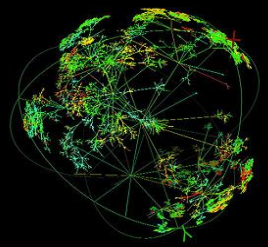
SC19 Advanced Networking For Data Intensive Science Experiments and Demonstrations

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), Director, StarLight International/National Communications Exchange Facility (www.startap.net/starlight),
PI IRNC: RXP: StarLight SDX, Co-PI Chameleon, PI-iGENI, PI-OMNINet

International Conference On High Performance Computing Networking, Storage and Analytics
Denver, Colorado
November 17-22, 2019





iCAIR

STARLIGHT SDX

Selected Applications



GENI
www.geni.net



Open Storage Network
www.openstorage network.org



OSG
www.openscience grid.org



Polar Geospatial Center
www.pgc.umn.edu



LSST
www.lsst.org



ISS
www.nasa.gov/ station



Virgo
www.virgo-gw.eu



LHC
home.cern/science/ accelerators/large-hadron-collider



GLEON
www.gleon.org



GRP
theglobalresearch platform.net/



IceCube
icecube.wisc.edu



Pierre Auger Observatory
www.auger.org



LIGO
www.ligo.caltech. edu



SKA
www.skatelescope. org





LHCONE
twiki.cern.ch/twiki/bin /view/LHCONE/ WebHome



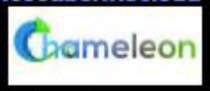
USGS EROS
www.usgs.gov/ centers/eros



Blue Waters
bluewaters.ncsa. illinois.edu



PRP
pacificresearch platform.org



Chameleon
www.chameleon cloud.org

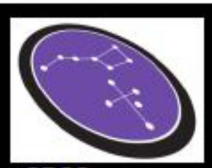
Jetstream
www.jetstream-cloud.org



Belle II
www.belle2.org



XENON
xenon.astro. columbia.edu



SDSS
www.sdss.org



XSEDE
www.xsede.org



CHASE-CI
www.calit2.net/ newsroom/artic le.php?id=2910



LHCOPN
twiki.cern.ch/twiki/bin /view/LHCOPN/ WebHome



PRAGMA
www.pragma-grid.net



NEON
www.neonscience. org



Genomic Science Program
genomicscience. energy.gov



LBNF/DUNE/ ProtoDUNE
lbnf.fnal.gov



NOvA
novaexperiment. fnal.gov



ALMA
www.alma observatory.org



CENTRA
www.global centra.org



SAGE2
sage2.sagecommons. org



IVOA
www.ivoa.net

StarLight – "By Researchers For Researchers"

StarLight: Experimental Optical Infrastructure/Proving Ground For Next Gen Network Services

Optimized for High Performance Data Intensive Science

Multiple 100 Gbps (60+ Paths)

StarWave

100 G Exchange

World's Most Advanced Exchange

Multiple First of a Kind Services and Capabilities



View from StarLight



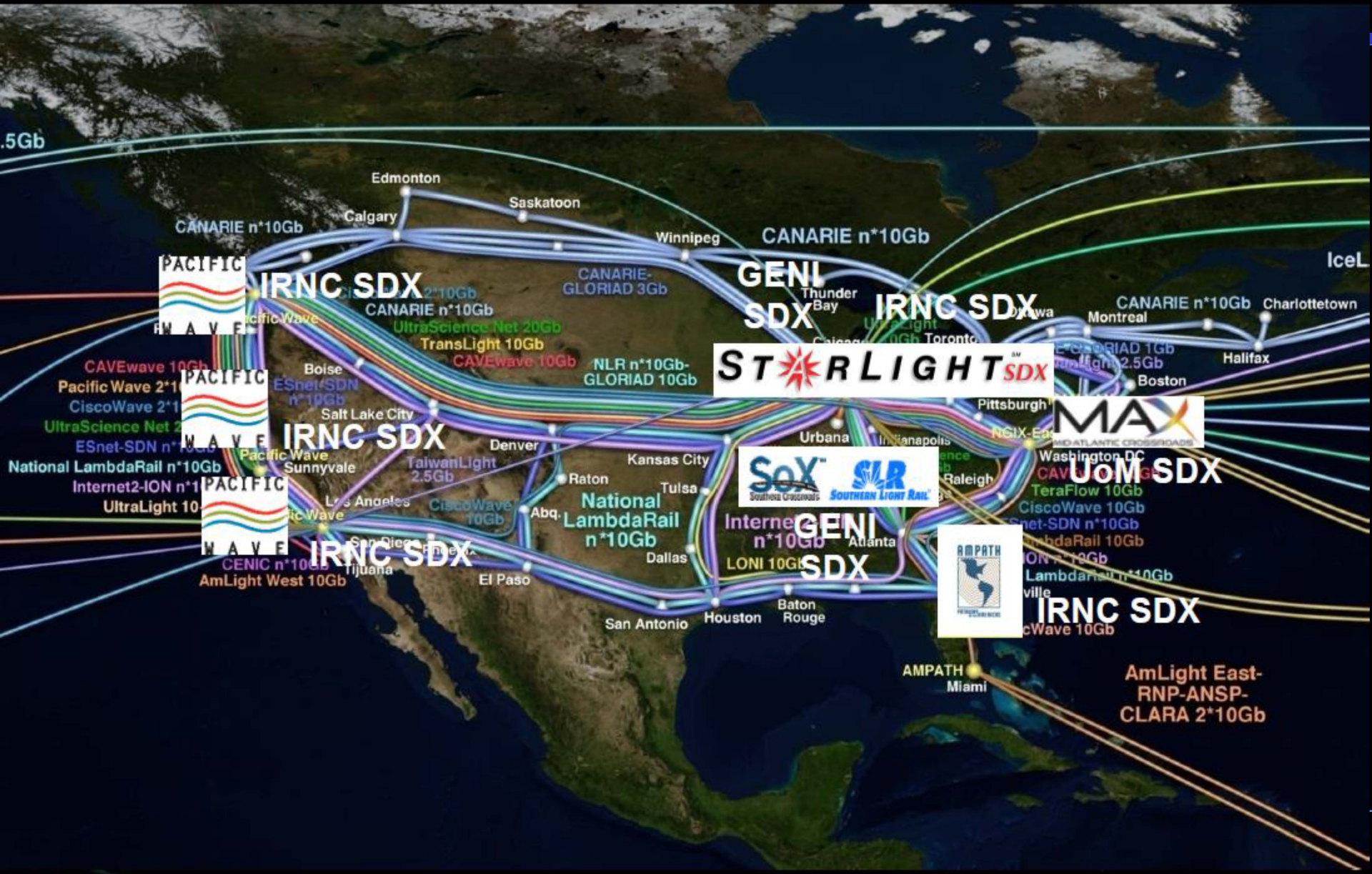
Abbott Hall, Northwestern University's Chicago Campus

Global Research Platform: Global Lambda Integrated Facility
Available Advanced Network Resources



Visualization courtesy of Bob Patterson, NCSA; data compilation by Maxine Brown, UIC.

Emerging US SDX Interoperable Fabric

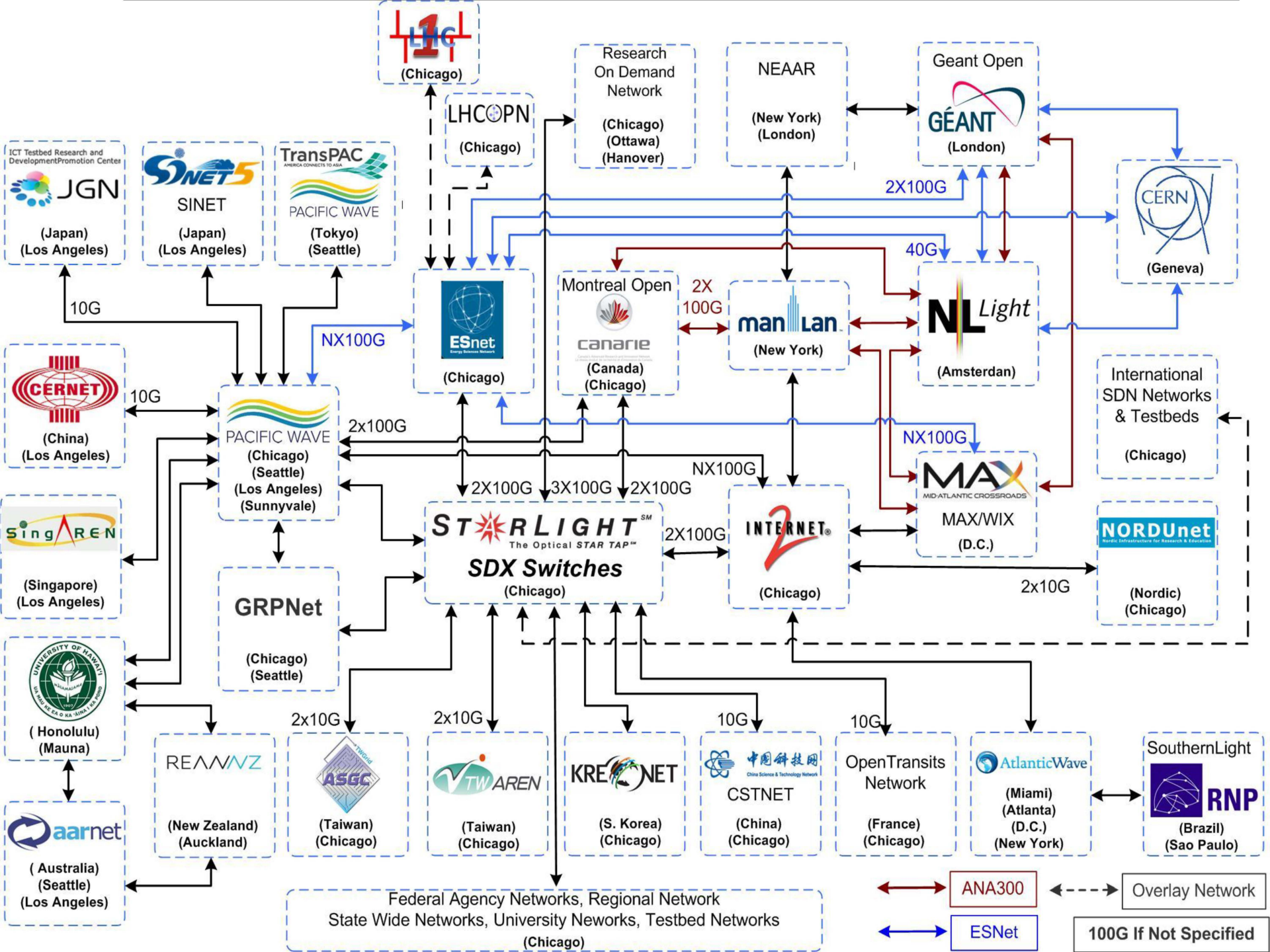


GLIF Automated GOLE Fabric 2018

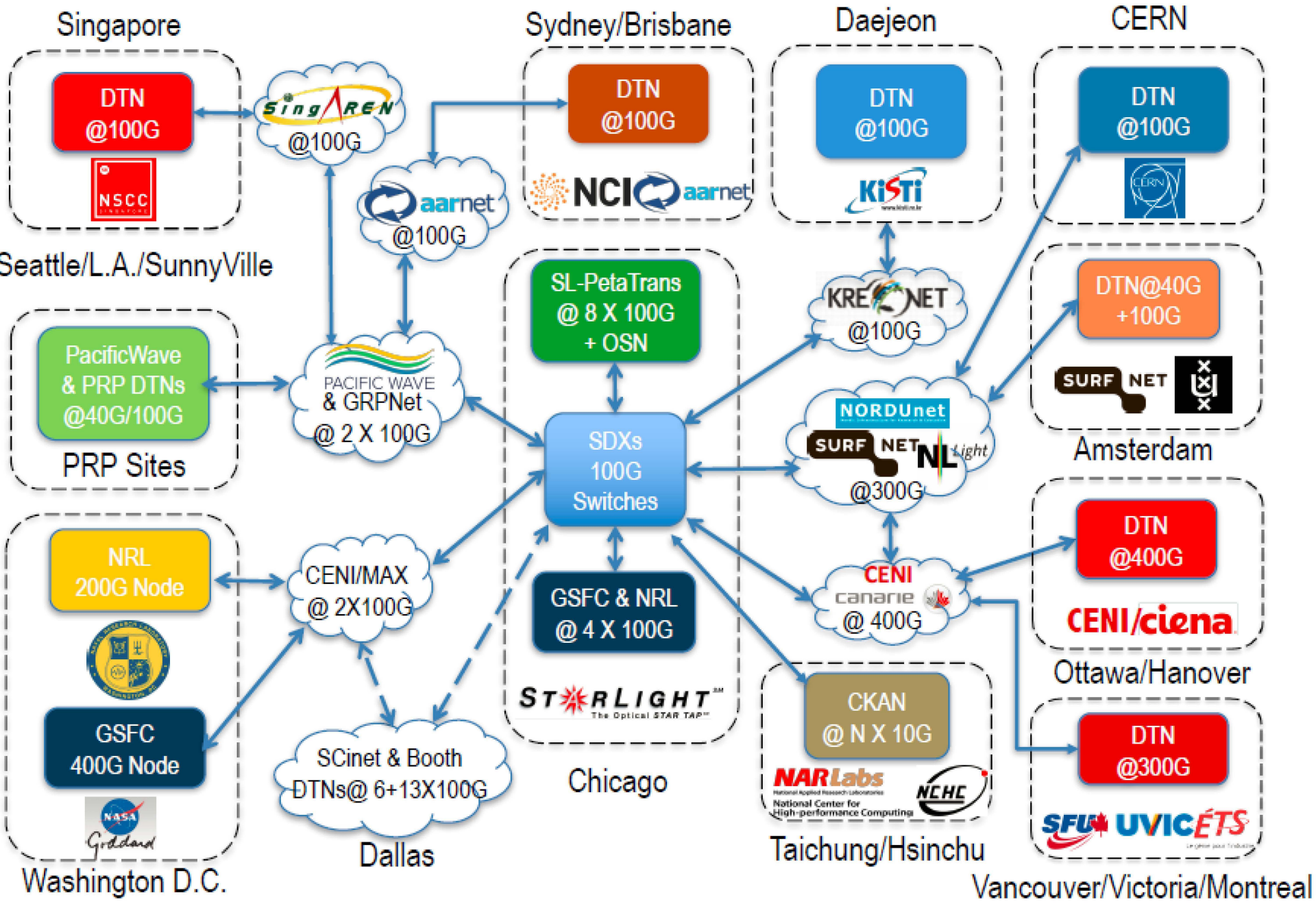


SURF

STARLIGHT℠


LHC
(Chicago)
LHCOPN
(Chicago)
Research On Demand Network
(Chicago)
(Ottawa)
(Hanover)
NEAAR
(New York)
(London)
Geant Open
GÉANT
(London)
CERN
(Geneva)
2X100G
40G
ICT Testbed Research and DevelopmentPromotion Center
JGN
(Japan)
(Los Angeles)
SINET
(Japan)
(Los Angeles)
TransPAC
PACIFIC WAVE
(Tokyo)
(Seattle)
10G
NX100G
ESnet
(Chicago)
Montreal Open
canarie
(Canada)
(Chicago)
2X 100G
manLan
(New York)
NL Light
(Amsterdan)
CERNET
(China)
(Los Angeles)
10G
PACIFIC WAVE
(Chicago)
(Seattle)
(Los Angeles)
(Sunnyvale)
2x100G
International SDN Networks & Testbeds
(Chicago)
NX100G
MAX
MID-ATLANTIC CROSSROADS
MAX/WIX
(D.C.)
2X100G
3X100G
2X100G
NX100G
SingAREN
(Singapore)
(Los Angeles)
STARLIGHT
The Optical STAR TAP
SDX Switches
(Chicago)
2X100G
INTERNET2
(Chicago)
2x10G
NORDUnet
(Nordic)
(Chicago)
GRPNet
(Chicago)
(Seattle)
(Honolulu)
(Mauna)
2x10G
2x10G
10G
10G
REANNZ
(New Zealand)
(Auckland)
ASGC
(Taiwan)
(Chicago)
TWAREN
(Taiwan)
(Chicago)
KREONET
(S. Korea)
(Chicago)
CSTNET
(China)
(Chicago)
OpenTransits Network
(France)
(Chicago)
AtlanticWave
(Miami)
(Atlanta)
(D.C.)
(New York)
SouthernLight
RNP
(Brazil)
(Sao Paulo)
aarnet
(Australia)
(Seattle)
(Los Angeles)
Federal Agency Networks, Regional Network
State Wide Networks, University Neworks, Testbed Networks
(Chicago)
ANA300
ESNet
Overlay Network
100G If Not Specified

PetaTrans: Petascale Sciences Data Transfer



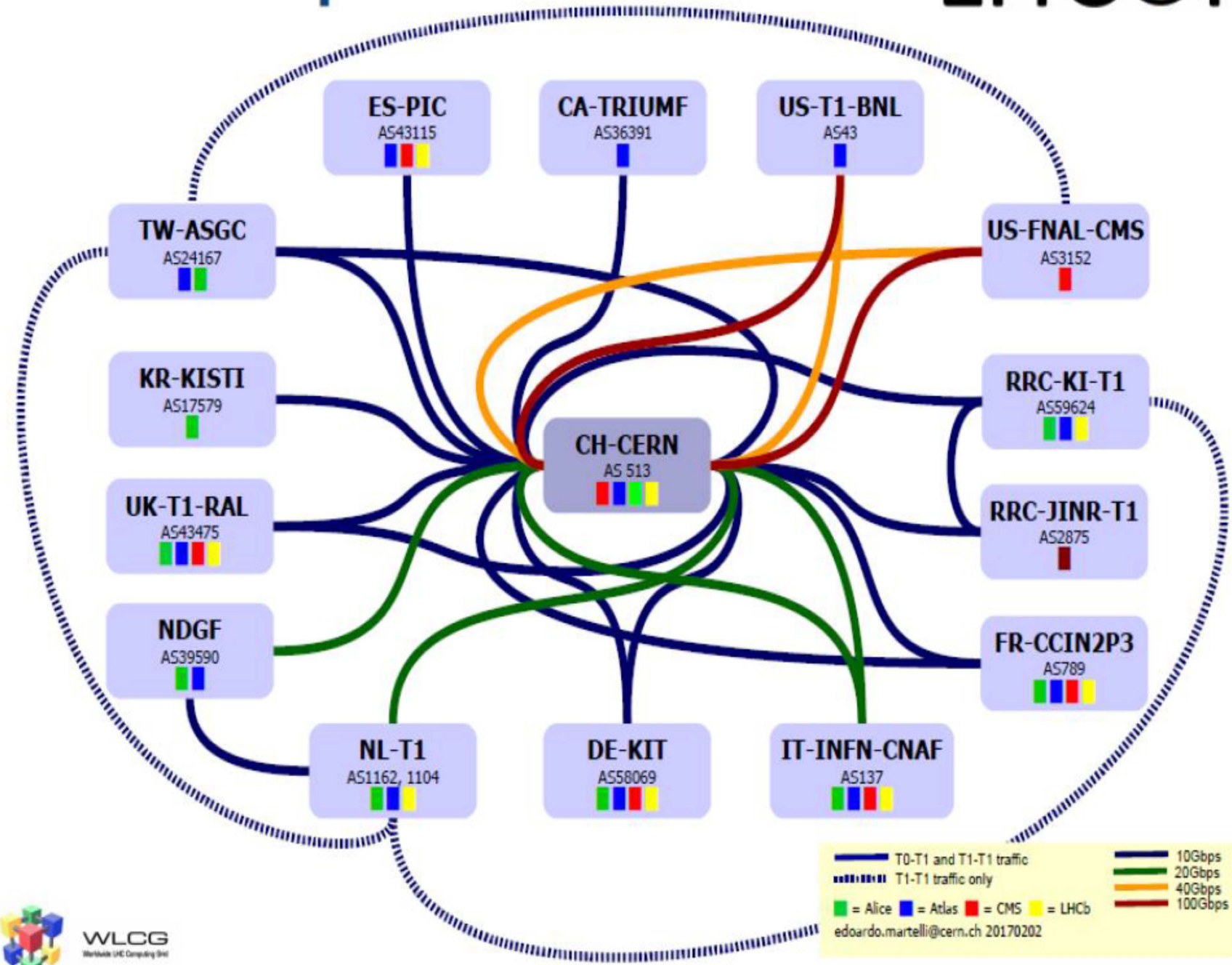
StarLight SDX Overview

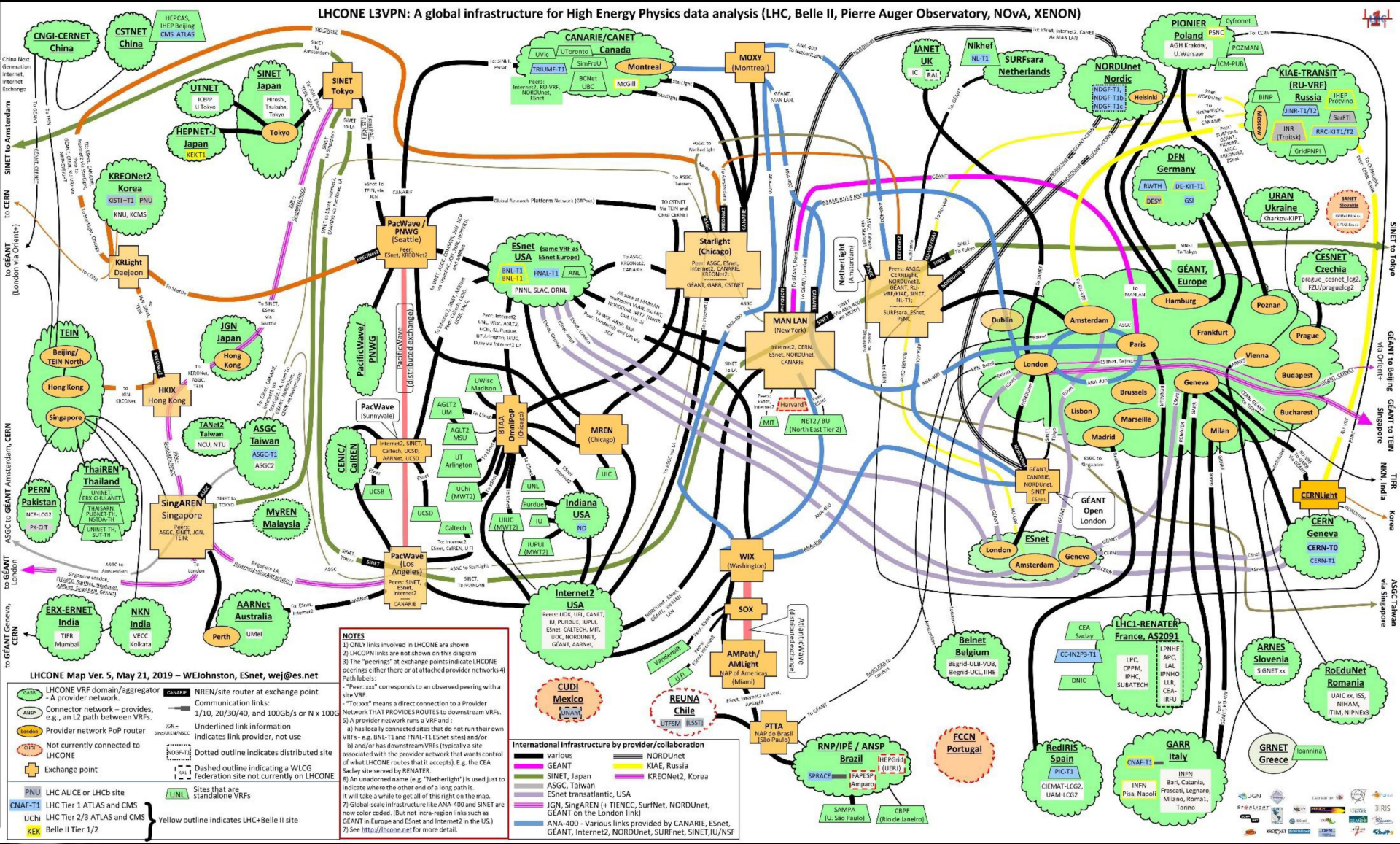
- This IRNC StarLight SDX Initiative Is Designing, Implementing, and Operating New Services For Global Data Intensive Sciences, Based On Emerging Next Generation Architecture and Technolgies, Including Virtualization, Orchestration, Segmentation (Slicing), Software Defined Resources, Programmability and Customization.
- These Macro Trends Enable Exchanges To Be Agile Platforms For Dynamic Services Provisioning, Real-Time Responsiveness, and Distributed Control Over Core Resources, Including By Applications, Edge Processes and Devices.

This Project Is Transitioning Network Exchanges To Open Innovation Platforms



LHCOPN map


ES-PIC
AS43115
CA-TRIUMF
AS36391
US-T1-BNL
AS43
TW-ASGC
AS24167
US-FNAL-CMS
AS3152
KR-KISTI
AS17579
RRC-KI-T1
AS59624
CH-CERN
AS 513
UK-T1-RAL
AS43475
RRC-JINR-T1
AS2875
NDGF
AS39590
FR-CCIN2P3
AS789
NL-T1
AS1162, 1104
DE-KIT
AS58069
IT-INFN-CNAF
AS137
T0-T1 and T1-T1 traffic
T1-T1 traffic only
= Alice = Atlas = CMS = LHCb
edoardo.martelli@cern.ch 20170202
10Gbps
20Gbps
40Gbps
100Gbps



STARLIGHT SM

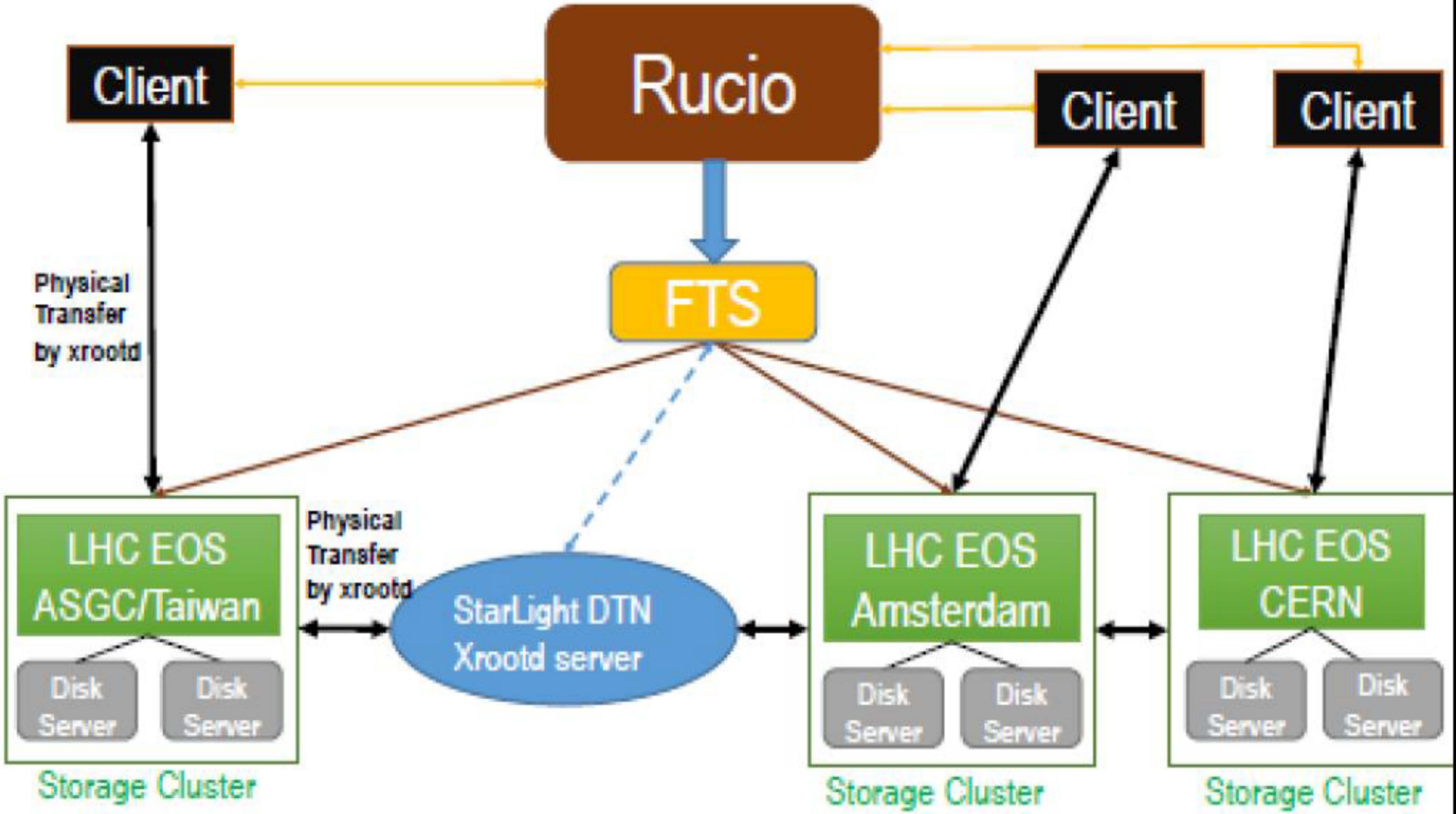
New Science Communities Using LHCONE

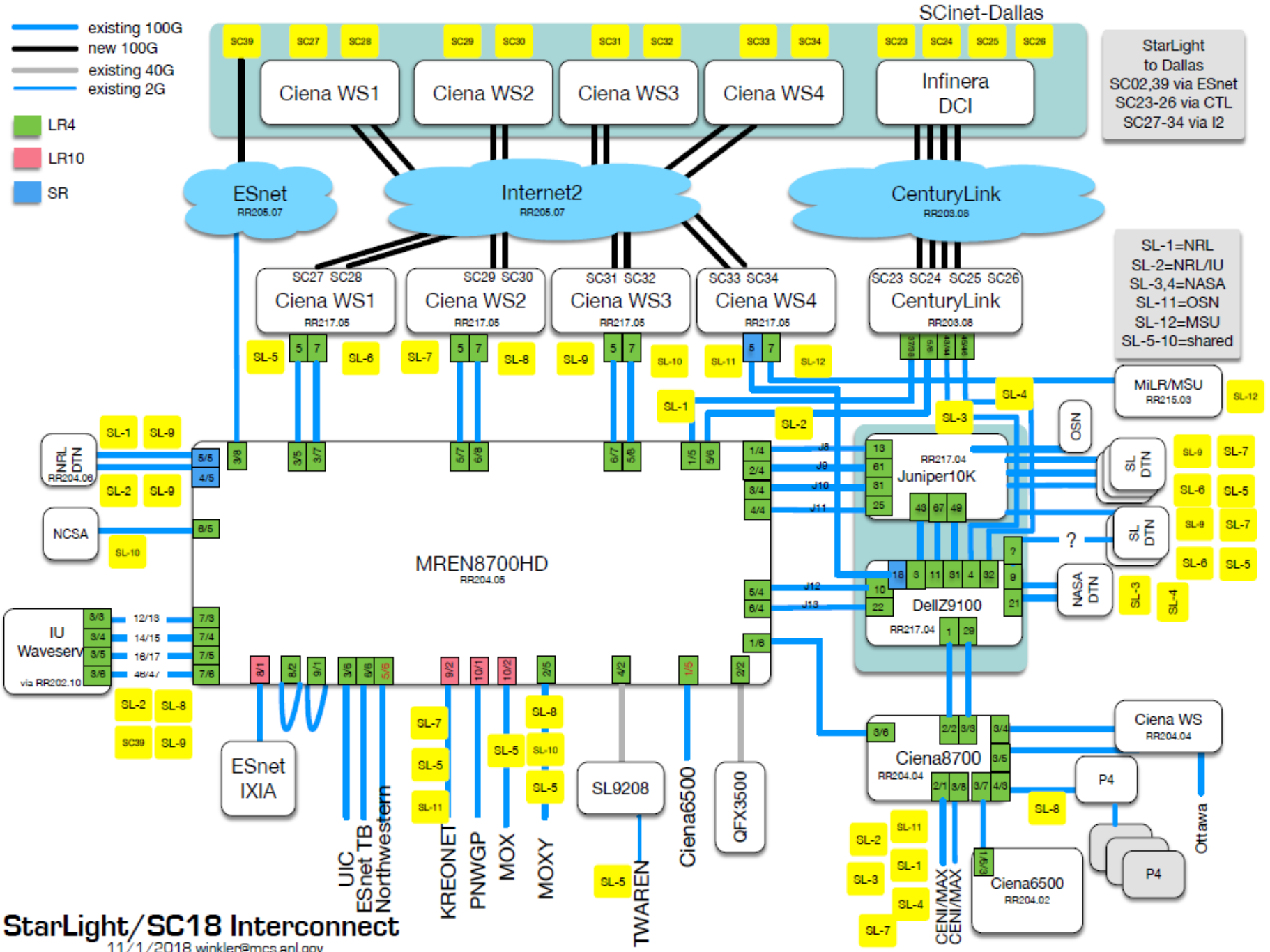
- Belle II Experiment, Particle Physics Experiment Designed To Study Properties of B Mesons (Heavy Particles Containing a Bottom Quark).
- Pierre Auger Observatory, Studying Ultra-High Energy Cosmic Rays, the Most Energetic and Rarest of Particles In the Universe.
- In August 2017 the PAO, LIGO and Virgo Collaboration Measured a Gravitational Wave Originating From a Binary Neutron Star Merger.
- The NOvA Experiment Is Designed To Answer Fundamental questions in neutrino Physics.
- The XENON Dark Matter Project Is a Global Collaboration Investing Fundamental Properties of Dark Matter, Largest Component Of The Universe.

New=> DUNE/ProtoDUNE – Deep Underground Nutrino Experiment



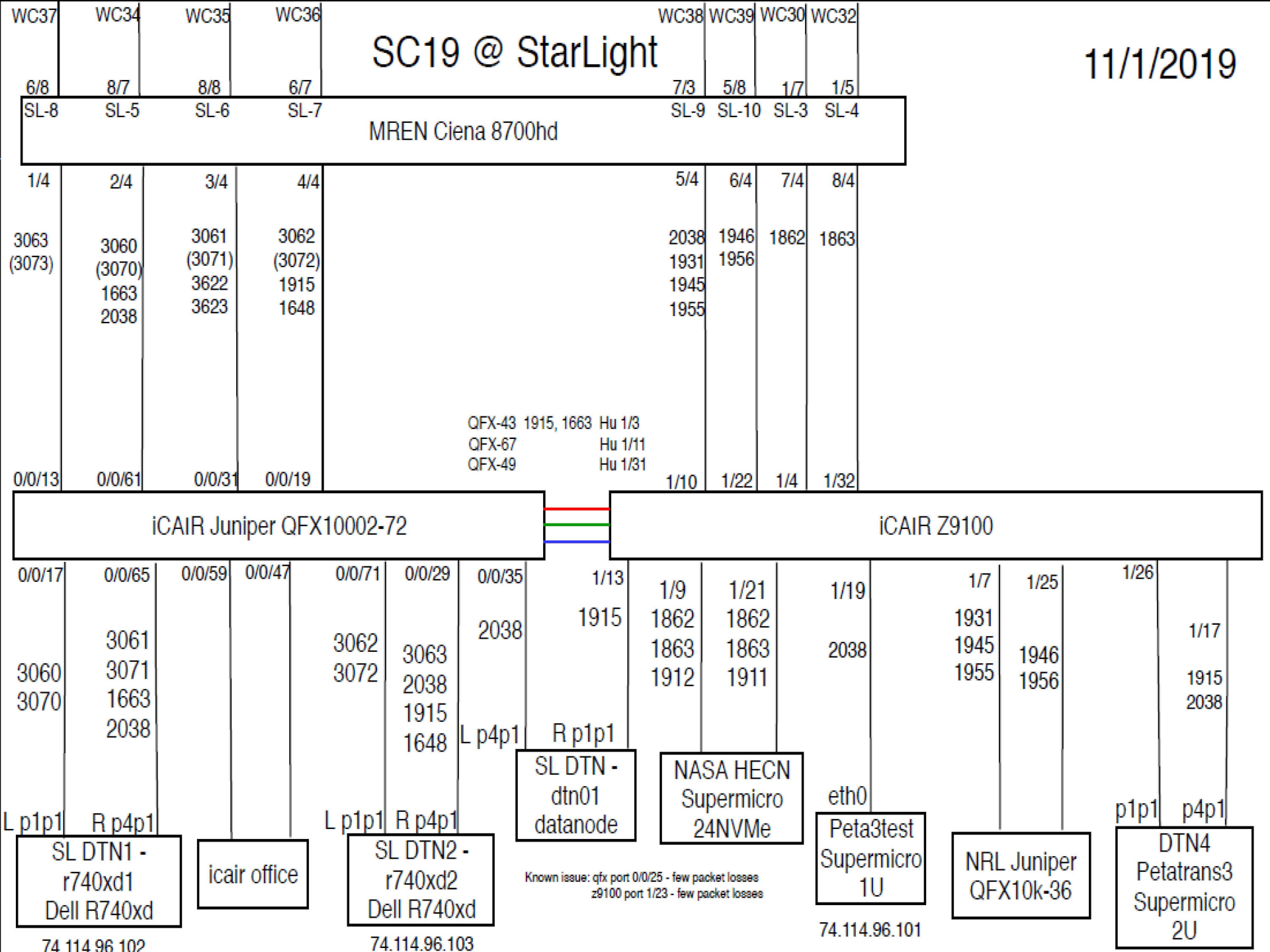
High Energy Physics SDX DTN Prototype Service

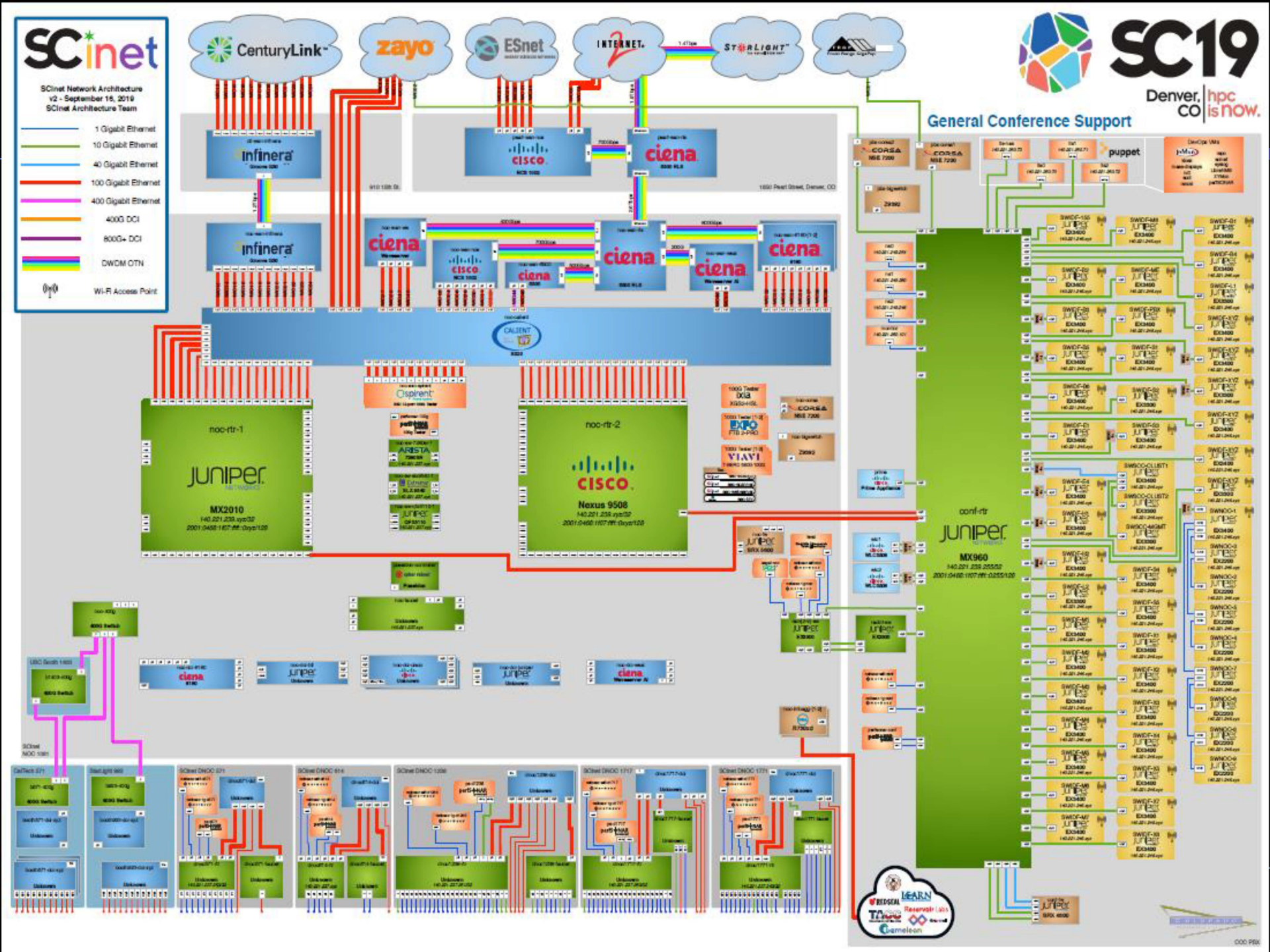


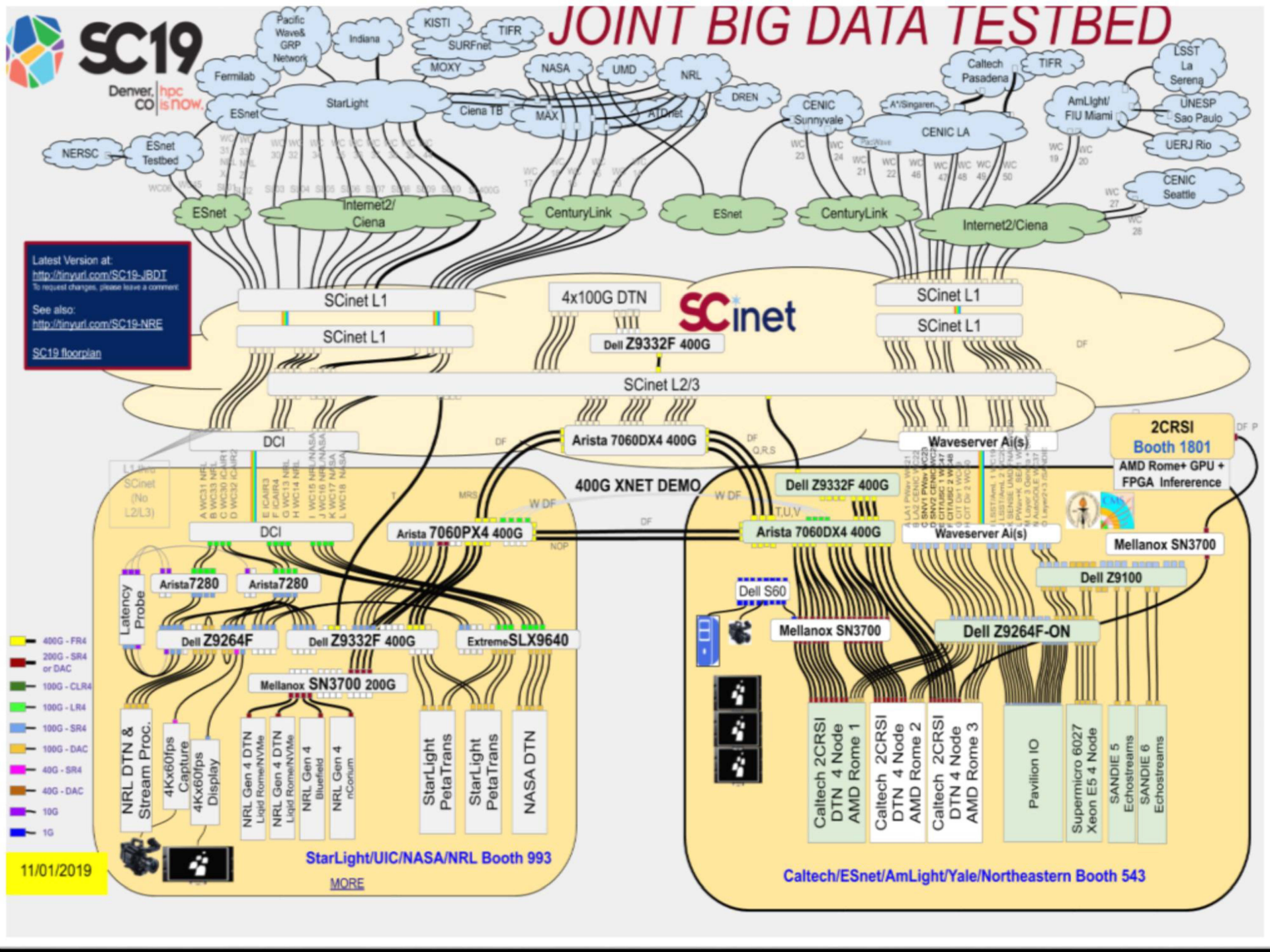

existing 100G
new 100G
existing 40G
existing 2G
LR4
LR10
SR
SCinet-Dallas
SC39 SC27 SC28 SC29 SC30 SC31 SC32 SC33 SC34 SC23 SC24 SC25 SC26
Ciena WS1
Ciena WS2
Ciena WS3
Ciena WS4
Infinera DCI
StarLight to Dallas
SC02,39 via ESnet
SC23-26 via CTL
SC27-34 via I2
ESnet
RR205.07
Internet2
RR205.07
CenturyLink
RR208.08
SC27 SC28 Ciena WS1 RR217.05
SC29 SC30 Ciena WS2 RR217.05
SC31 SC32 Ciena WS3 RR217.05
SC33 SC34 Ciena WS4 RR217.05
SC23 SC24 SC25 SC26 CenturyLink RR208.08
SL-1=NRL
SL-2=NRL/IU
SL-3,4=NASA
SL-11=OSN
SL-12=MSU
SL-5-10=shared
MiLR/MSU RR215.03
OSN
SL DTN
NASA DTN
NRL DTN RR204.06
NCSA
MREN8700HD RR204.05
RR217.04 Juniper10K
DellZ9100 RR217.04
IU Waveserv via RR202.10
ESnet IXIA
UIC
ESnet TB
Northwestern
KREONET
PNWGP
MOX
MOXY
SL9208
TWAREN
Ciena6500
QFX3500
Ciena8700 RR204.04
Ciena WS RR204.04
P4
Ottawa
Ciena6500 RR204.02
CENI/MAX
CENI/MAX
StarLight/SC18 Interconnect
11/1/2018 winkler@mcs.anl.gov
Linda Winkler

SC19 @ StarLight

11/1/2019




SCinet
SCinet Network Architecture
v2 - September 16, 2019
SCinet Architecture Team
1 Gigabit Ethernet
10 Gigabit Ethernet
40 Gigabit Ethernet
100 Gigabit Ethernet
400 Gigabit Ethernet
400G DCI
800G+ DCI
DWDM OTN
Wi-Fi Access Point
CenturyLink
zayo
ESnet
INTERNET2
STARLIGHT
Infinera
cisco
ciena
noc-rtr-1
JUNIPER
MX2010
noc-rtr-2
Nexus 9508
conf-rtr
MX960
SC19
Denver, CO
hpc is now.
General Conference Support
puppet


SC19
Denver, CO hpc is now.
JOINT BIG DATA TESTBED
Pacific Wave& GRP Network
Indiana
KISTI
TIFR
SURFnet
MOXY
Fermilab
ESnet
StarLight
Ciena TB
NASA
UMD
NRL
MAX
AtlPnet
DREN
Caltech Pasadena
TIFR
LSST La Serena
CENIC Sunnyvale
A*/Singaren
CENIC LA
AmLight/ FIU Miami
UNESP Sao Paulo
UERJ Rio
NERSC
ESnet Testbed
CENIC Seattle
ESnet
Internet2/ Ciena
CenturyLink
ESnet
CenturyLink
Internet2/Ciena
Latest Version at:
http://tinyurl.com/SC19-JBDT
To request changes, please leave a comment
See also:
http://tinyurl.com/SC19-NRE
SC19 floorplan
SCinet L1
SCinet L1
4x100G DTN
SCinet
Dell Z9332F 400G
SCinet L1
SCinet L1
SCinet L2/3
DCI
Arista 7060DX4 400G
Waveserver Ai(s)
2CRSI
Booth 1801
AMD Rome+ GPU + FPGA Infererence
400G XNET DEMO
DCI
Arista 7060PX4 400G
Dell Z9332F 400G
Arista 7060DX4 400G
Waveserver Ai(s)
Mellanox SN3700
Latency Probe
Arista7280
Arista7280
Dell Z9264F
Dell Z9332F 400G
ExtremeSLX9640
Mellanox SN3700 200G
Dell S60
Mellanox SN3700
Dell Z9100
Dell Z9264F-ON
NRL DTN & Stream Proc.
4Kx60fps Capture
4Kx60fps Display
NRL Gen 4 DTN Liqid Rome/NVMe
NRL Gen 4 DTN Liqid Rome/NVMe
NRL Gen 4 Bluefield
NRL Gen 4 nConium
StarLight PetaTrans
StarLight PetaTrans
NASA DTN
Caltech 2CRSI DTN 4 Node AMD Rome 1
Caltech 2CRSI DTN 4 Node AMD Rome 2
Caltech 2CRSI DTN 4 Node AMD Rome 3
Pavilion IO
Supermicro 6027 Xeon E5 4 Node
SANDIE 5 Echostreams
SANDIE 6 Echostreams
400G - FR4
200G - SR4 or DAC
100G - CLR4
100G - LR4
100G - SR4
100G - DAC
40G - SR4
40G - DAC
10G
1G
11/01/2019
StarLight/UIC/NASA/NRL Booth 993
MORE
Caltech/ESnet/AmLight/Yale/Northeastern Booth 543

StarLight SC19 400 Gbps WAN/LAN With SCinet

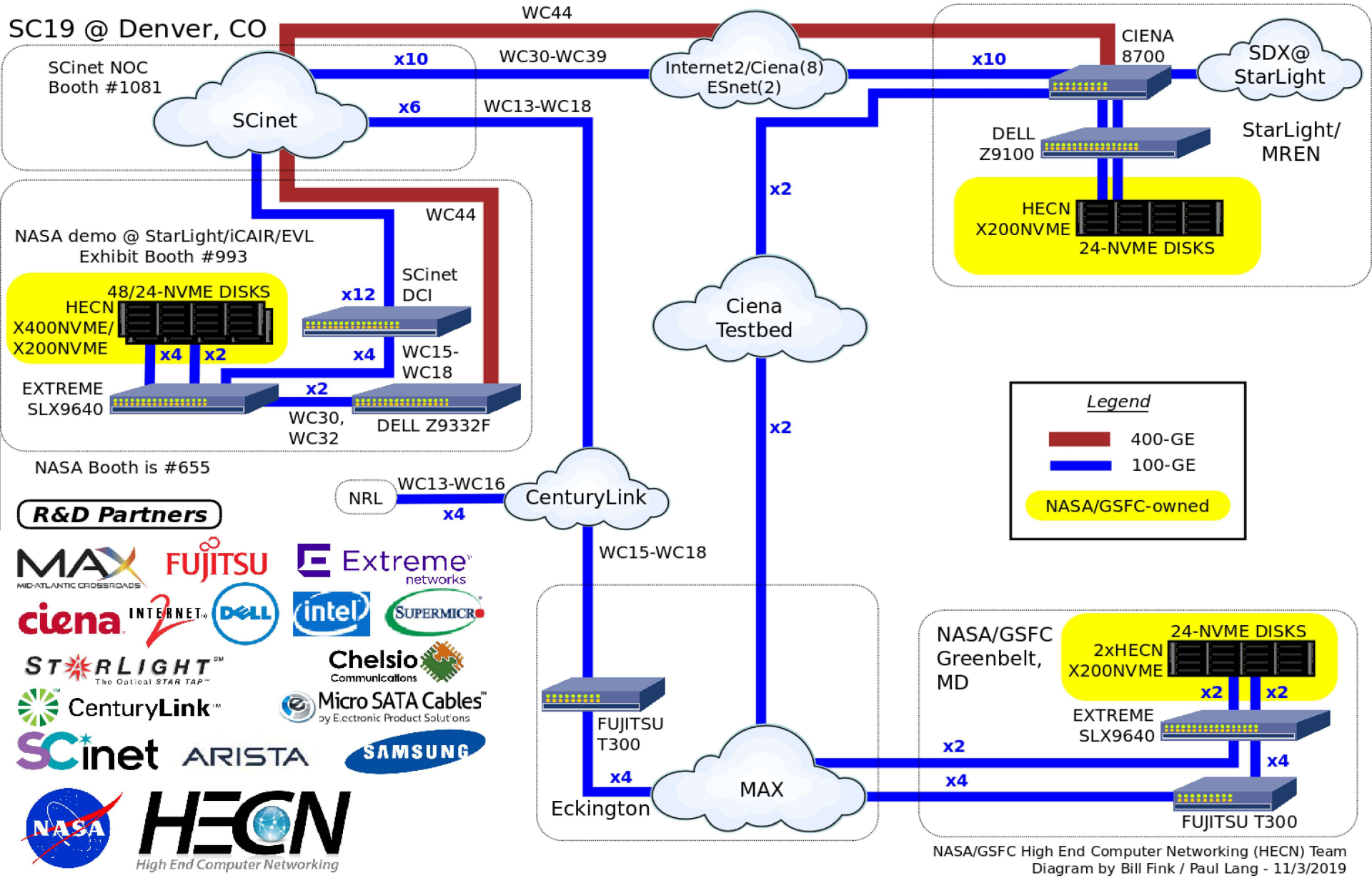
- 400 Gbps E2E
- 400 Gbps WAN
- 400 Gbps LAN
- 400 Gbps WAN/LAN Interface
- Resource Mapping



SC19

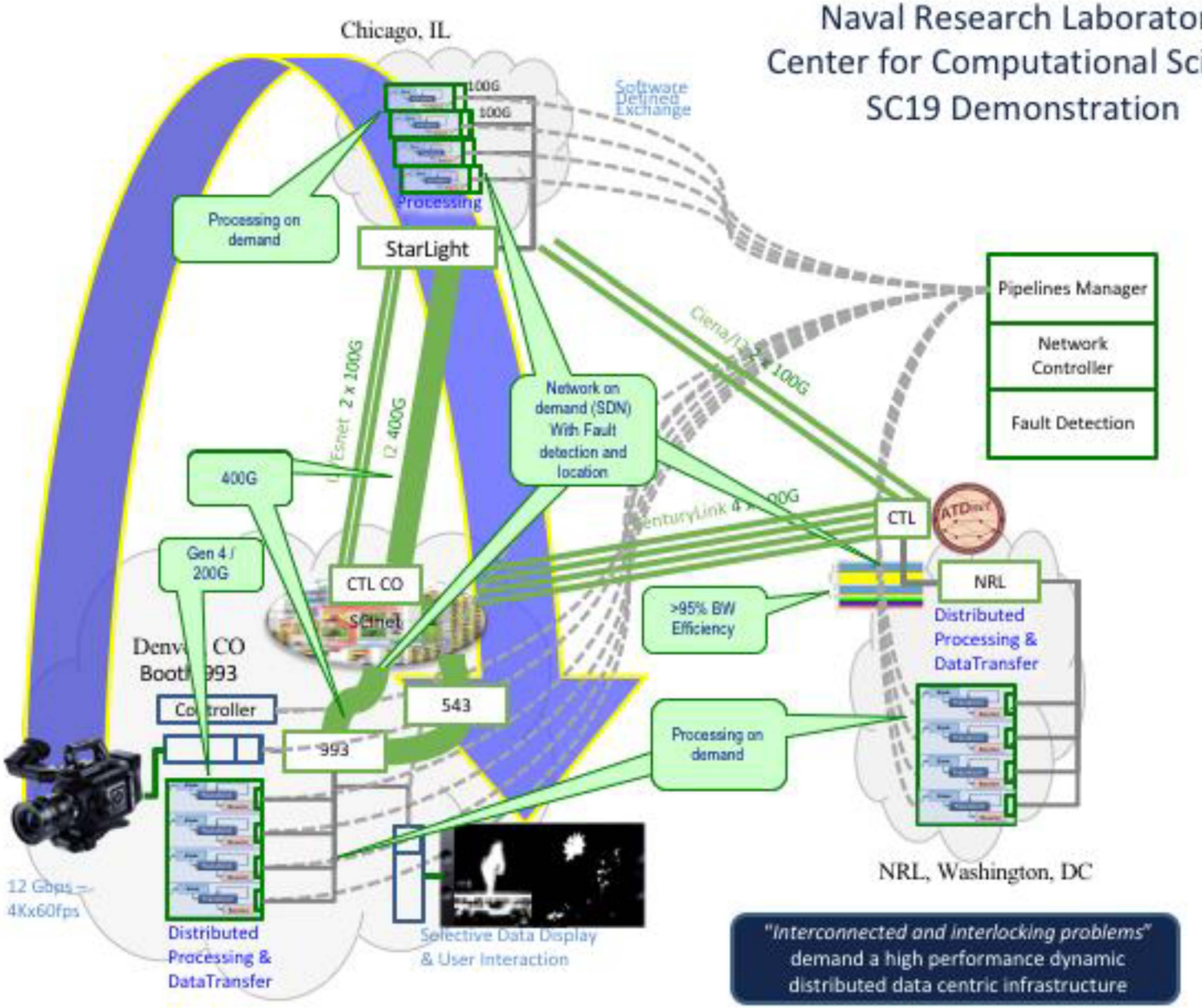
Demonstrations of 400 Gbps Disk-to-Disk WAN File Transfers using RDMA and NVMe Drives

An SC19 Collaborative Initiative Among NASA and Several Partners



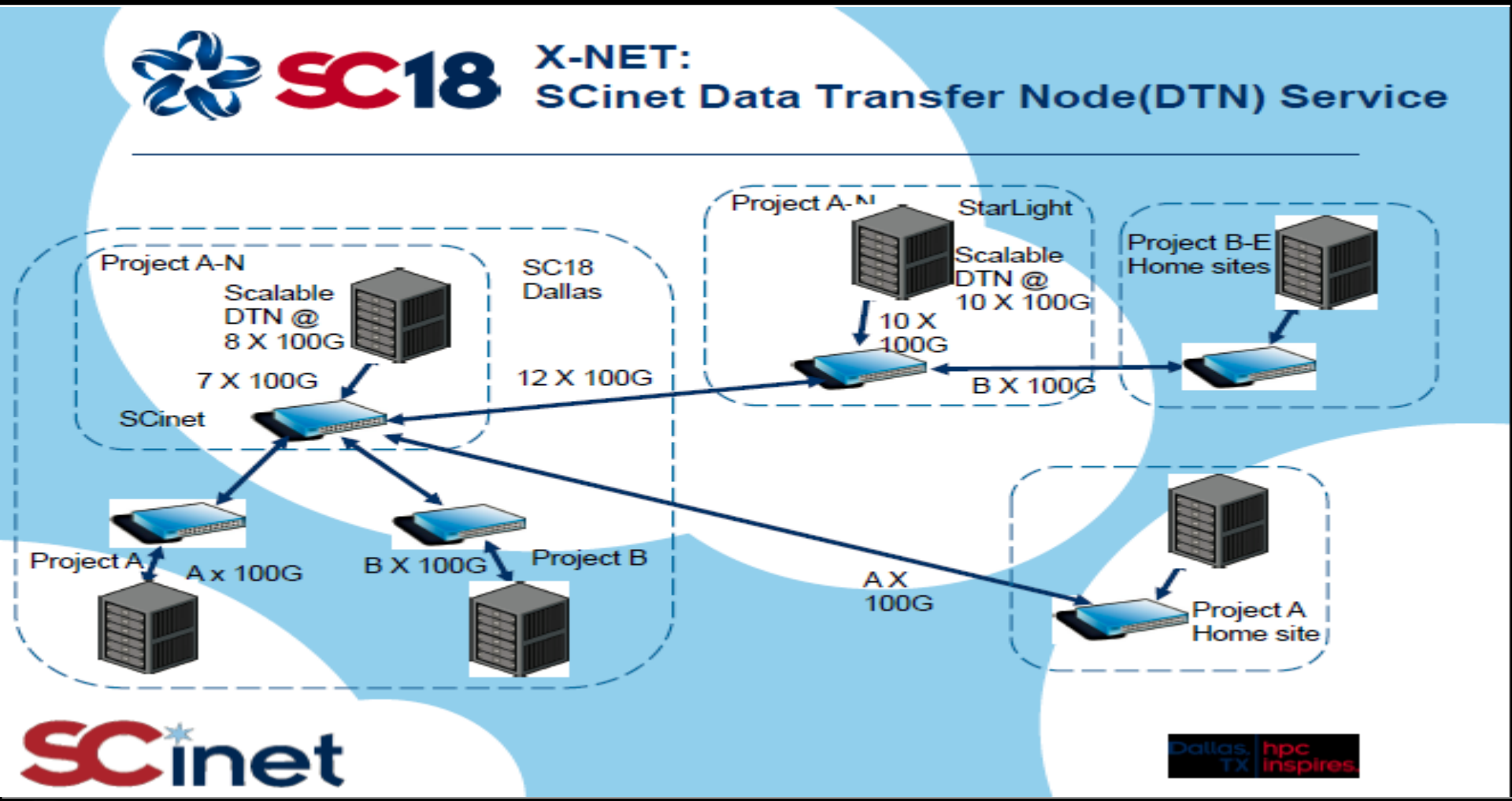
NASA/GSFC High End Computer Networking (HECN) Team
Diagram by Bill Fink / Paul Lang - 11/3/2019

Naval Research Laboratory
Center for Computational Science
SC19 Demonstration



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

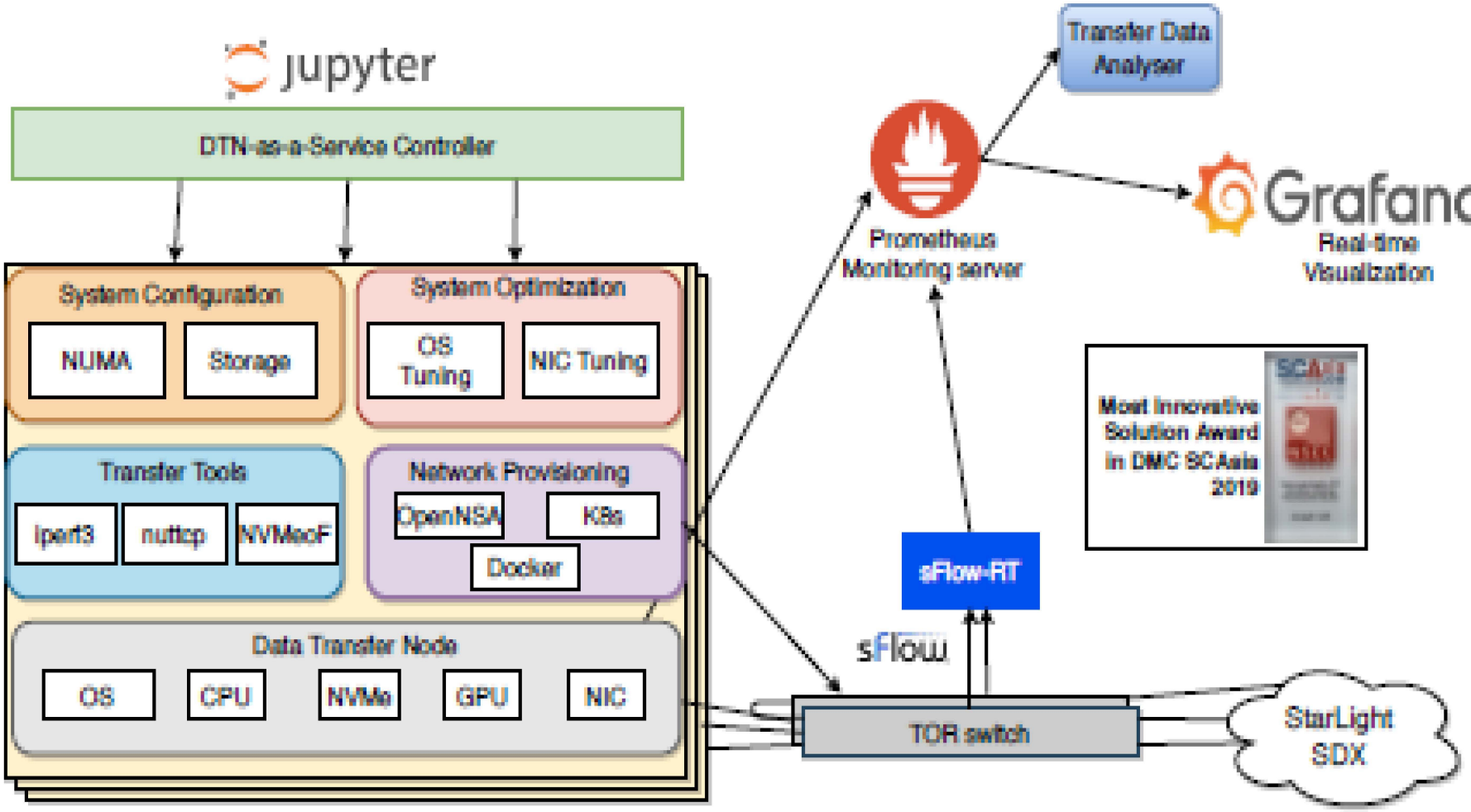
Implementing a SCinet DTN (DTN-as-a-Service, DaaS)



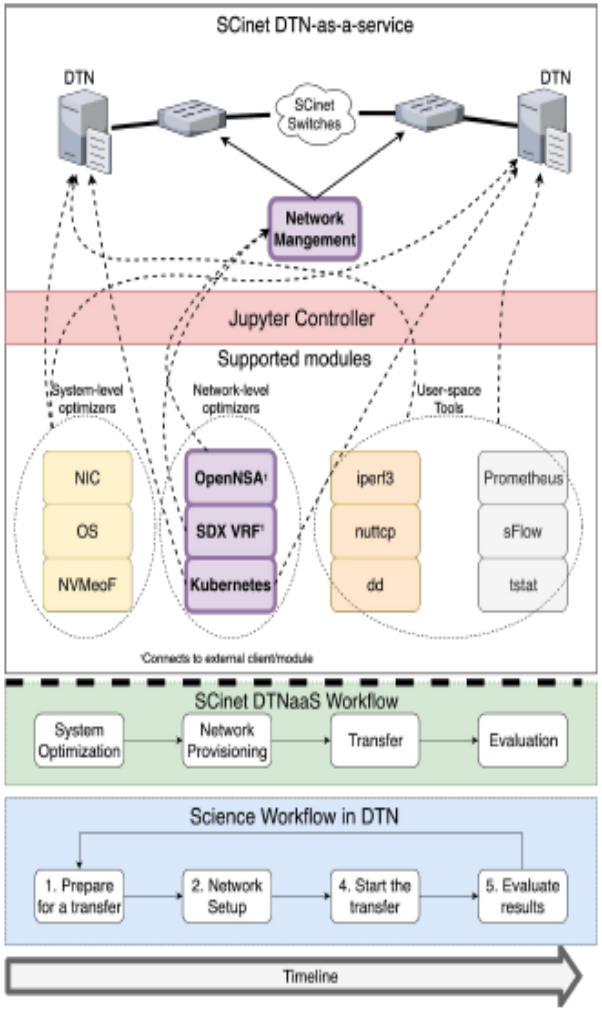
Source; Jim Chen, iCAIR

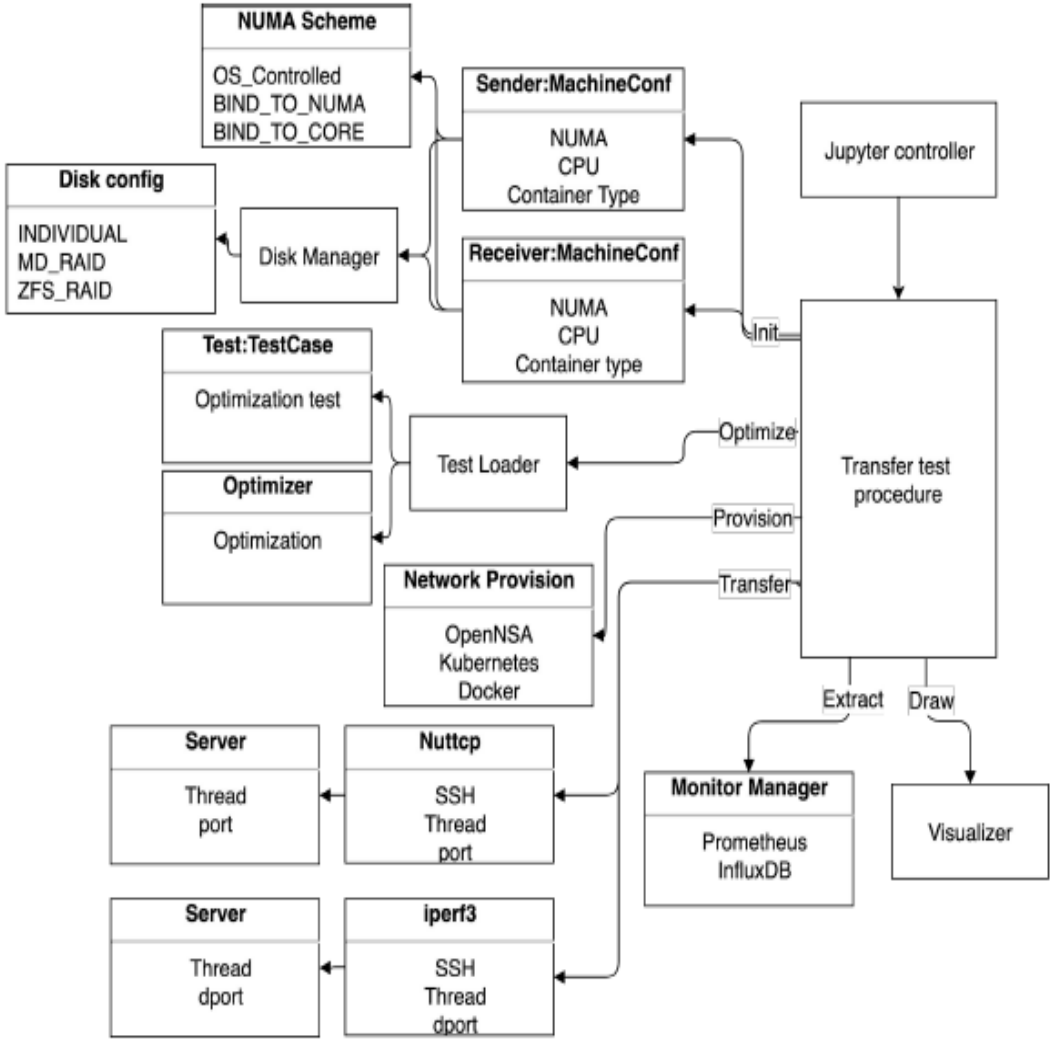
STARLIGHT SM

STARLIGHT SDX DTN-as-a-Service



Design and implementation

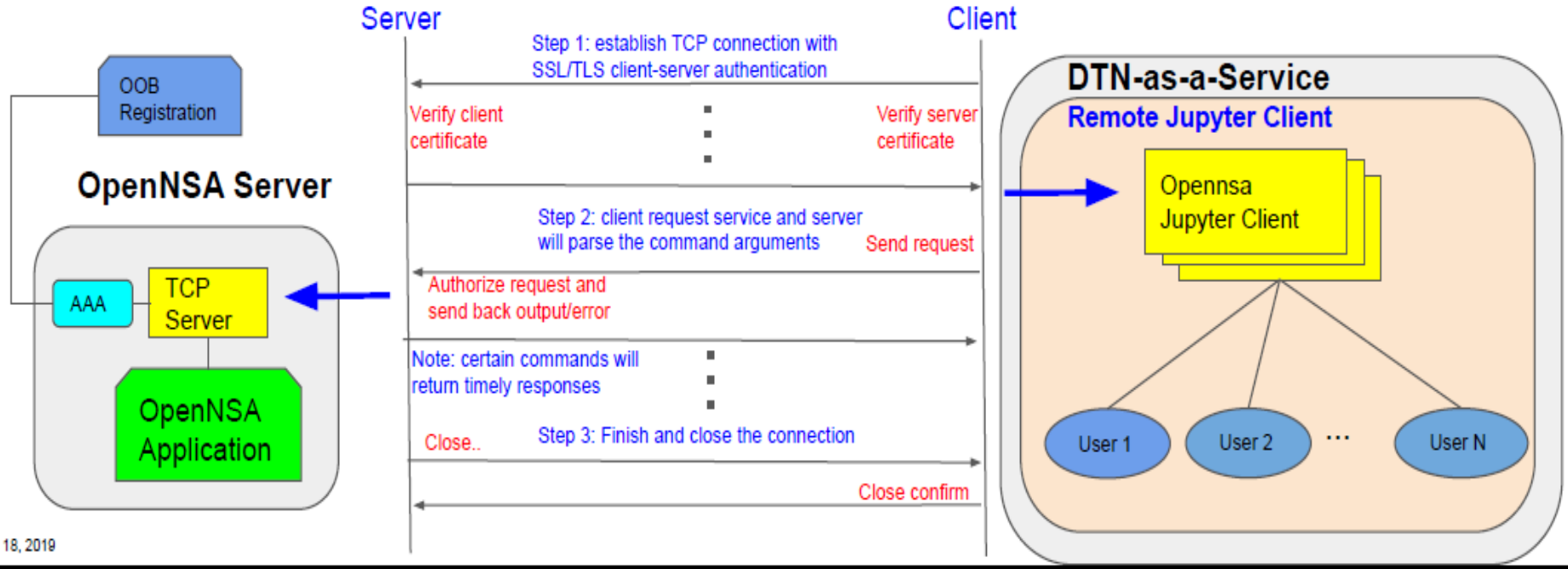




Jupyter Client for NSI OpenNSA Integration

Securely allow users to run NSI OpenNSA services(i.e. dynamically stitch layer 2/3 circuits based on technologies e.g. VLAN.)

- Features: **Authentication, Authorization, Accounting**
- **Authentication**: SSL/TLS operation authenticate both server and client. (certificates need to be exchanged prior unless using public certificates)
- **Authorization**: server will allow access users to request services on certain ports/VLANs based on user identification. I.e. Request command arguments will be parsed and authorized if allowed.
- **Accounting**: for future requirement.
- Additional Feature:Asynchronous(non-blocking operation allows multiple users to request services simultaneously)



Jan 18, 2019

Source: Se Young Yu, iCAIR

STARLIGHT SM

BigData Express: Toward Predictable, Schedulable, High-performance Data Transfer

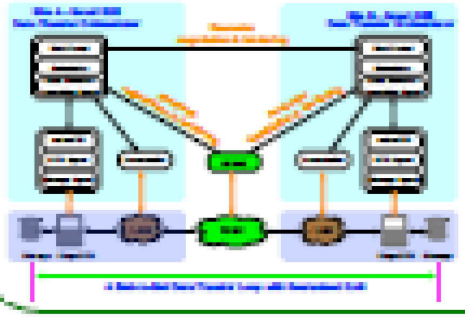
Fermilab

ENERGY

What Is BigData Express?

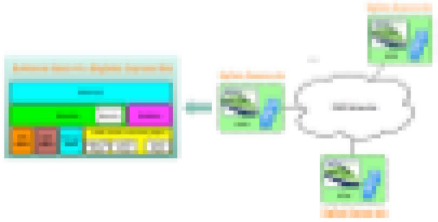
- A schedulable, predictable, and high-performance data transfer service
  - Peer-to-peer, scalable, and extensible data transfer model
  - Visually appealing, easy-to-use web portal
  - High-performance data transfer engine
  - On-demand provisioning of network paths with guaranteed QoS
  - Robust and flexible error handling
  - CILogon-based security
  - A rich set of REST APIs to support scientific workflows
- Project web site: http://bigdataexpress.fnal.gov

End-to-End Transfer Loop

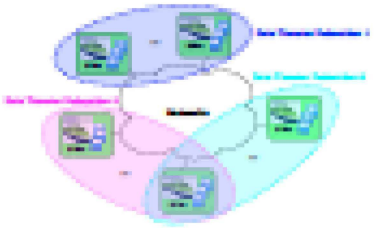


- Application-aware network service
  - On-demand programming
- Fast-provisioning of end-to-end network paths with guaranteed QoS
- Distributed resource negotiation & brokering

Distributed, Peer-to-Peer Model

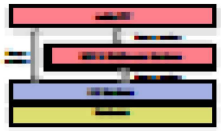


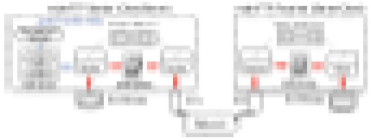
Each site independently provides data transfer service



Flexible data transfer federations

A High-performance Data Transfer Engine – mdtmFTP

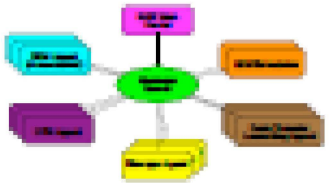




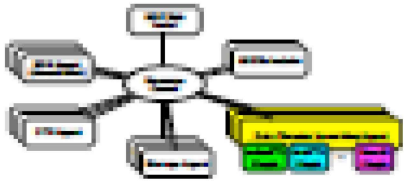


- A pipelined I/O centric design to streamline data transfer
- Multicore-aware data transfer middleware optimizes use of underlying hardware
- Extremely efficient in transferring of Lots Of Small Files

Scalable & Extensible Design



- Scheduler manages site resources through agents
- Use MQTT as message bus



- Extensible Plugin framework to support various data transfer protocols
  - mdtmFTP, GridFTP, XrootD

Miscellaneous

- Apache 2.0 License
- Docker release is available at publicregistry.fnal.gov
- Manuals:
  - Admin: https://bigdataexpress.fnal.gov/admin_manual/index.html
  - Web portal user: https://bigdataexpress.fnal.gov/portal_manual/index.html
- Deployment and collaborations

KISTI KREONET MAX ciena AutoLight KSTAR STARLIGHT NCI

STARLIGHT℠

A Cross-Pacific SDN Testbed

Fermilab

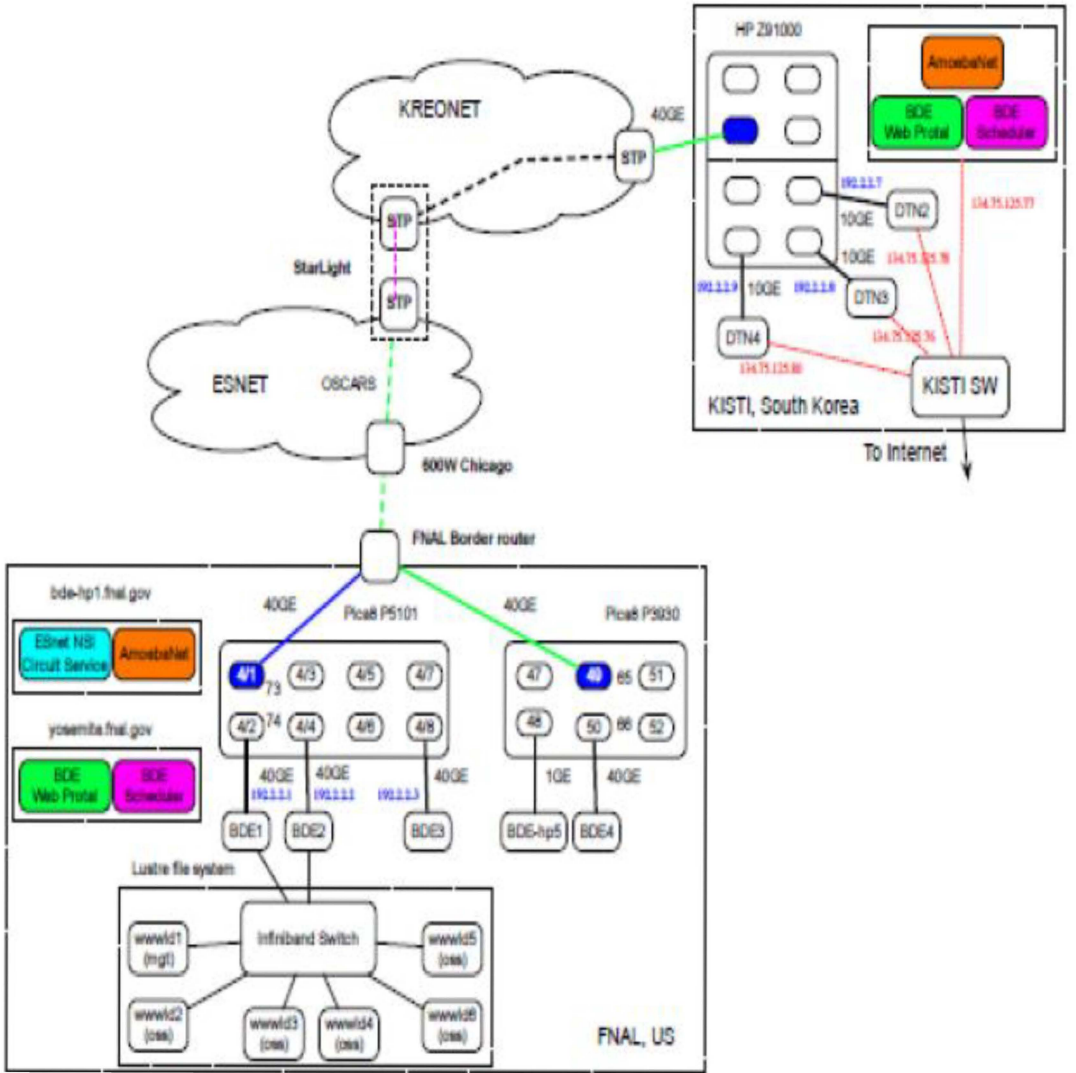
OAK RIDGE National Laboratory

ESnet

KREONET

KiSTi www.kisti.re.kr

STARLIGHT SM The Optical STAR TAP

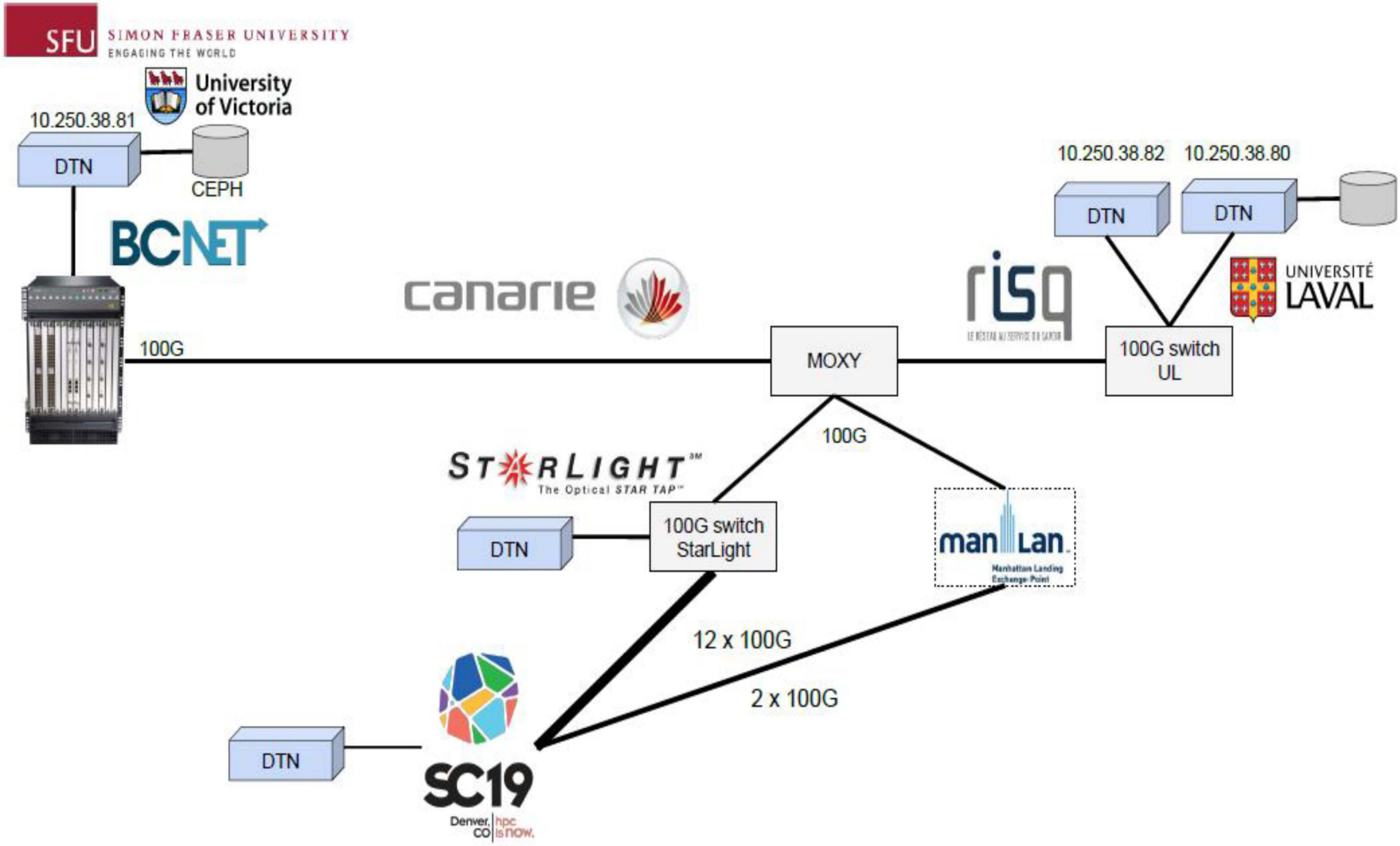


Argonne National Laboratory Advanced Photon Source

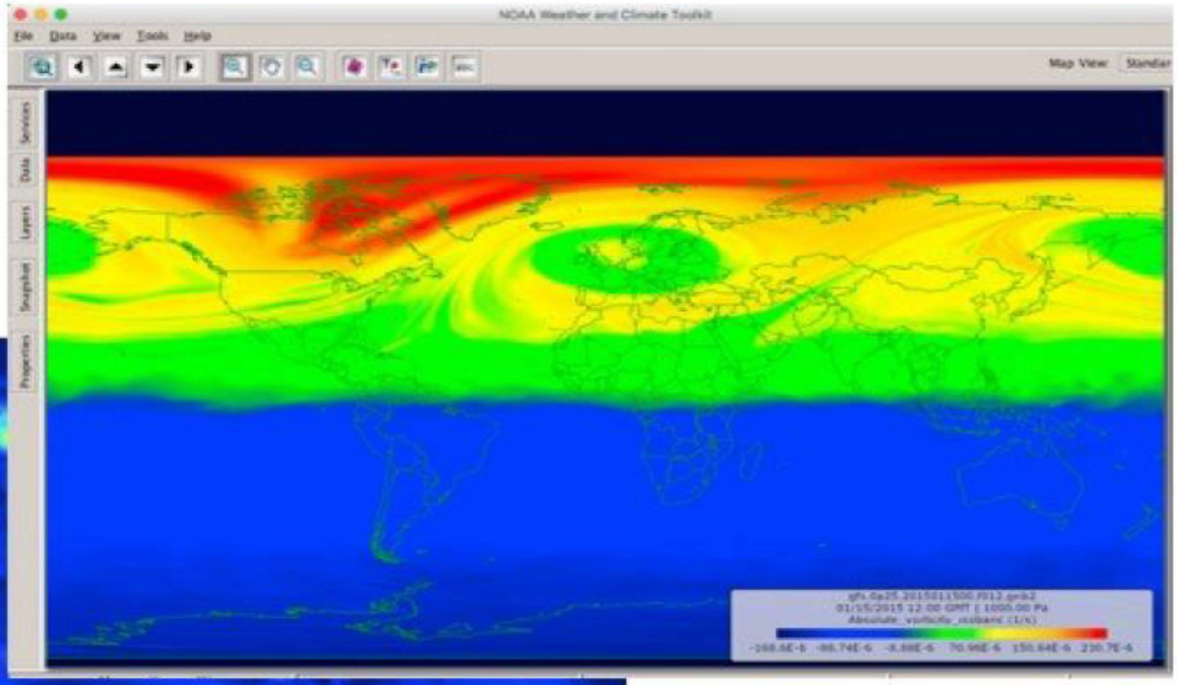


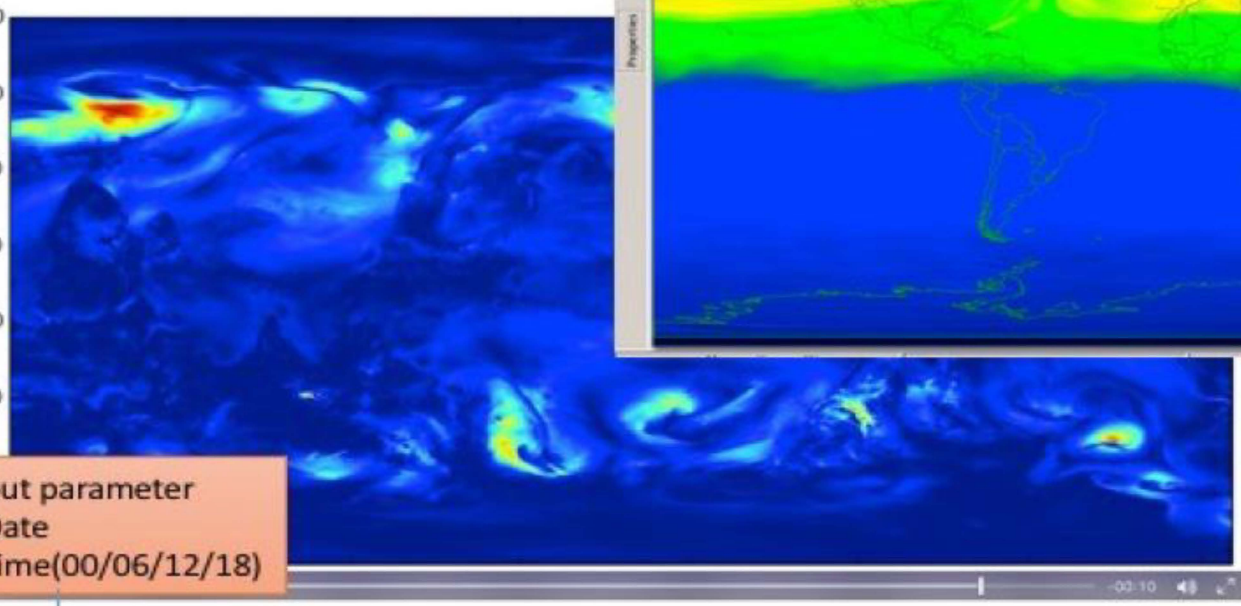
ANL Elastic Data Experiment Prototype Preparing for High Luminosity APS

STARLIGHT℠


SFU
SIMON FRASER UNIVERSITY
ENGAGING THE WORLD
University of Victoria
10.250.38.81
DTN
CEPH
BCNET
canarie
100G
MOXY
10.250.38.82
10.250.38.80
DTN
DTN
risq
UNIVERSITÉ LAVAL
100G switch UL
100G
STARLIGHT
The Optical STAR TAP
DTN
100G switch StarLight
manLan
Manhattan Landing Exchange Point
12 x 100G
2 x 100G
DTN
SC19
Denver, CO | hpc is now.

StarLight SDX Geoscience Research Workflow





Input parameter
1.Date
2.Time(00/06/12/18)

Download To SDX DTN → grib2 → Data convert1 → netCDF → Data convert2 → png → Convert to video → mp4 → Display → Select Data → Data Transfer To Other DTN

Composable Platform as a Service

Instrument for Deep Learning & Visualization (COMPaaS DLV)

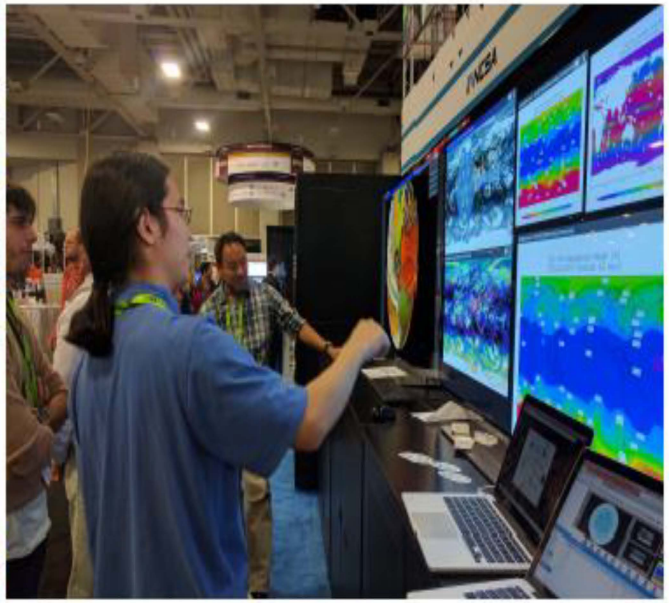
Using the *Liqid* Composable Infrastructure System for...



...on-demand post-processing, AI, deep-leaning, CFD, and visualization.



...interoperable Science DMZs and Data Transfer Nodes (DTNs) at a national/ international scale, using best strategies from the Pacific Research Platform (PRP) over the GLIF infrastructure.



...SAGE2™, the *de facto* operating system for managing Big Data on tiled-display walls, enabling local and remote collaborators to access, share and interact with a variety of digital media.

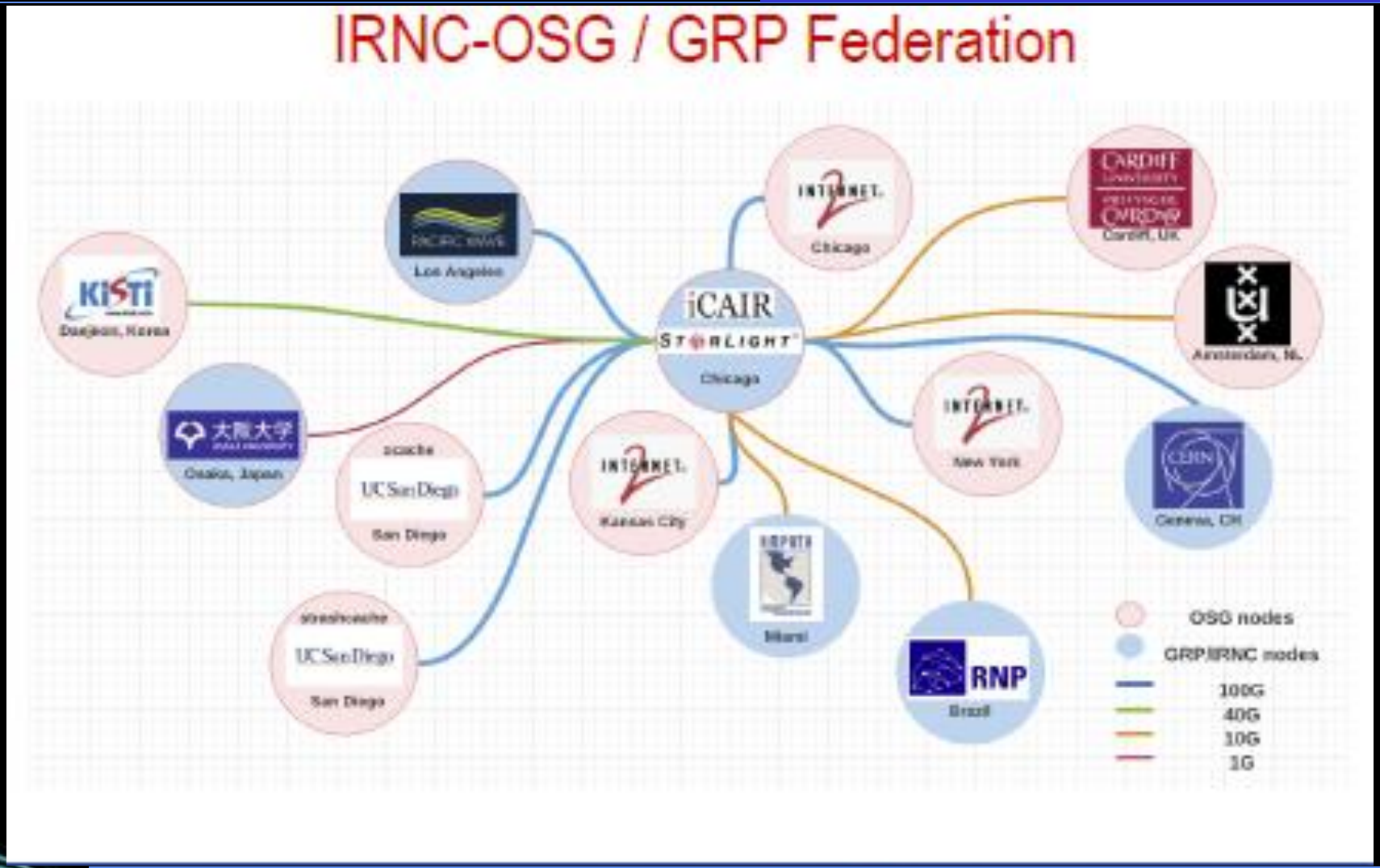
evl
Electronic Visualization Laboratory
University of Illinois at Chicago
https://www.evl.uic.edu

NSF award CNS-1828265 to UIC/EVL for COMPaaS DLV
https://www.evl.uic.edu/compaas-dlv

SAGE2 is trademark of the University of Illinois Board of Trustees

Liqid: https://www.liqid.com/
PRP: https://prp.ucsd.edu
GLIF: https://www.glif.is
SAGE2: http://sage2.sagecommons.org/

IRNC-OSG Global Fabric



International Federated Testbeds As Instruments for Computer Science/Network Science

- The StarLight Communications Exchange Facility Supports ~ 25-30 Network Research Testbeds (Instruments For Computer Science/Networking Research)
- Software Defined Networking Techniques Are Used To Segment ("Slice") Distributed Virtual Environments
- StarLight Supports Two Software Defined Exchanges (SDXs), An NSF IRNC SDX & a Network Research GENI SDX (Global Environment for Network Innovations)
- The GENI SDX Supports National and International Federated Testbeds



Chameleon

www.chameleoncloud.org

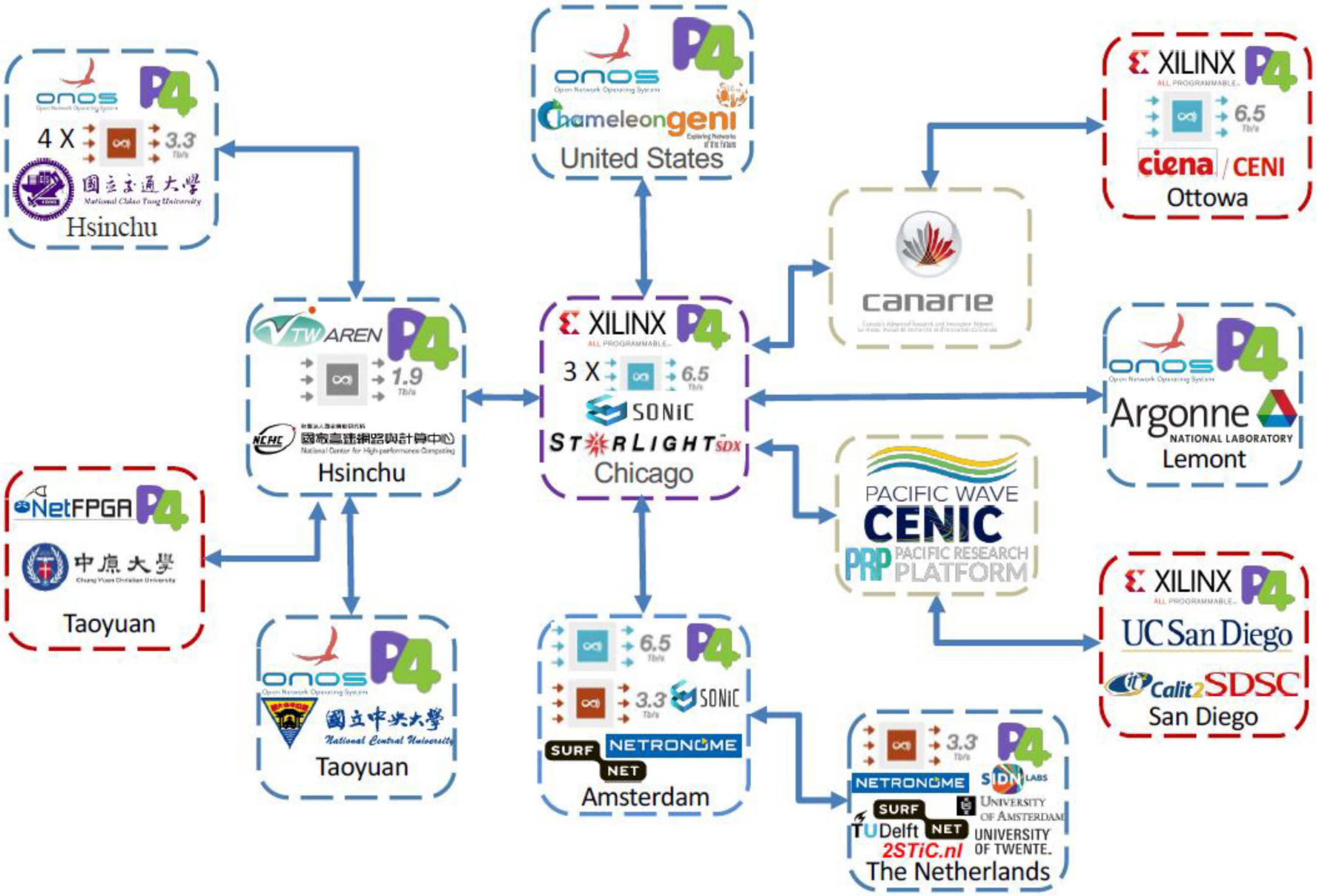
CHAMELEON: A LARGE SCALE, RECONFIGURABLE EXPERIMENTAL INSTRUMENT FOR COMPUTER SCIENCE

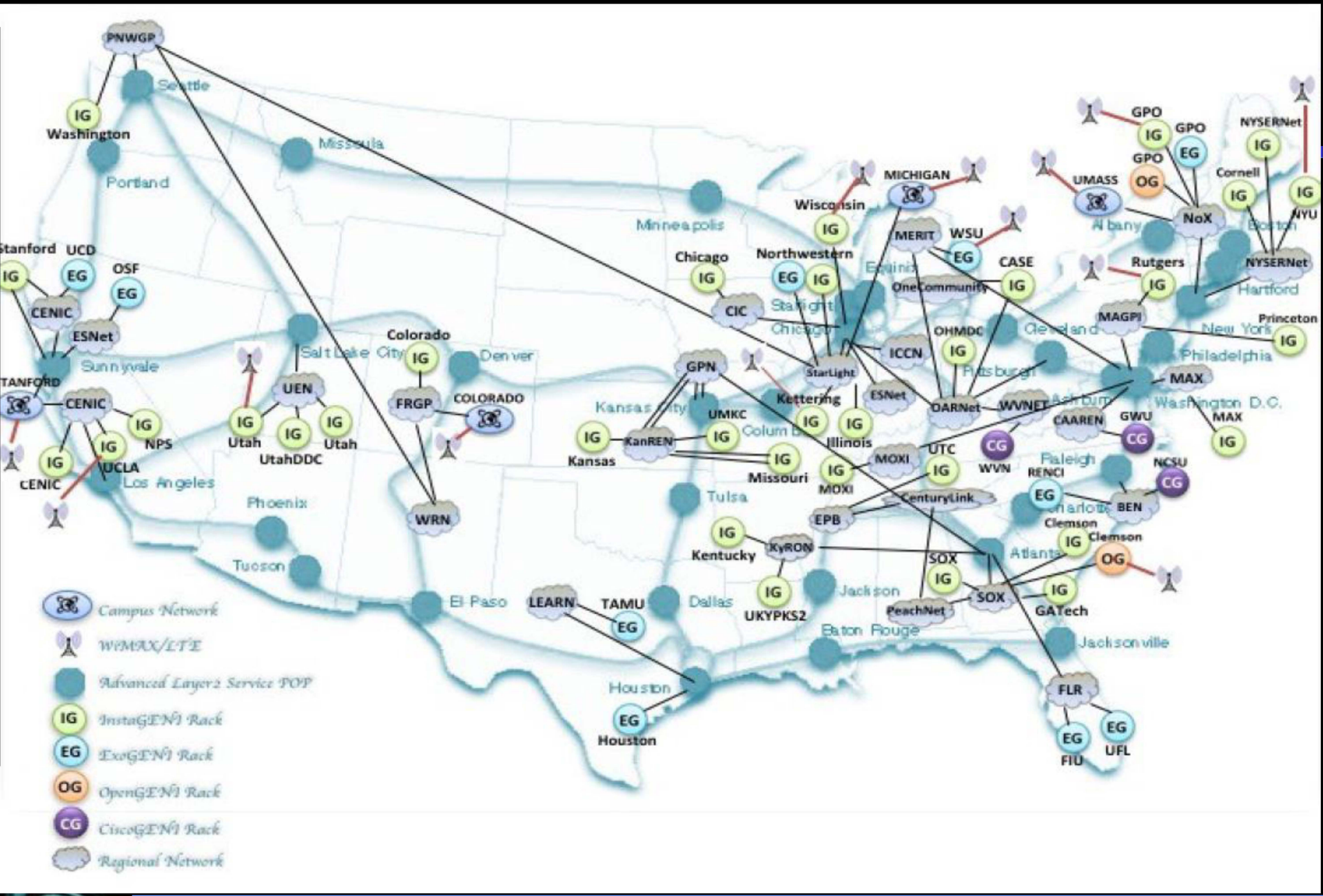
Kate Keahey

Joe Mambretti, Pierre Riteau, Paul Ruth, Dan Stanzione

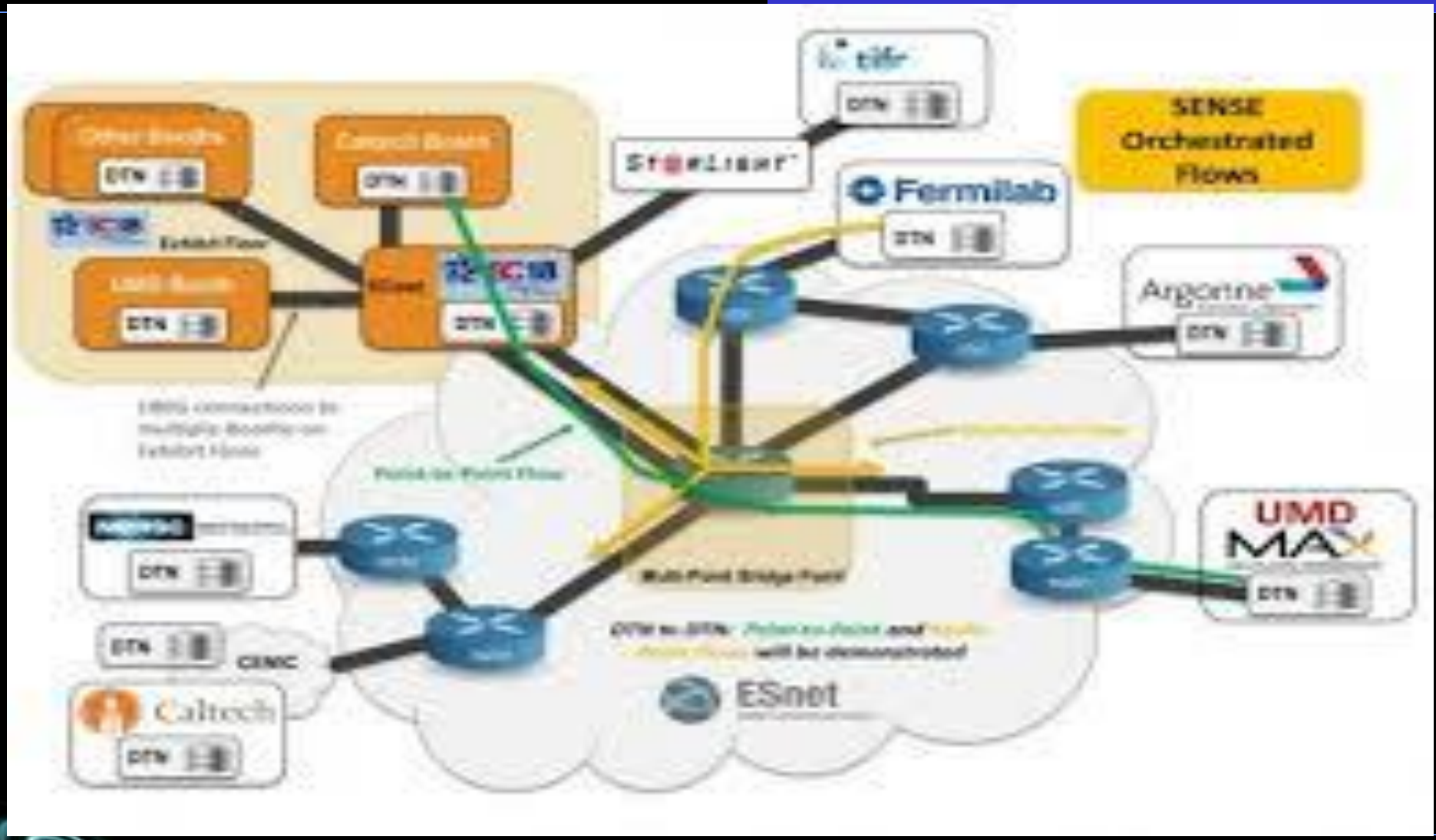
THE UNIVERSITY OF CHICAGO TACC renci NORTHWESTERN UNIVERSITY NSF

International P4 Experimental Networks (i-P4EN)

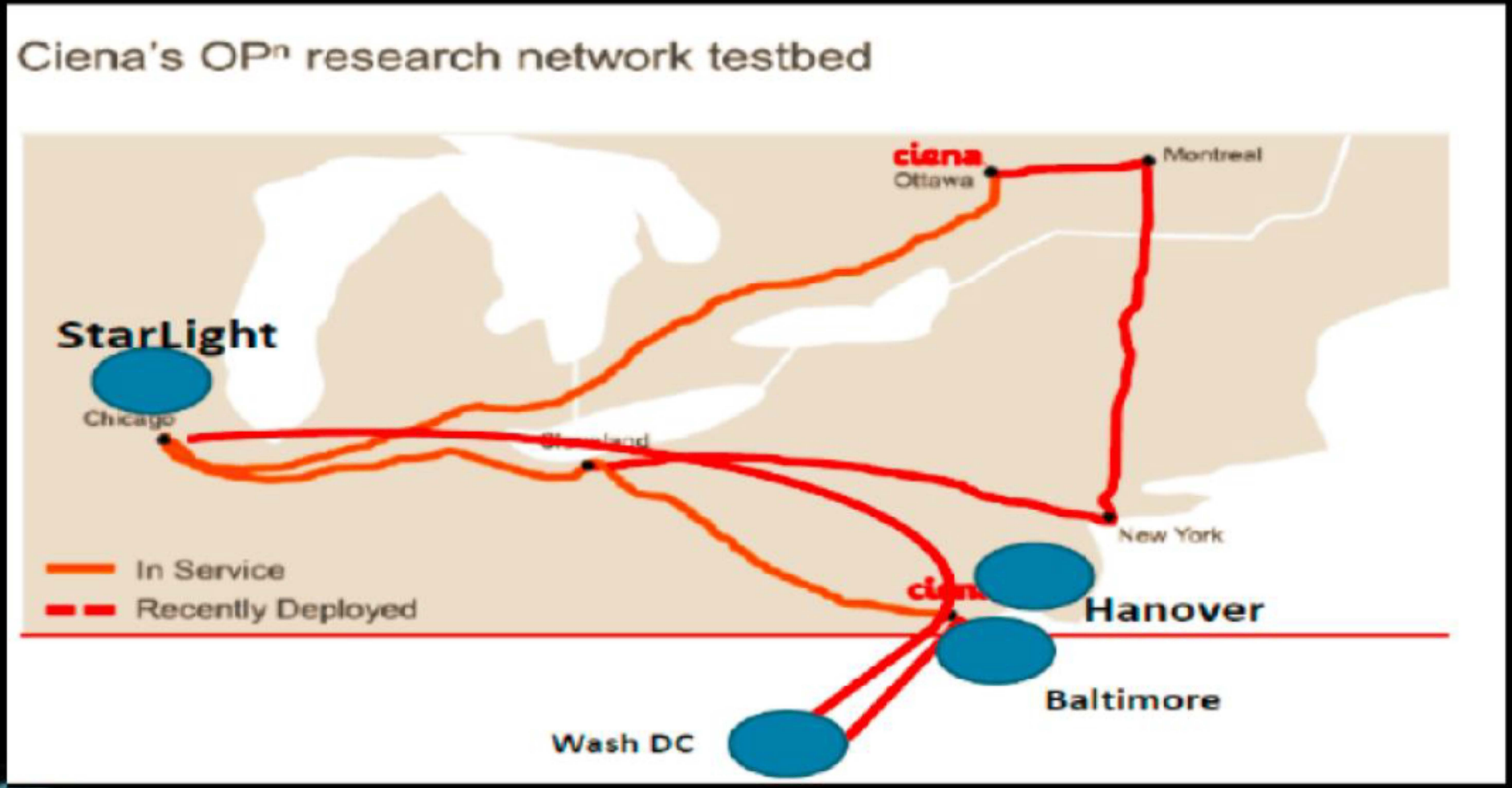



PNWGP
Seattle
IG
Washington
Portland
Missoula
Minneapolis
Stanford
UCD
IG
EG
OSF
EG
CENIC
ESNet
Sunnyvale
TANFORD
CENIC
IG
NPS
IG
UCLA
cENIC
Los Angeles
Salt Lake City
UEN
IG
Utah
IG
UtahDDC
IG
Utah
Colorado
IG
FRGP
COLORADO
Denver
WRN
Phoenix
Tucson
El Paso
LEARN
TAMU
EG
Dallas
Houston
EG
Houston
Wisconsin
IG
MICHIGAN
Chicago
IG
Northwestern
EG
IG
CIC
StarLight
Chicago
MERIT
WSU
EG
CASE
IG
OneCommunity
GPN
IG
Kansas
KanREN
UMKC
IG
IG
Missouri
Kettering
IG
Illinois
ESNet
ICCN
OHMDC
IG
StarLight
DARNet
WVNET
MOXI
IG
MOXI
UTC
IG
CG
WVN
EPB
CenturyLink
IG
Kentucky
XyRON
IG
UKYPKS2
SOX
IG
PeachNet
SOX
IG
GATech
Atlanta
Jackson
Baton Rouge
Jacksonville
FLR
EG
FIU
EG
UFL
GPO
IG
GPO
EG
GPO
OG
UMASS
Albany
NoX
NYSERNet
IG
Cornell
IG
IG
NYU
Boston
NYSERNet
Rutgers
IG
Hartford
MAGPI
Princeton
IG
New York
Philadelphia
MAX
Washington D.C.
MAX
IG
CAAREN
GWU
CG
Ashburn
Raleigh
RENCI
EG
NCSU
CG
BEN
Clemson
IG
Clemson
OG
IG
Campus Network
WiMAX/LTE
Advanced Layer2 Service POP
InstaGENI Rack
ExoGENI Rack
OpenGENI Rack
CiscoGENI Rack
Regional Network

SENSE Orchestration Testbed



100 Gbps DTN Optical Testbed

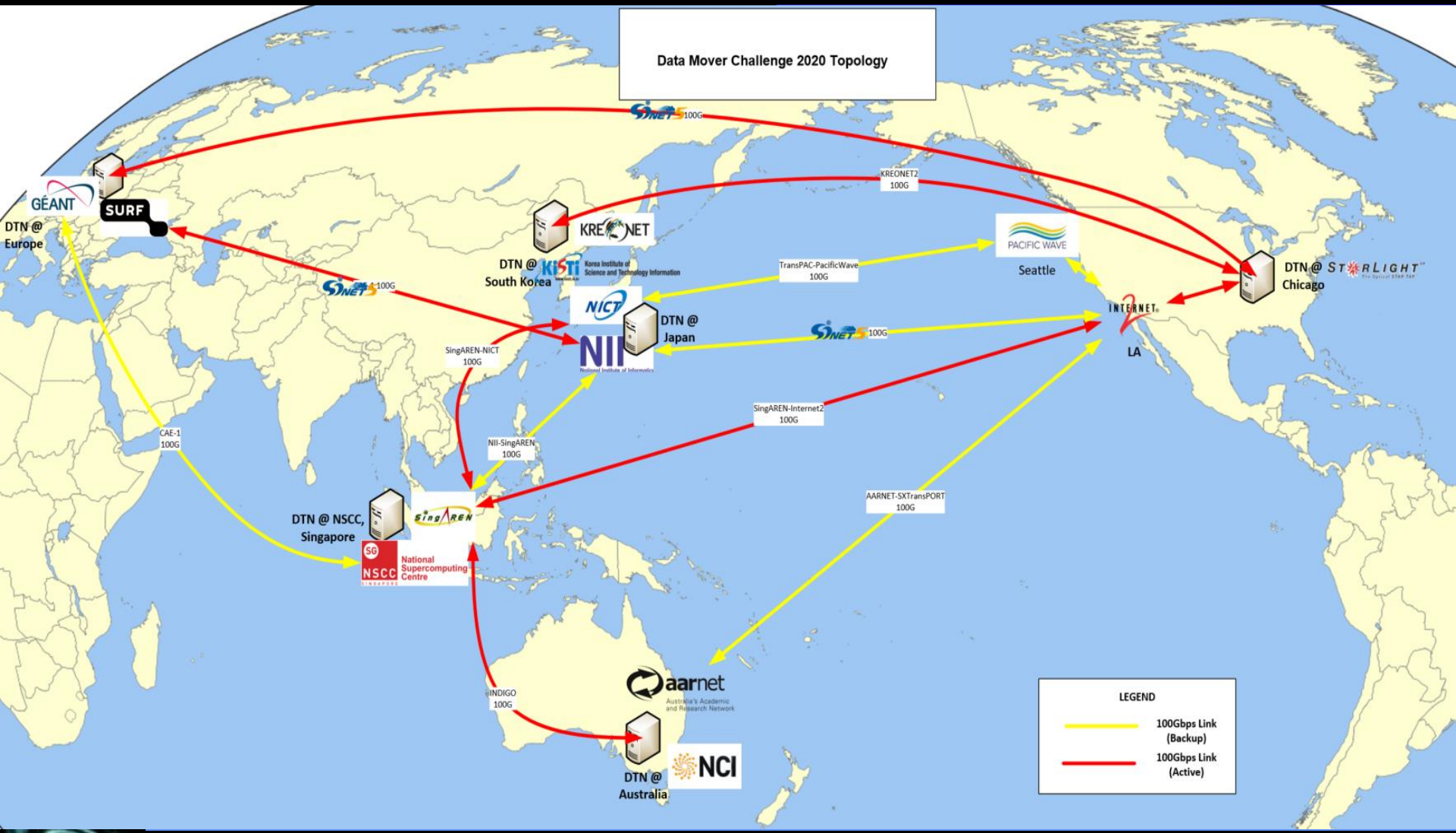


Future Directions

- SCA 2020
- Quantum Communications and Quantum Networking
- Tbps Networks
- Global Research Platform



Supercomputing Asia Data Mover Challenge 2020



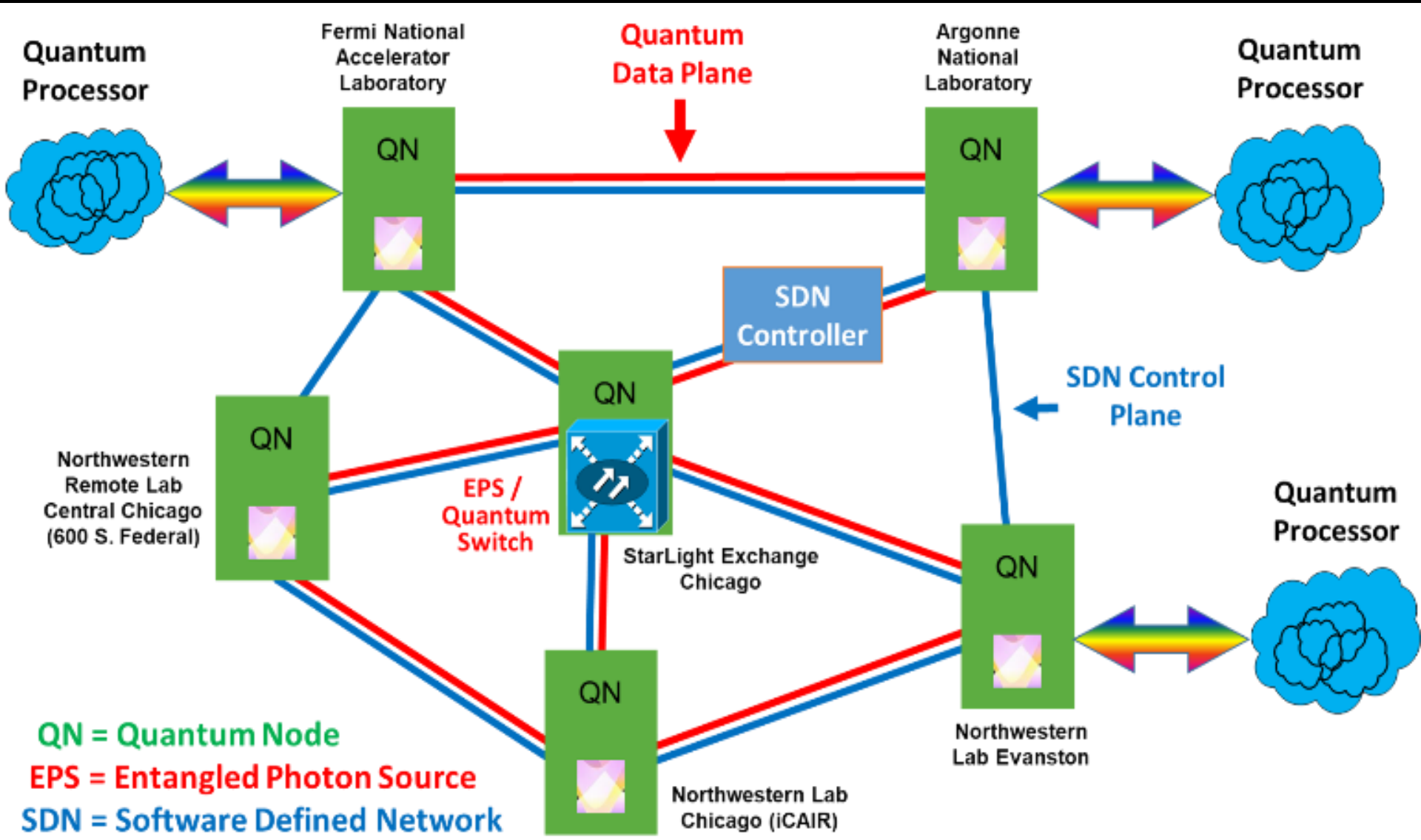
STARLIGHT SM

National Quantum Science Initiative

- Multiple Federal Agency Initiatives
- Multiple Consortium Projects
- Multiple Topics
  - Physics
  - Quantum Science
  - Information Theory
  - Information Science
  - Optimization Theory
  - Materials Science
  - Quantum Components
  - Quantum Computing
  - Quantum Communications
  - Quantum Networking
- Multiple Projects: Chicago Quantum Exchange, Illinois Quantum Express Et Al



Planned Chicago Quantum Exchange Testbed



Ilya Baldine PI, RENCI: FABRIC



Next Generation Distributed Enviroment For Global Science



STARLIGHT℠


PACIFIC WAVE
INTERNATIONAL PEERING EXCHANGE
Global Research Platform Network
100 Gbps
StarLight
Global Research Platform

www.startap.net/starlight



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

STARLIGHT℠

"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

