

## **The StarLight International Software-Defined Network Exchange (SDX): Emerging Innovations In Global Interdomain SDN Services, Architecture, Capabilities and Technologies**

Joe Mambretti, Director, ([j-mambretti@northwestern.edu](mailto:j-mambretti@northwestern.edu))

International Center for Advanced Internet Research ([www.icaair.org](http://www.icaair.org))

Northwestern University

Director, Metropolitan Research and Education Network ([www.mren.org](http://www.mren.org))

Co-Director, StarLight, PI-iGENI, PI-OMNINet ([www.startap.net/starlight](http://www.startap.net/starlight))

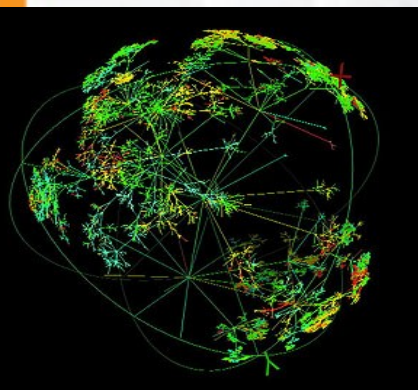
Co-PI Chameleon ([www.chameleoncloud.org](http://www.chameleoncloud.org))

Jim Chen, Associate Director, International Center for Advanced Internet Research  
([www.icaair.org](http://www.icaair.org))

Northwestern University

GLIF Technical Working Group Meeting

March 26-27, 2015



- GENI StarLight SDX Project Goal: To Provide Implementation Of Key Software and Hardware Components of Layer 2 SDN/OpenFlow Exchange Between GENI L2 Network Resources and Other Research Networks.
- Provide Tools for Experimenters To Request and Receive Resources From the Exchange That Are Fully Integrated With GENI Standard Interfaces Such As the GENI Clearinghouse, the GENI AM API, GENI Stitching AMs, and the GENI Commercial Software Defined Exchange Point.
- Integrate Provided GENI tools With Experimenter Tools From Other Participating Networks.
- Demonstrate One Functioning Exchange With at Least Two Research Network Partners, Two Data-Intensive Science Campuses, and Multiple Experimenters On Multiple Participating Layer 2 Networks.

- With the Increasing Deployment of SDN In Production Networks, the Need for an SDN Exchange (SDX) Has Been Recognized.
- Current SDN Architecture Is Single Domain Centralized Controller Oriented – Many SDN Islands Being Created
- Many Motivations Exist for SDXs – e.g., Connecting SDN Islands, Enhanced Control Over Specific Traffic Flows, Resource Optimization, Network Function Virtualization --  
\*\*New Services\*\*
- Required Capabilities: Multi-Domain Distributed SDN Resource Discovery, Signaling Provisioning, Operations, and Fault Detection and Recovery.

- Transitions
  - From Static L3 BGP Exchanges
  - To Multi-Layer Multi-Services Exchanges
  - Foundations
    - Grid Networking Federation
    - OGF - GLIF NIS CS
    - Global Environment For Network Innovations (GENI)
  - Enabling Multiple Layers/Services and Hybrids
  - Deep Visibility Into All Flows
  - Direct Control Over All Flows
  - Significant New Capabilities for Customized Exchanges!
  - Including For Domain Sciences...





Compilation By Maxine Brown

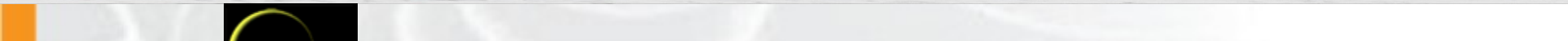


Sponsored by the National Science Foundation

March 26, 2015



Many Ye  
Service In  
nable Pat





# Automated GOLE Fabric



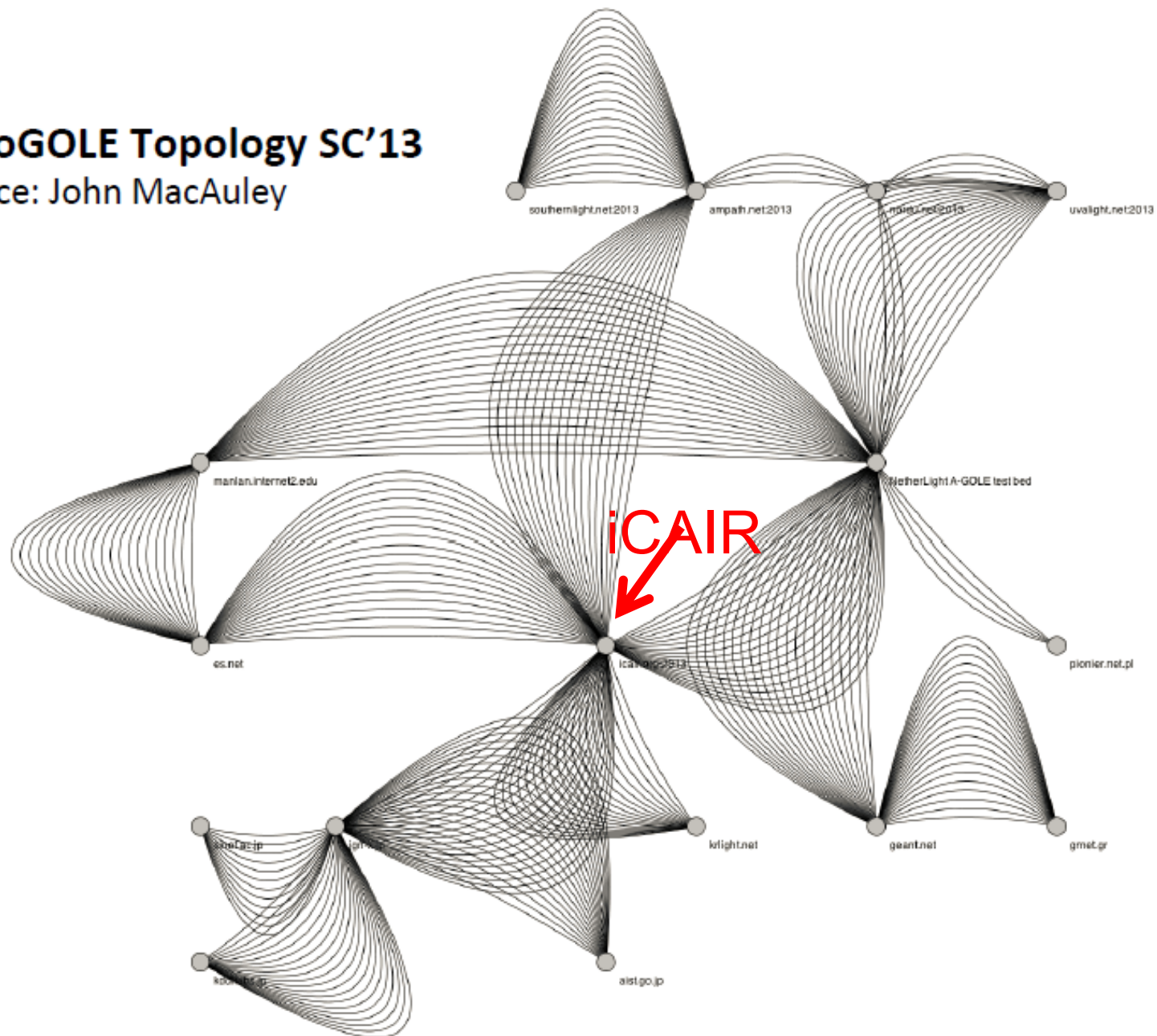






# AutoGOLE Topology SC'13

Source: John MacAuley



# Tasks/Goals For 2014 Expansion To Incorporate SDN/OF

## Work items 2014

| Item                             | Description   | Due     | Leading organization   |
|----------------------------------|---|---------|--|
| Authentication / Authorization   | Creating a AAI framework that allows secure setup of services   | TNC2014 | SURFnet<br>(Hans Trompert)   |
| Topology Exchange                | Creating a mechanism that exchanges topology descriptions of GOLEs automatically  | SC'14   | ESnet, UvA<br>(Chin Guok, Miroslav Zivkovic)                             |
| Retagging capabilities           | Describing what's necessary to implement retagging capabilities inside the AutoGOLE fabric – also creating a plan for implementing  | SC'14   | Group effort   |
| SDN/OpenFlow inside the AutoGOLE | It's foreseen that AutoGOLE NRMs could be talking OpenFlow to actual hardware. This item results in deployment of an OpenFlow controller speaking NSlv2 inside the AutoGOLE | Q4      | iCAIR<br>(Jim Chen, Joe Mambretti)                                       |
| Operational items                | Creating concepts on strengthening operations, implementing these   | Q4      | Tanguy Courouarn to look for someone to lead (uniform) perational issues |







StarLight is an experimental  
optically based infrastructure and  
proving ground for network  
services optimized for  
high-performance applications

Multiple  
10GE+100 Gbps  
StarWave  
Multiple 10GEs  
Over Optics –  
World’s “Largest”  
10G/100G Exchange  
First of a Kind  
Enabling Interoperability  
At L1, L2, L3



View from StarLight

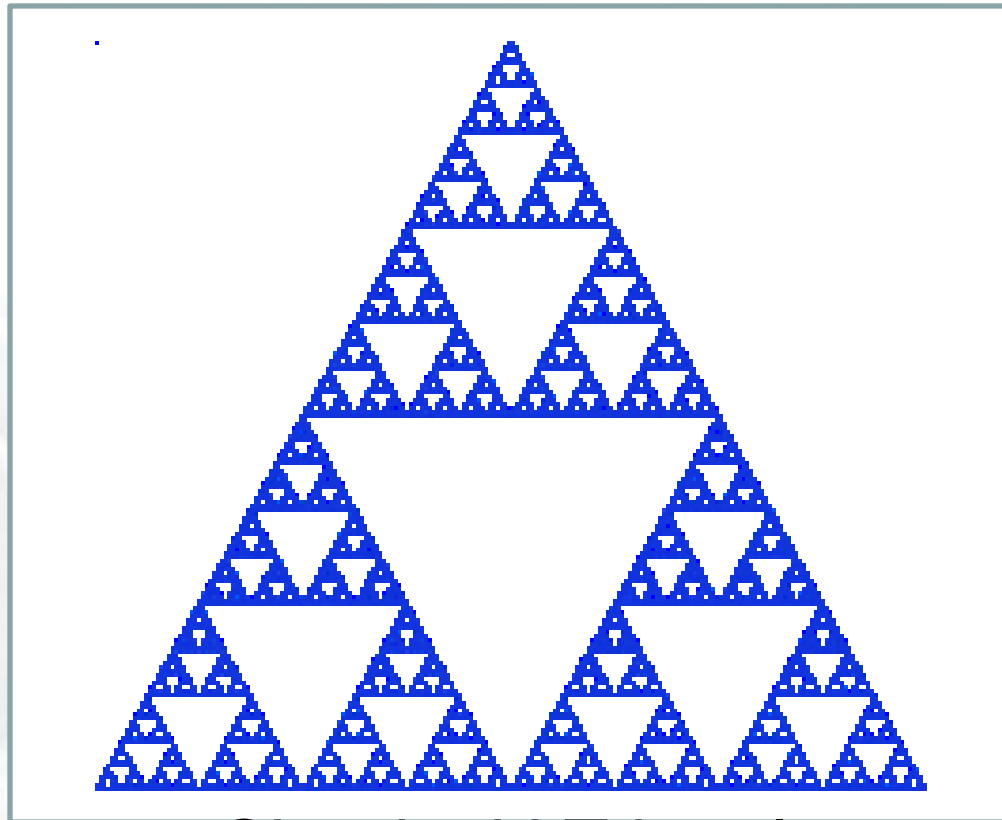


Abbott Hall, Northwestern University's  
Chicago Campus

Currently Undertaking Major Infrastructure Renovation

- At Highest Level, Appears As a Very Large Scale Virtual Switch
- Resources Can Be Segmented/Partitioned
- Architecture Is Informed By NSI, GENI, Related International Network Testbeds, Emerging Concepts
- Based On An Underlying Foundation of Programmable Resources
- Includes Specialized APIs for Provisioning
- This SDX is Being Showcased Through GLIF, GENI and SC14 Demonstrations

# SDX As Recursive Virtual Switch

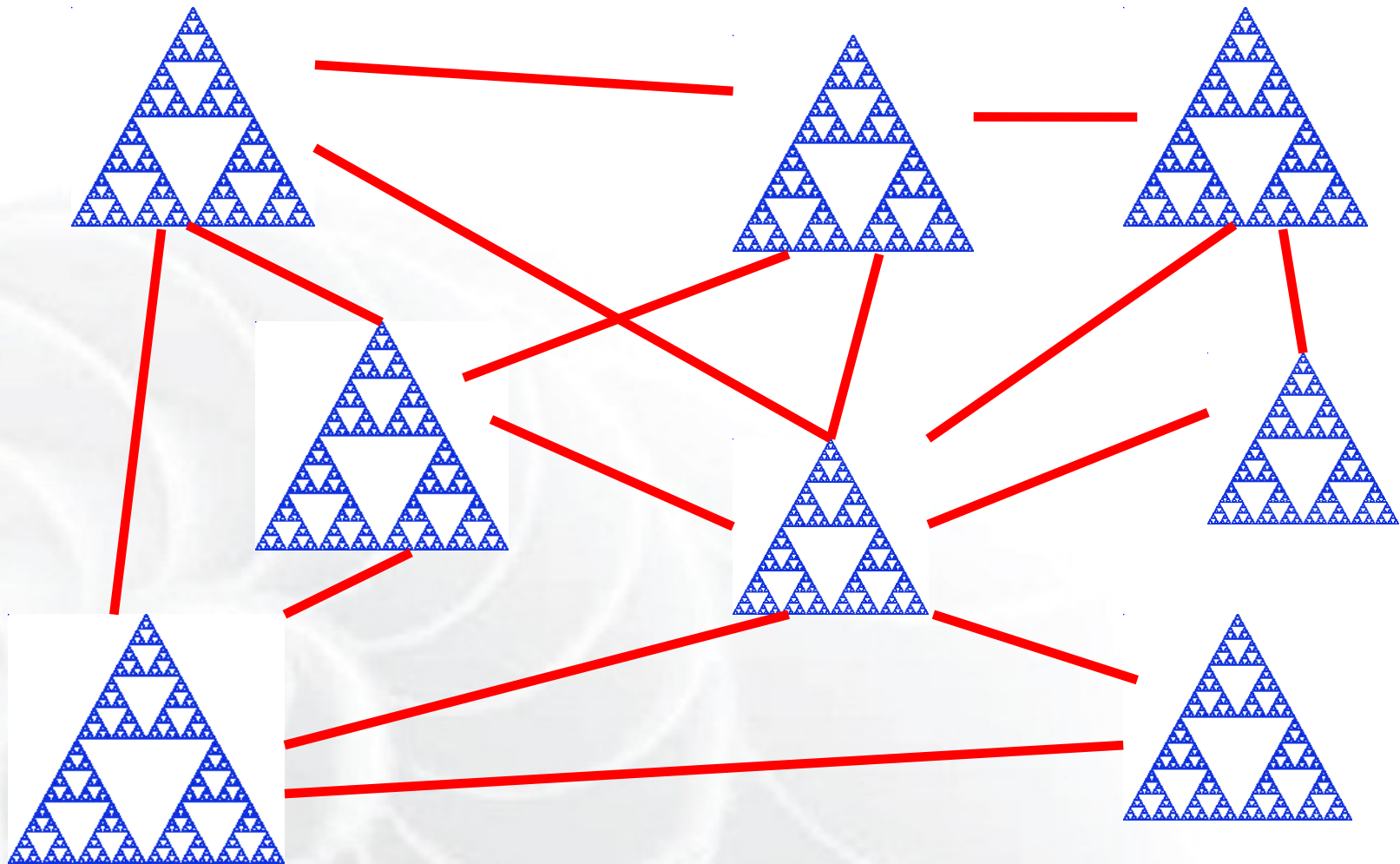


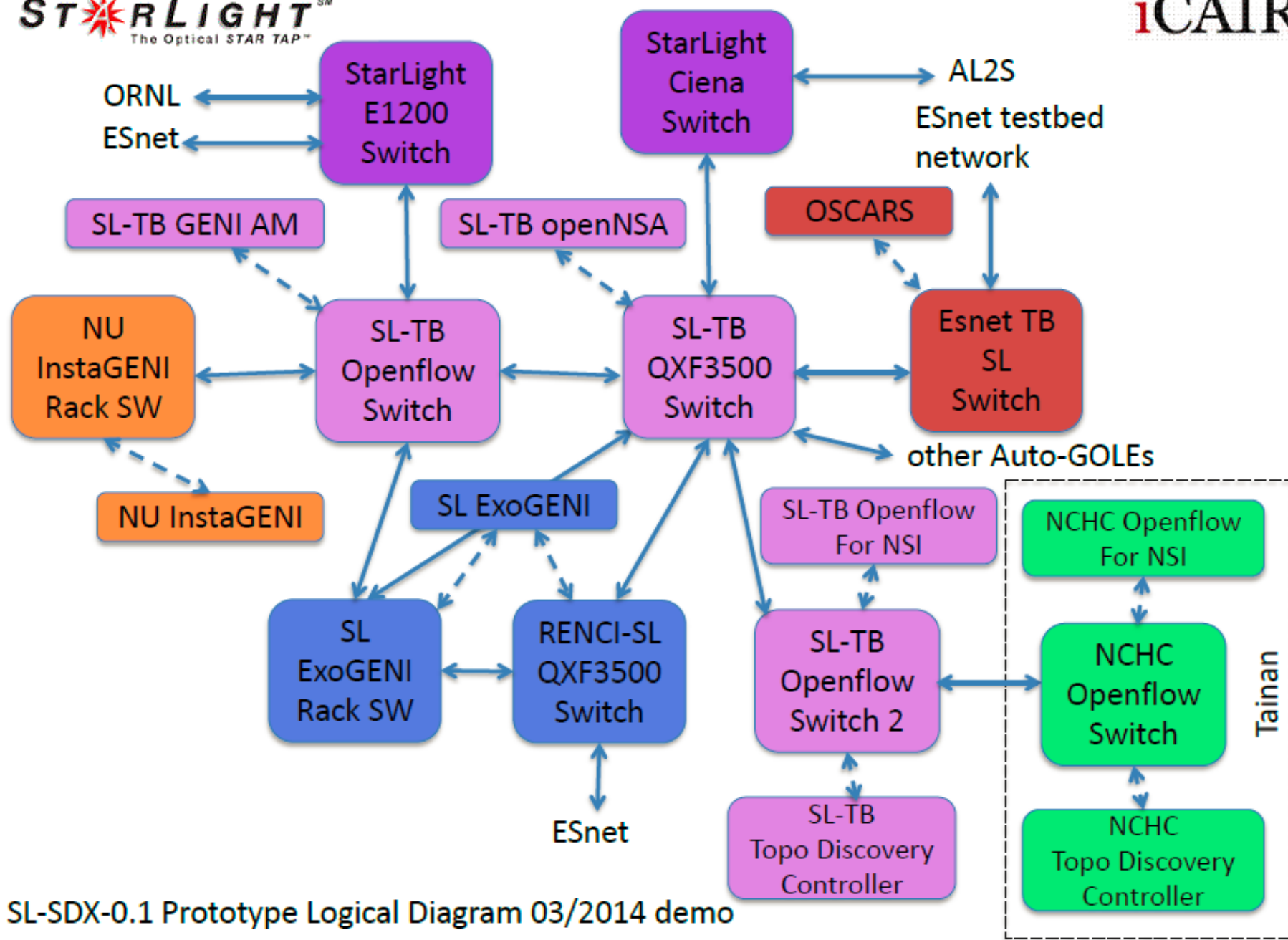
Sierpinski Triangle

Unlimited Number of Customized  
Virtual Switches Within Macro  
Virtual Switch

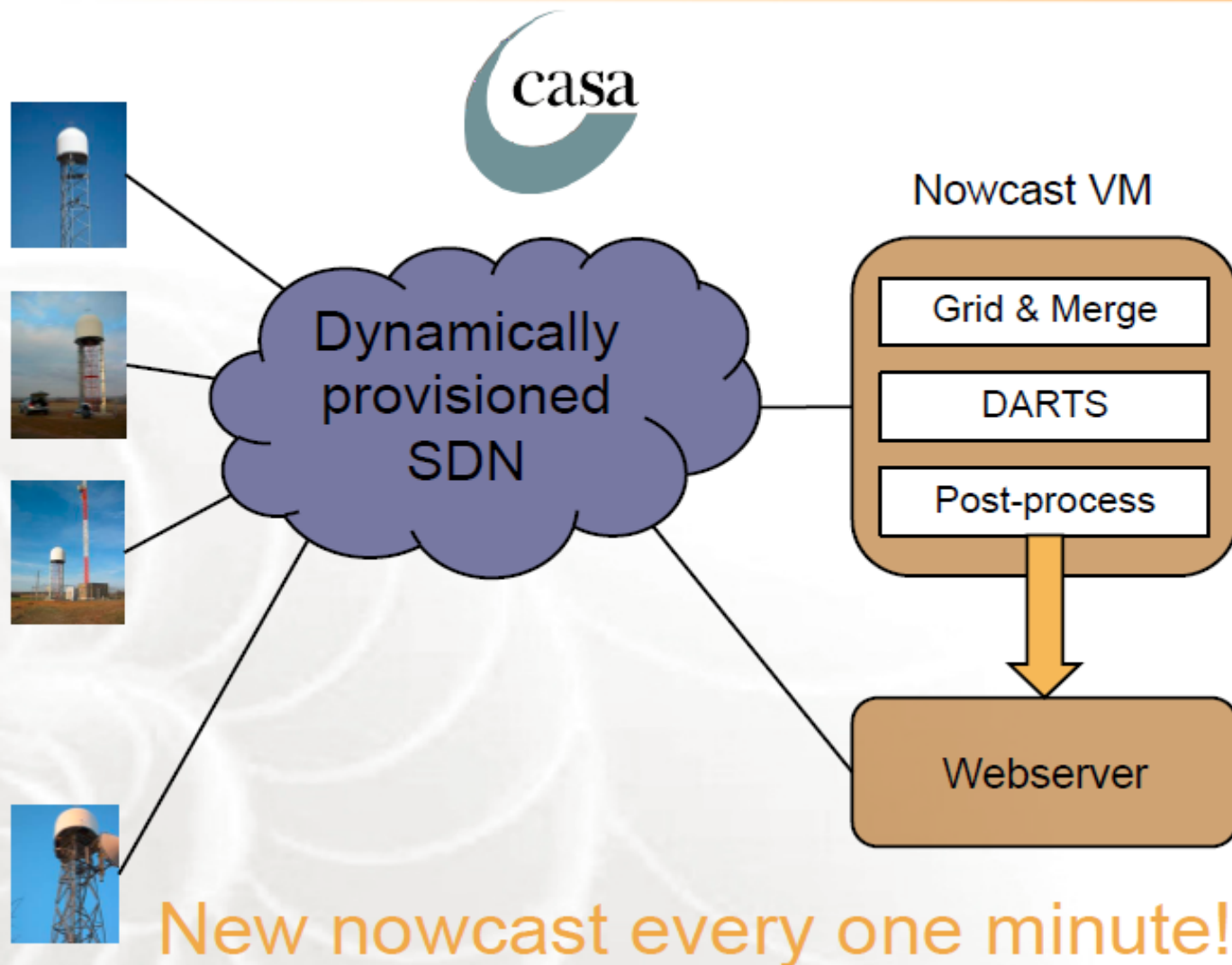


# GLIF Based On SDXs Supporting Slice Exchanges





SL-SDX-0.1 Prototype Logical Diagram 03/2014 demo

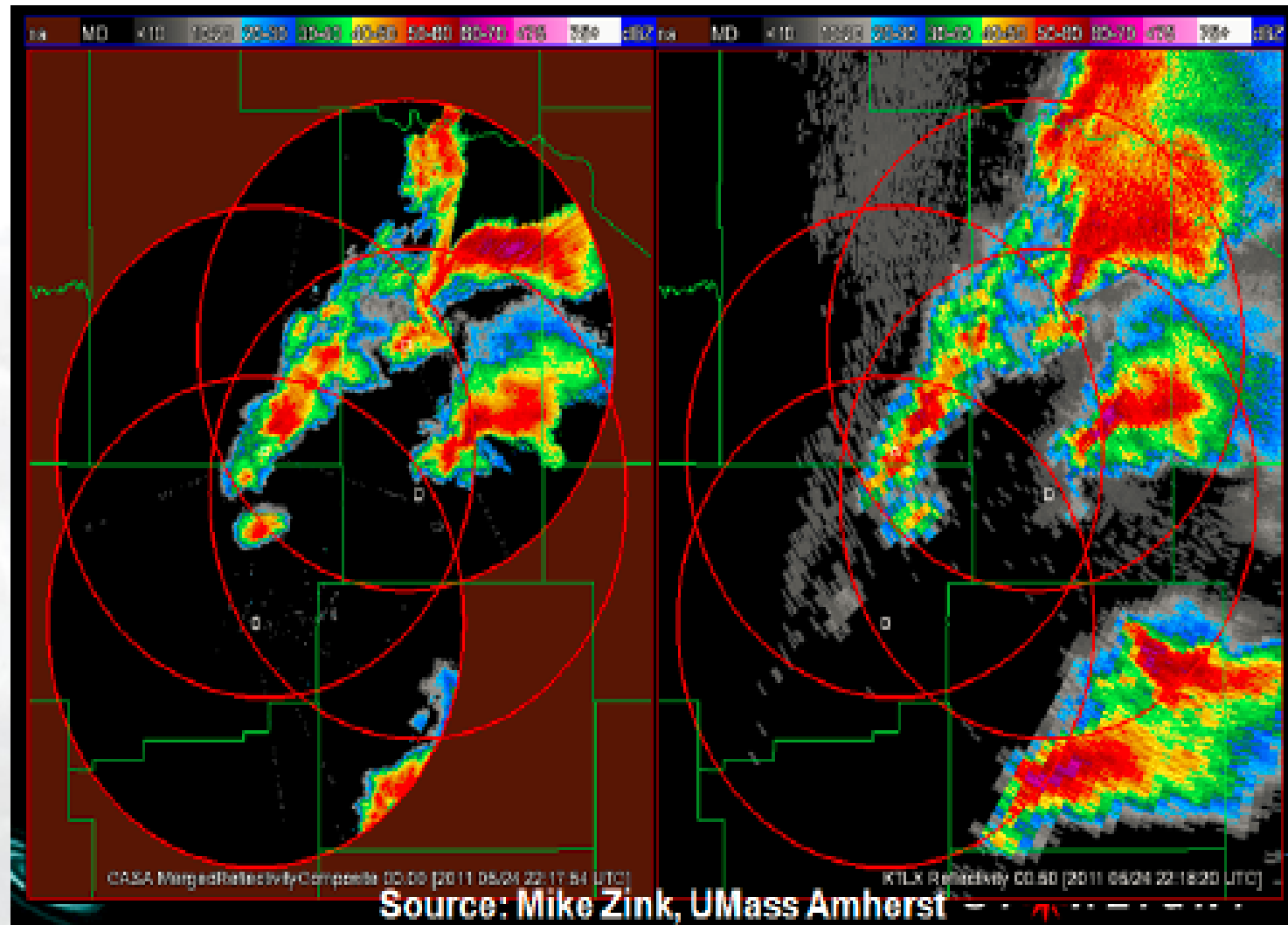


New nowcast every one minute!



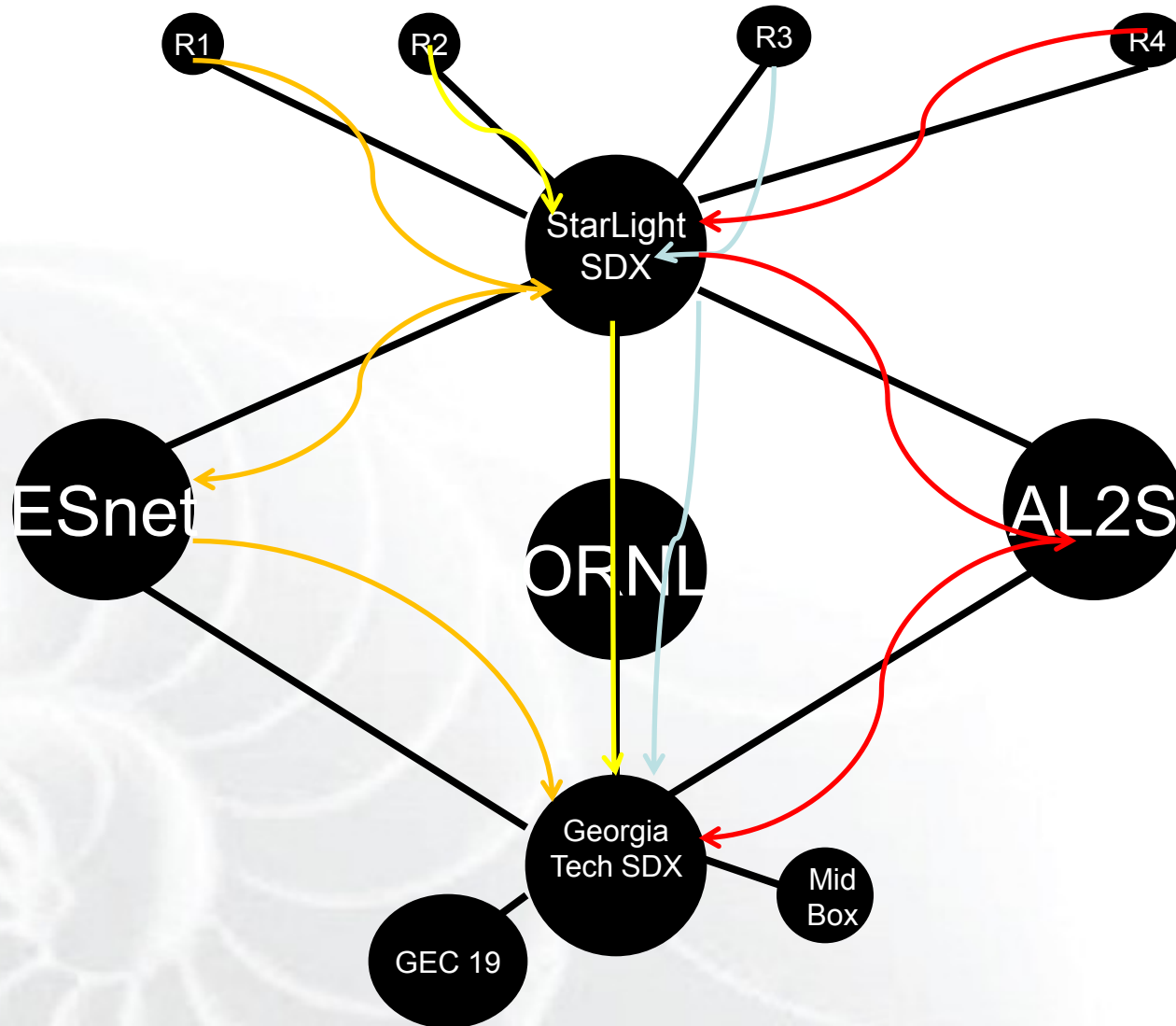
# Comparison With Existing System

*Slide by Mike Zink, UMass Amherst*



# GENI SDX Demo Scenario 2

Simulated  
Radar (4)



# SDX StarLight ↔ NetherLight

**STARLIGHT™**  
The Optical STAR TAP™

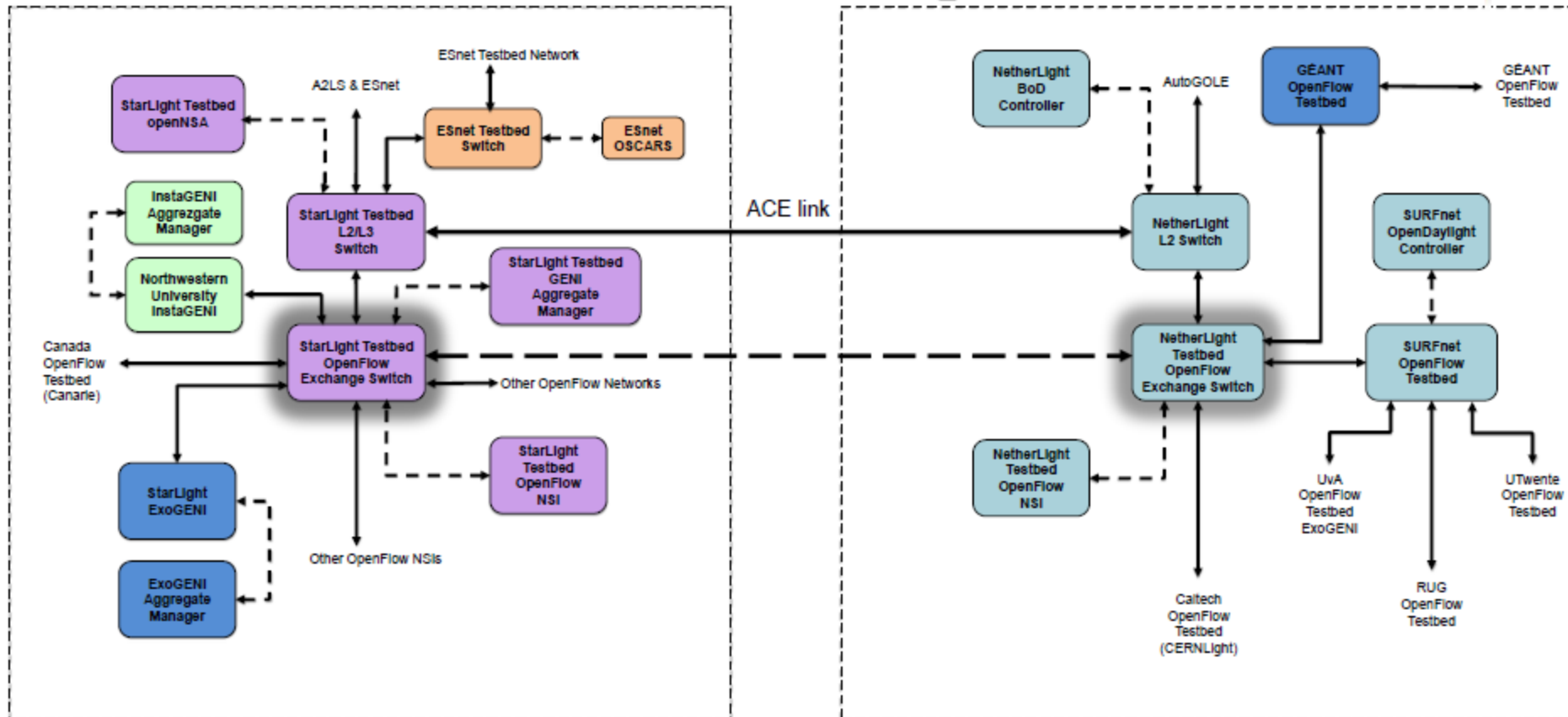
**StarLight, Chicago**

**iCAIR**

**SURF NET**

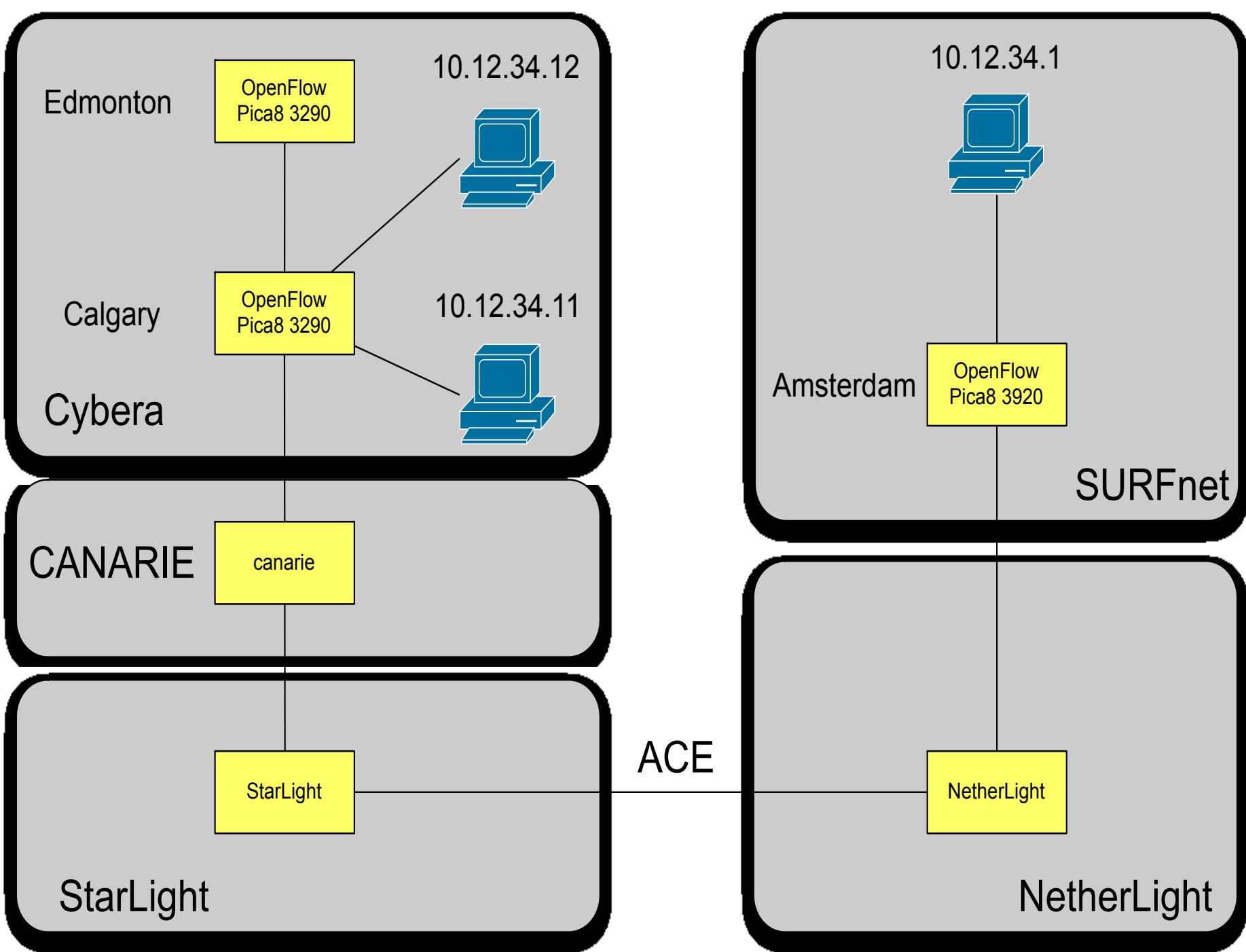
**NetherLight, Amsterdam**

**NL** Light



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





# International Software-Defined Network Exchanges (iSDXs): A Demonstration of Global Capabilities

Joe Mambretti, Jim Chen, Fei Yeh

International Center for Advanced Internet Research

Northwestern University, USA

Mike Zink, Divyashri Bhat

University of Massachusetts, Amherst, USA

Ronald Van der Pol

Surfnet, Netherlands

Grace Lee, WunYuan Huang, Te-Lung Liu

NARLabs, National Center for High Performance Computing, Taiwan

Thomas Tam, Herve Guy,

CANARIE, Canada

Alex Valiushko, John Shillington,

Cybera, Canada

Buseung Cho, KISTI

Republic of Korea

Michiaki Hayashi, KDDI Labs, Japan

Toshiaki Tarui, Hitachi, Japan

Aki Nakao, University of Tokyo, Japan

Steve Cotter, T. Charles Yun, Jamie Curtis, Andrej Ricnik

REANNZ, New Zealand

Josh Bailey, Google, New Zealand

Artur Binczewski Belter Bartosz Miłosz Przywecki Piotr Rydlichowski

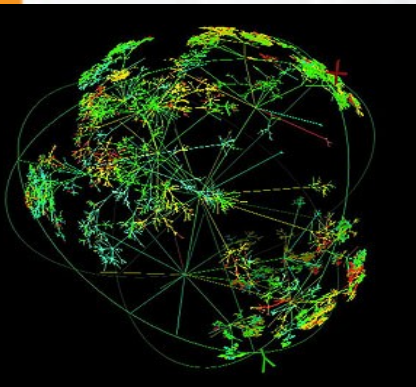
Poznan Supercomputing and Networking Center, Poland

Russ Clark, Georgia Tech, USA

Global LambdaGrid Workshop

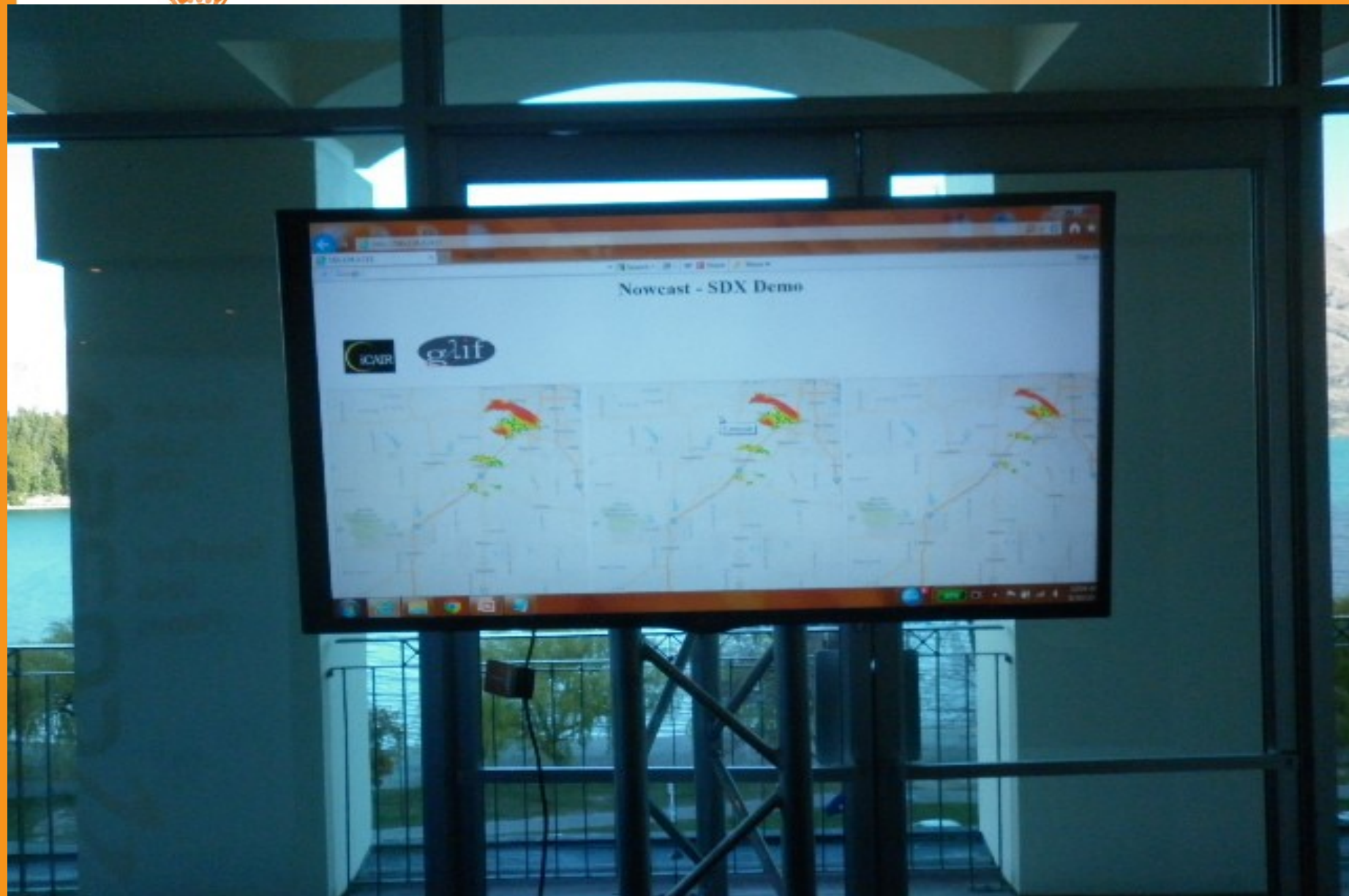
Queenstown, New Zealand

September 30-October 1, 2014



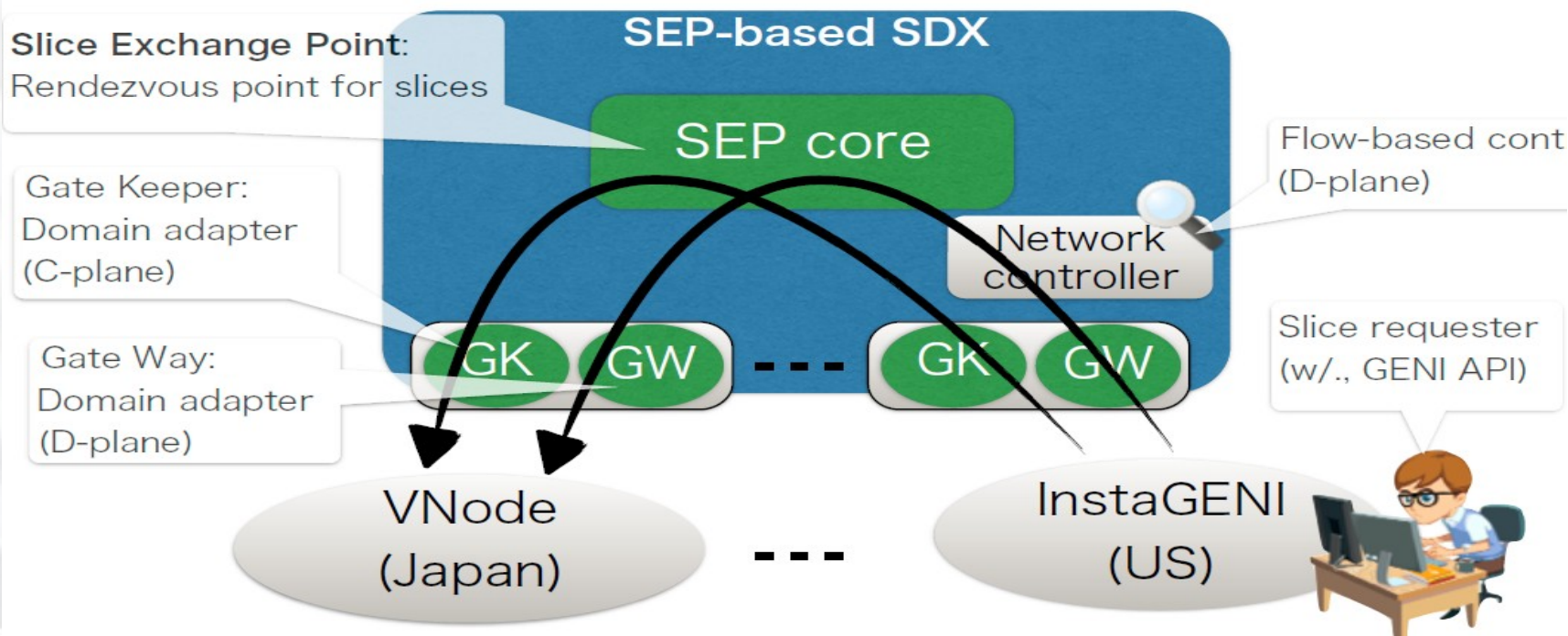






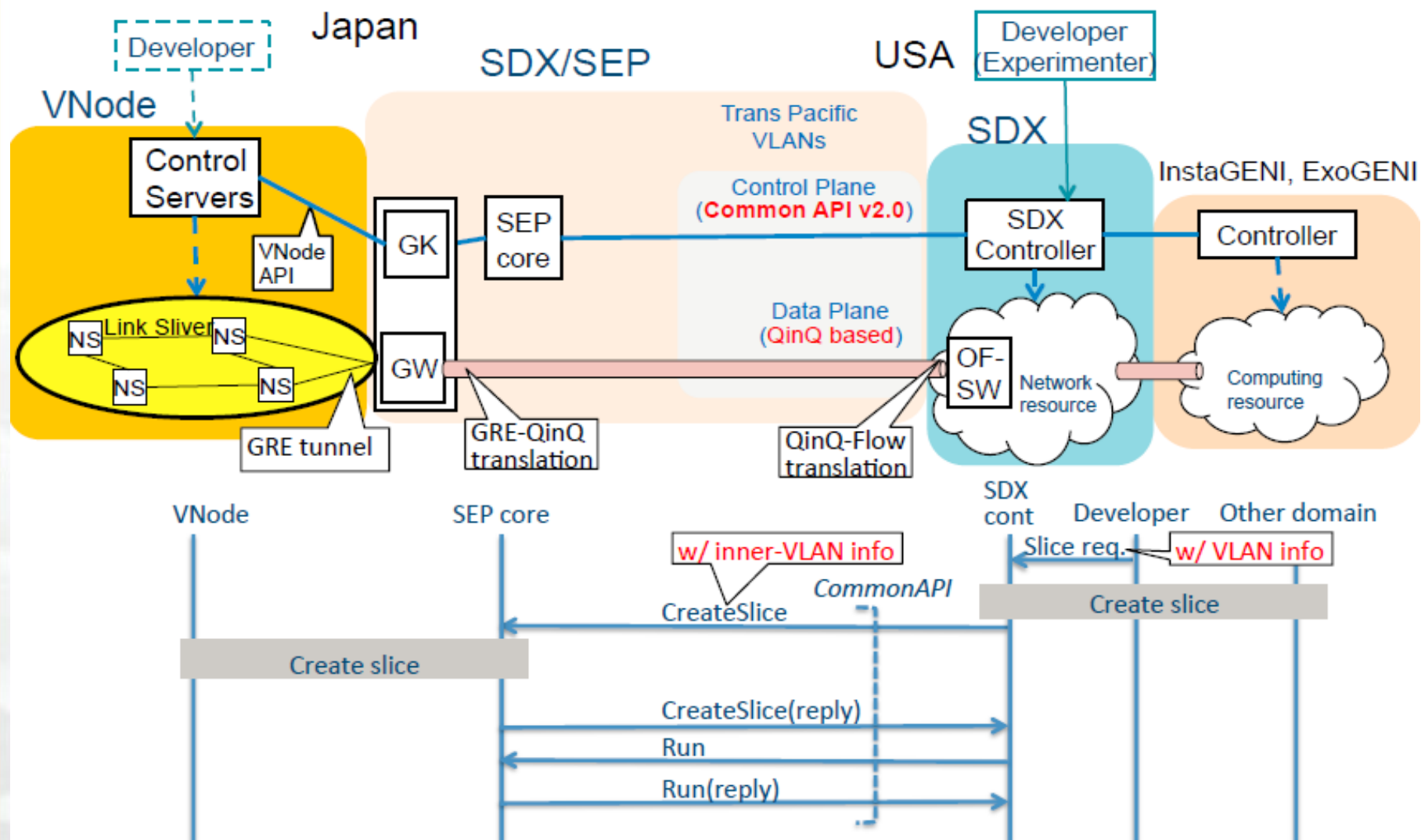


## Slice Exchange Architecture



# Inter-SDX federation for GEC21

## Multi-architecture Federation



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

2

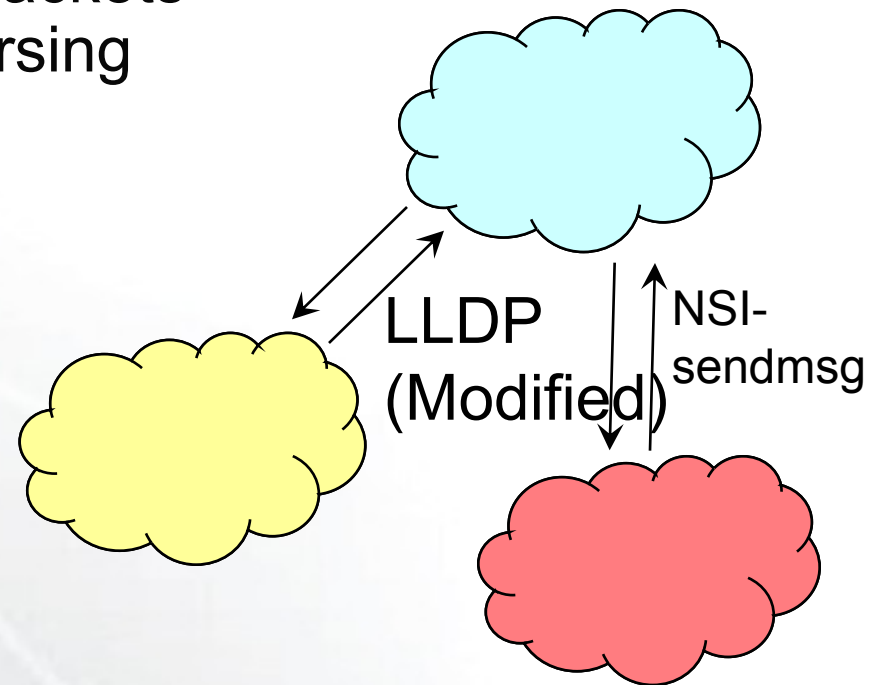


- **Modified LLDP** (also **NDP** and **BDDP**)  
Was Used To Exchange Neighbor  
Controller Information Among SDN  
Domains Automatically

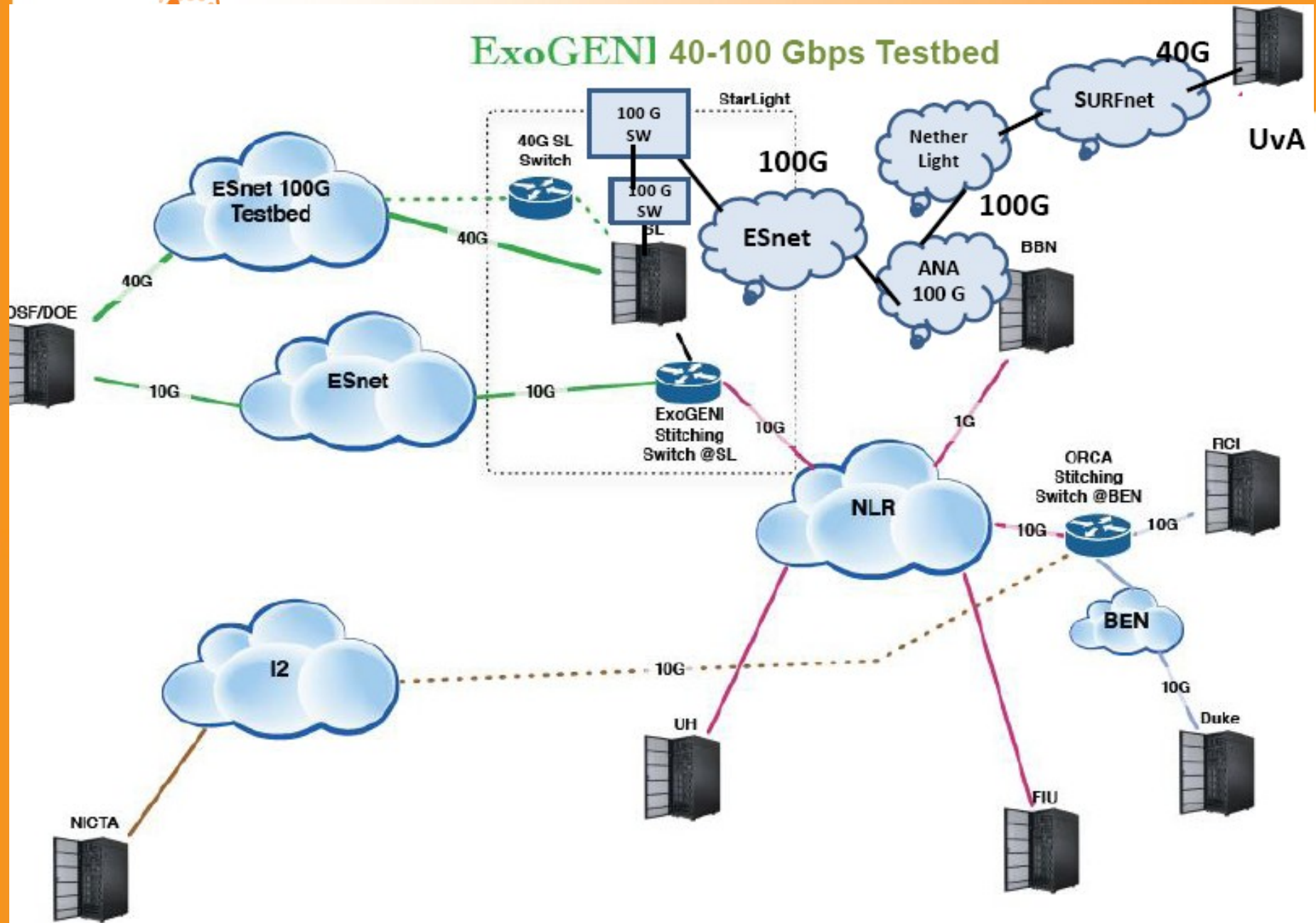
- However, These LLDP-like Packets  
May Be Filtered When Traversing  
Domain Boundaries

- For That Case, We Utilize **NSI**  
To Share Information Among  
NSAs

- A New **sendmsg** Service Is  
Implemented
  - In Addition To Topology  
Discovery, Service  
Announcements Could Also  
Be Delivered



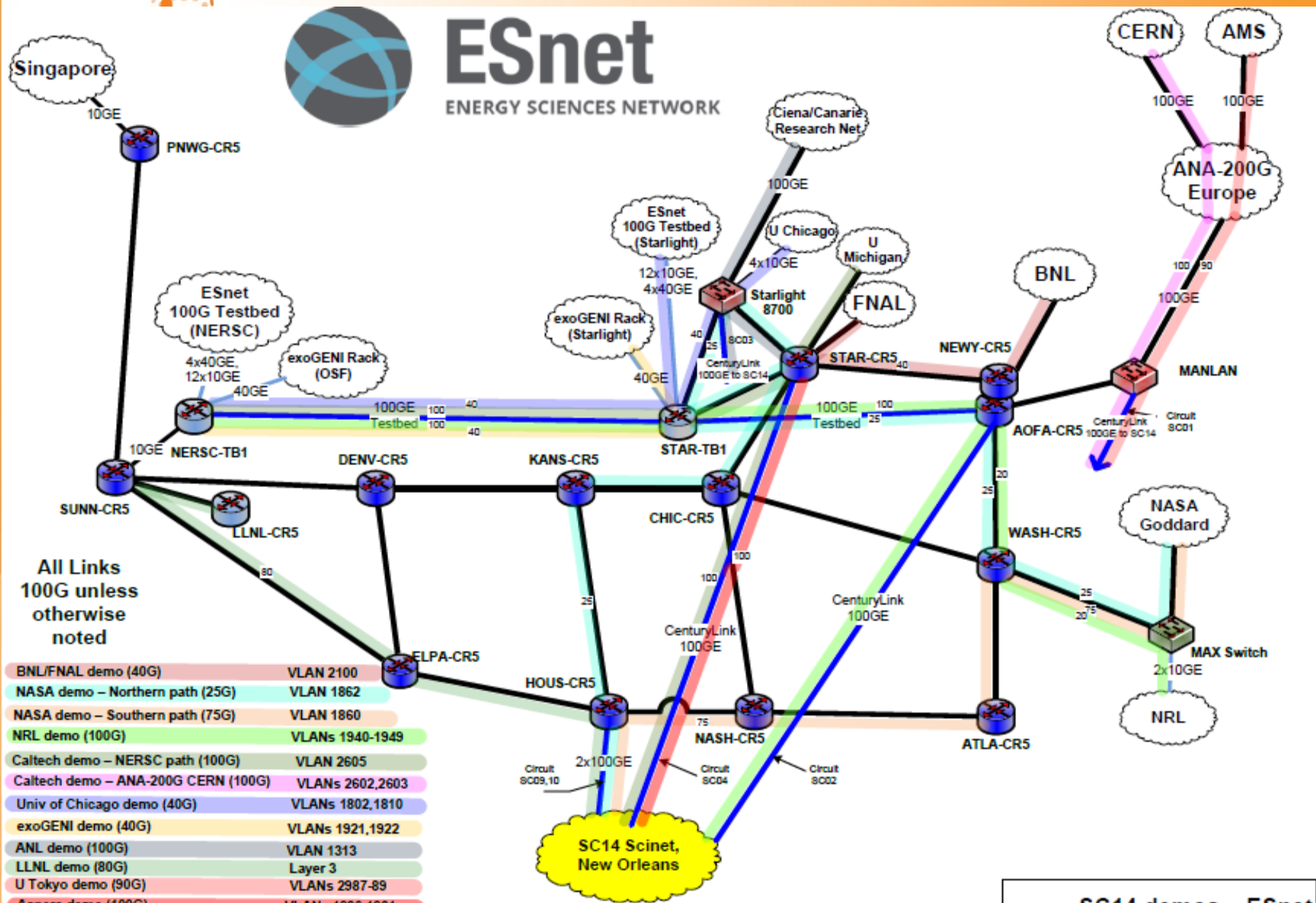
# ExoGENI 40-100 Gbps Testbed





# ESnet

ENERGY SCIENCES NETWORK



## SC14 demos – ESnet

Brian Tierney, ESnet 11/10/2014

FILENAME

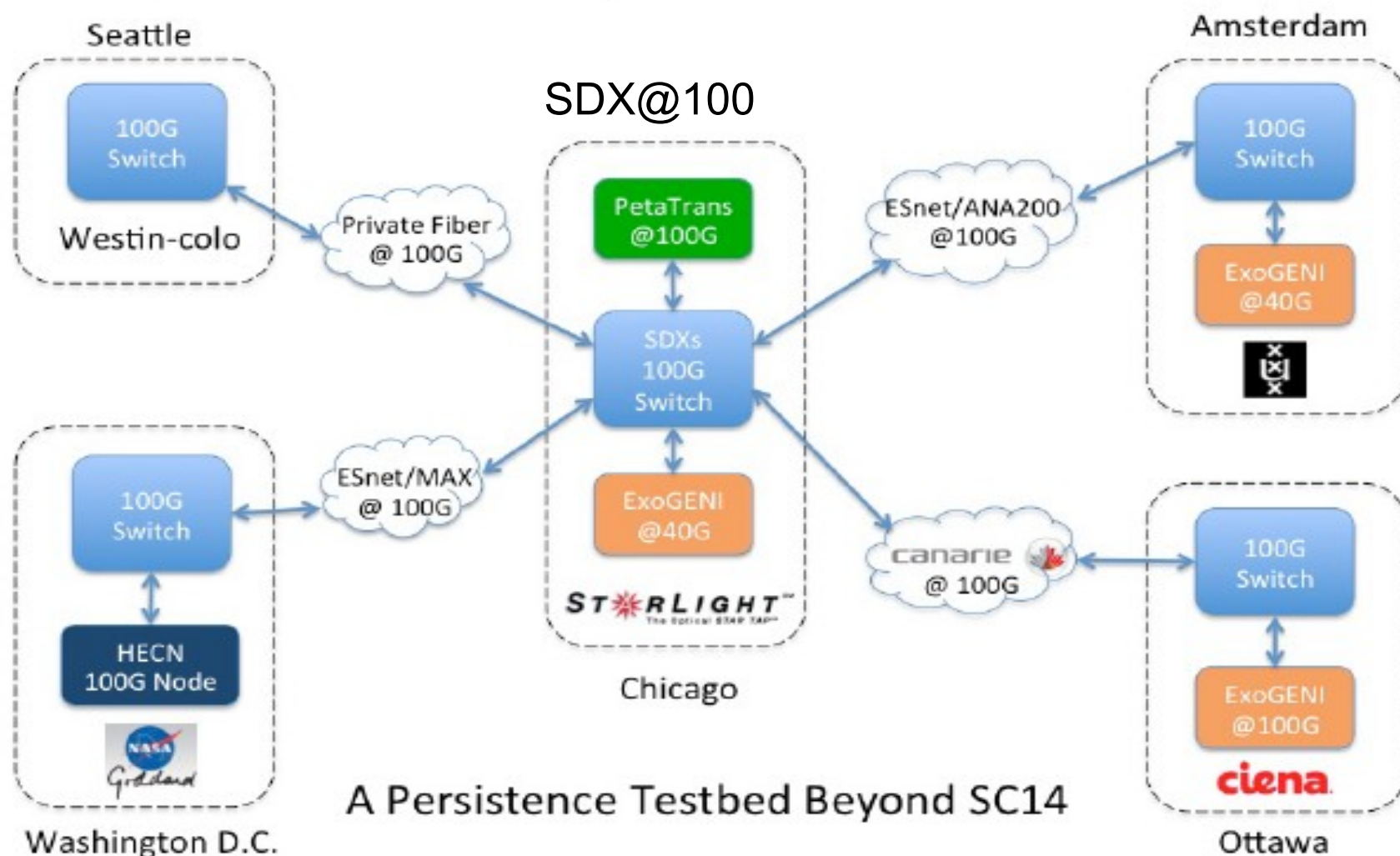
SC14-DEMOS-V14.VSD







# PetaTrans: Peta Byte Science Data Transfer



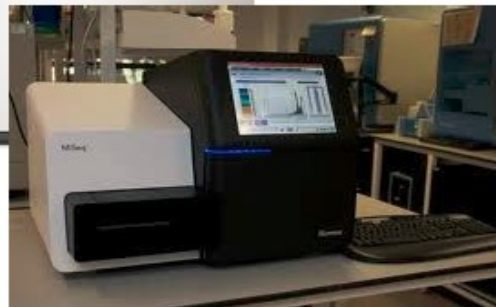
Global Software-Defined Dynamic Circuits for Data Intensive Science  
(PhEDEx – ANSE – PANDA – OpenDayLight)

- Haploid Human Genome (23 Chromosomes) = Approximately **3.2 Billion** Bases Long
- Contains 20,000–25,000 Distinct Protein-Coding Genes.
- 6 Billion Base Pairs Per Diploid Cell.
- Kilobase (kb): Molecular Biology Measurement Unit Equal To 1000 Base Pairs of Deoxyribonucleic Acid (DNA) or Ribonucleic acid (RNA).
- Information Flow : From DNA Through RNA To Proteins
- Soon: 85 Petabytes of Data
- 1<sup>st</sup> Sequence: 13 Years
- Today: < 30 Minutes

## Sequencers



©2012 Illumina, Inc. All rights reserved.



Source: Don Pruess





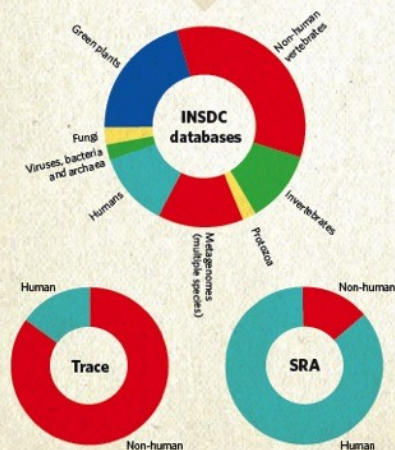
## THE SEQUENCE EXPLOSION

At the time of the announcement of the first drafts of the human genome in 2000, there were 8 billion base pairs of sequence in the three main databases for 'finished' sequence: GenBank, run by the US National Center for Biotechnology Information; the DNA Databank of Japan; and the European Molecular Biology Laboratory (EMBL) Nucleotide Sequence Database. The databases share their data regularly as part of the International Nucleotide Sequence Database Collaboration (INSDC). In the subsequent first post-genome decade, they have added another 270 billion bases to the collection of finished sequence, doubling the size of the database roughly every 18 months. But this number is dwarfed by the amount of raw sequence that has been created and stored by researchers around the world in the Trace archive and Sequence Read Archive (SRA).

See Editorial, page 649, and human genome special at [www.nature.com/humangenome](http://www.nature.com/humangenome)

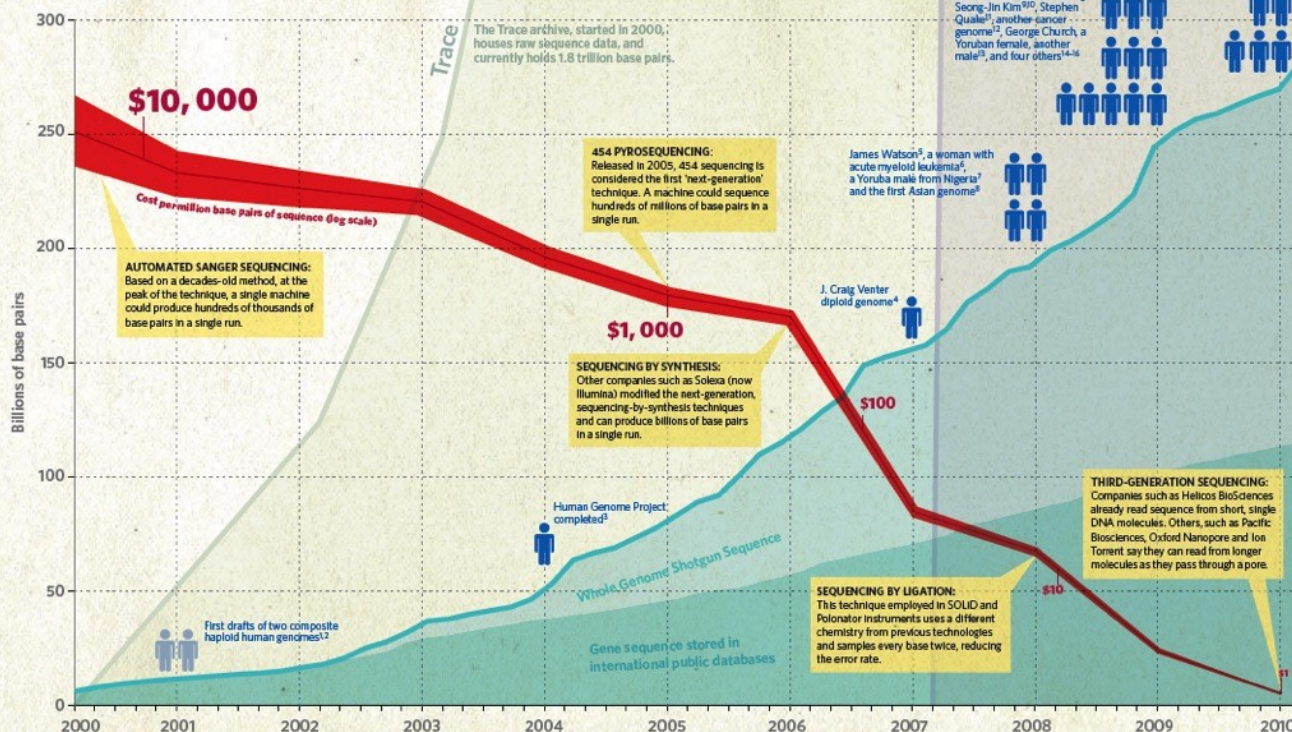
### DNA SEQUENCES BY TAXONOMY

**International Nucleotide Sequence Database Collaboration:** The main repositories of 'finished' sequence span a wide range of organisms, representing the many priorities of scientists worldwide.



**Trace Archive:** Developed to house the raw output of high-throughput sequencers built in the late 1990s, the trace archive spans a wide range of taxa.

**Sequence Read Archive:** Houses raw data from next-generation sequencers. Dominated by human sequence, including multiple coverage for more than 170 people.



### HOW MANY HUMAN GENOMES?

The graphic shows all published, fully sequenced human genomes since 2000, including nine from the first quarter of 2010. Some are resequencing efforts on the same person and the list does not include unpublished completed genomes.

- Venter, J. C. et al. *Science* **291**, 1304-1351 (2000).
- International Human Genome Sequencing Consortium *Nature* **409**, 860-921 (2000).
- International Human Genome Sequencing Consortium *Nature* **431**, 931-945 (2004).
- Levy, S. et al. *PLoS Biol.* **5**, e254 (2007).
- Wheeler, D. A. et al. *Nature* **452**, 872-876 (2008).
- Ley, T. J. et al. *Nature* **456**, 66-72 (2008).
- Bentley, D. R. et al. *Nature* **456**, 53-59 (2008).
- Wang, J. et al. *Nature* **456**, 60-65 (2008).
- Ahn, S.-M. et al. *Genome Res.* **19**, 1622-1629 (2009).
- Kim, J.-I. et al. *Nature* **460**, 1011-1015 (2009).
- Pushkarev, D., Neff, N. F. & Quake, S. R. *Nature Biotechnol.* **27**, 847-850 (2009).
- Mardis, E. R. et al. *N. Engl. J. Med.* **10**, 1058-1066 (2009).
- Lupski, J. R. et al. *N. Engl. J. Med.* **10**, 1058-1066 (2009).
- Drmanac, R. et al. *Science* **327**, 78-81 (2009).
- McKernan, K. J. et al. *Genome Res.* **19**, 1527-1541 (2009).
- Pleasance, E. D. et al. *Nature* **463**, 191-196 (2010).
- Pleasance, E. D. et al. *Nature* **463**, 184-190 (2010).
- Clark, M. J. et al. *PLoS Genet.* **6**, e1000832 (2010).
- Rasmussen, M. et al. *Nature* **463**, 757-762 (2010).
- Schuster, S. C. et al. *Nature* **463**, 943-947 (2010).
- Lupski, J. R. et al. *N. Engl. J. Med.* **10**, 1056/NEJMc0908094 (2010).
- Roach, J. C. et al. *Science* **327**, 1010/10126/science.1186802 (2010).





# Beyond Today's Internet Experiencing a Smart Future



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

Robert Grossman – University of Chicago  
Joe Mambretti – Northwestern University  
Piers Nash – University of Chicago  
Jim Chen – Northwestern University  
Allison Heath – University of Chicago



THE UNIVERSITY OF  
**CHICAGO**  
MEDICINE

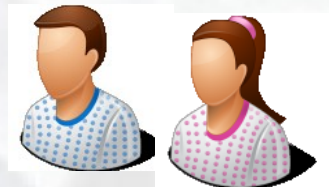


**NORTHWESTERN**  
UNIVERSITY

# Precision Medicine Enabled By Precision Networks

- Precisely match treatments to patients and their specific disease
- Genomic data promises optimal matching.
- 1.7 million cancer cases diagnosed in America each year.
- A single RNA-seq file is 10-20 GB, Whole genome raw data files are > 100 GB.
- Analysis has become the bottleneck and data size is an issue.
  - 2,000,000 genomes  $\approx$  1 Exabyte (1,000,000,000,000 MB)
  - Cost to sequence 1 genome less than \$5,000 and falling fast.
  - Cost to analyze 1 genome is approx. \$100,000 and rising.
- A key step towards Algorithm-assisted Personalized medicine is building Data Commons/Cloud analytics and the \*Programmable\* Networks & Communication Exchanges (SDXs) for high performance, flexible data transport.

# Future Vision: A Nationwide Virtual Comprehensive Cancer Center



Patients



Hospitals,  
Doctors



Cloud Computation  
Genomic Data Commons

Output: Data-Aware,  
Analytics-Informed  
Diagnosis,  
Prognosis,  
Optimal Treatment

## BIONIMBUS PROTECTED DATA CLOUD

Secure cloud services for the scientific community

### What is the Bionimbus PDC?

The Bionimbus Protected Data Cloud (PDC) is a collaboration between the Open Science Data Cloud (OSDC) and the IGSB (IGSB), the Center for Research Informatics (CRI), the Institute for Translational Medicine (ITM), and the University of Chicago Comprehensive Cancer Center (UCCCC). The PDC allows users authorized by NIH to compute over human genomic data from dbGaP in a secure compliant fashion. Currently, selected datasets from the The Cancer Genome Atlas (TCGA) are available in the PDC.

### How can I get involved?

- Apply for an Bionimbus PDC account and use the Bionimbus PDC to manage, analyze and share your data.
- Partner with us and add your own racks to the Bionimbus PDC (we will manage them for you).
- Help us develop the open source Bionimbus PDC software stack.

You can contact us at [info@opencloudconsortium.org](mailto:info@opencloudconsortium.org).

### How do I get started?

First, apply for an account. Once your account is approved, you can login to the console and get started. Support questions can be directed to [support@opensciencedatacloud.org](mailto:support@opensciencedatacloud.org).

Apply for the PDC Now

Login to the PDC Console

- Petabyte-scale, secure compliant biomedical cloud that interoperates with dbGaP controlled access data at NIH.





# Opportunity: Close Integration of Research Workflows and Foundation Networks

- Opportunity: Using GENI To Develop Innovative Techniques for Extremely Close Integration of Research WorkFlows and Dynamic Programmable Network Resources, Enabling Precision Networking
- Network Foundation Architecture: GENI + Innovative Customized Software Defined Networking Exchange (SDX)
- For This Demonstration: Specifically To Meet The Requirements of Bioinformatic Workflows

Data Repository A (West Coast)

Data Repository B  
(South)

**Required Resources (Data & Tools) Are  
Highly Distributed**

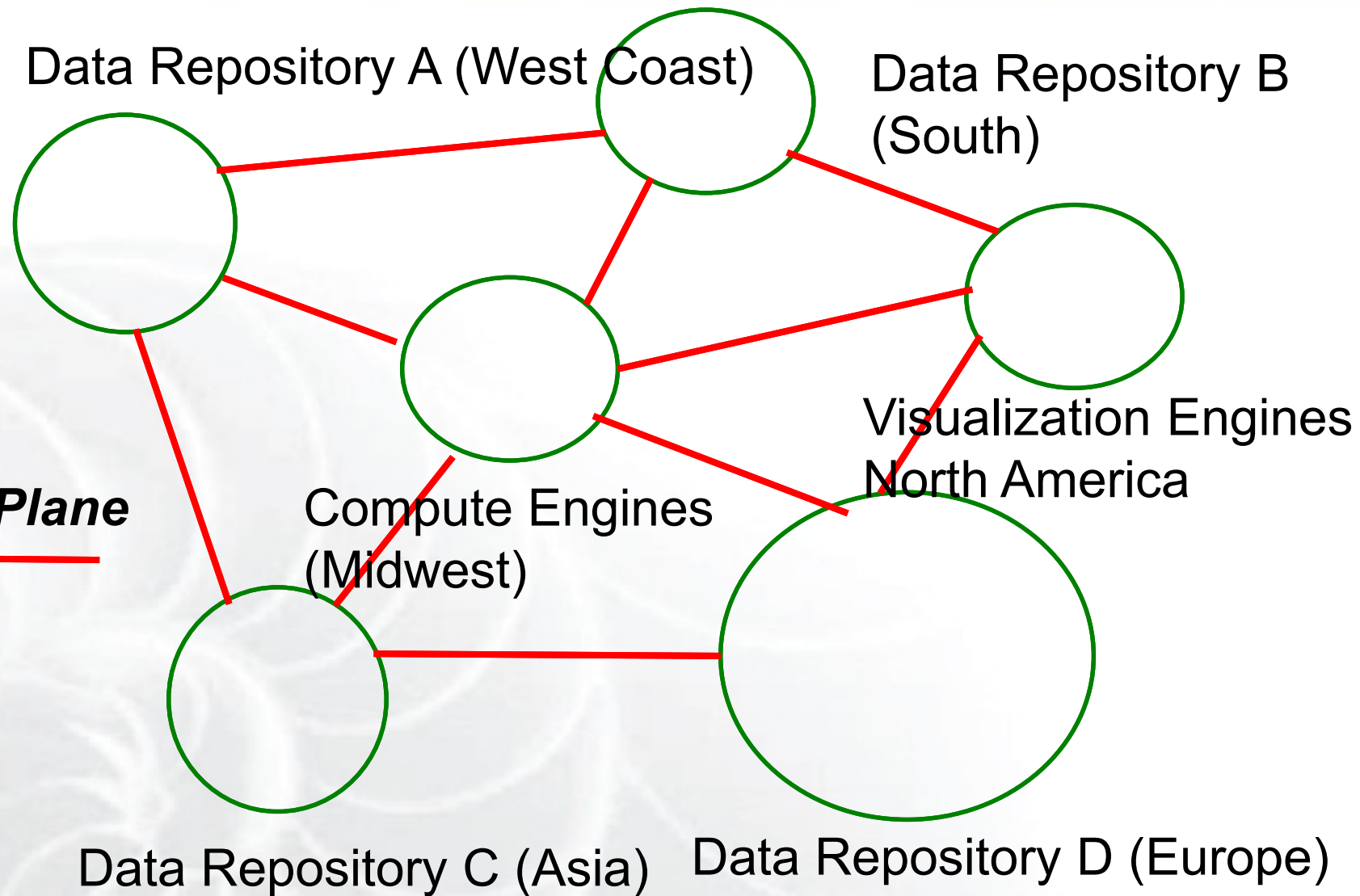
Visualization Engines  
North America

Compute Engines  
(Midwest)

Data Repository C (Asia)

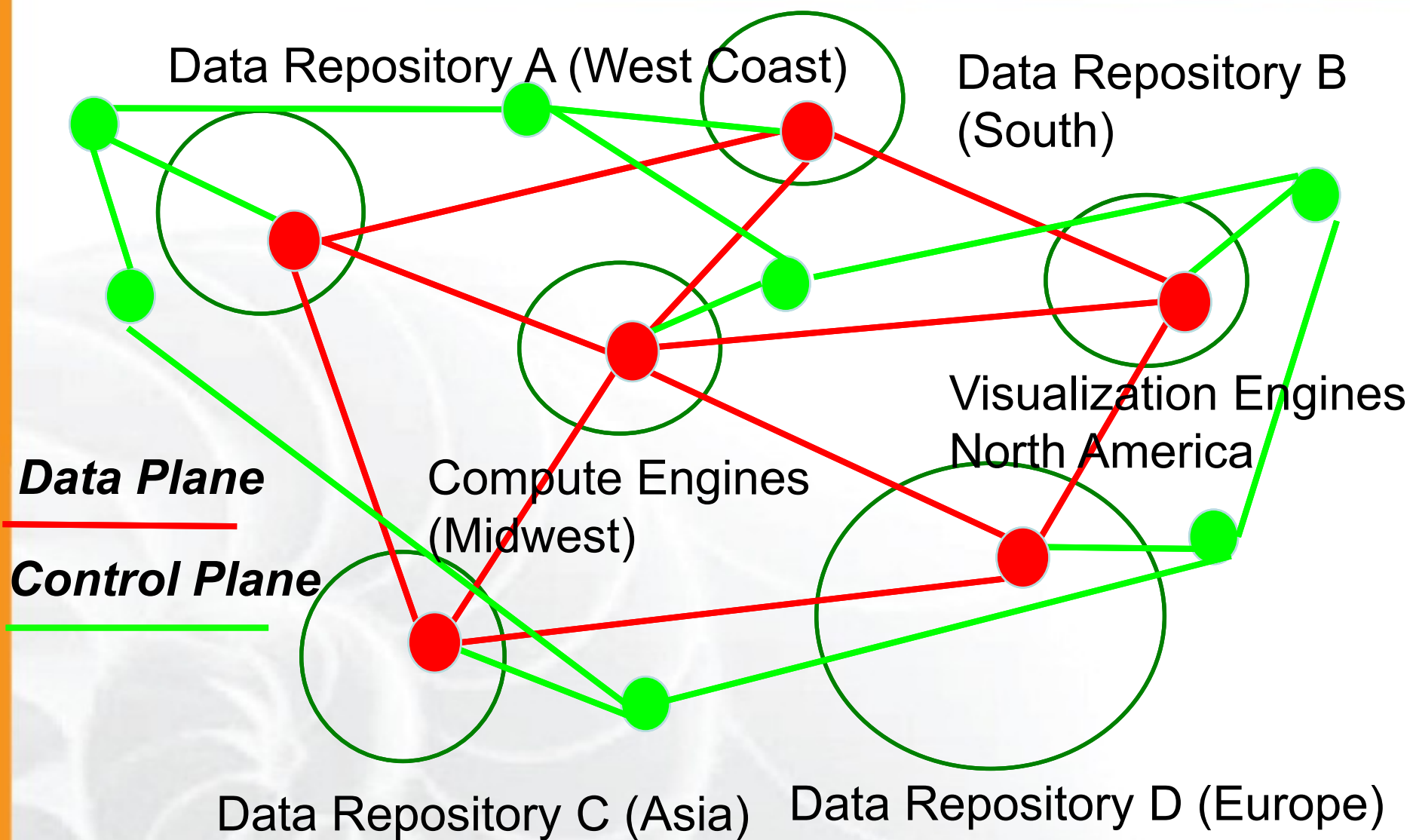
Data Repository D (Europe)

# Biomedical Data Commons: Flow Orchestration: Data Plane

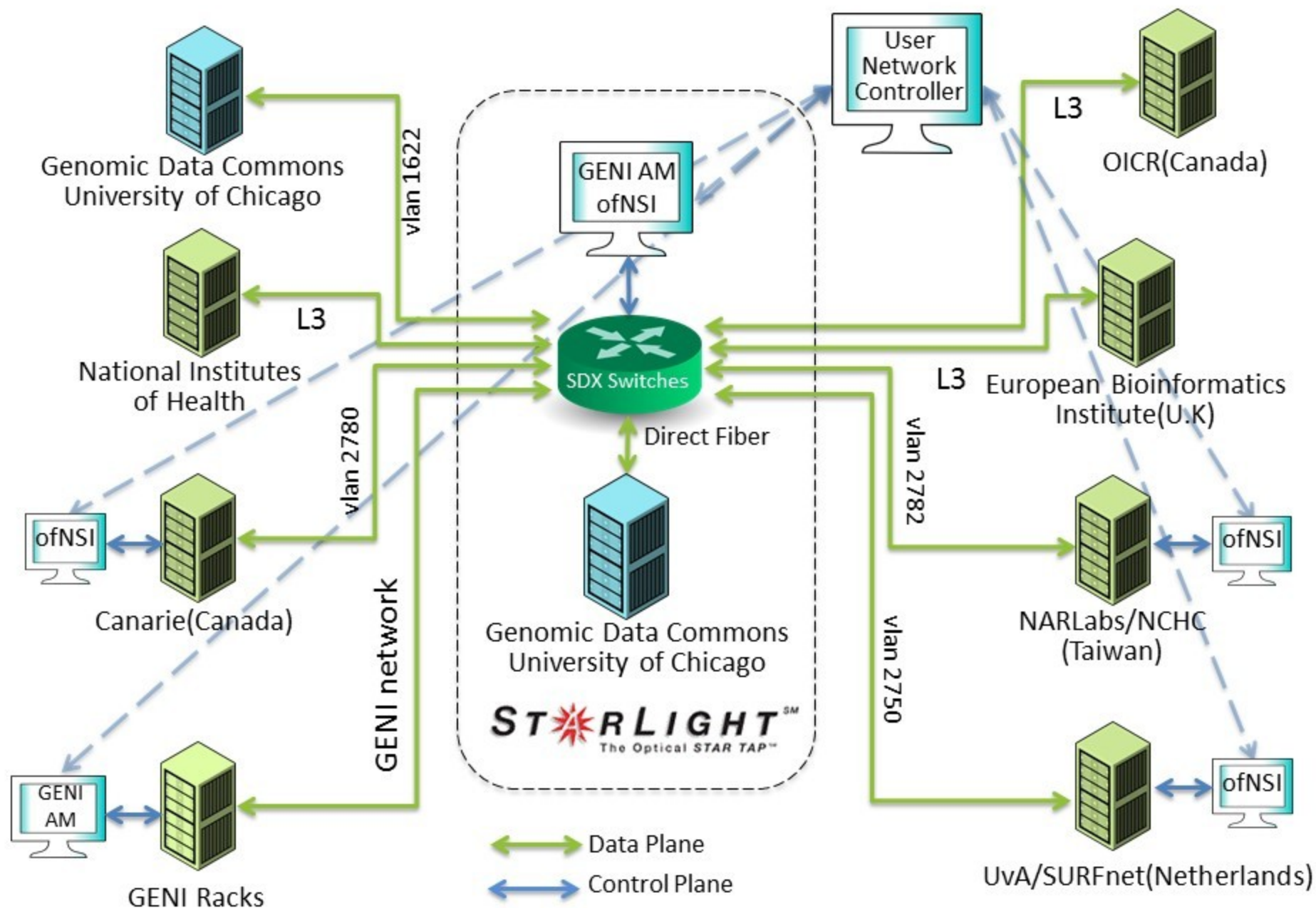




# Biomedical Data Commons: Flow Orchestration: Control Plane + Data Plane



# GEC22 Bioinformatics SDXs Demo Network



# Bioinformatics SDX Demonstration

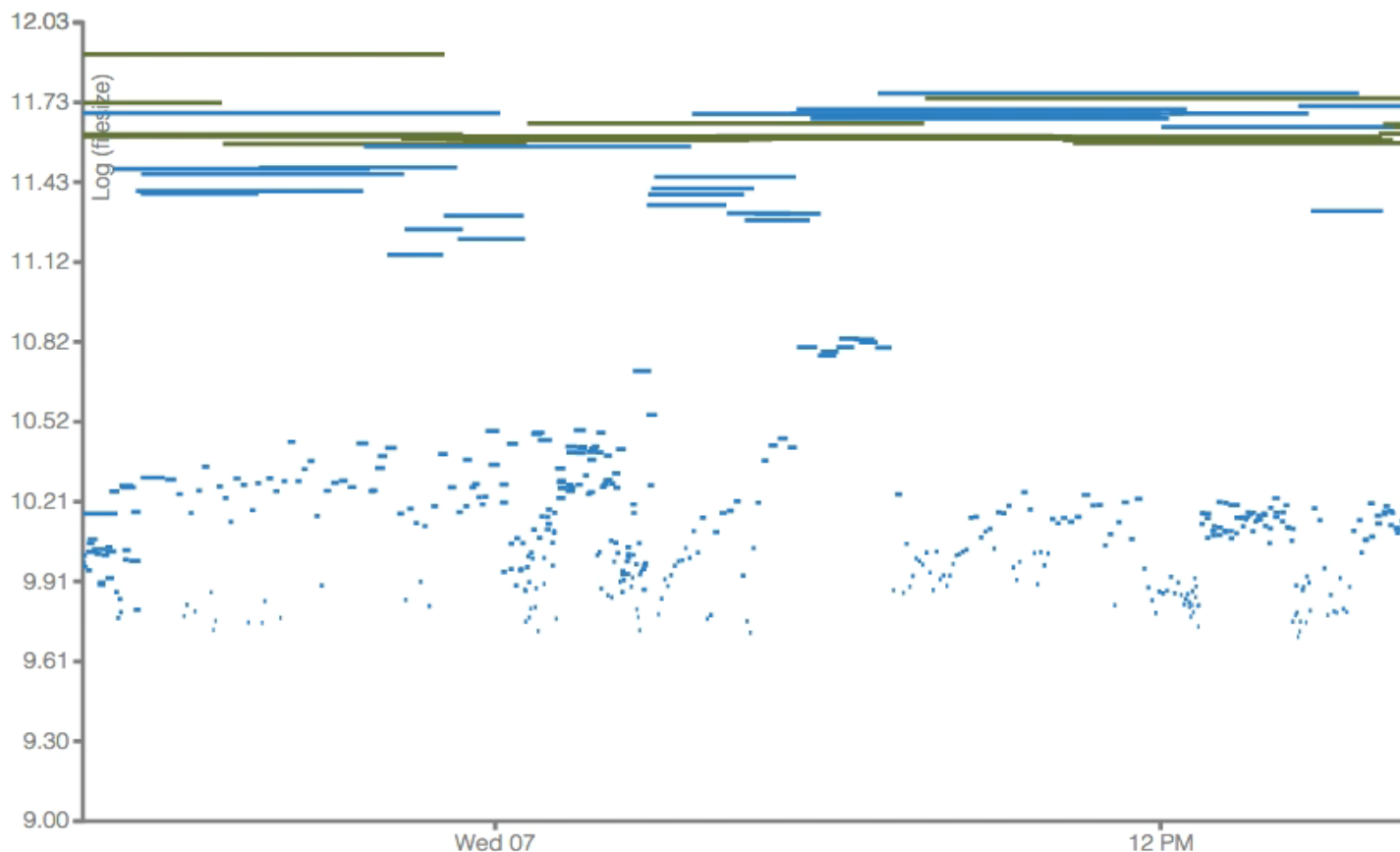
- A) Dynamically Moving Core Data Files Among Multiple Sites Around the World Via StarLight SDX
- B) Moving RNA-seq Data Files From NCI (Bethesda, MD) and EBI (Hinxton, UK) Through SDX Switch/Routers to The University of Chicago.
  - Analysis By Comparison To Known Data Correlated To Drug Response.
  - Determine Possible Actionable Therapeutic Options.
  - Return Viable Treatment Options To the Originating Site.

## Data Commons Compute Status

[animate](#)
[stop](#)
■ ceph-TARGET

■ ceph-TCGA

■ cleversafe-TCGA

■ cleversafe-TARGET




- Precision Medicine Requires Data Commons That Scale To Hundreds of Petabytes, With Programmable Networks and Data Peering To Support Data Sharing.
- Speed Discovery and Support Analytics-Driven Healthcare To Recommend Treatment.
- Large Scale Data Analysis and Dynamic Pipelines For Workflows Are Essential For Determining Optimal Results.

- An Innovative Approach To Advanced Knowledge Discovery and Medical Treatment: **Precision Medicine Supported By Precision Networking**
- Precision Mapping Of Communication Services To BI Workflow Requirements Across the World Using Advanced Analytics, the Biomedical Data Commons, and a Programmable Dynamic SDX
- Looking Forward:
  - A) Further Development/Refinement of Basic Capabilities
  - B) Transition to **Actual Production Services**
  - **C) The Biomedical Data Commons and Bionimbus Protected Data Cloud Are Being Developed As a Key Production Knowledge Discovery/Transformational Medical Treatment Facility**

# Using GENI To Invent the Future...

## Thank You!



**NORTHWESTERN  
UNIVERSITY**







[www.chameleoncloud.org](http://www.chameleoncloud.org)

## Another SDX Opportunity...

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

Principal Investigator: Kate Keahey

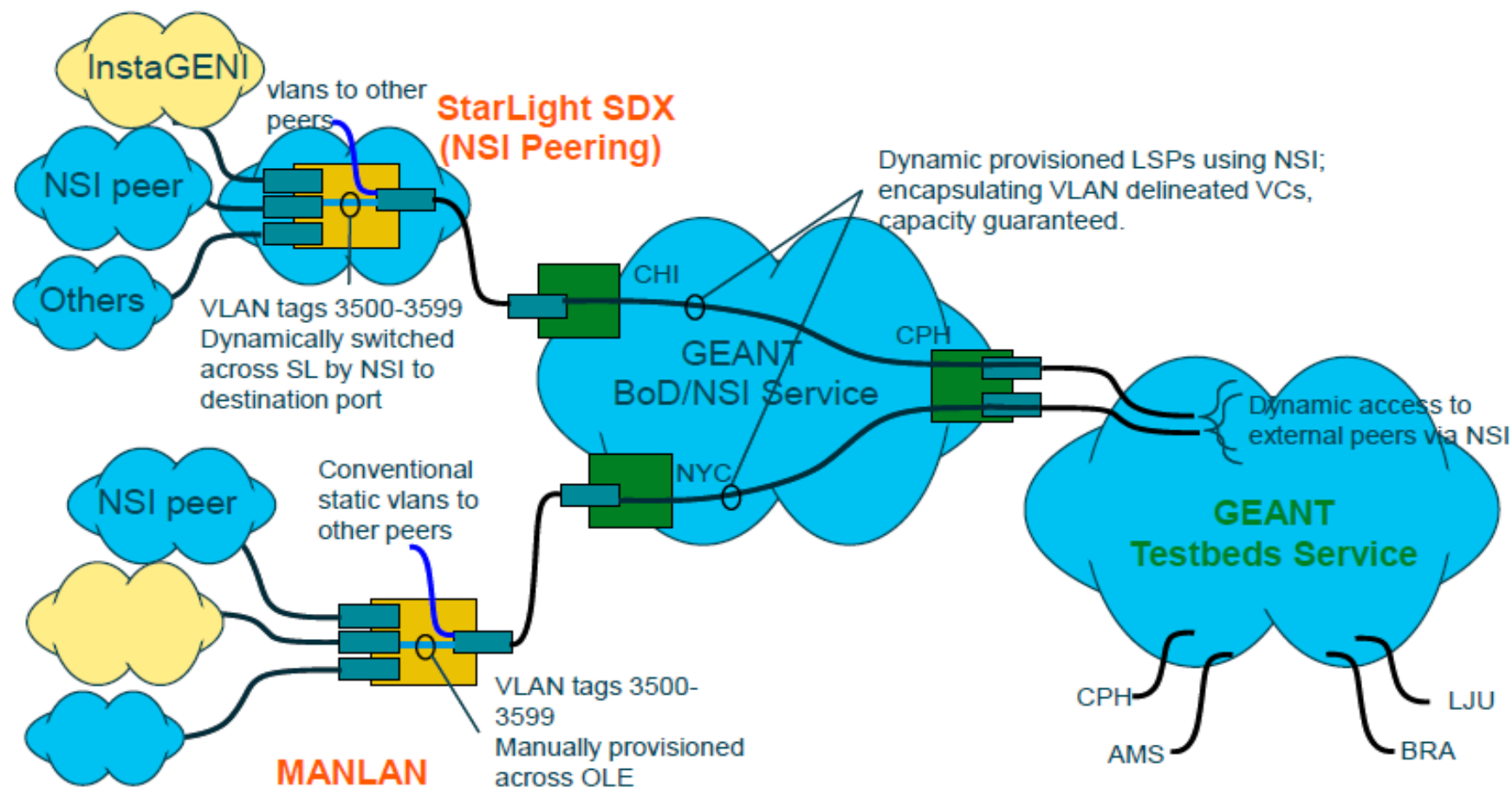
Co-PIs: J. Mambretti, D.K. Panda, P. Rad, W. Smith, D. Stanzione

AUGUST 29, 2014



# GTS connectivity to GENI Facilities

## (step 2: Dynamic transit provisioning)



- Science DMZ Exchange Services
- Emerging Software Defined Infrastructure (SDI)
- Network Function Virtualization (NFV)
- SDX Services Via “App Store”
- Etc...
- Ref: Forthcoming Information On [www.startap.net/starlight](http://www.startap.net/starlight)

- National Science Foundation International Research Network Connections Program: StarLight International SDX: A Software Defined Networking Exchange for Global Science Research and Education





Thanks to the NSF, The GENI GPO,  
DOE, DARPA  
Universities, National Labs,  
International Partners,  
and Other Supporters

