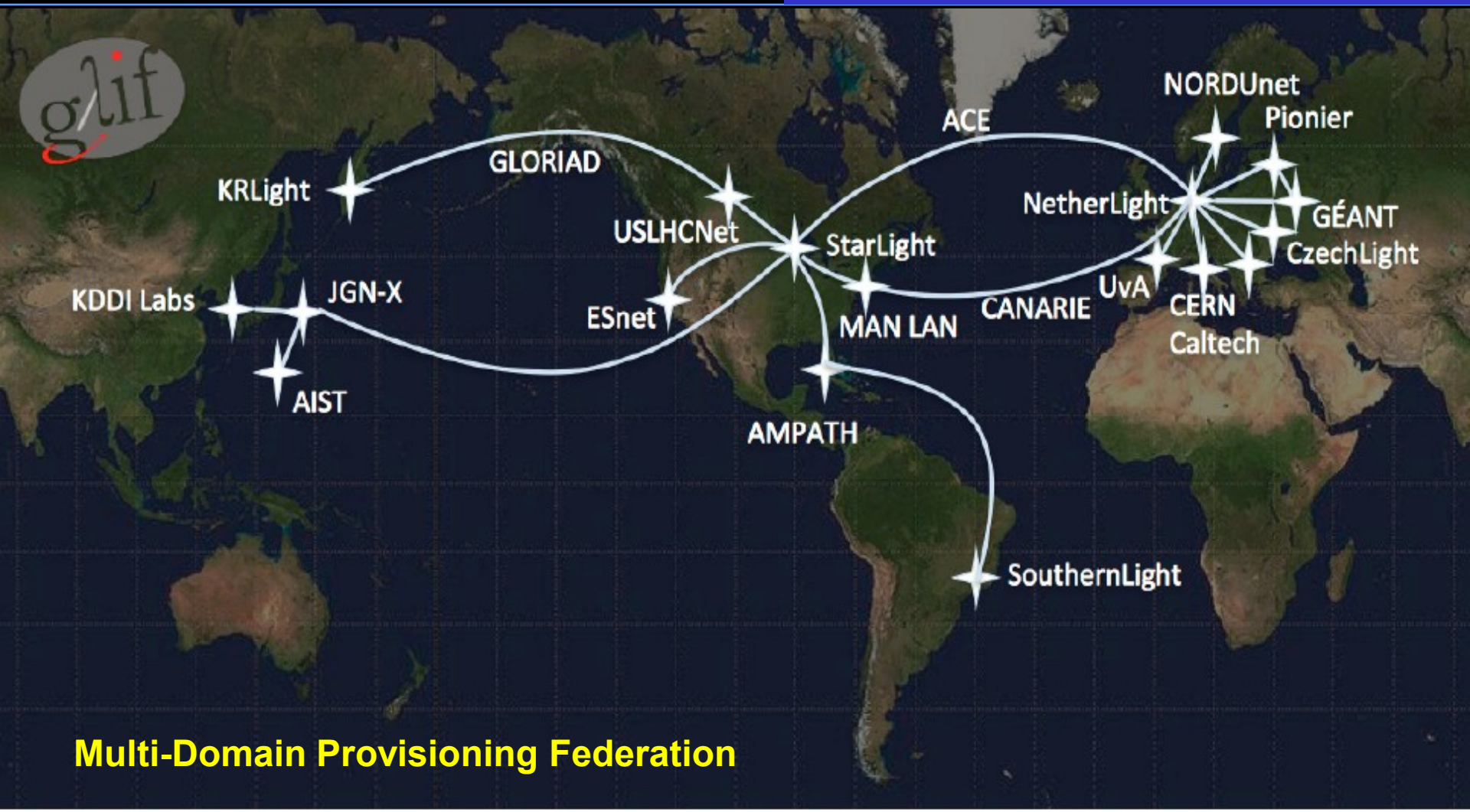
Also, StarWave Multi-100 Gbps Exchange

View from StarLight

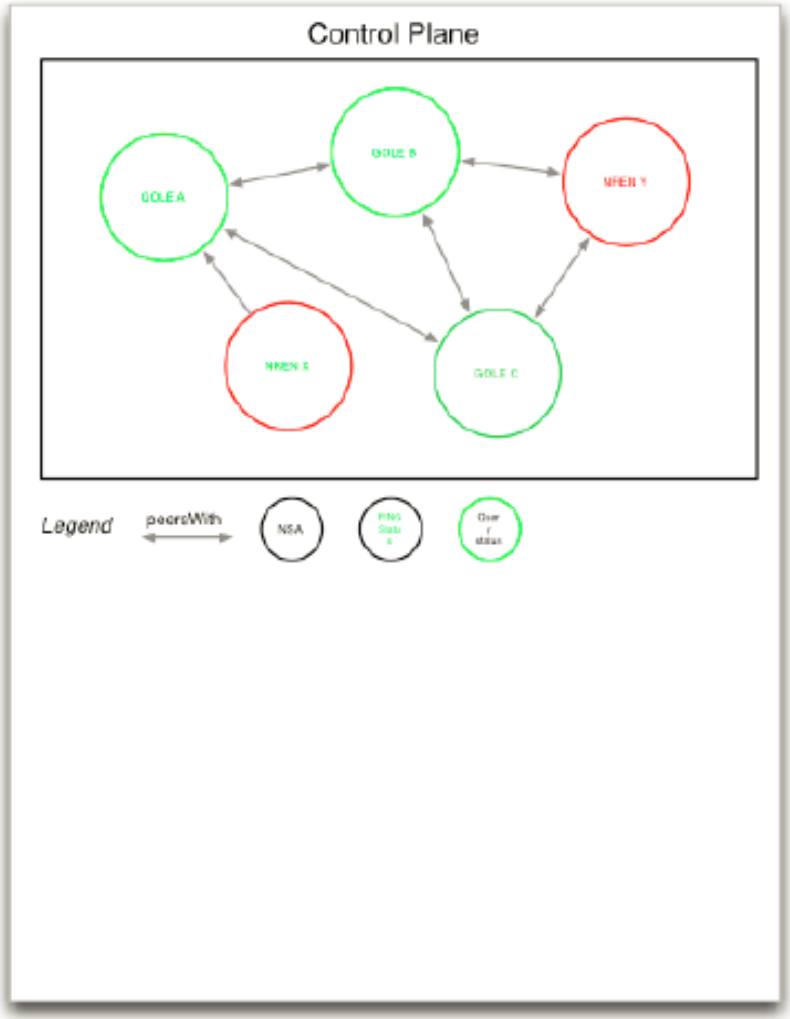
Abbott Hall, Northwestern University's Chicago Campus

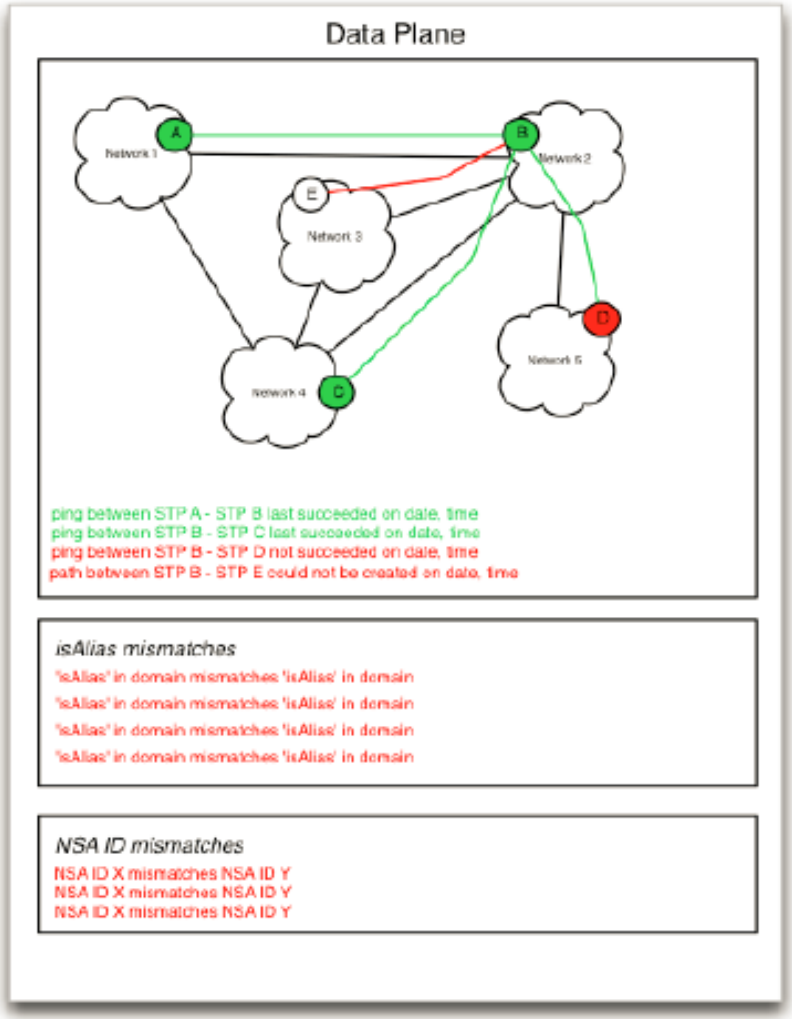

KUAS
NetFPGA
NOX
OpenFlow
NCKU
NetFPGA
NOX
OpenFlow
ETRI
NOX
OpenFlow
NetFPGA
KiSTi
University of Victoria
TOMATO
OpenFlow
InstaGENI
CRC
NOX
OpenFlow
NetFPGA
cenic
InstaGENI
OpenFlow
JGN-X
NSI
東京大学
THE UNIVERSITY OF TOKYO
VNode
PACIFIC NORTHWEST GIGAPOP
aarnet
NICTA
ExoGENI
OpenFlow
NCHC
NetFPGA
NOX
OpenFlow
ofNSI
KRLight
NSI
canarie
APPLIED RESEARCH CENTER FOR COMPUTER NETWORKS
OpenFlow
GLORIAD
NSI
SingLight
SingAREN
OpenFlow
NetFPGA
Northern Wave
SoX
Southern Crossroads
GENI AM
OpenFlow
InstaGENI
renci
ExoGENI
OpenFlow
MREN
STARLIGHT
The Optical STAR TAP
GENI AM
ExoGENI
OpenFlow
NSI
ofNSI
INTERNET2
GENI AM
OpenFlow
BEN
Breakable Experimental Network
ExoGENI
Duke
UNIVERSITY
ExoGENI
OpenFlow
CNC
OpenFlow
NORTHWESTERN UNIVERSITY
NetFPGA
NOX
OpenFlow
TOMATO
InstaGENI
AMPATH
NSI
OpenFlow
MAGPI
uen
UTAH EDUCATION NETWORK
GENI AM
OpenFlow
THE UNIVERSITY OF UTAH
VNode
OpenFlow
InstaGENI
CPqD
OpenFlow
RNP
NSI
NLight
NSI
NOX
OpenFlow
NetFPGA
UNIVERSITY OF AMSTERDAM
NSI
ExoGENI
Princeton University
OpenFlow
InstaGENI
USP
InstaGENI
TOMATO
TECHNISCHE UNIVERSITÄT KAISERSLAUTERN
DFN
Deutsches Forschungsnetz
GEANT2
NSI
ja.net
University of BRISTOL
OpenFlow
University of Essex
OpenFlow
grnet
NSI
Belnet
dedicated connectivity
iMinds
CONNECT.INNOVATE.CREATE
InstaGENI

AutoGOLE Initiative



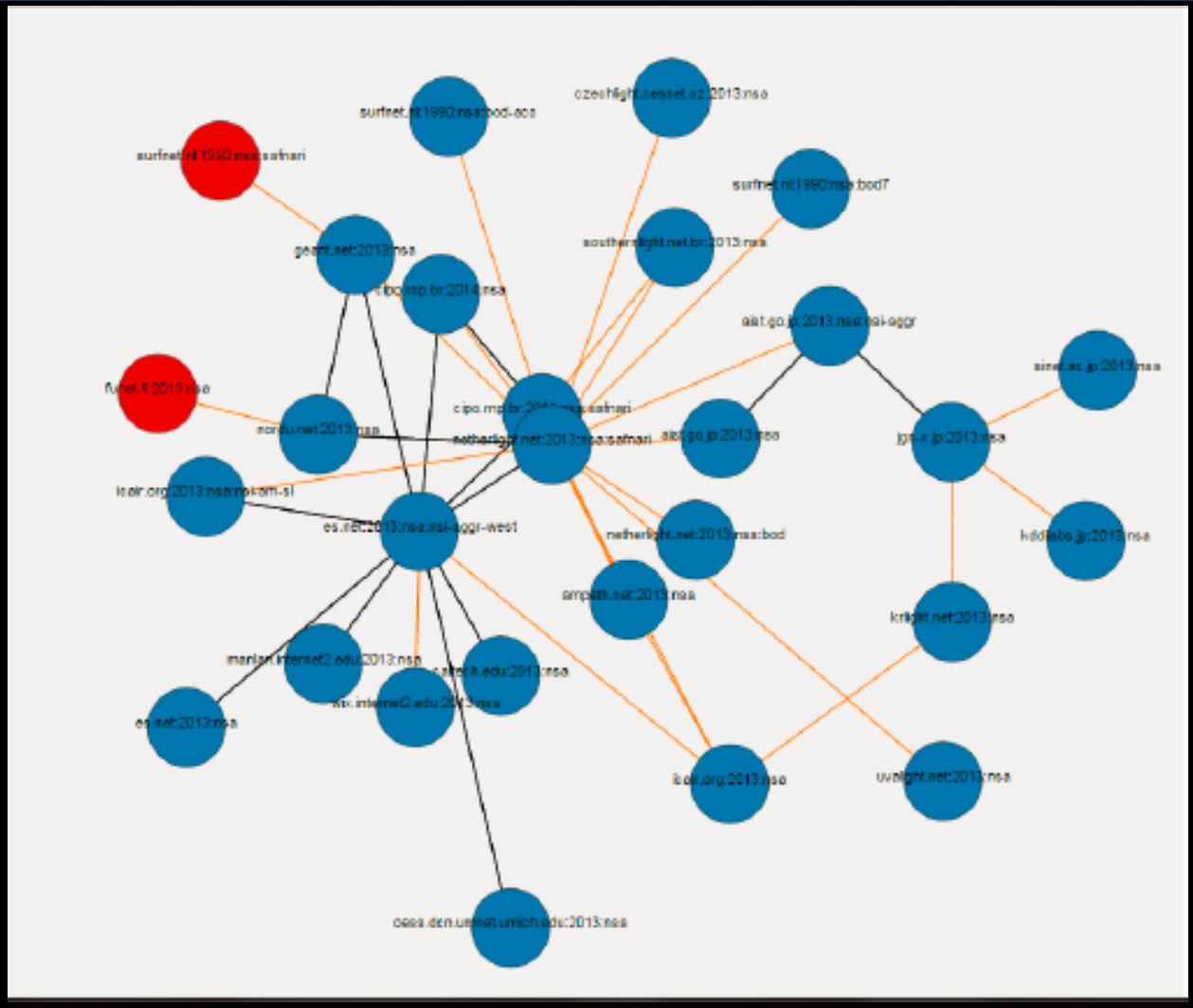
AutoGOLE Dashboard (In Development)



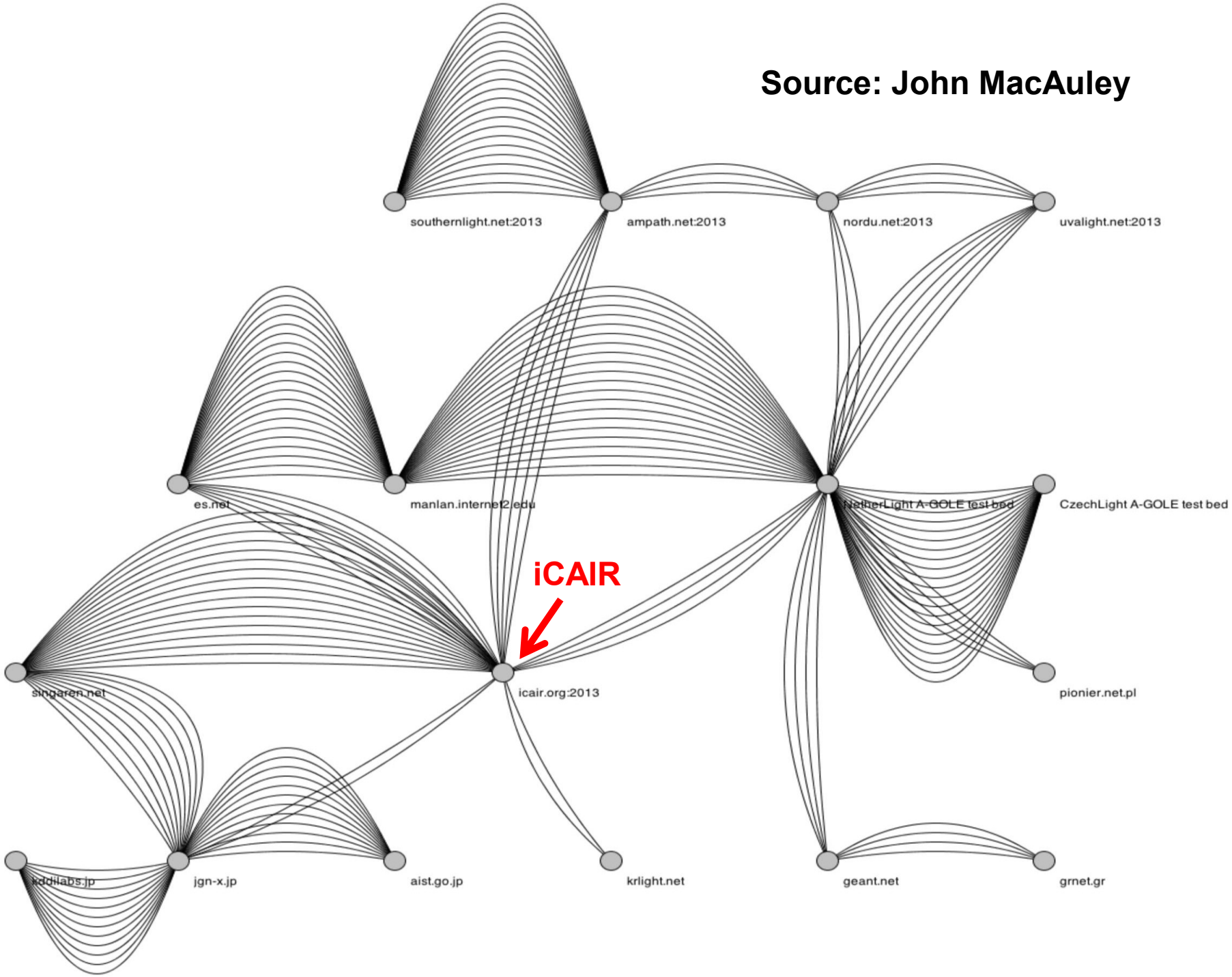


Source: Gerben van Malenstein, SURFnet

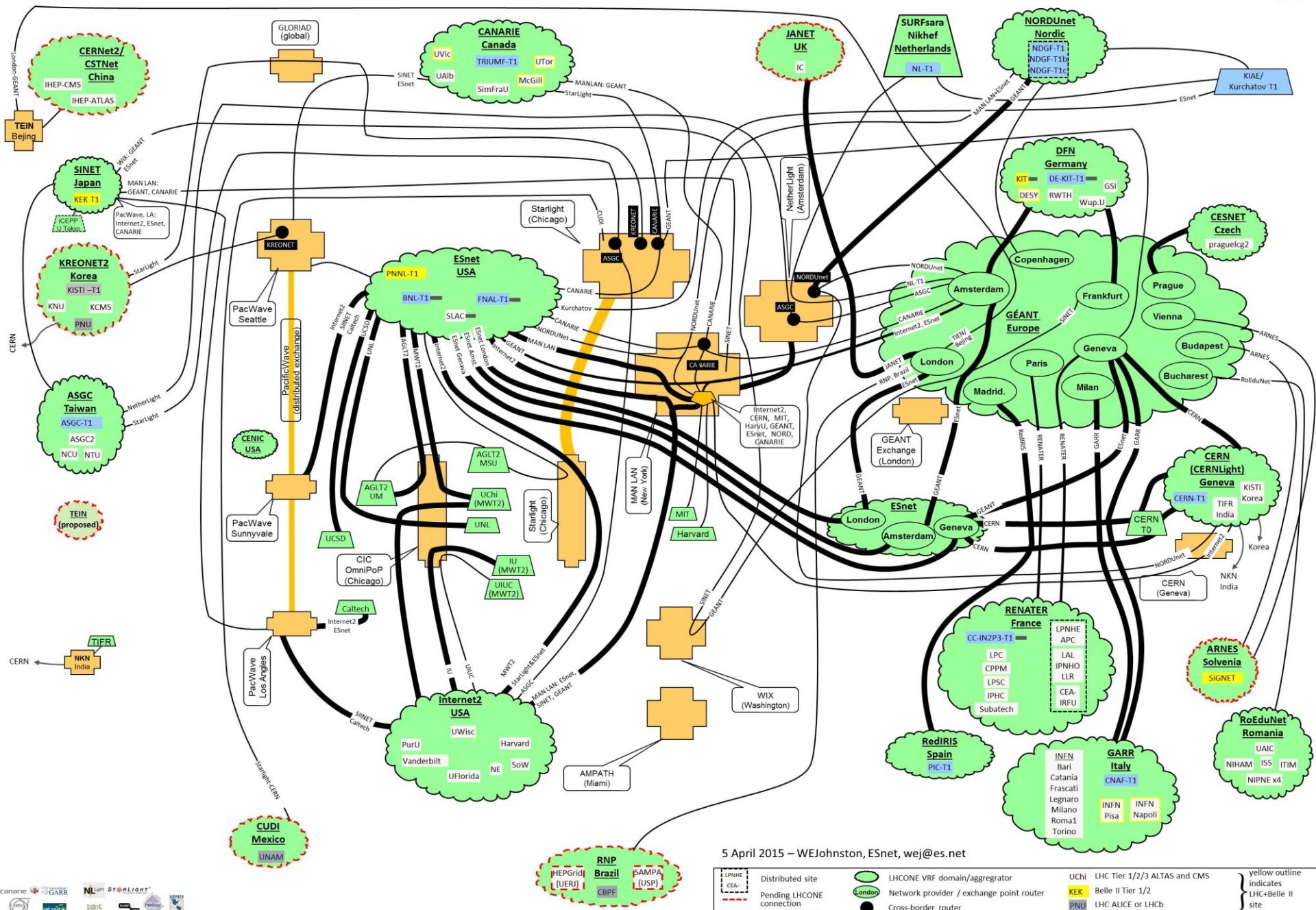
Current Version of Dashboard



Source: John MacAuley



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



5 April 2015 – WEJohnston, ESnet, wej@es.net

Also see http://lhcone.net for details.

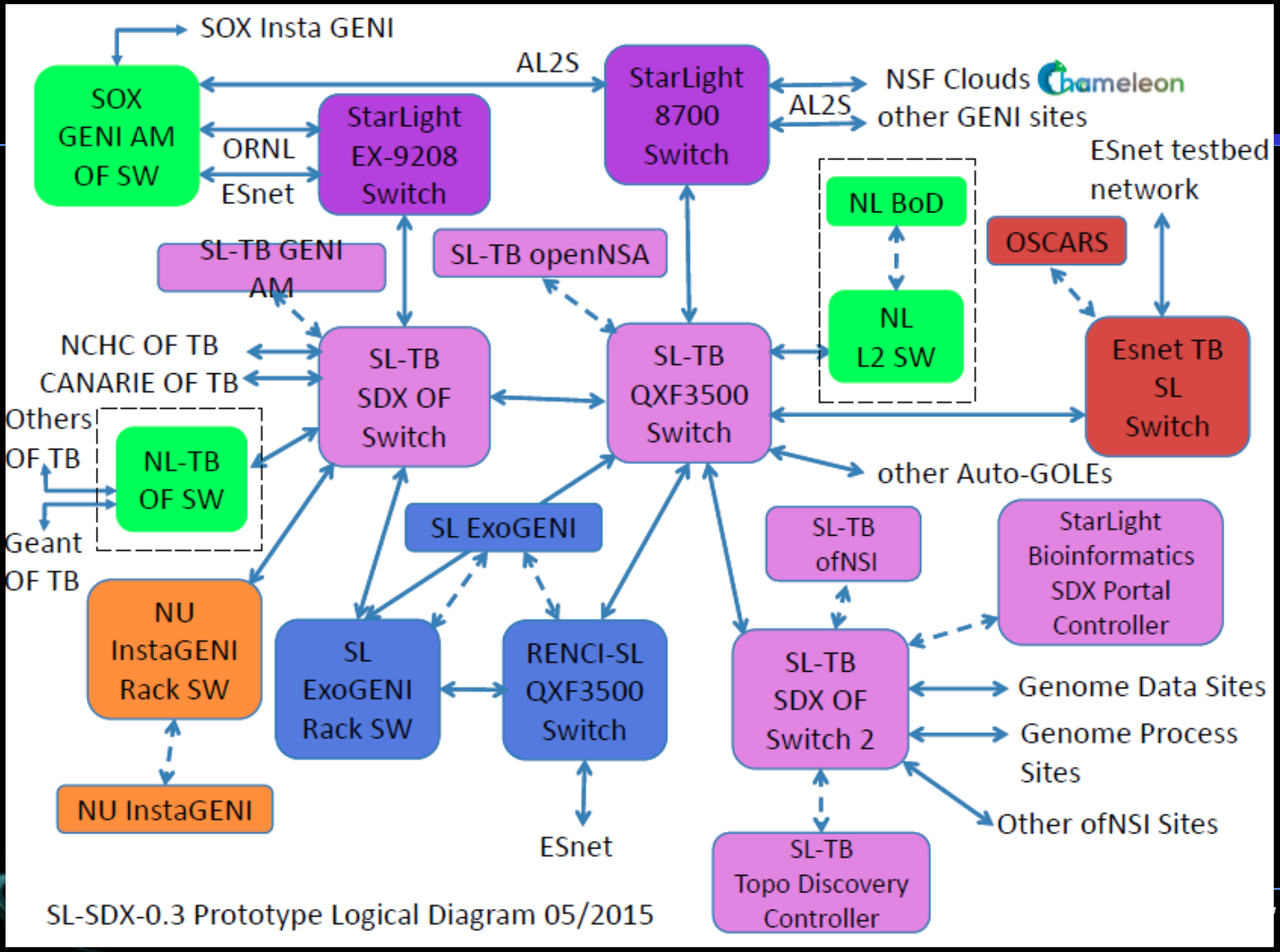
Distributed site
Pending LHCONE connection
Sites connected at 40G-100G
Broadcast VLAN
LHCONE VRF domain/aggregator
Network provider / exchange point router
Cross-border router
Regional R&E communication nexus w/ switch providing VLAN connections
UChi LHC Tier 1/2/3 ALTAS and CMS
KEK Belle II Tier 1/2
PNU LHC ALICE or LHCb
UNL Sites that are standalone VRFs,
yellow outline indicates LHC+Belle II site
Communication links: 1/10, 20/30/40, and 100Gb/s

Software Defined Networking Exchanges (SDXs)

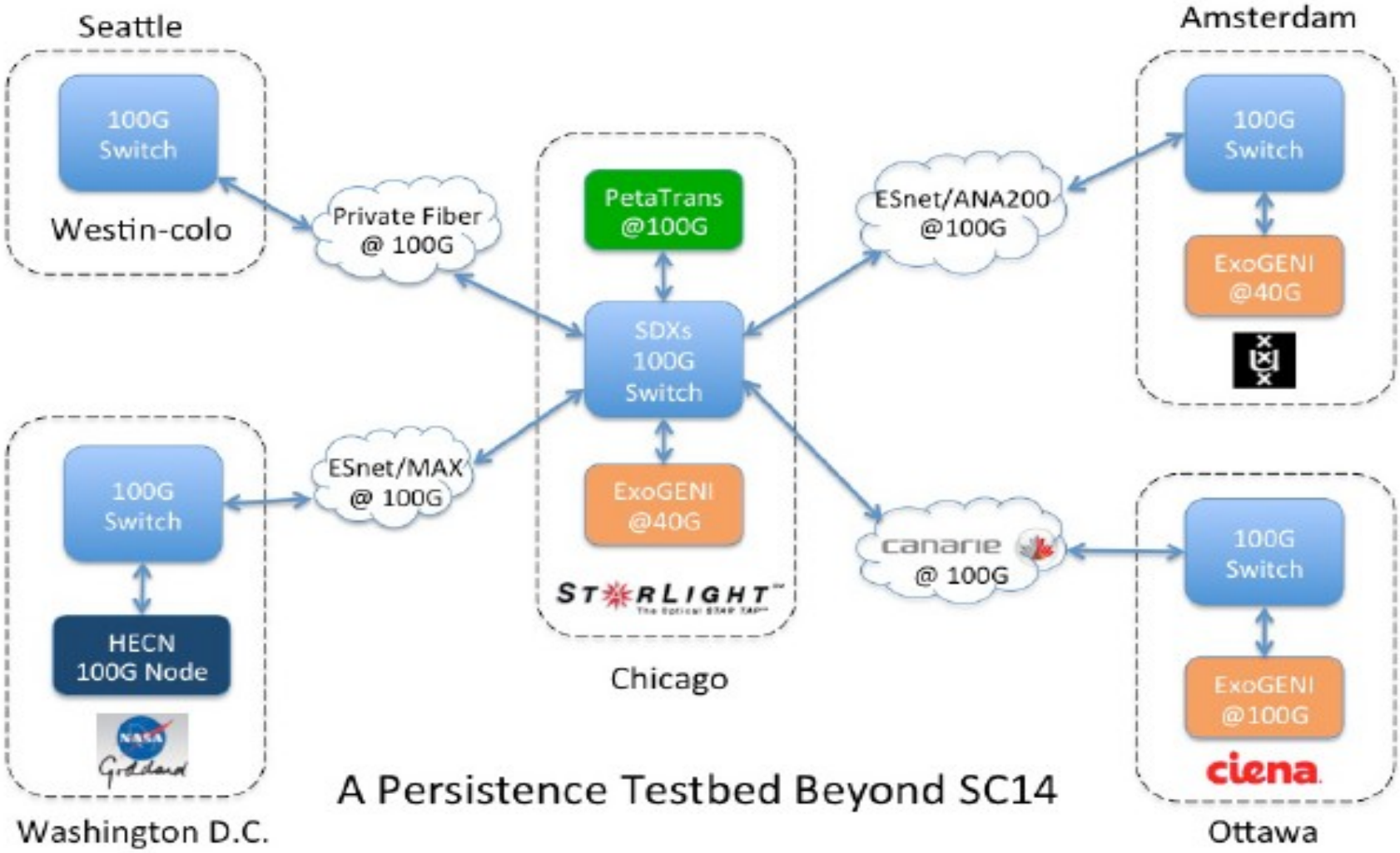
- With the Increasing Deployment of SDN In Production Networks, the Need for an SDN Exchange (SDX) Has Been Recognized.
- Many Motivations Exist for SDXs
  - Bridging SDNs (Which Are Single Domain & Centralized Controller Oriented)
  - Granulated Engineering Over Flows
  - High Degrees Of Exchange Customization
- Required: Capabilities for Multi-Domain Distributed SDN Resource Discovery, Signaling, Provisioning, Federation, Operational Functions, Fault Detection and Recovery
- These Are Fairly Challenging Issues

Selected SDX Architectural Attributes

- Control and Network Resource APIs
- Multi Domain Integrated Path Controllers (With Federation)
- Controller Signaling, Including Edge Signaling
- SDN/OF Multi Layer Traffic Exchange Services
- Multi Domain Resource Advertisement/Discovery
- Topology Exchange Services
- Multiple Highly Customized Services At All Layers
- Granulated Resource Access (Policy Based), Including Through Edge Processes, Including To individual Streams
- Foundation Resource Programmability
- Various Types of Gateways To Other Network Environments
- Integration of OF and Non-OF Paths, Including 3rd Party Integration
- Programmability for Large Scale Large Capacity Streams


SOX Insta GENI
SOX GENI AM OF SW
AL2S
StarLight 8700 Switch
NSF Clouds
Chameleon
AL2S
other GENI sites
ORNL
ESnet
StarLight EX-9208 Switch
ESnet testbed network
NL BoD
OSCARS
SL-TB GENI AM
SL-TB openNSA
NL L2 SW
NCHC OF TB
CANARIE OF TB
SL-TB SDX OF Switch
SL-TB QXF3500 Switch
Esnet TB SL Switch
Others OF TB
NL-TB OF SW
other Auto-GOLEs
SL ExoGENI
SL-TB ofNSI
StarLight Bioinformatics SDX Portal Controller
Geant OF TB
NU InstaGENI Rack SW
SL ExoGENI Rack SW
RENCI-SL QXF3500 Switch
SL-TB SDX OF Switch 2
Genome Data Sites
Genome Process Sites
Other ofNSI Sites
NU InstaGENI
ESnet
SL-TB Topo Discovery Controller
SL-SDX-0.3 Prototype Logical Diagram 05/2015

PetaTrans: Petascale Science Data Transfer

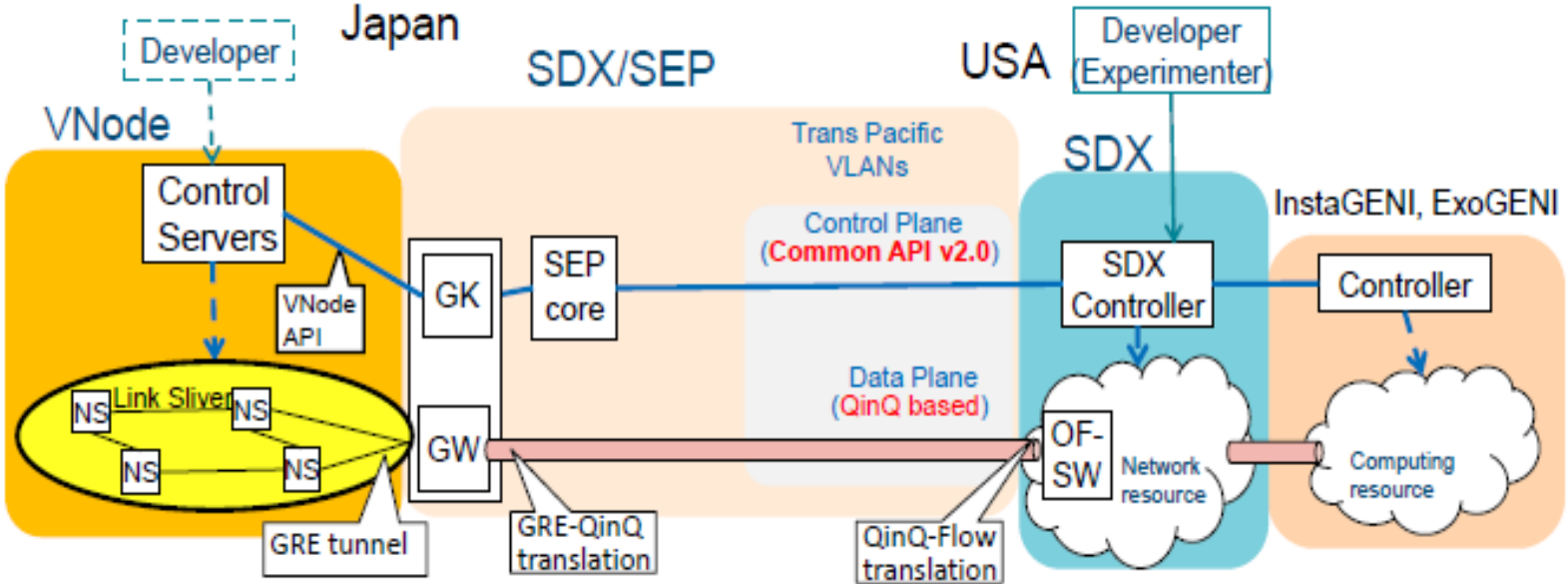


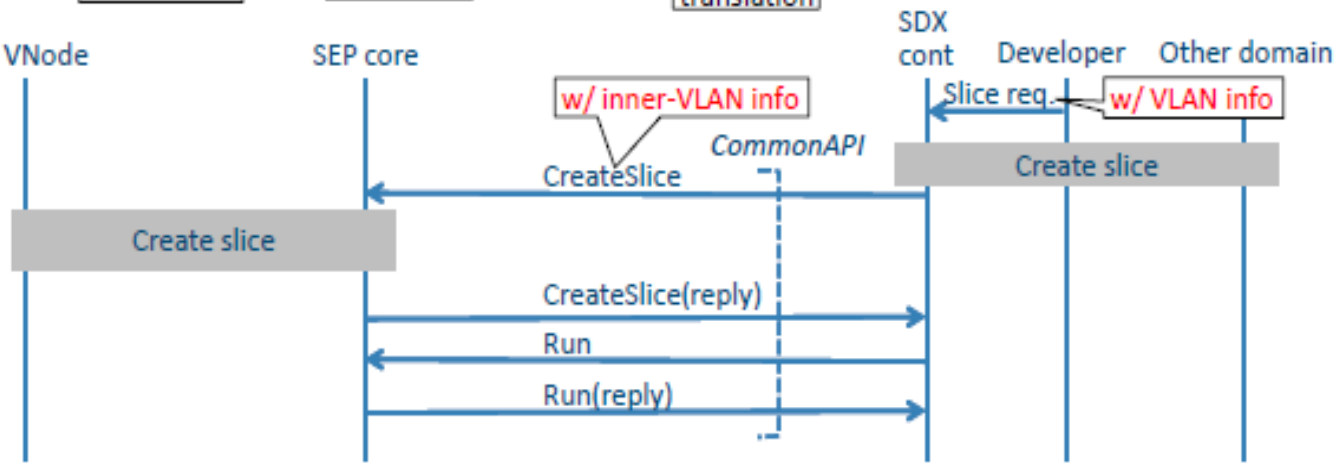
A Persistence Testbed Beyond SC14

Global Software-Defined Dynamic Circuits for Data Intensive Science
(PhEDEx – ANSE – PANDA - OpenDayLight)

vNode/SEP & StarLight Inter-SDX federation 2014

Multi-architecture Federation





© KDDI R&D Laboratories Inc, HITACHI LTD, The University of Tokyo

StarLight Software Defined Networking Exchange (SDX)

- The StarLight SDX Will Provide The Services, Architecture, and Technologies Designed To Provide Scientists, Engineers, and Educators With Highly Advanced, Diverse, Reliable, Persistent, and Secure Networking Services, Enabling Them to Optimally Access Resources in North America, South America, Asia, South Asia (including India), Australia, New Zealand, Europe, the Middle East, North Africa, And Other Sites Around the World.
- The StarLight SDX Initiative Undertakes Continued innovation and Development of Advanced Networking Services and Technologies.
- Potential For Providing Federation-as-a-Service

www.startap.net/starlight

iCAIR

STARLIGHT℠