Global Research Platform Workshop: An Overview

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

Global Research Platform Workshop
Co-Located With Supercomputing Asia Conference
Singapore
February 27- March 2, 2023

Next Generation Distributed Environment For Global Science

STARLIGHT℠

NSF's Cyberinfrastructure Framework for the 21st Century (CIF21)

- ***"Across the full range of NSF--supported fields increasingly sophisticated instrumentation and expanded computational resources are opening new windows onto phenomena from the universe to the human brain, from the largest scales to the smallest. Across all domains, data play the key role in a profound transformation of the culture and conduct of science and society.***
- ***This Revolution Will Transform Research, Practice, And Education In Science and Engineering As Well As Advance Innovation In Society***
- ***This vision of the near future shows clearly the urgent need for a comprehensive, scalable, cyberinfrastructure that bridges diverse scientific communities and integrates high--performance computing, data, software, and facilities in a manner that brings theoretical, computational, experimental, and observational approaches together to advance the frontier."***

Global Collaborative Research Communities

- Science Is Global
- Open Information Sharing, A Cornerstone of The Science Process Is A Key Motivation For This Forum
- Concepts, Experiments, Instruments, Methods, Techniques, Data, Technologies And Results Are Openly Communicated and Shared Among Collaborative Science Communities World-Wide
- The Global Research Platform Is An International Collaborative Partnership Creating A Distributed Environment for International Data Intensive Science
- The GRP Facilitates High Performance Data Gathering, Analytics, Transport (100 Gbps-Tbps E2E), Computing, And Storage
- www.theglobalresearchplatform.net

Note On Science

- Although Not A Perfect Path To Knowledge, Science Is The Best Path, A Tradition Of Continuous Curiosity, Questioning, Openess
- In A Crisis, Society Depends On Science As A Source Of Truth, As A Source Of Solutions
- However, Fundamental Science Pursues Knowledge For Its Own Sake, Not A Mere Utilitarian Tool
- Fundamental Science Has A Long Term vs Short Term Horizon
- "The Really Profound Changes In Human Life All Have Their Ultimate Origins In Knowledge Pursued For Its Own Sake" *Alfred North Whitehead, Mathematician, Logician, Philosopher*
- The Four Pillars Of Science Motivate Advanced Cyberinfrastructure:
  - Theory
  - Experimentation
  - Modeling/Simulation
  - Data Analytics

Cyberinfrastructure Ecosystem Design and Implementation

- Projection/Definition of Future Requirements, Architecture, Services, Techniques, Technologies, Processes Described In Cyberinfrastructure "Blueprints"
- Addressing Cambrian Explosion Of Requirements and Innovations
- Techniques and Technologies Emerge from Multiple Sources
  - a Academic Research
  - b Commercial Technology Research
  - c Government Labs
  - d Utilitarian Imperatives (e.g., Commercial Clouds)
  - e Specialized Requirements and Ecosystems
- Challenge Of Selecting Themes And Topics
- Macro-Trend: "Software Eating The World" - Software Defined Everything
- ***Multiple Software Building Blocks For Data-Intensive Science (Modules/Components) Are Emerging: Including Those Highlighted In This Workshop***

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

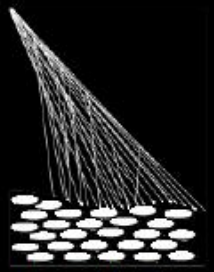
Virgo
www.virgo-gw.eu

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

SKA
www.skatelescope. org

LIGO
www.ligo.caltech. edu

USGS EROS
www.usgs.gov/ centers/eros

Blue Waters
bluewaters.ncsa. illinois.edu

PRAGMA
www.pragma-grid.net

NEON
www.neonscience. org

CENTRA
www.global centra.org

SAGE2
sage2.sagecommons. org

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

Jetstream
www.jetstream-cloud.org

Chameleon
www.chameleon cloud.org

PRP
pacificresearch platform.org

Genomic Science Program
genomicscience. energy.gov

Belle II
www.belle2.org

LBNF/DUNE/ ProtoDUNE
lbnf.fnal.gov

XENON
xenon.astro. columbia.edu

NOvA
novaexperiment. fnal.gov

SDSS
www.sdss.org

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

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

ALMA
www.alma observatory.org

IVOA
www.ivoa.net

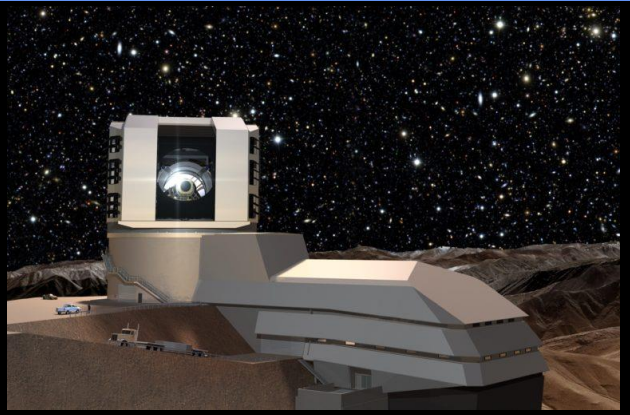
Compilation by Maxine Brown and Joe Mambretti

STARLIGHT℠

Instruments: Exebytes Of Data

High Luminosity LHC

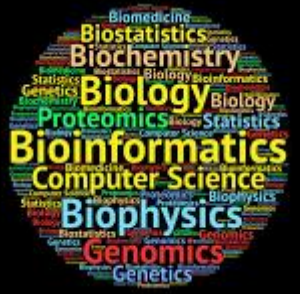
SKA Australia Telescope Facility



Vera Rubin Observatory

KSTAR Korea Superconducting Tokamak

Next Gen Advanced Photon Source

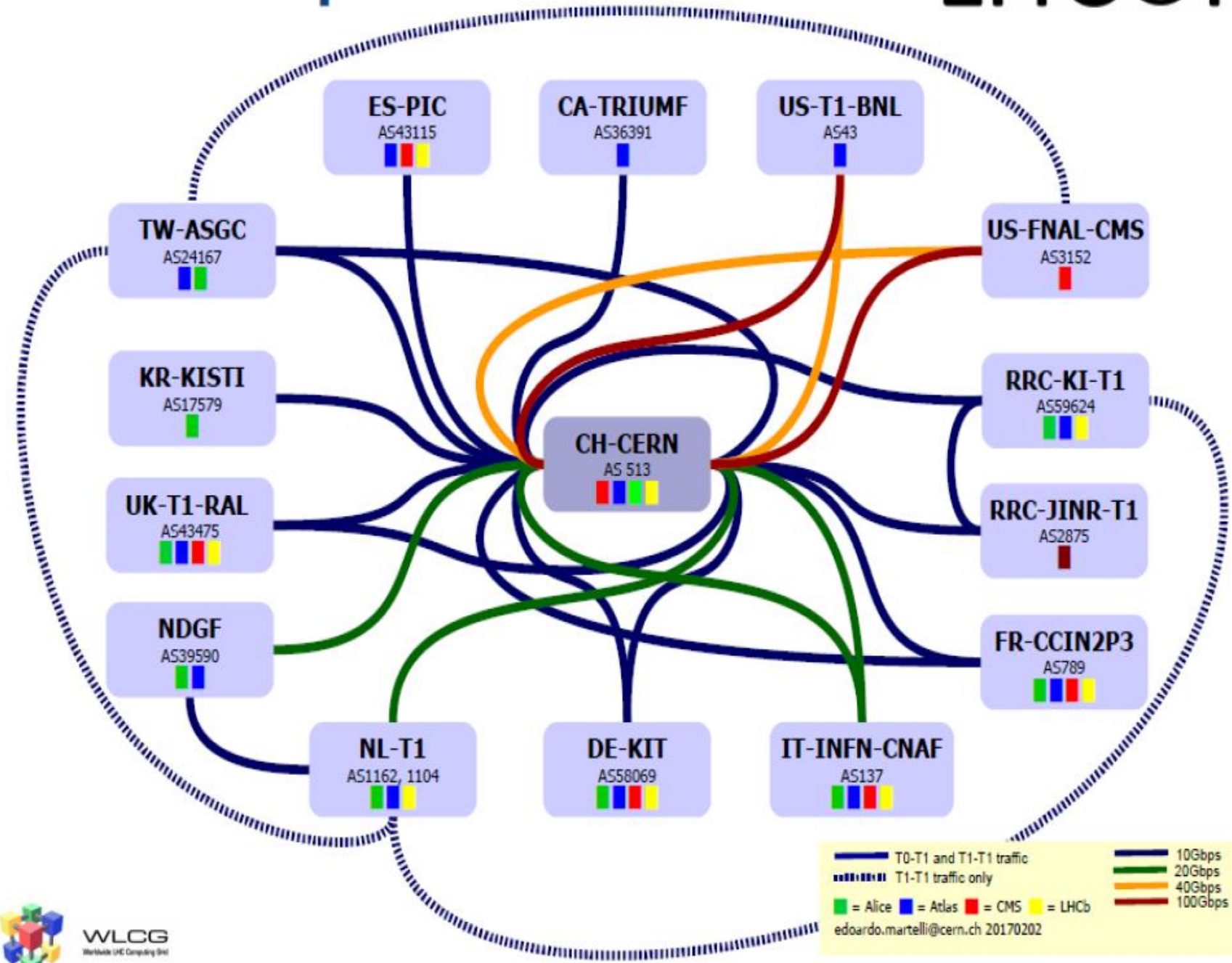


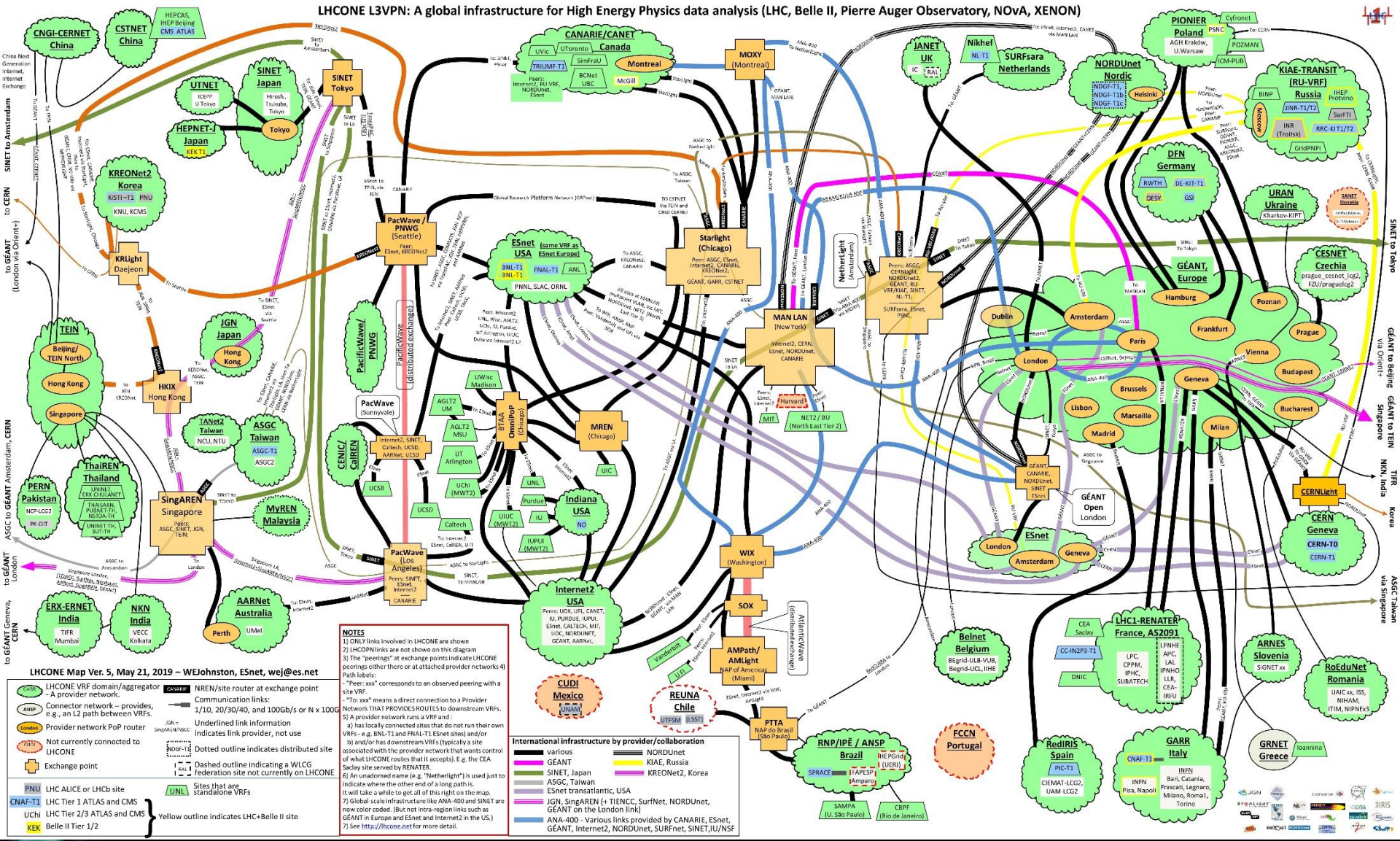
Bioinformatics/Genomics

Theme 1: Large Scale Global Science

- Science Domains Create Cyberinfrastructure Ecosystems, Some Distributed World Wide, Some Devoted To Domains, Some Shared Among Domains
- Minimal Opportunities For Information Sharing On Cyberinfrastructure Architecture, Implementation, Technologies and Operations Among Projects
- Such Opportunities Are Especially Useful For Cross Disciplinary Research
- Example Ecosystem: HEP - LHC

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



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

Theme 2: Next Generation Research Platforms

- *“a comprehensive, scalable, cyberinfrastructure that bridges diverse scientific communities and integrates high--performance computing, data, software, and facilities in a manner that brings theoretical, computational, experimental, and observational approaches together to advance the frontier”*
- **Large Scale Science DMZs**
- **Super Facilities**
- **National Research Platforms**
- **Continental Research Platforms**

Theme 3: Orchestration Among Multiple Domains

- Instrumentation and Analytic, Storage Resources Are Highly Distributed Among Multiple Domains Interconnected With High Performance Networks
- A Key Issues Is Discovering Resources, Claiming Them, Integrating Them, Utilizing Them and Releasing Them
- Increasingly, New Software Defined Infrastructure Architecture, Services, Techniques And Technologies Are Addressing These Issues

Theme 4: Large-Scale High Capacity Data WAN Transport

- Large-Scale High Capacity Data WAN Transport Has Always Been And Remains A Major Challenge, Especially Over Global Paths
- This Issue Is Emphasized By A Next Generation Of Instrumentation That Will Generate Exponentially Large Volumes Of Data That Has To Be Distributed Across the Globe
- Often, This Issue Is Considered Reductively Only In Terms Of Network Capacity
- However, Actually It Is More An E2E Issue, Especially Given Advances In Core Optical Networking Technologies

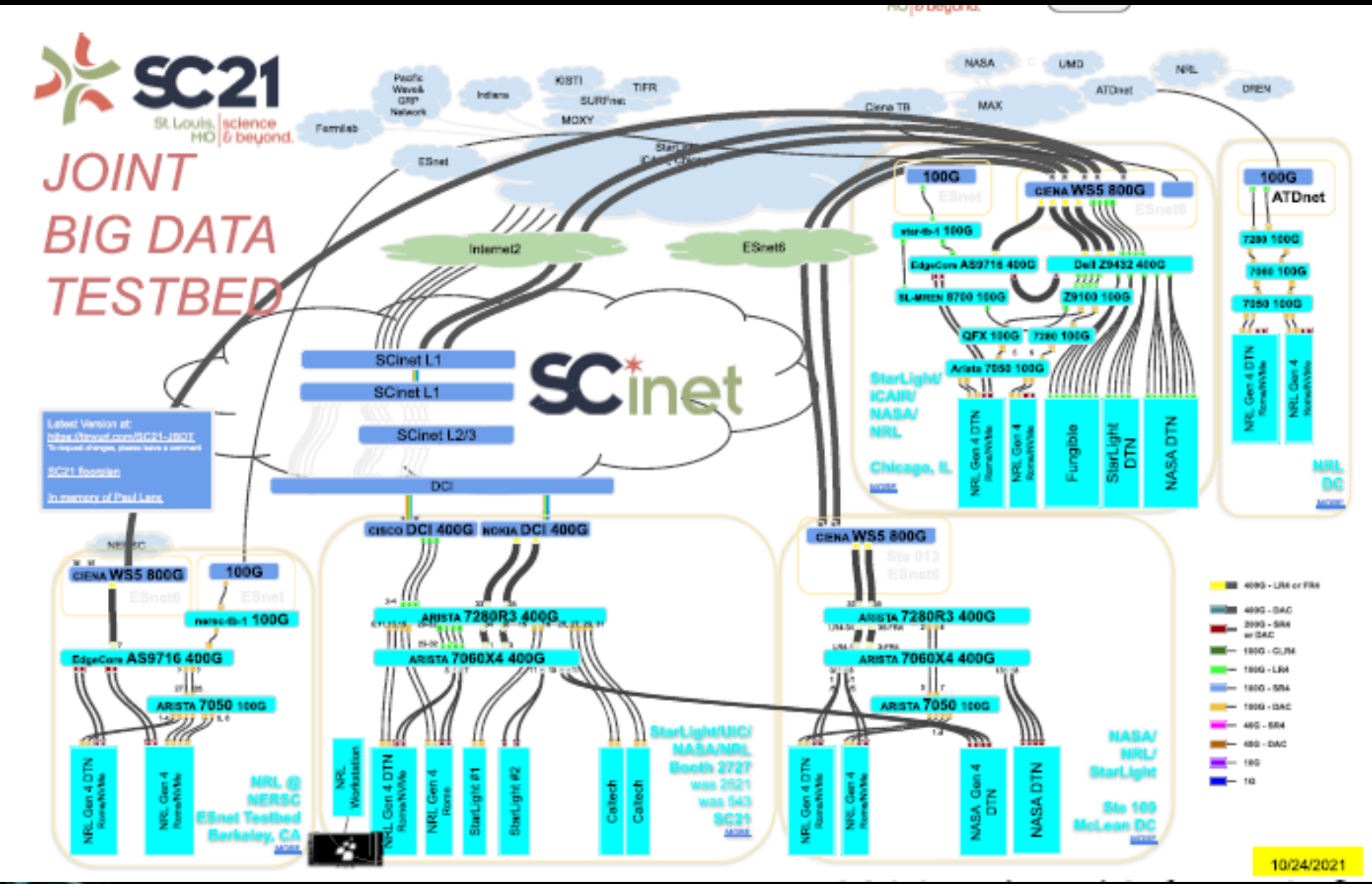
Theme 5: High-Fidelity Data Flow Monitoring, Visualization, Analytics, Diagnostic Algorithms, Event Correlation AI/ML/DL

- A Major Opportunity For Data Transport Optimization Is Being Provided By New Methods For Directly Detecting And Analyzing All Data Flows And Their Characteristics
- Because These Techniques Enable High-Fidelity Views Of All Flows, Real Time, Dynamic Traffic Engineering Is Possible With Much More Sophistication Than Traditional Approaches
- These Techniques Can Be Significant Enhanced Using AI/ML/DL, Which (Although Still Emerging) Are Becoming Critically Important Tools In The Near Term

Theme 6: International Testbeds for Data-Intensive Science

- Given The Challenging Requirements Of Anticipated Large Scale Science Projects Along With Accelerated Rates Of Ongoing Innovation, International Testbeds Are Required for Pre-Production Investigations And Prototyping Of New Technologies And Techniques Specifically Related To Data Intensive Science
- Such Global Experimental Research Testbeds Exist Today, And They Are Being Developed With Enhanced Capacities, Sites, And Capabilities

400 Gbps, 800 Gbps, 1.2 Tbps Paths

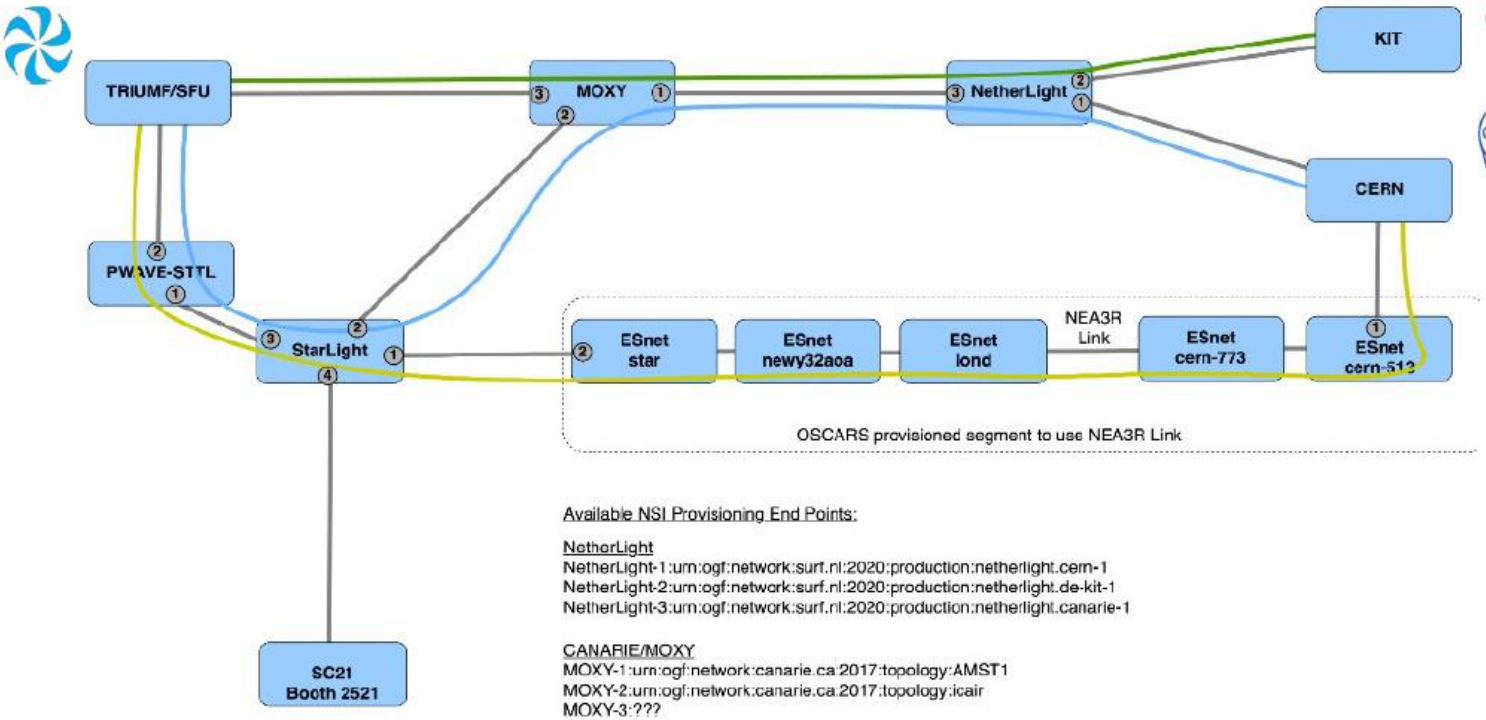


Source: Linden Mercer

STARLIGHT SM

CERN: Network Optimized for the Transport of Experimental Data (NOTED)

Dynamic circuits



Available NSI Provisioning End Points:

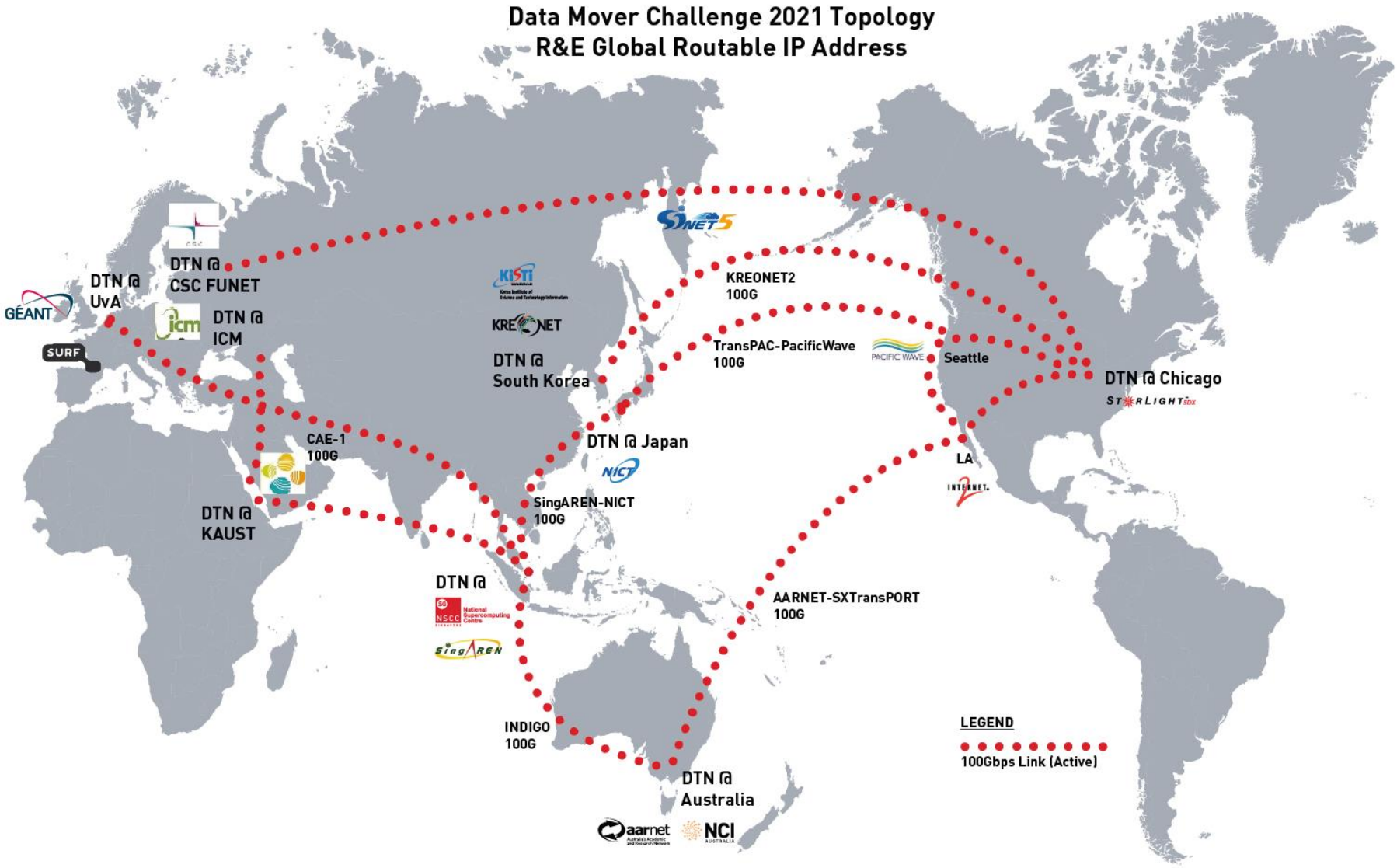
NetherLight
NetherLight-1:urn:ogf:network:surf.nl:2020:production:netherlight.cern-1
NetherLight-2:urn:ogf:network:surf.nl:2020:production:netherlight.de-kit-1
NetherLight-3:urn:ogf:network:surf.nl:2020:production:netherlight.canarie-1

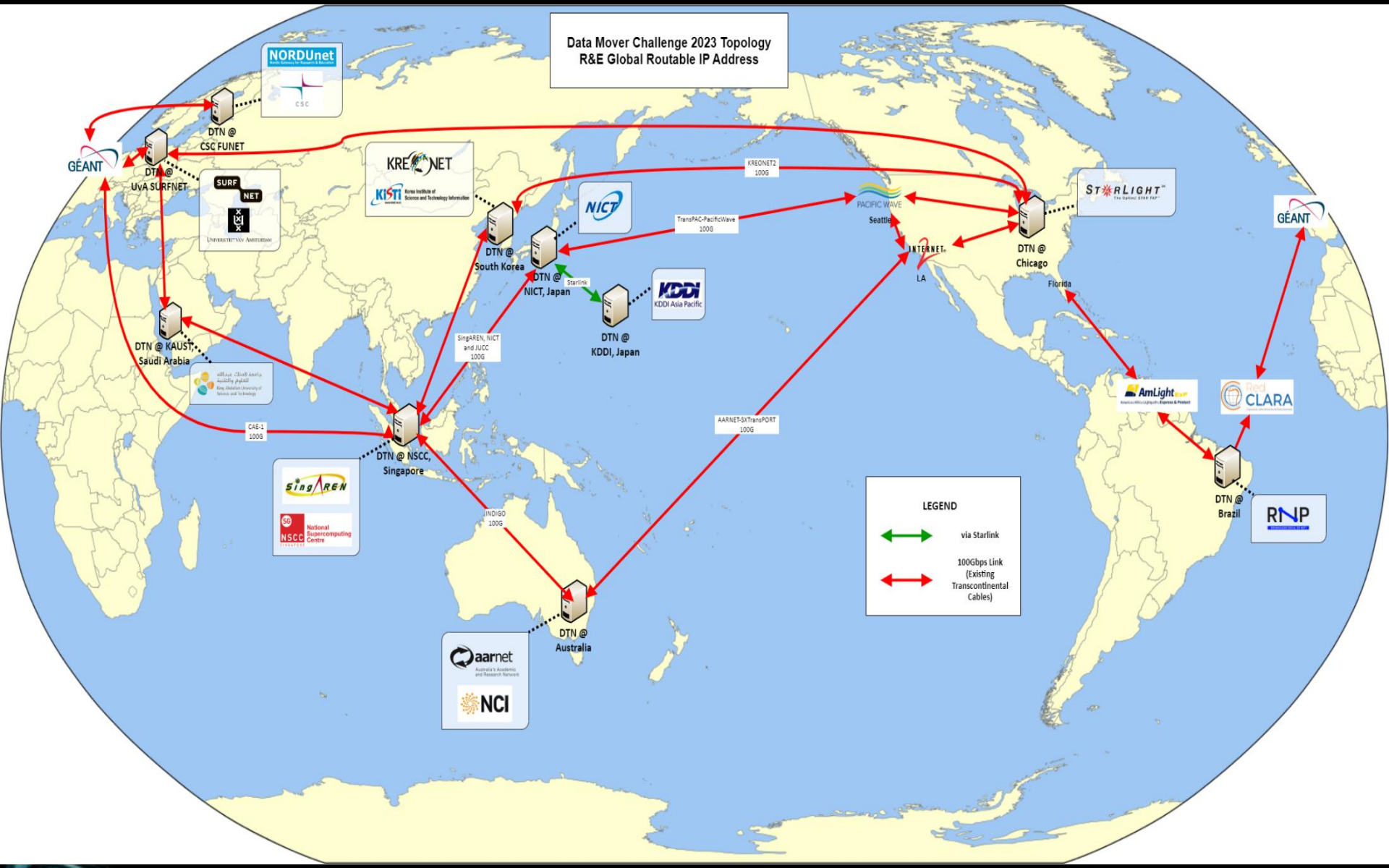
CANARIE/MOXY
MOXY-1:urn:ogf:network:canarie.ca:2017:topology:AMST1
MOXY-2:urn:ogf:network:canarie.ca:2017:topology:icair
MOXY-3:???

ESnet
ESnet-1:urn:ogf:network:es.net:2013::cern-513-cr5:lag-2:+
ESnet-2::urn:ogf:network:es.net:2013::star-cr55:2_1_c3_1:+

StarLight
StarLight-1:urn:ogf:network:icair.org:2013:mren8700:esnet
StarLight-2:urn:ogf:network:icair.org:2013:mren8700:canarie
StarLight-3:urn:ogf:network:icair.org:2013:mren8700:pwave-grp
StarLight-4:???

Supercomputing Asia DMC International Testbed



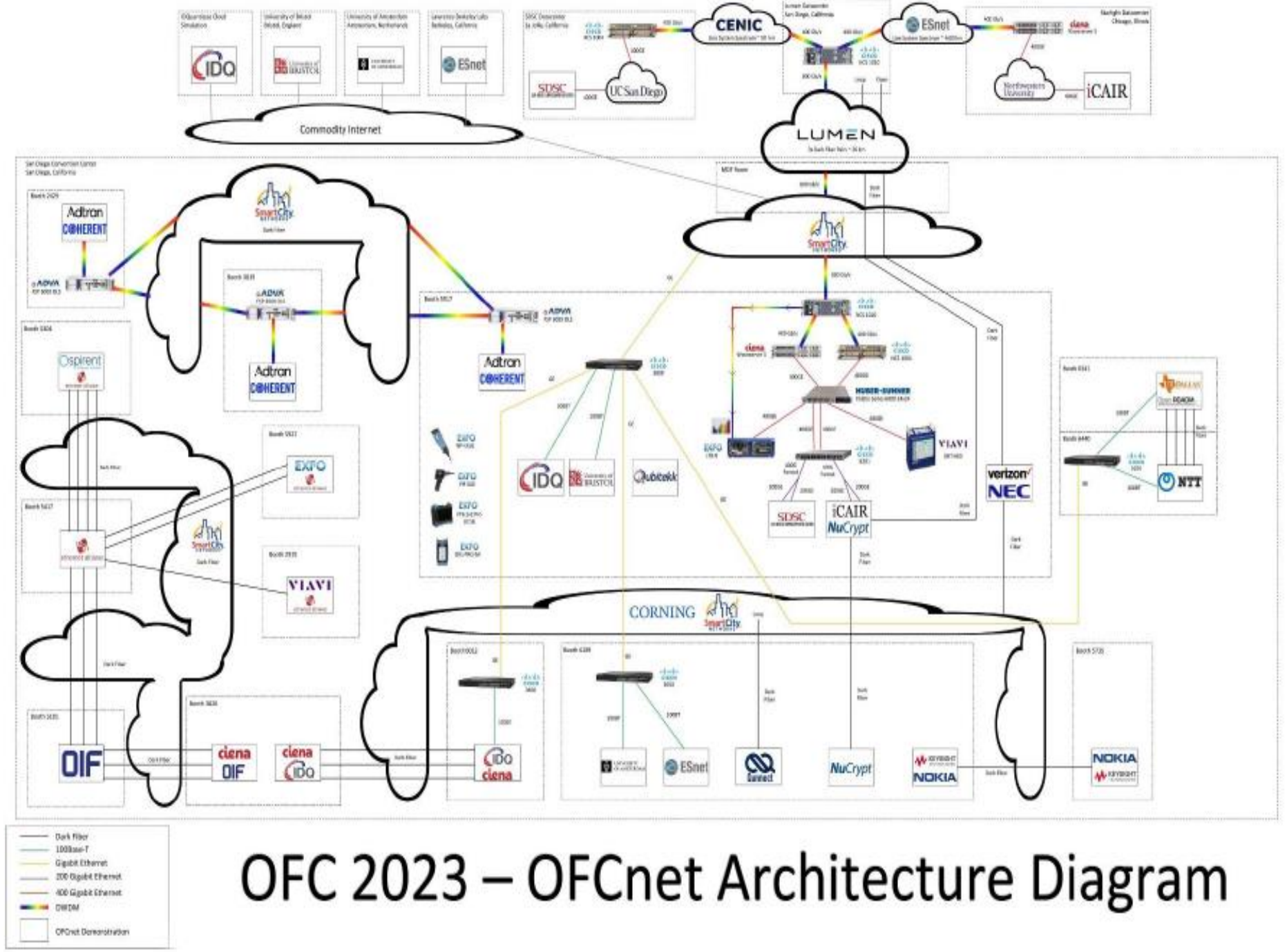

Data Mover Challenge 2023 Topology
R&E Global Routable IP Address
NORDUnet
DTN @ CSC FUNET
GÉANT
DTN @ UvA SURFNET
SURFNET
KREONET
KISTI
DTN @ South Korea
NICT
DTN @ NICT, Japan
Starlink
DTN @ KDDI, Japan
KDDI Asia Pacific
KREONET2 100G
TransPAC-PacificWave 100G
PACIFIC WAVE
Seattle
INTERNET2
LA
DTN @ Chicago
STARLIGHT
Florida
GÉANT
DTN @ KAUST, Saudi Arabia
CAE-1 100G
SingAREN, NICT and JUCC 100G
DTN @ NSCC, Singapore
SingAREN
NSCC National Supercomputing Centre
INDIGO 100G
AARNET-SXTransPORT 100G
AmLight
RedCLARA
DTN @ Brazil
RNP
DTN @ Australia
aarnet
NCI
LEGEND
via Starlink
100Gbps Link (Existing Transcontinental Cables)

Support For Large Scale Demonstrations At OFC

San Diego California

March 6-9, 2023

STARLIGHT℠



OFC 2023 – OFCnet Architecture Diagram

Ilya Baldine PI, RENCI: FABRIC

Core = 3*400 Gbps

STARLIGHT SM

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

Annual Global Research Platform Workshop – Co-Located With IEEE International Conference On eScience Oct 9-10

www.startap.net/starlight

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

STARLIGHT℠