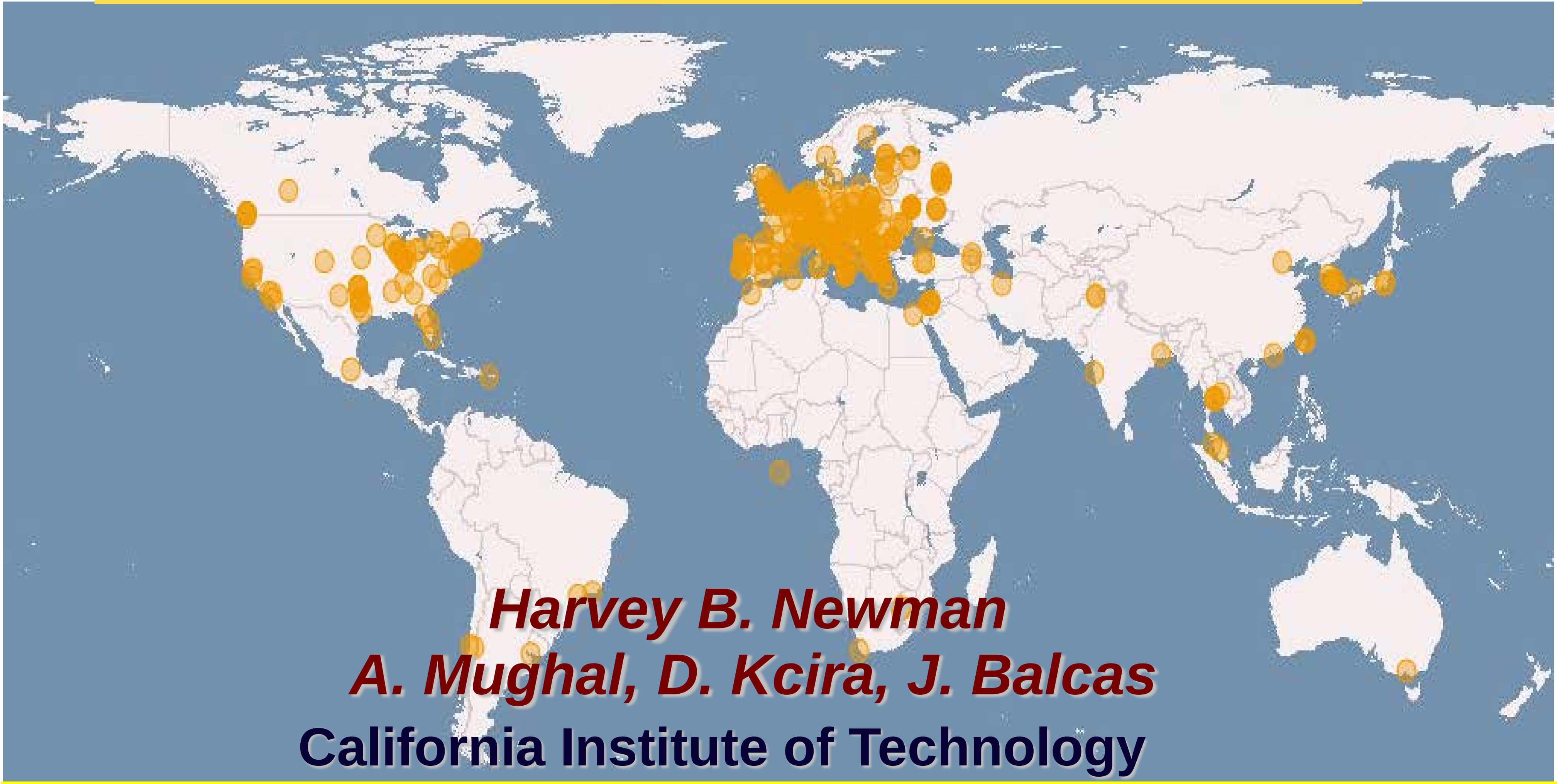
ICFA Standing Committee on Interregional Connectivity (SCIC)

Global Networks for HEP in 2016-17



Harvey B. Newman

A. Mughal, D. Kcira, J. Balcas

California Institute of Technology

Presentation and Reports at http://icfa-scic.web.cern.ch/

Discovery of a Higgs Boson
July 4, 2012; Nobel Prize 2013

Physicists Find Elusive Particle Seen as Key to Universe

The New York Times

Englert | Higgs

Theory : 1964

LHC + Experiments Concept: 1984

Construction: 2001

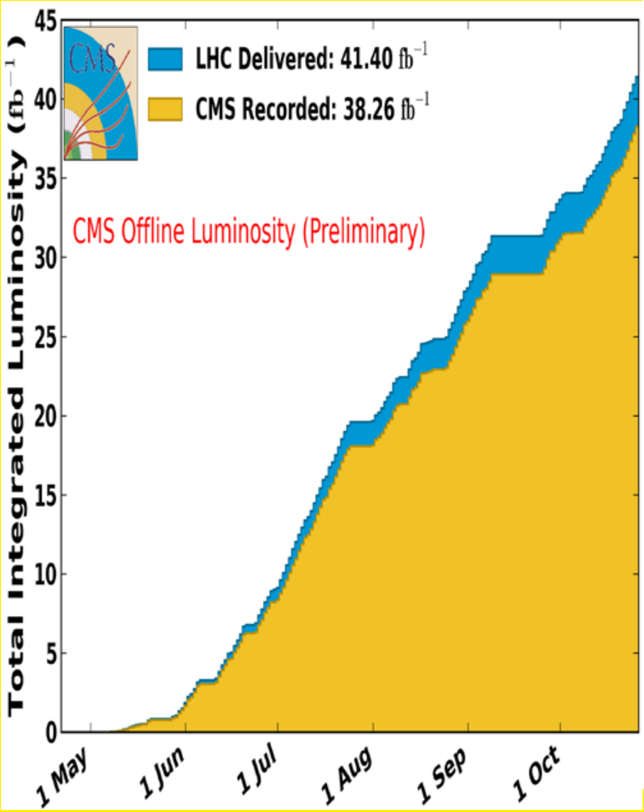
Operation and Discovery: 2009-12

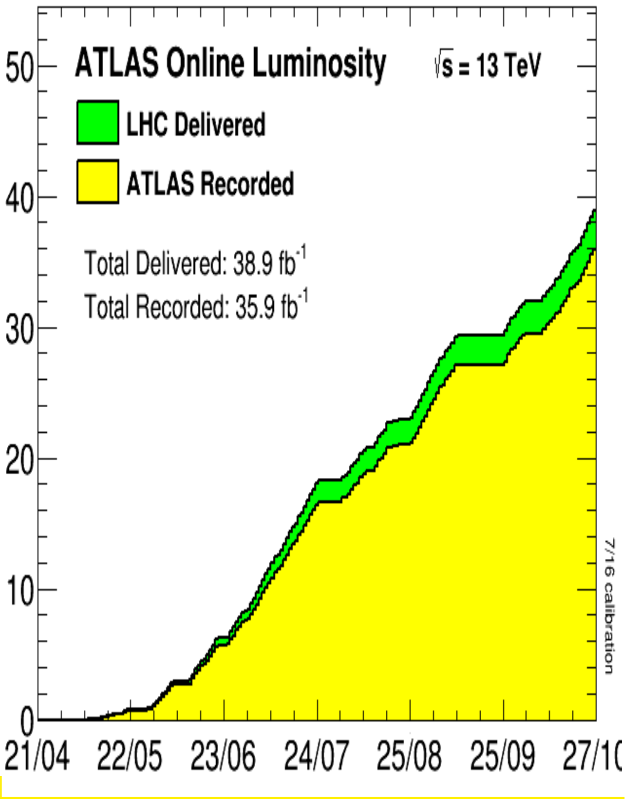
The Economist
In praise of charter schools
Britain's banking scandal spreads
Volkswagen overtakes the rest
A power struggle at the Vatican
When Lonesome George met Nora
JULY 7TH-13TH 2012
Economist.com
A giant leap for science
Finding the Higgs boson

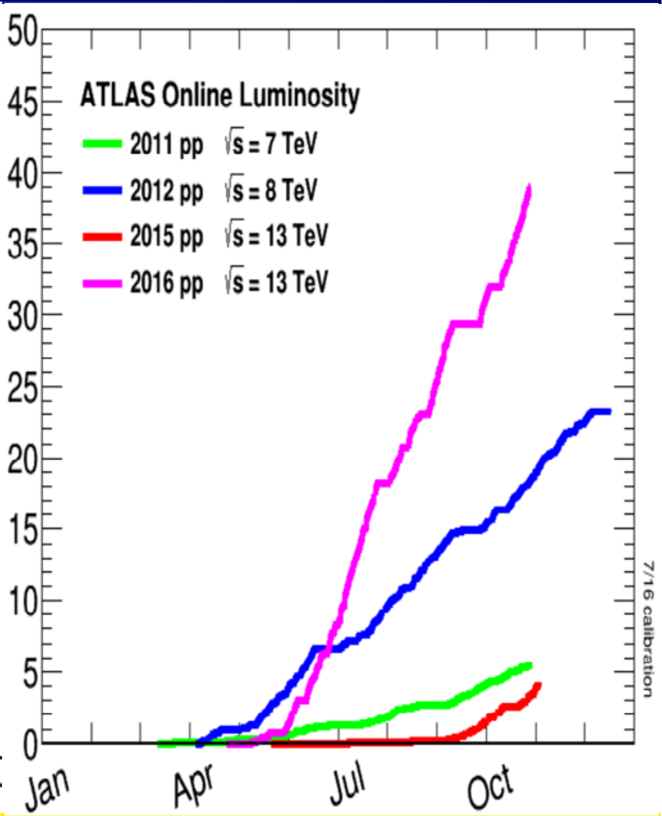
Highly Reliable Advanced Networks Were Essential to the Higgs Discovery and Every Ph.D Thesis of the last 20+ Years

They will be Essential to Future Discoveries, and Every Ph. D Thesis to Come

2016 LHC pp Luminosity to 50% Above Design Higher (to 90% Above ?) in 2017







40 Inverse Femtobarns Delivered ! 92+% Recorded

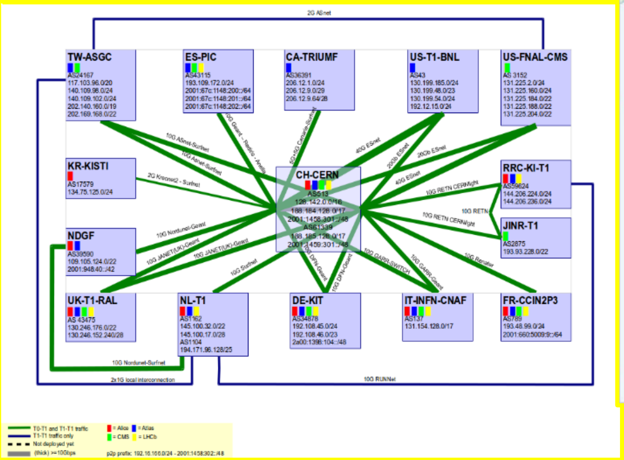
1.4-1.5 X 10^{34}/cm^2/sec Peak; μ to ~50 ! (Test to 95)

2017 Outlook: to 1.9 X 10^{34}/cm^2/sec, 56/fb with β* = 33 cm ?

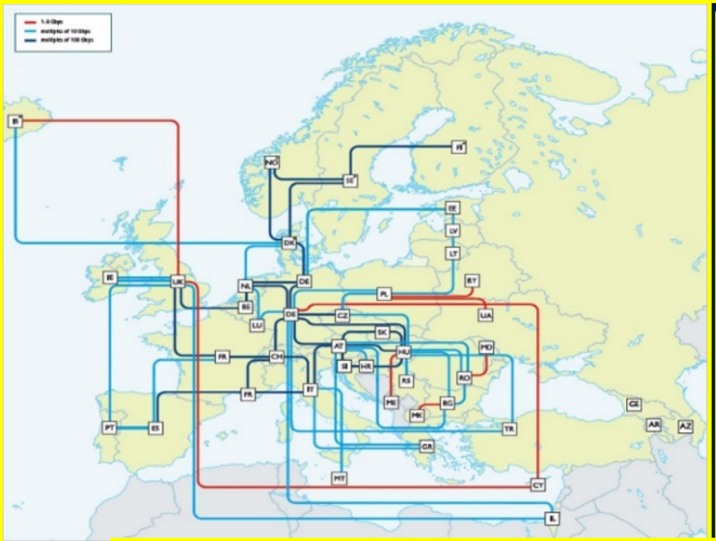
Accelerated Challenges: Data Volumes Vs. Available Storage, CPU and Networks starting in 2017-2018

Core of LHC Networking LHCOPN, LHCONE, GEANT, Esnet, Internet2

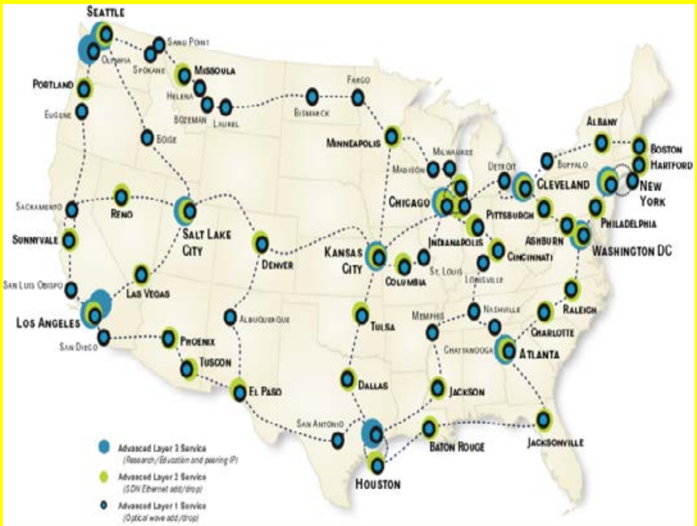
LHCOPN: Simple and Highly Reliable, for Tier0+1 Operations



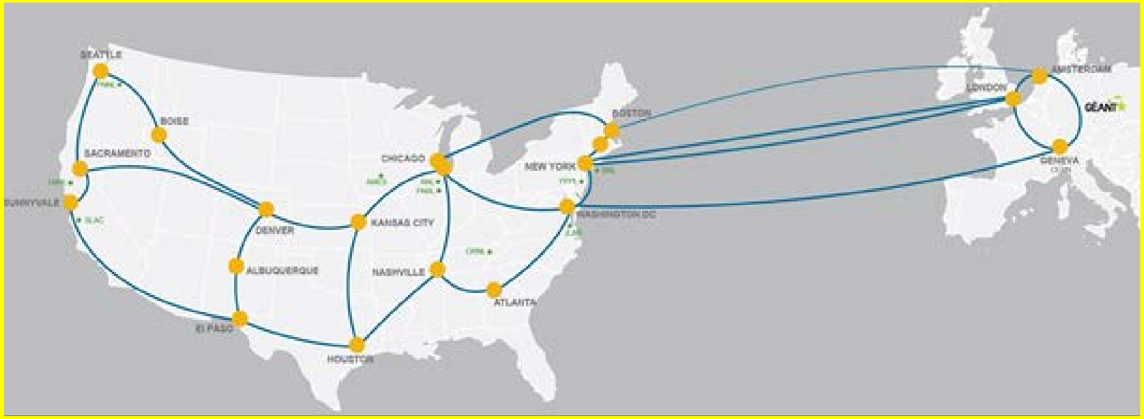
GEANT



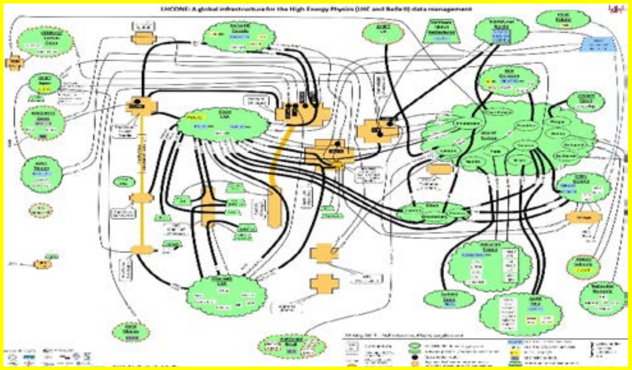
Internet2



ESnet (with EEX)



LHCONE



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

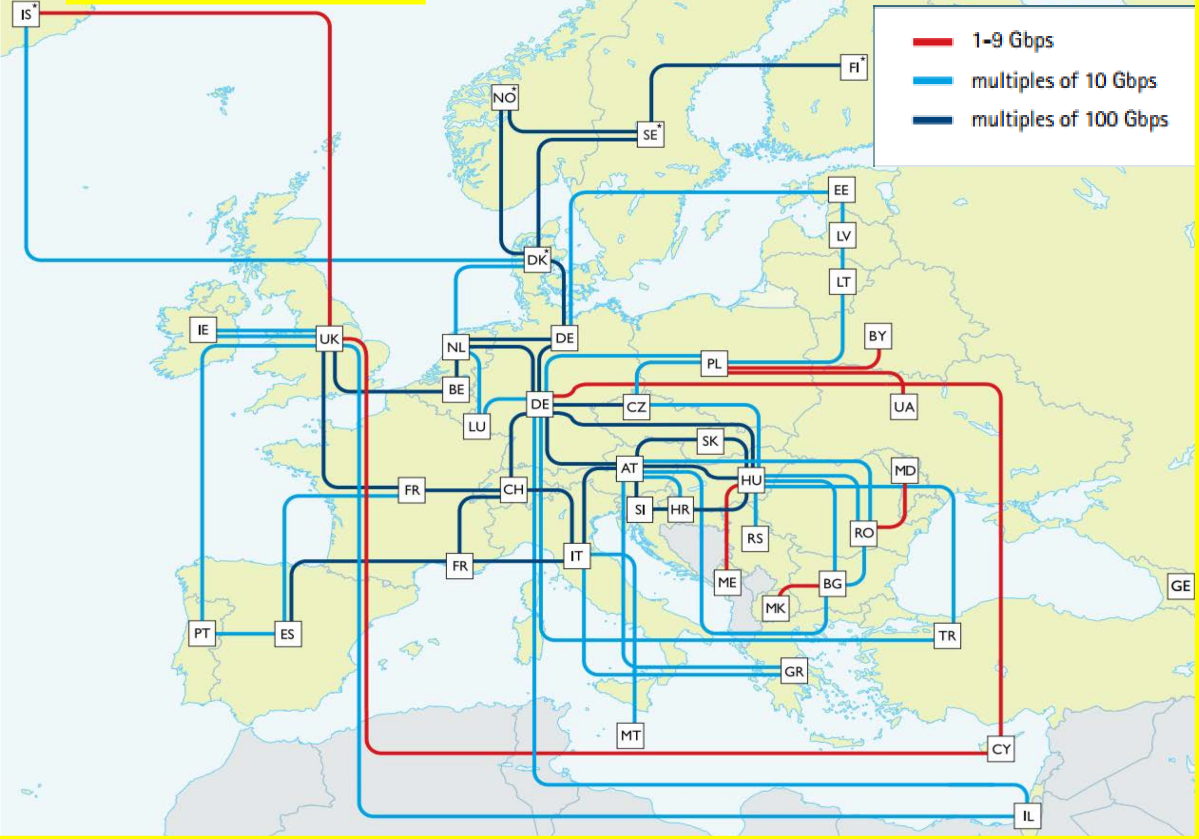
+ NRENs in Europe, Asia, Latin America, Au/NZ; US State Networks

GÉANT Pan-European Backbone
50M Users at 10k Institutions

www.geant.net

50k km backbone fully migrated to 100G in 2013
CERN – Wigner Data Center (HU) 2nd 100G in Service from 2016

January 2017



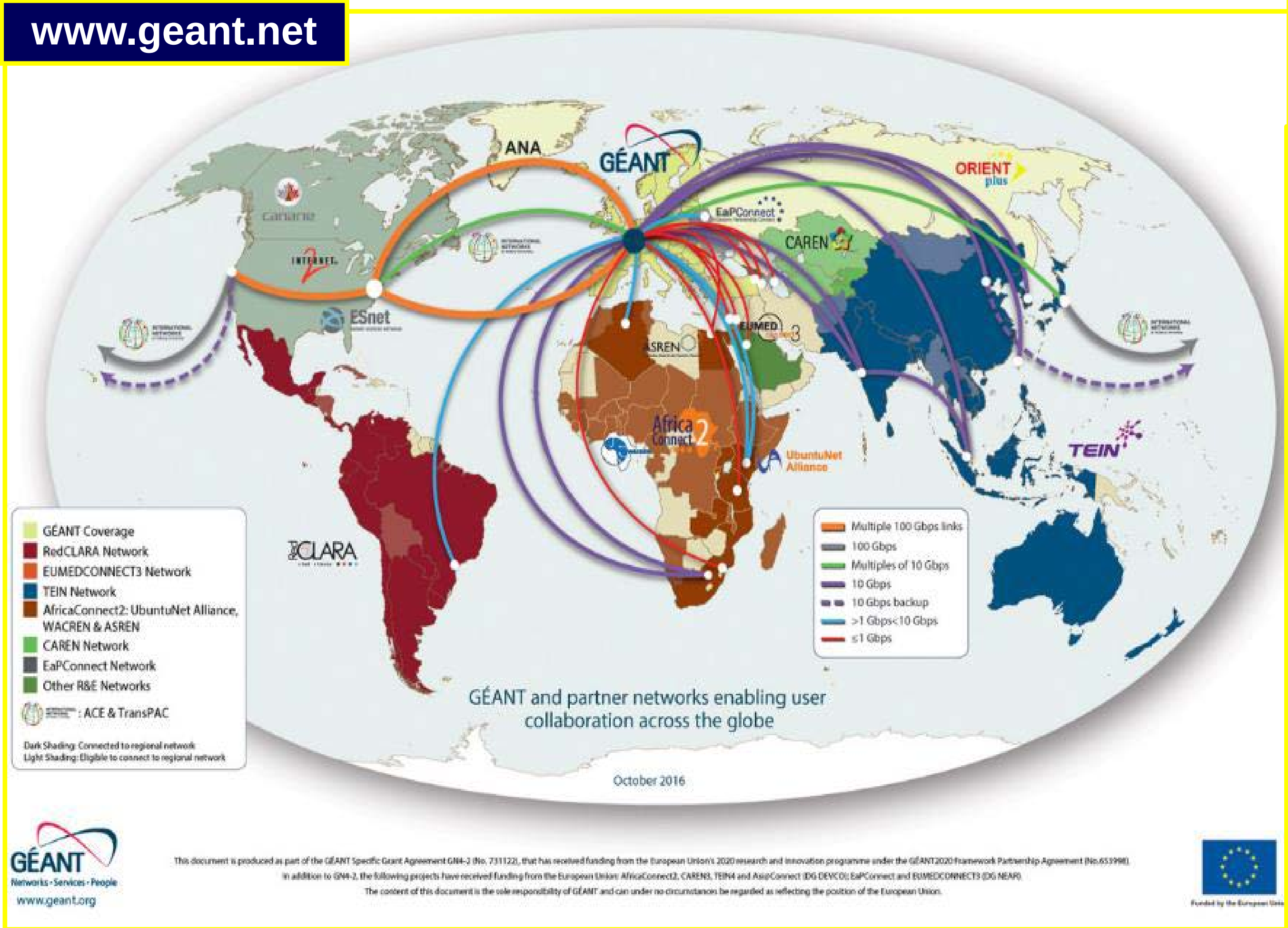
41 NREN Partners

- 12000 km Dark Fiber Core
  17 Major Cities; 16 Countries
- 500G Superchannels:
  26 100G Links
- 100G available
  between any two NRENs
- 17 NRENs directly on
  N X 100G backbone
- Service Availability 99.92%
  on Avg.; Core Nodes 100%
- Dynamic Circuits:
  NSI development
- *2016 Traffic: 1.4 Exabytes*
  4 Petabytes/day Avg.

https://www.geant.org/Projects/GEANT_Project_GN4/Documents/GEANT%20Project%20Highlights%20GN4-1.pdf

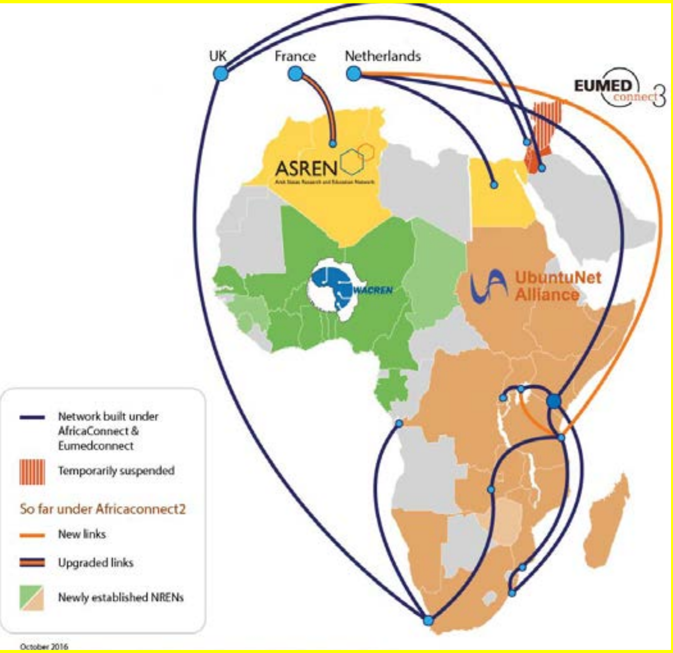
GÉANT At the Heart of Global R&E Networking

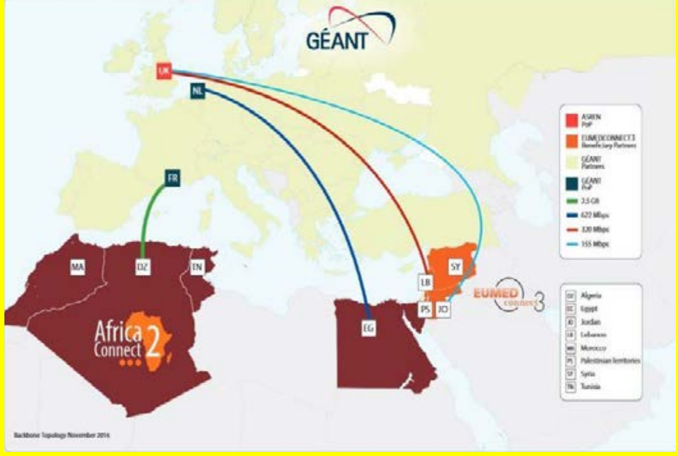
www.geant.net



This document is produced as part of the GÉANT Specific Grant Agreement GN4-2 (No. 731122), that has received funding from the European Union's 2020 research and innovation programme under the GÉANT2020 Framework Partnership Agreement (No.653998).
In addition to GN4-2, the following projects have received funding from the European Union: AfricaConnect2, CAREN3, TEIN4 and AsiaConnect (DG DEVCO); EaPConnect and EUMEDCONNECT3 (DG NEAR).
The content of this document is the sole responsibility of GÉANT and can under no circumstances be regarded as reflecting the position of the European Union.

AfricaConnect and EUMedConnect October 2016



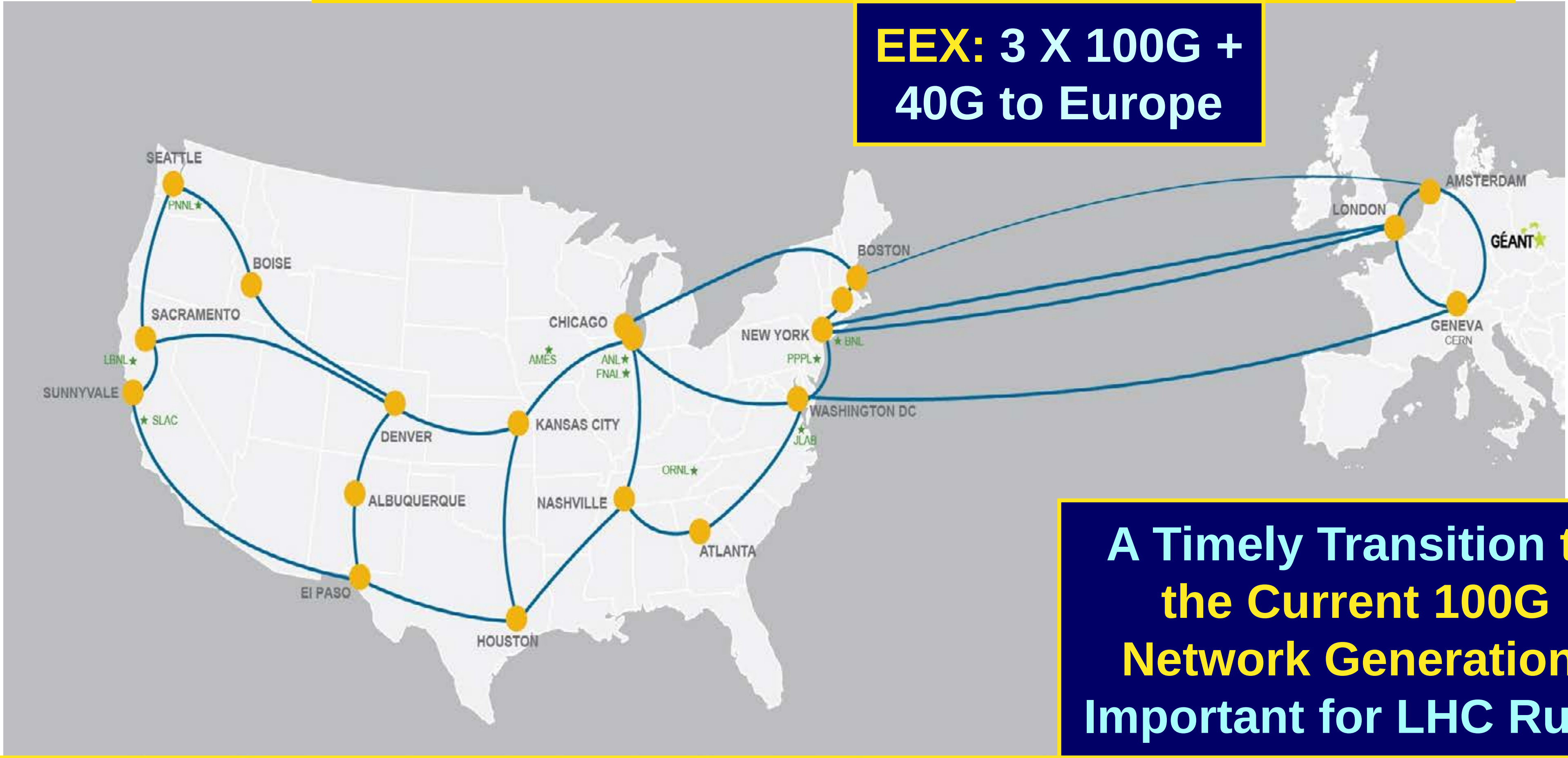


Energy Sciences Network: ESnet

100G Backbone Completed in 2012

ESnet EEX to Europe: Completed in Dec. 2014

EEX: 3 X 100G + 40G to Europe



A Timely Transition to the Current 100G Network Generation, Important for LHC Run2

2 X 100G to BNL and 2 x 100G to Fermilab; 17 Hubs with N X 100G
100G Dark Fiber Testbed; Share 100G ANA-300 TA Research Links

R&E Network Trends in 2015-16

100G Generation Maturing; 400G on the Horizon

We are midway in the 7-8 Year generational cycle of 100G networks

❑ 100G core backbones now mature: Internet2 and ESnet core completed in 2012; GEANT 100G completed in 2013-14; 100G endsites proliferating !

✷ Transatlantic transition: ESnet EEX (340G) from 2015; ANA-300 from 2016

❑ 100GE links spreading in Europe and Asia: e.g. Netherlands, Japan, Romania, Czech Republic, Hungary, Poland, China, Korea

❑ 100G Links to Tier2 Centers: ~Complete in US; increasing in other regions

❑ TransPacific: Multiple 100G Links to Major 100G US Networks: Multiple 100G R&E Transpacific Links: Japan Tokyo, Singapore, Korea + Guam Exchange

Pacific Wave, Internet2, ESnet, Starlight; Connection to 100G Transatlantic

❑ 2015-16: 32-48 X 100G top of rack switches, low cost 100GE server NICs, high performance SSDs ; 100 to 2 X 100GE servers now a reality

❑ Higher WAN Throughput: 350G+ at SC15+16; to 1 Tbps Local: Caltech, StarLight, FIU, Grid UNESP, Fermilab, etc.

❑ Software Defined Networks (Openflow; ODL or ONOS; ALTO): Move to built in intelligence: a major focus of the global community and industry

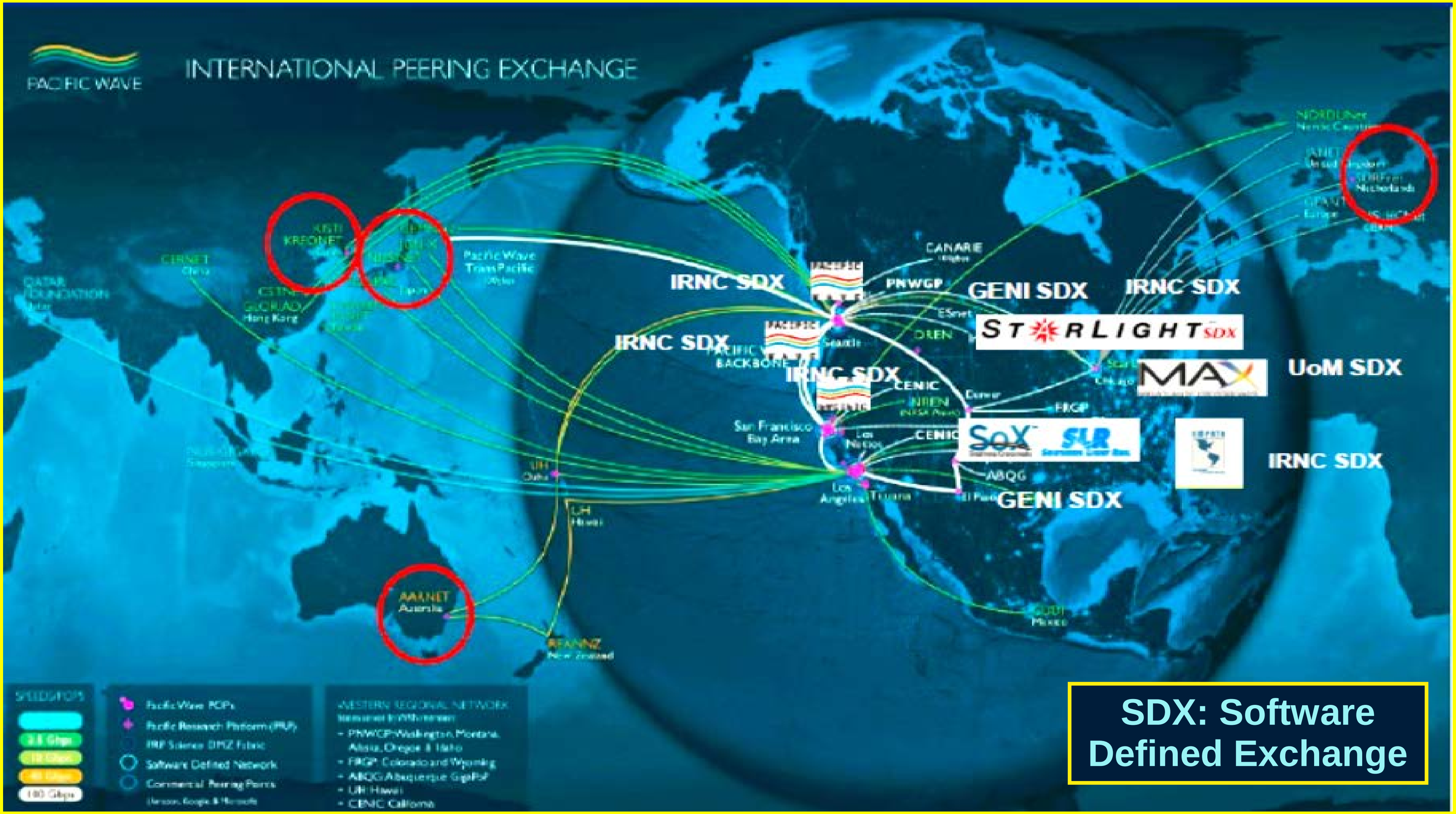
Issue: Will next generation 400G networks be affordable in time for Run3 ?

Major Development Examples
in R&E Networks in Europe

- Czech Republic (CESnet): Move to resilient N X 100G optical net
- Slovakia (SANET): 100G in 2016; Move to N X 100G optical net
- France (RENATER): Multi-100G backbone from 2015;
  [400G production trial Lyon-Paris already in 2013]
- Italy (Garr-X): 0.5 to 1 Tbps Superchannel Hybrid Core;
  Increasing Tier1+2 40-100G access foreseen
- Germany (DFN): Building on N X 100G optical platform since 2013
- Poland (PIONIER): N X 100G core connecting HPC centers + Metro 100G access and 100G Poznan – CERN in 2016
- Japan (SINET5): 200G+ backbone + 100G full nat'l mesh in 2016
- Netherlands (SURFNet): Renewal of the photonic layer in 2016
- Nordic Countries (NorduNet): N X 100G Core completed in 2017
- Greece (GRNET-4): Completing 100G optical + carrier service + IP service, including dynamic optical paths at 1/10/40/100G

Next Step: Global Research Platform

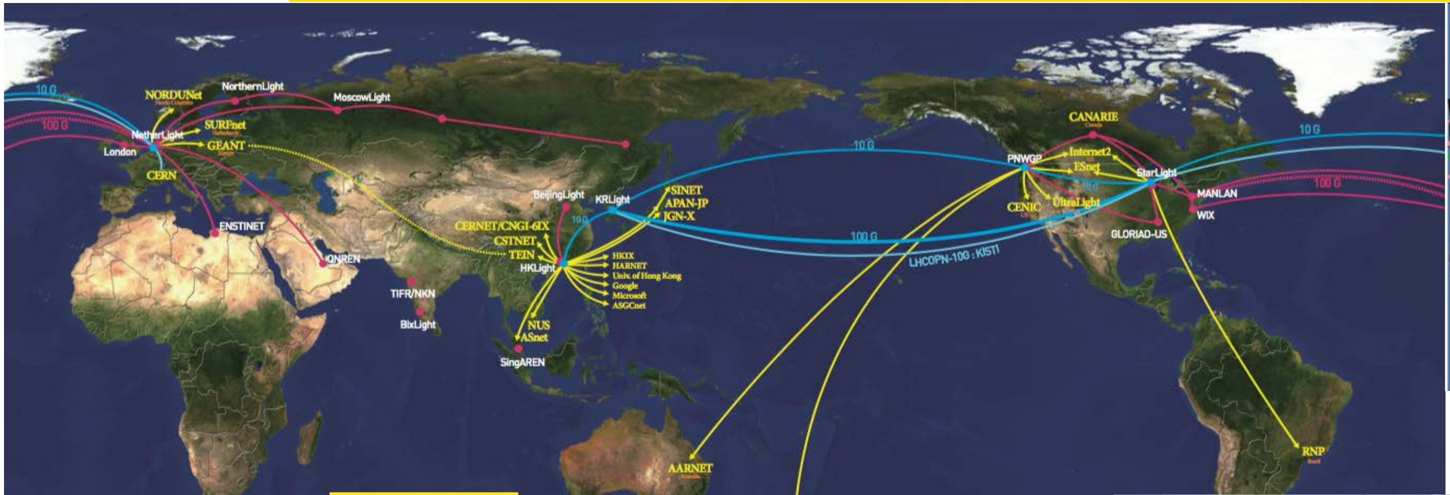
Building on CENIC/Pacific Wave and GLIF



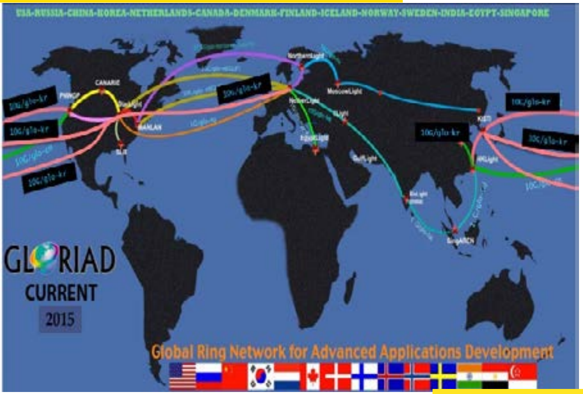
SDX: Software Defined Exchange

Multiple 100G R&E Transpacific Links: Japan, Singapore, Korea, Hawaii + Guam Exchange in 2017

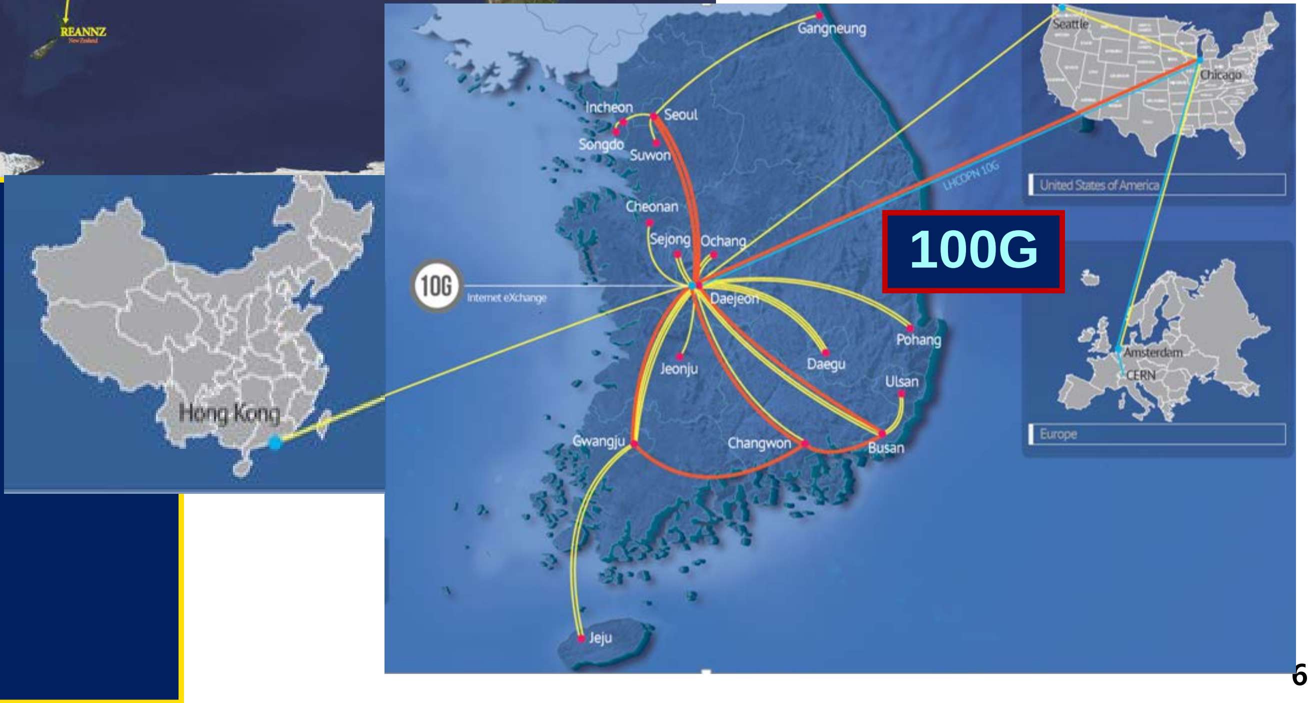
KREONet2 and GLORIAD-KR And SDN Deployment (KREONET-S)



2016


USA-RUSSIA-CHINA-KOREA-NETHERLANDS-CANADA-DENMARK-FINLAND-ICELAND-NORWAY-SWEDEN-INDIA-EGYPT-SINGAPORE
GLORIAD
CURRENT
2015
Global Ring Network for Advanced Applications Development

2015



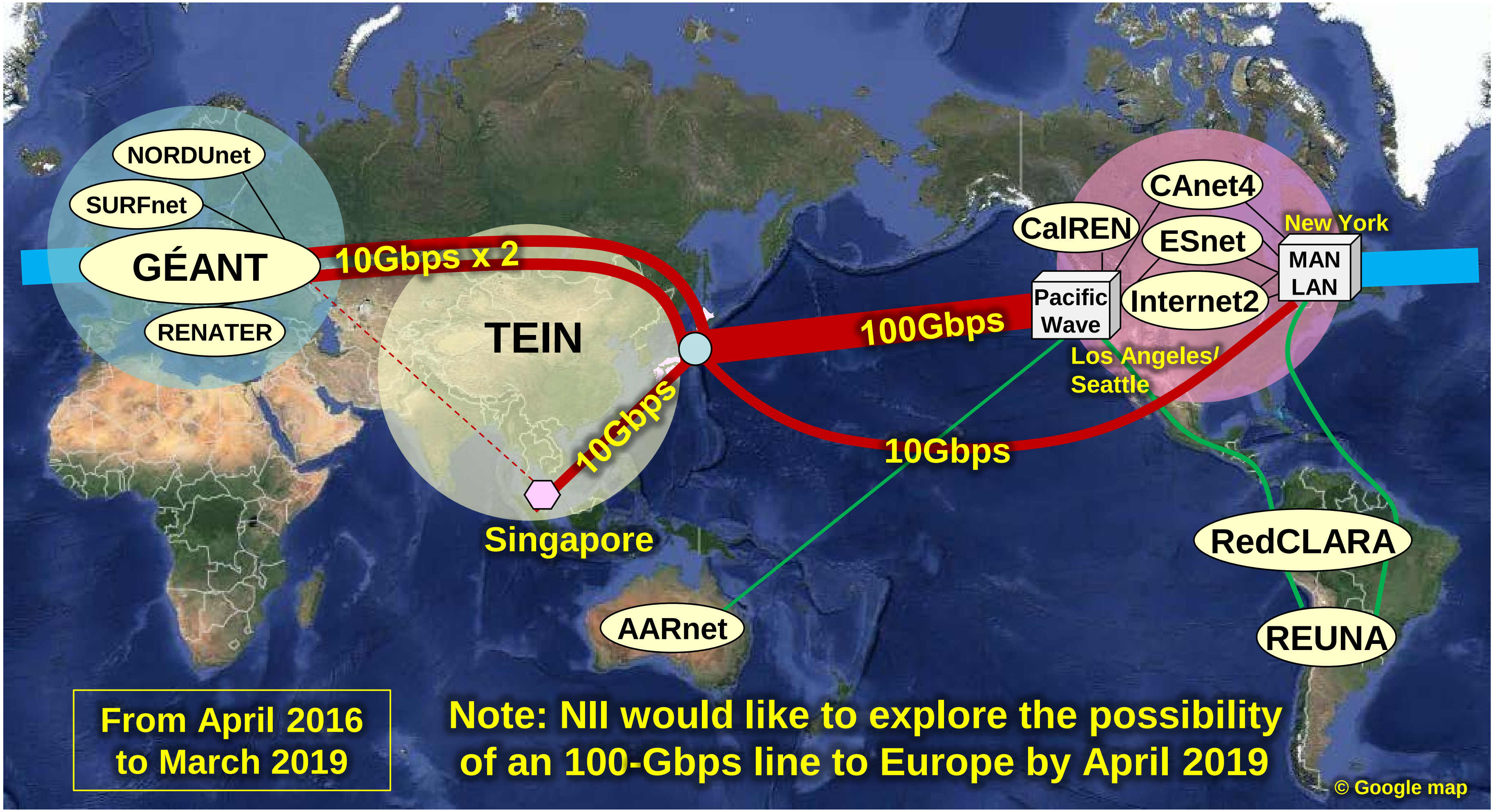
100G

2015-2016 Highlights

1. 100G from Daejon to Chicago/StarLight
2. 100G Ring linking major cities
3. 17 GigaPoPs with 1G, 10G or 40G

Japan: SINET5 Will Have Direct International Links to USA, Europe and TEIN/ASIA

- USA: 100 Gbps line to Los Angeles/ Seattle and 10-Gbps line to New York
  Europe: Two 10 Gbps lines to London (and New York line as backup route)
- TEIN/Asia: 10-Gbps line to Singapore



From April 2016 to March 2019

Note: NII would like to explore the possibility of an 100-Gbps line to Europe by April 2019

© Google map

SINET5: Nationwide Academic Network

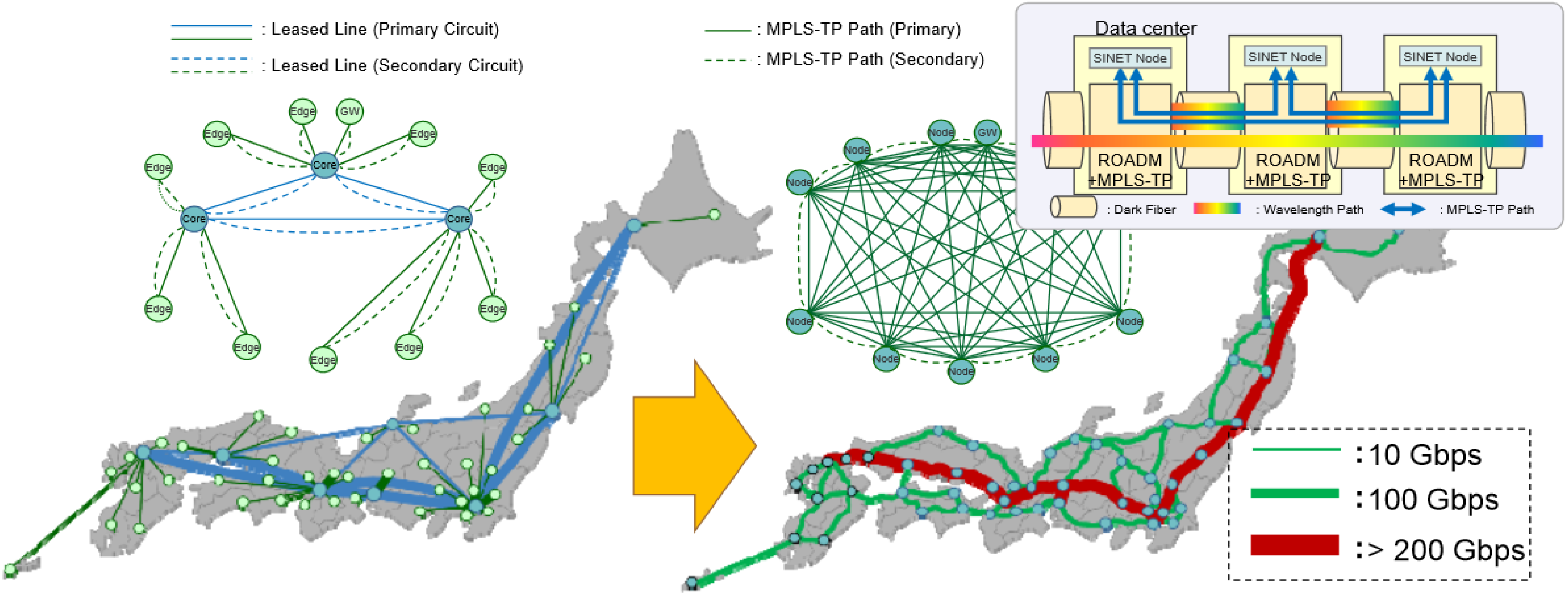
- ◆ 2016 SINET5 connects all the SINET nodes in a fully-meshed topology and minimizes the latency between every pair of the nodes using nationwide dark fiber
- ◆ MPLS-TP devices connect a pair of the nodes by primary and secondary MPLS-TP paths.

SINET4 present

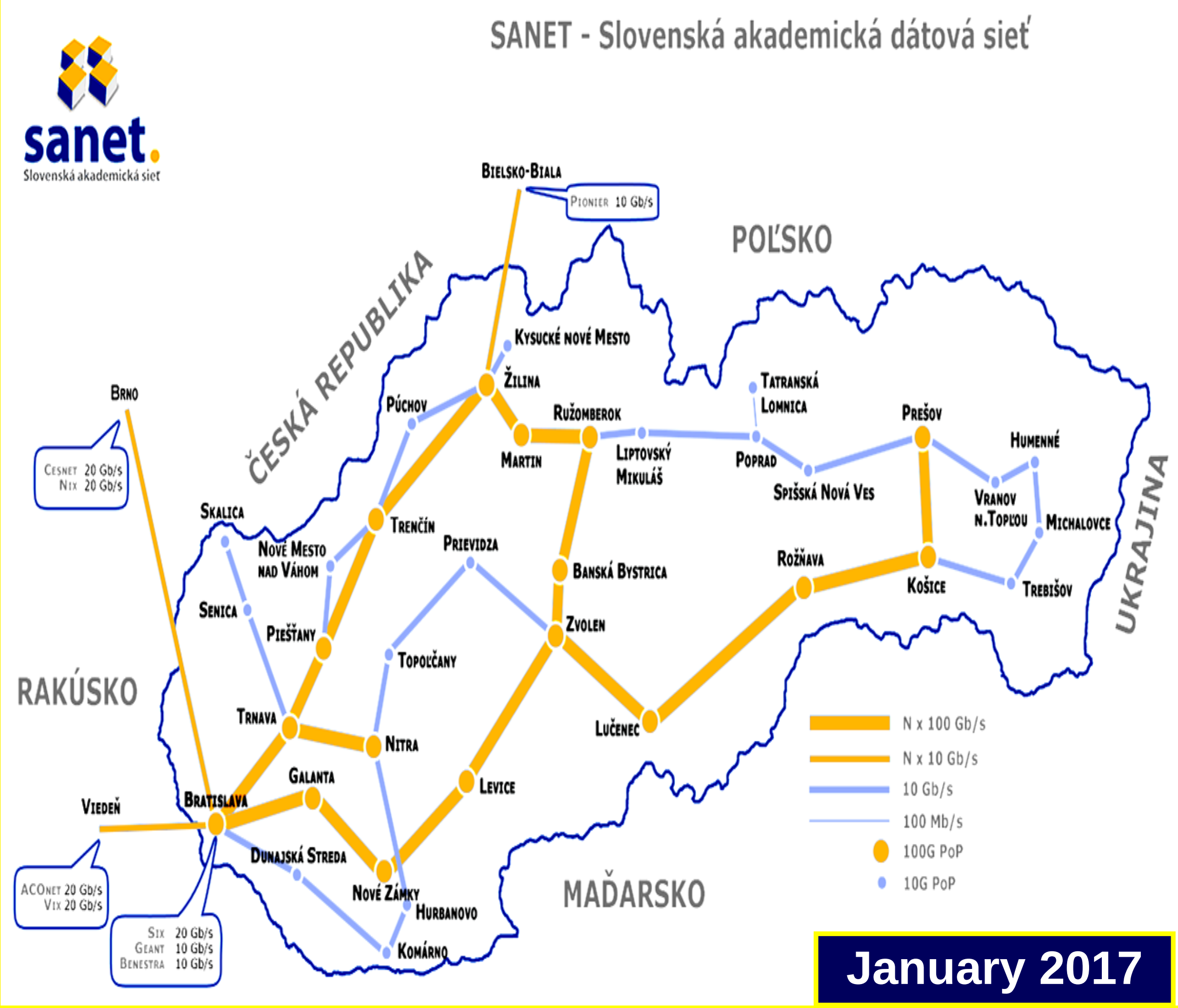
- Connects nodes in a star-like topology
- Secondary circuits of leased lines need dedicated resources

SINET5 2016

- Connects all the nodes in a fully-meshed topology with redundant paths
- Secondary paths do not consume resources



SANET (Slovakia) Status and Plan



January 2017

2002: single highest bandwidth link was 4 Mbps

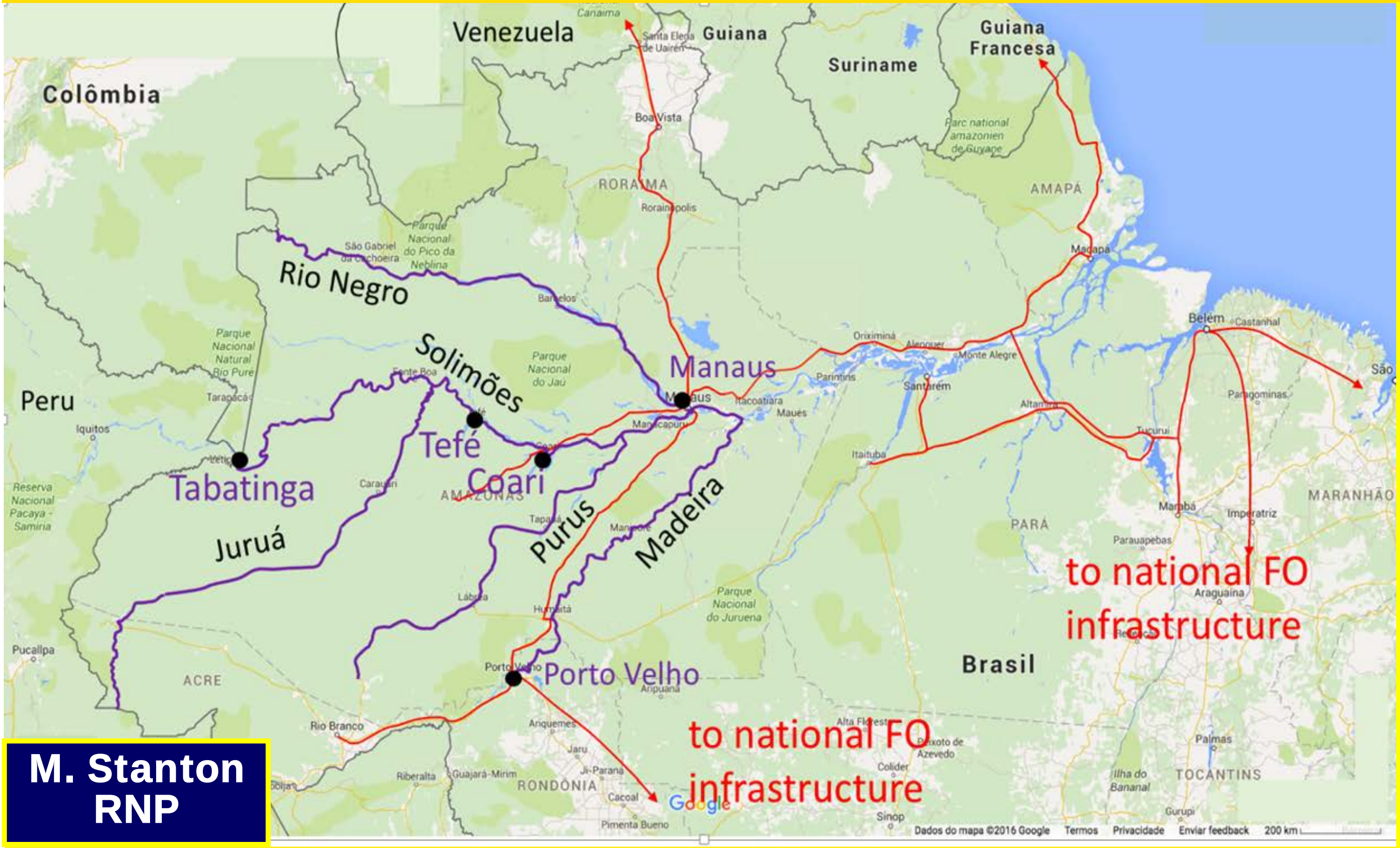
Overall improvement: ~100,000 times in 15 Years

- ❒ SANET network infrastructure consists of several rings
- ❒ Provides full redundancy covering all Slovak universities and research institutions in 37 towns
- ❒ In 2016 SANET completed a major national infrastructure upgrade, providing N x 100GE capacity
- ❒ SANET is planning to install an Infinera DWDM system on additional links
  - ❒ To establish a fully resilient N x 100GE backbone

RNP and the Brazilian Army: Amazonia Conectada Project

http://www.amazoniaconectada.eb.mil.br/eng/

7000 km of Data Highways (Infovias) planned along the Negro, Solimoes, Jurua, Purus and Madeira Rivers



M. Stanton
RNP

Blue lines are proposed subfluvial fiber

First 240 km section: Caori – Tefe completed April 2016

Next two sections: Manaus – Caori and Tefe – Tabatinga

ICFA SCIC

Reports and Trends

SCIC in 2016-17

http://cern.ch/icfa-scic

A Worldview of *Networks for and from HEP*

Focus on the LHC Program during Run2 and Beyond

- ***2017 Presentation**: "Networking for HEP"* [HN, A. Mughal, D. Kcira, J. Balcas]: Updates on the Digital Divide, World Network Status, Transformative Trends in the Global Internet
- **32 Annexes for 2016-17 [22 New]: A World Network Overview** *Status and Plans of International, Nat'l & Regional Networks, HEP Labs, and Advanced Network Projects*
- ***2016 Monitoring Working Group Report*** [S. McKee, R. Cottrell, et al]: Quantifying the Digital Divide: PingER Data from worldwide monitor set PerSONAR and WLCG Monitoring Efforts
- ✷ Also See: http://internetlivestats.com: Worldwide Internet Use
- *GEANT (formerly TERENA) Compendia* (https://compendium.geant.org/compendia): R&E Networks in Europe
- Telegeography.com; Interactive Submarine Cable Map: http://submarinecablemap.com

SCIC Work Areas

- *Closing the Digital Divide*
  - Monitoring the world's networks, with a focus on the Divide; work towards greater equality of scientific opportunity
  - Work on throughput improvements; problem solutions
  - Encouraging the development of national advanced network infrastructures: *through knowledge sharing, and joint work*
- Advanced network technologies and systems
  - New network concepts and architectures: Creation and development; with many network partners
    - *LHCOPN, LHCONE*
    - Software defined networking and OpenFlow; OpenDaylight
  - Integration of advanced network methods with experiments' mainstream data distribution and management systems
  - High throughput methods; + community engagement to apply the methods in many countries, for the LHC and other major programs (HEP, LIGO, AMS, et al.)

ICFA SCIC in 2017

- We are continuing our work in many countries to Close the Digital Divide
  - Both in the physics community and in general
- To make physicists from all world regions full partners in the scientific discoveries
- We are learning to help do this effectively, in partnership with advanced networks, many agencies and HEP groups:
  - ➔ Brazil (RNP), Asia Pacific (APAN), Mexico (CUDI)
  - ➔ AmLight (FIU): US – Latin America
  - ➔ GLORIAD Ring Around the Earth, Including to Russia, China, Middle East and India
- But we are indeed leaving other countries and regions behind, for example: Africa, the Rest of Latin America, Most of the Middle East, South and SE Asia
- A great deal of work remains:
- *Support for the PingER Monitoring Effort at SLAC is a vital part*

Conclusions
Recommendations
and Requests to ICFA

SCIC Conclusions for 2017: 1 of 3

- Extensive, efficient use of the world's national, continental and transoceanic networks by the HEP community continues to be a key factor in the key measurements and search for new physics at Run2 and throughout our field
- The exceptional performance of the LHC presents new challenges
  - Our field’s use of networks continues to grow exponentially
  - Our field’s awareness of our impact on the world’s R&E networks is essential to our future success
- Beyond being major users, through the SCIC and other leading representative organizations in our field, we are now among the world’s leading network developers
  - Working in a global partnership to enable the current and next generation of major science programs and discoveries
- These developments provide a strong foundation for the next round of Computing Models, including at the High Luminosity LHC
  - Based on coordinated, agile use of growing larger but still limited network, computing and storage resources

SCIC Conclusions for 2017: 2 of 3

- While changes in the LHC Computing Models are well underway, the common vision of the next generation Model(s) has yet to come into focus
- ICFA has an important potential role to play in overseeing that the necessary studies are undertaken
- Proposal: A new common project should be formed, aimed at meeting future needs in the context of the emerging paradigm of intelligent networks, and coordinated use of resources
- The need for attention in 2017-18 is heightened by:
  - Rising competition for the use networks from other data intensive fields
  - Exponential growth of our own use, at a rate faster than the growth of affordable network capacity
- Engaging in these common developments will have profound benefits, not only for our field, but for many fields of data intensive science, in terms of:
  - Working efficiency
  - Discovery potential
  - Budgets

SCIC Conclusions for 2017: 3 of 3

- The PingER work of the Monitoring WG, led by Cottrell, is of special, central importance to the work of the SCIC, and to the field as whole.
- The Digital Divide activities of the SCIC rely on the singular effort of Les Cottrell and the students and visitors working with him in the PingER project, including
  - Tracking the world’s network connectivity and obtainable throughput in all regions
  - Providing information and training on network monitoring and advanced methodologies
- The impact of the work of this group is great, both within and beyond the bounds of the HEP community
- The financial needs, while relatively modest, have not been met
- The work of the Monitoring Group WG, which is a vital part of the work of the SCIC in meeting the charge given by ICFA, is now at risk
- We request ICFA's help in solving this ongoing problem

SCIC Recommendations and Requests to ICFA

- We request that ICFA consider and encourage the development of a new paradigm for network-integrated worldwide distributed computing for our field, leveraging the profound and rapid developments in networking described in the SCIC reports for 2015-16.
- We request that ICFA consider effective ways to build an inter-regional, interdisciplinary collaborative effort in support of this goal, and to achieve the greater goal of more effective worldwide systems supporting the science goals of many data intensive fields.
- We request ICFA's support and guidance in finding ways to improve the connectivity to several regions of the world that continue to lag behind, as made clear in this Report and the report of the Monitoring Working Group, in order to achieve greater equality of access to the data and results. This is essential to give physicists in all world regions the opportunity to be strong partners in the global process of search and discovery, and to develop strong HEP groups for this purpose, thereby strengthening our field as a whole.
- We request that ICFA helps the SCIC find ways to provide the financial support needed by the Monitoring Working Group, so that its work can continue.

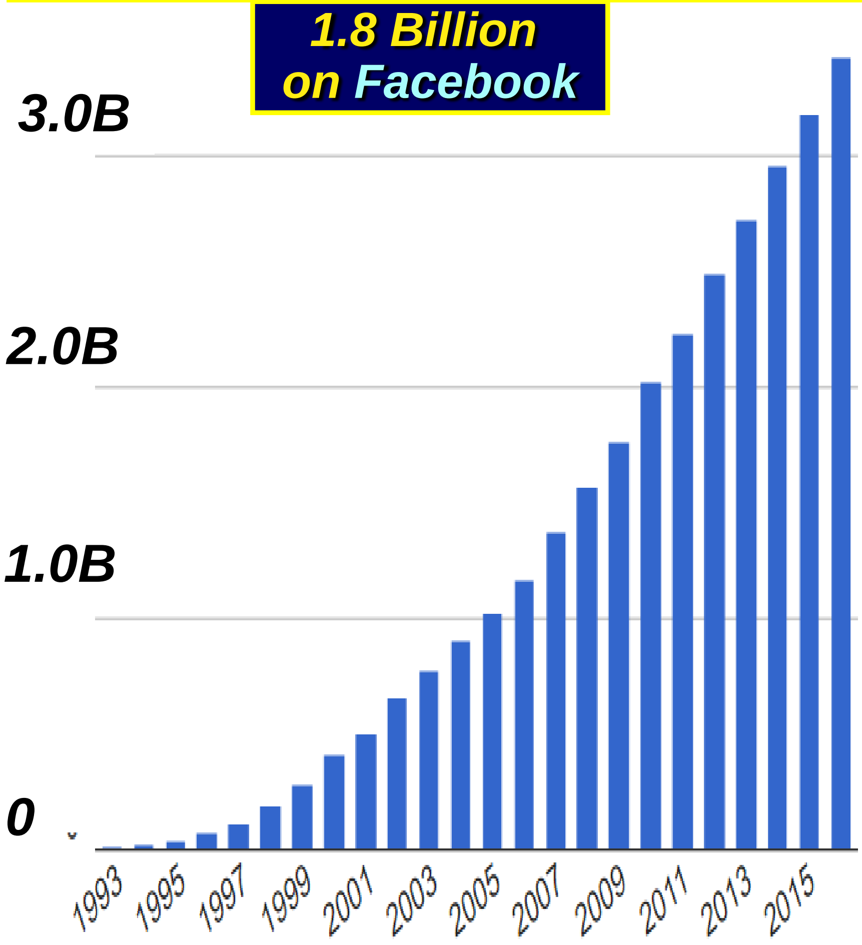
Global Trends

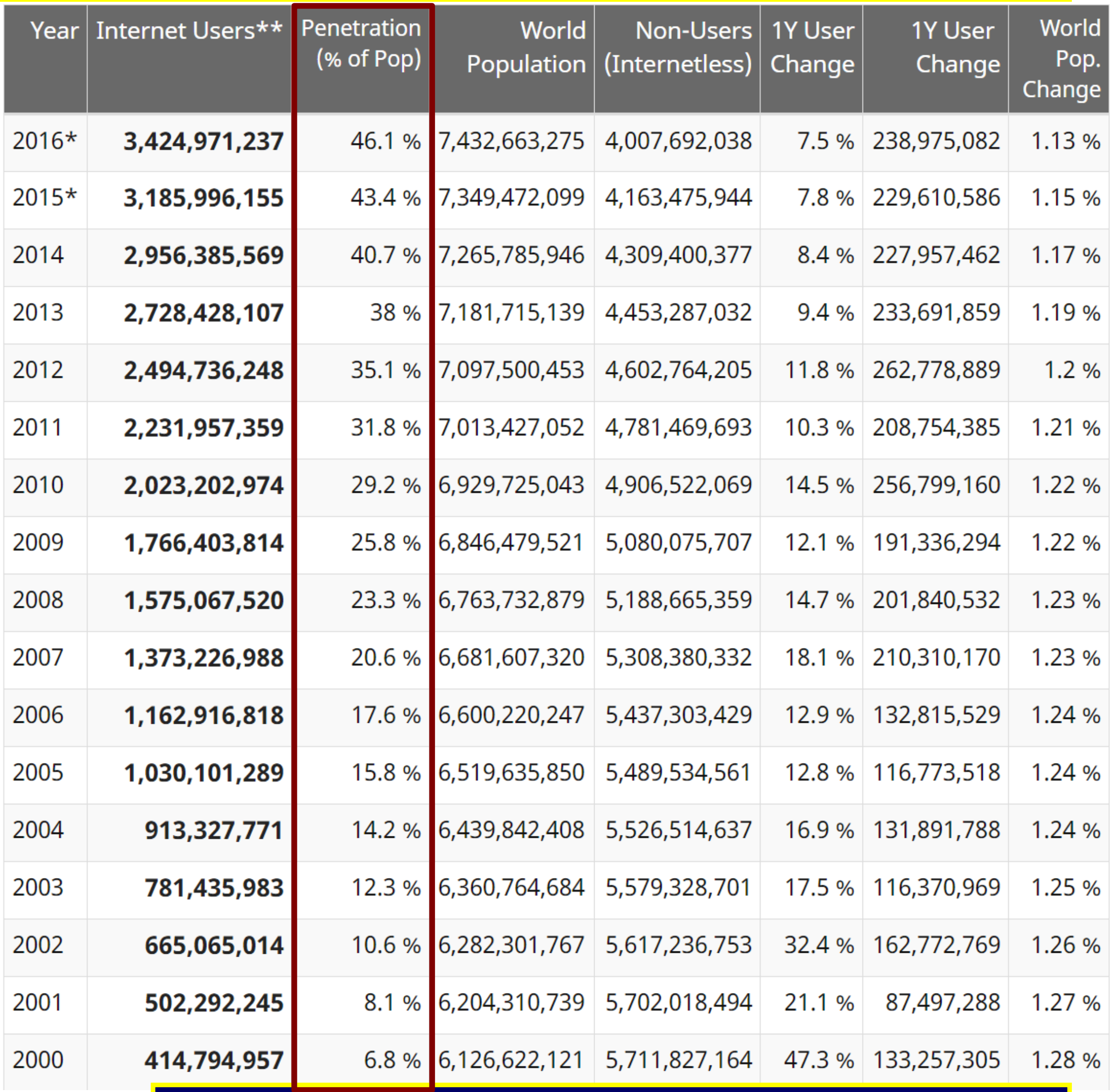
The Internet and

International Networks

3.4 Billion Internet Users; 732M in China
Penetration 46% [14% in 2004; 1% in 1995]; + 8%/Year

Internet Users in the World


1.8 Billion on Facebook

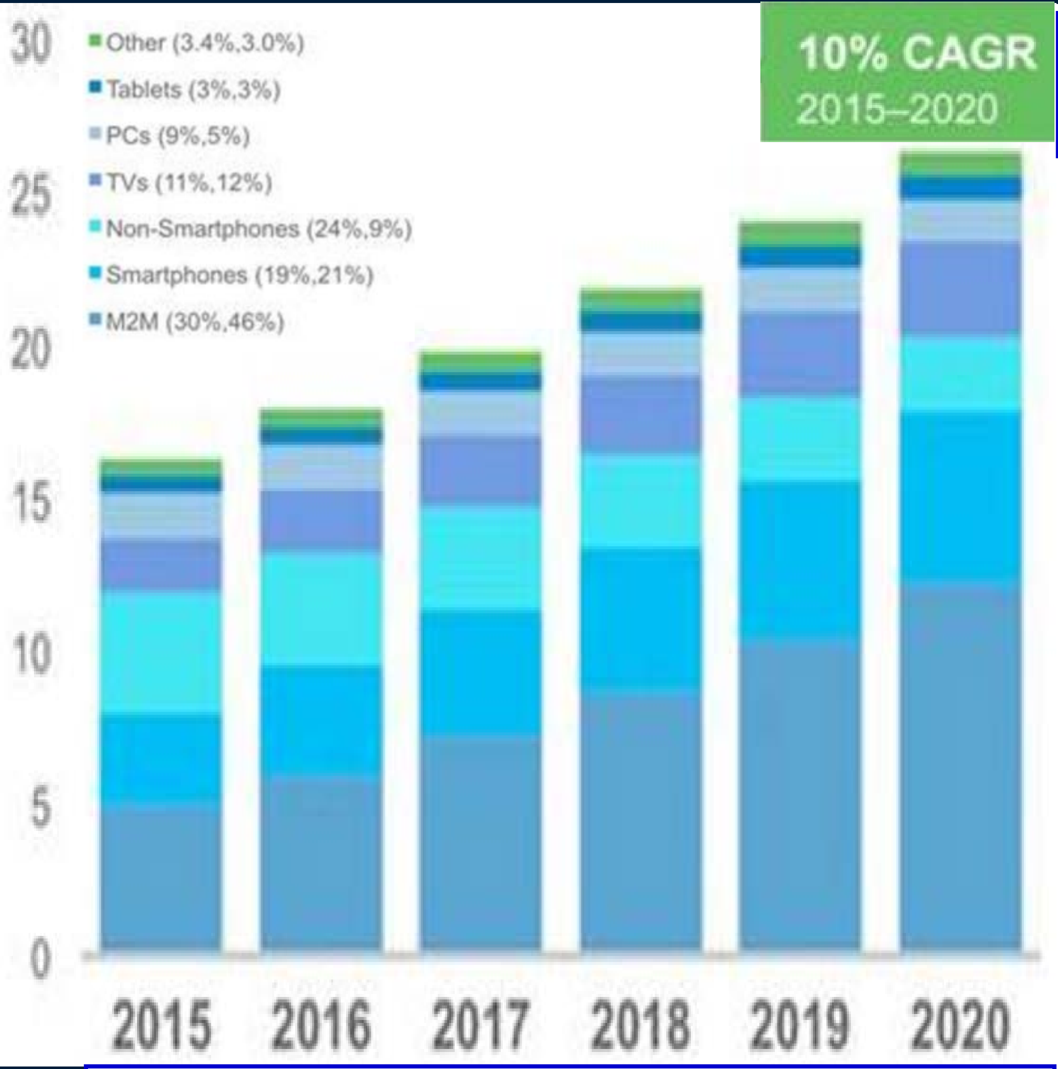


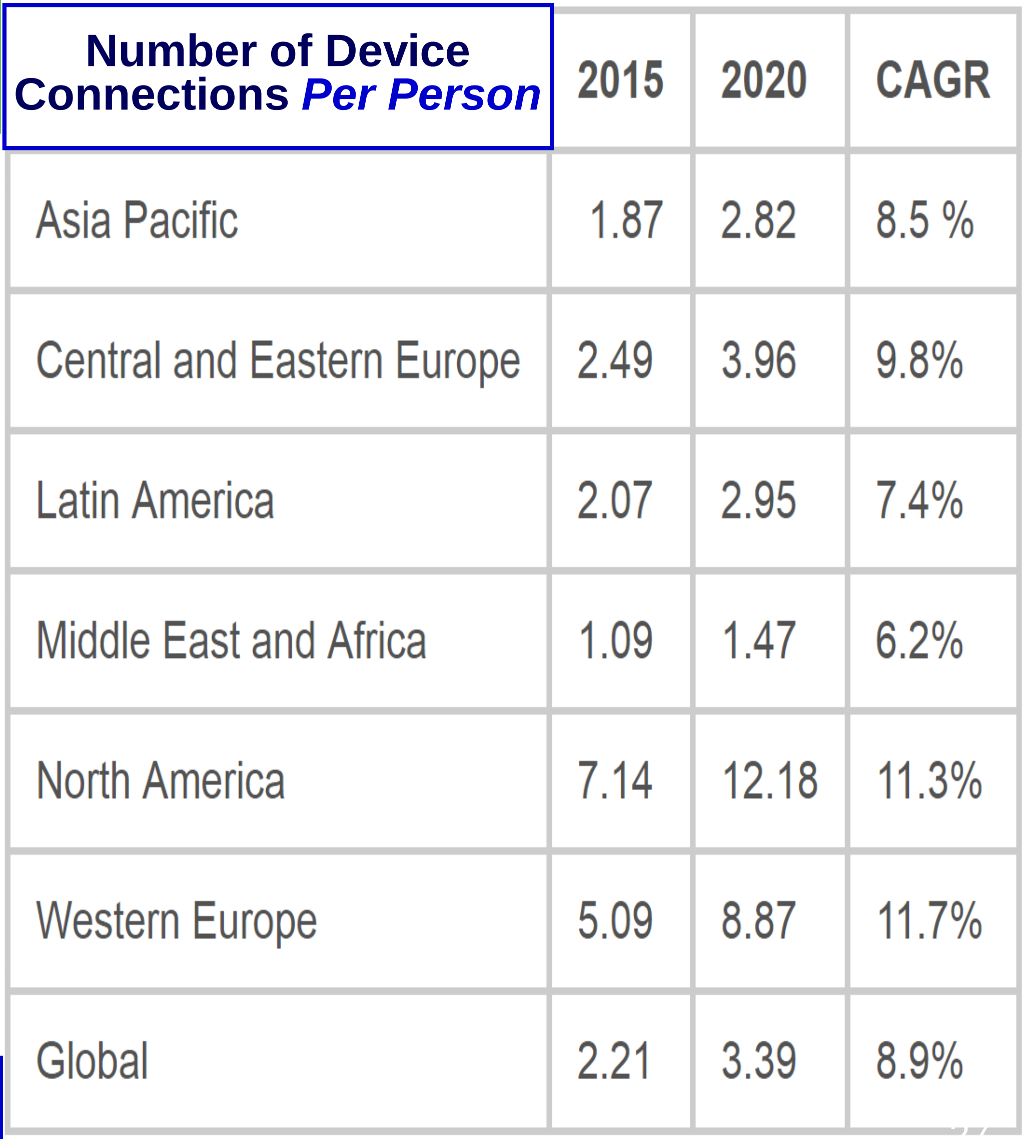
Year	Internet Users**	Penetration (% of Pop)	World Population	Non-Users (Internetless)	1Y User Change	1Y User Change	World Pop. Change
2016*	**3,424,971,237**	46.1 %	7,432,663,275	4,007,692,038	7.5 %	238,975,082	1.13 %
2015*	**3,185,996,155**	43.4 %	7,349,472,099	4,163,475,944	7.8 %	229,610,586	1.15 %
2014	**2,956,385,569**	40.7 %	7,265,785,946	4,309,400,377	8.4 %	227,957,462	1.17 %
2013	**2,728,428,107**	38 %	7,181,715,139	4,453,287,032	9.4 %	233,691,859	1.19 %
2012	**2,494,736,248**	35.1 %	7,097,500,453	4,602,764,205	11.8 %	262,778,889	1.2 %
2011	**2,231,957,359**	31.8 %	7,013,427,052	4,781,469,693	10.3 %	208,754,385	1.21 %
2010	**2,023,202,974**	29.2 %	6,929,725,043	4,906,522,069	14.5 %	256,799,160	1.22 %
2009	**1,766,403,814**	25.8 %	6,846,479,521	5,080,075,707	12.1 %	191,336,294	1.22 %
2008	**1,575,067,520**	23.3 %	6,763,732,879	5,188,665,359	14.7 %	201,840,532	1.23 %
2007	**1,373,226,988**	20.6 %	6,681,607,320	5,308,380,332	18.1 %	210,310,170	1.23 %
2006	**1,162,916,818**	17.6 %	6,600,220,247	5,437,303,429	12.9 %	132,815,529	1.24 %
2005	**1,030,101,289**	15.8 %	6,519,635,850	5,489,534,561	12.8 %	116,773,518	1.24 %
2004	**913,327,771**	14.2 %	6,439,842,408	5,526,514,637	16.9 %	131,891,788	1.24 %
2003	**781,435,983**	12.3 %	6,360,764,684	5,579,328,701	17.5 %	116,370,969	1.25 %
2002	**665,065,014**	10.6 %	6,282,301,767	5,617,236,753	32.4 %	162,772,769	1.26 %
2001	**502,292,245**	8.1 %	6,204,310,739	5,702,018,494	21.1 %	87,497,288	1.27 %
2000	**414,794,957**	6.8 %	6,126,622,121	5,711,827,164	47.3 %	133,257,305	1.28 %

1st Billion in 2005; 2nd in 2010
3rd Billion by the End of 2014

http://www.internetlivestats.com/

Rise of the Internet of Things
27 Billion Devices by 2020





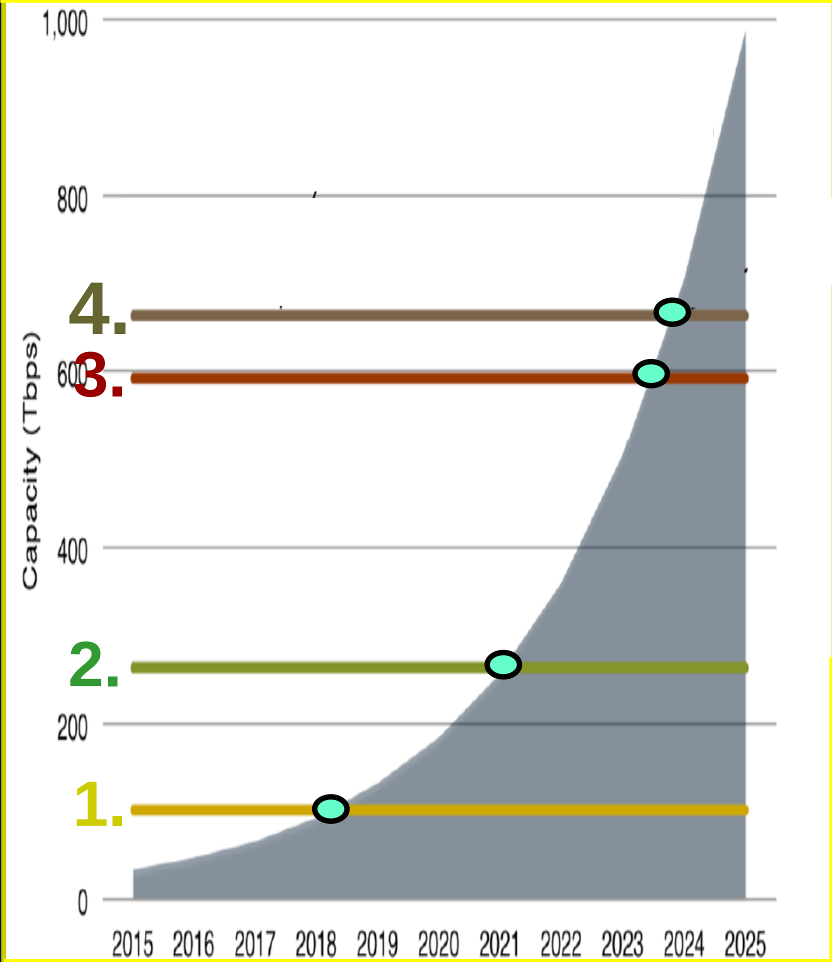
Number of Device Connections *Per Person*	2015	2020	CAGR
Asia Pacific	1.87	2.82	8.5 %
Central and Eastern Europe	2.49	3.96	9.8%
Latin America	2.07	2.95	7.4%
Middle East and Africa	1.09	1.47	6.2%
North America	7.14	12.18	11.3%
Western Europe	5.09	8.87	11.7%
Global	2.21	3.39	8.9%

From the Internet of Things to the Internet of Everything; Not If but When

Possible Transpacific Bandwidth Exhaustion

40% Growth Scenario

Lit and Potential TransPacific Subsea Capacity 2015-2025



Outlook: 400G Subsea links will be needed within 5-7 years

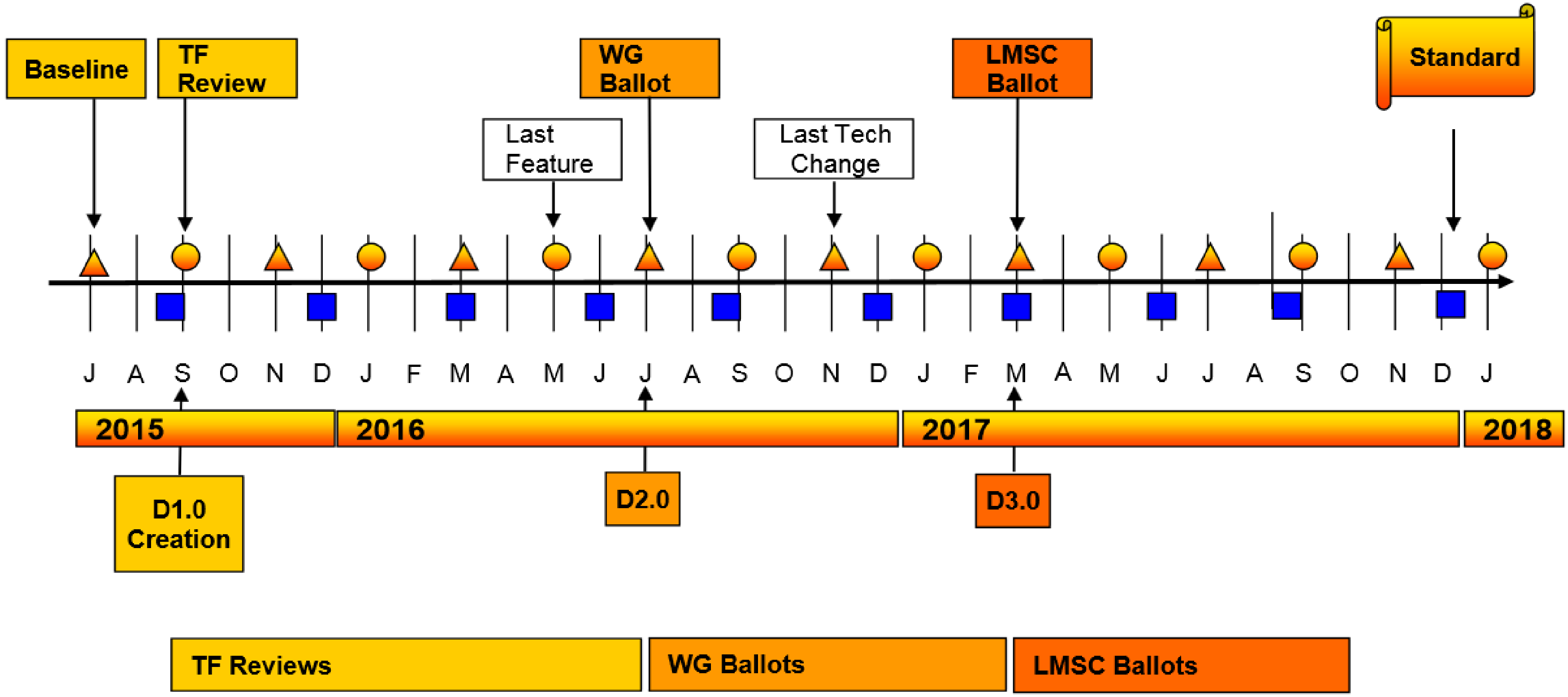
Case	Exhaustion
1. Existing Cables	Mid-2018
2. + Planned Cables: FASTER, SEA-US, NCP	Begin 2021
3. + If older cables could support 100 X 100G each	Begin 2023
4. + Adding 1 New 72 Tbps cable	Mid 2023

Telegeography; A. Maulding, PTC 16

NOTE: ***400G*** **6000 km subsea field trials by Huawei and Tata in 2014**
http://www.lightwaveonline.com/articles/2014/05/tata-huawei-complete-400g-long-haul-subsea-network-field-trial.html

And Alcatel-Lucent on ACE submarine link (Africa Coast – Europe) in Jan. 2015
http://subseaworldnews.com/2015/01/16/alcatel-lucent-to-boost-ace-submarine-link/

400G Ethernet Timeline for Completion of the IEEE Standard: by 2018



http://www.ieee802.org/3/bs/

Feb. 10, 2017: Grass Roots Movement by Cloud Equipment Provider Arista: 800G Ethernet in MegaData Centers *Asap*

Networking for High Energy Physics

A 30+ Year Retrospective

ICFA and Global Networks for HENP (Retrospective)

- 1981 Start: International Networking for HEP
- ICFA Visionary Statement of 1996
- *2004 (Paris): National and International Networks, with sufficient (rapidly increasing) capacity and seamless end-to-end capability, are essential for*
  - ➔ *The formation of worldwide collaborations*
  - ➔ *The daily conduct of collaborative work in both experiment and theory*
  - ➔ *Detector development & construction on a global scale*
  - ➔ *Grid systems supporting analysis by involving physicists in all world regions*
  - ➔ *The conception, design and implementation of next generation facilities as "global networks"*
  - *"Collaborations on this scale would never have been attempted, if they could not rely on excellent networks"*

[TA Network WG, Larry Price et al 2001.]

31 Years of BW Growth of Int'l HENP Networks (US-CERN Example)

- ***Rate of Growth > Moore's Law.* (US-CERN Example)**

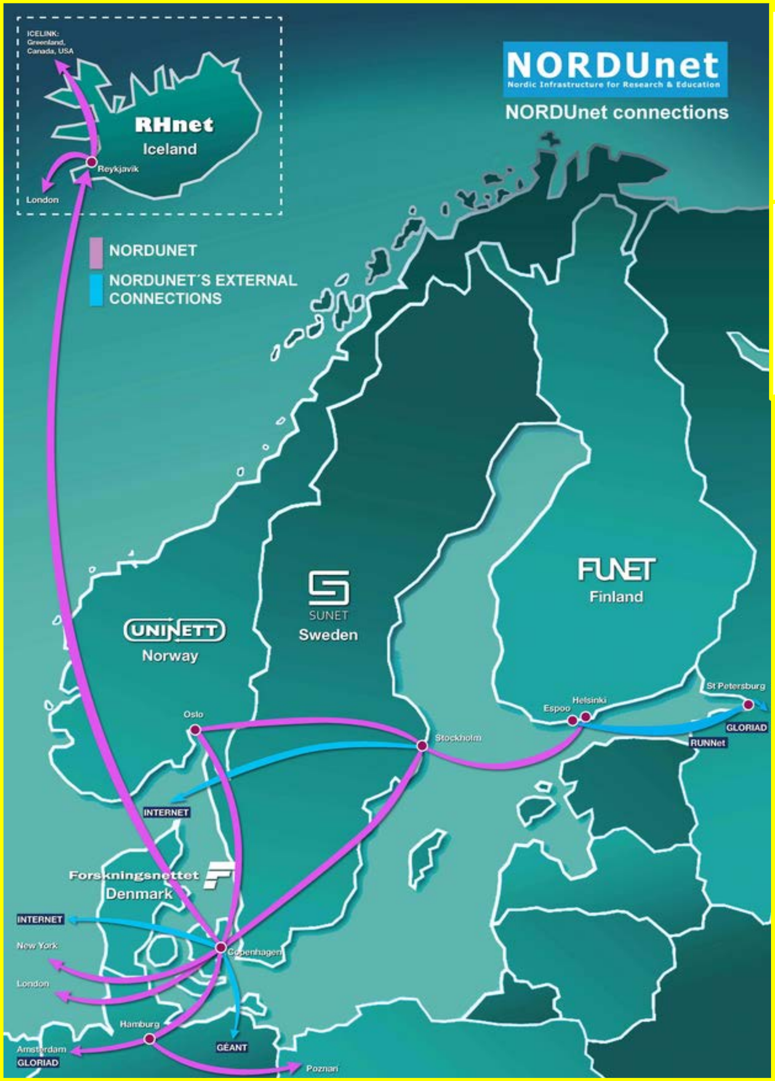
9.6 kbps Analog	1985 Radio Suisse, RCA [1]	
64-256 kbps Digital {X.25: IP: DECNet}	1989 – 1994	[X 7 – 27]
1.5 Mbps Shared	1990-3; IBM	[X 160]
2-4 Mbps	1996-1998	[X 200-400]
12-20 Mbps {ATM}	1999-2000	[X 1.2k-2k]
155-310 Mbps {OC3}	2001-2	[X 16k – 32k]
622 Mbps {OC12}	2002-3	[X 65k]
2.5 G λ	2003-4	[X 250k]
10 G λ {OC192; 10GE}	2005	[X 1M]…
~600G {100G Waves}	Today	[X ~60M]

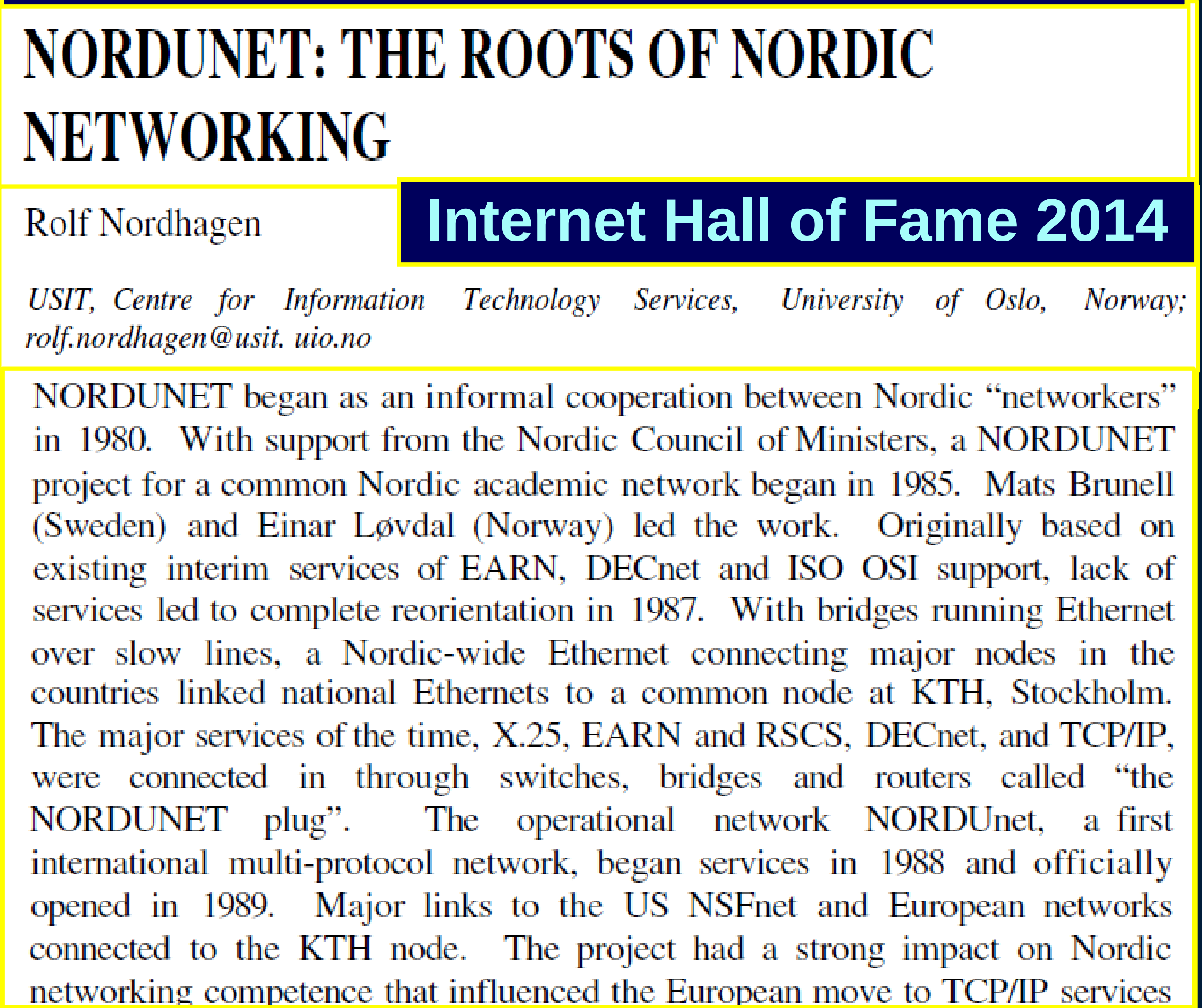
- ***HEP has become a leading applications driver, and a co-developer of global networks***

Note the Growth Trends: A factor of ~1M over 1985-2005 (~5k during 1995-2005 alone); *Only 60X Since 2005*

NORDUnet

1980 – 2016: From Informal Cooperation to Global Reach for R&E





NORDUNET: THE ROOTS OF NORDIC NETWORKING

Rolf Nordhagen

USIT, Centre for Information Technology Services, University of Oslo, Norway; rolf.nordhagen@usit. uio.no

NORDUNET began as an informal cooperation between Nordic "networkers" in 1980. With support from the Nordic Council of Ministers, a NORDUNET project for a common Nordic academic network began in 1985. Mats Brunell (Sweden) and Einar Løvdal (Norway) led the work. Originally based on existing interim services of EARN, DECnet and ISO OSI support, lack of services led to complete reorientation in 1987. With bridges running Ethernet over slow lines, a Nordic-wide Ethernet connecting major nodes in the countries linked national Ethernets to a common node at KTH, Stockholm. The major services of the time, X.25, EARN and RSCS, DECnet, and TCP/IP, were connected in through switches, bridges and routers called "the NORDUNET plug". The operational network NORDUnet, a first international multi-protocol network, began services in 1988 and officially opened in 1989. Major links to the US NSFnet and European networks connected to the KTH node. The project had a strong impact on Nordic networking competence that influenced the European move to TCP/IP services

Internet Hall of Fame 2014

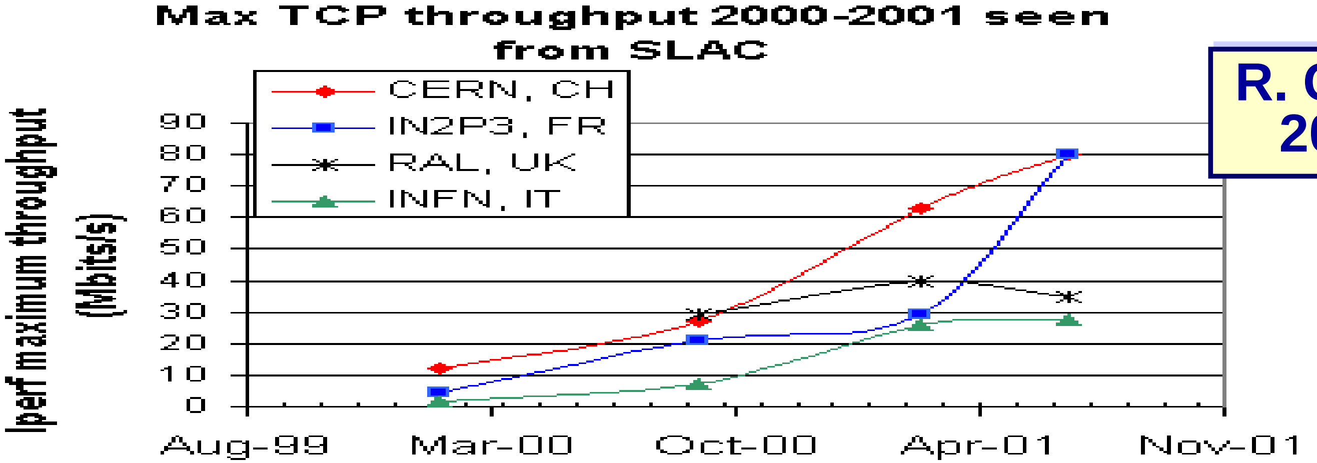
The First International Multi-Protocol Network: from 1988

Many parallels to the HEP experience in networking: LEP3Net

ICFA Retrospective: Networks for HENP, Conclusions in 2004 (Paris)

- **Current generation of 2.5-10 Gbps network backbones and major Int’l links arrived in 2002-2004 [US+Europe+Japan]**
  - **Capability Increased from ~4 to several hundred times [e.g. Slovakia], i.e. much faster than Moore’s Law**
  - **This is a direct result of the continued precipitous fall of network prices for 2.5 or 10 Gbps links in these regions**
- **Bandwidth Usage is growing by 80-100% Per Year**
- **Grids may accelerate this growth and the demand for seamless high performance**
- **Technological progress may drive BW higher, unit price lower**
  - **More wavelengths on a fiber; Cheap, widespread Gbit Ethernet**
- ***Some regions are moving to owned or leased dark fiber***
- ***The rapid rate of progress is confined mostly to the US, Europe, Japan and Korea, as well as the major Transatlantic routes; this*** **threatens to** ***cause the Digital Divide to become a Chasm***

HEP is Learning How to Use Gbps Networks Fully: Factor of ~50 Gain in Max. Sustained Throughput in 2 Years, On Some US+Transoceanic Routes



R. Cottrell
2000-1

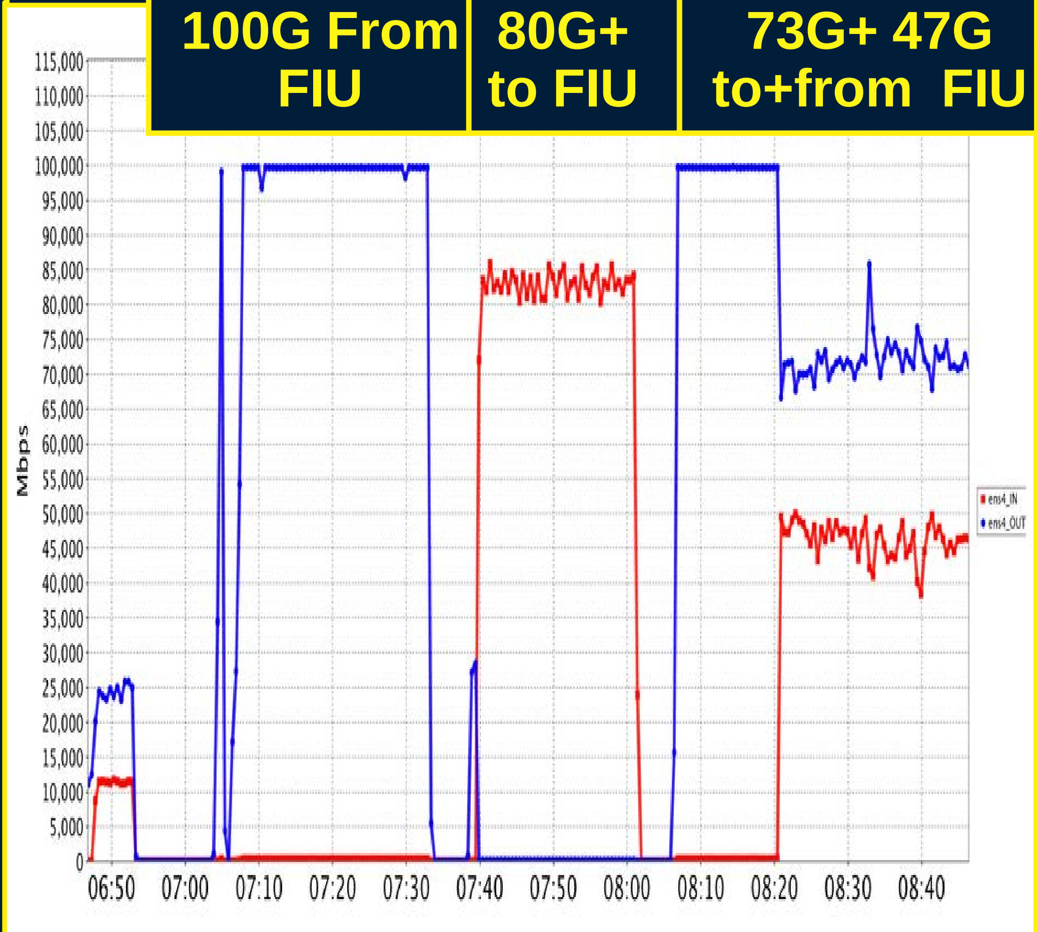
Caltech,
SLAC
and CERN

- 9/01 105 Mbps 30 Streams: SLAC-IN2P3; 102 Mbps 1 Stream CIT-CERN
- 5/20/02 450-600 Mbps SLAC-Manchester on OC12 with ~100 Streams
- 6/1/02 290 Mbps Chicago-CERN One Stream on OC12
- 9/02 850, 1350, 1900 Mbps Chicago-CERN 1,2,3 GbE Streams, 2.5G Link
- 11/02 [LSR] 930 Mbps in 1 Stream California-CERN, and California-AMS
  FAST TCP 9.4 Gbps in 10 Flows California-Chicago
- 2/03 [LSR] 2.38 Gbps in 1 Stream California-Geneva (99% Link Use)
- 5/03 [LSR] 0.94 Gbps IPv6 in 1 Stream Chicago- Geneva
- TW & SC2003: 5.65 Gbps (IPv4), 4.0 Gbps (IPv6) in 1 Stream Over 11,000 km

Milestones: FAST TCP from 2002 (Above); FDT from 2006

At SC15 Conference: Mellanox and Qlogic 100G NICs; 100G and N X 100G Results at SC15

FIU – Caltech Booth – Dell Booth



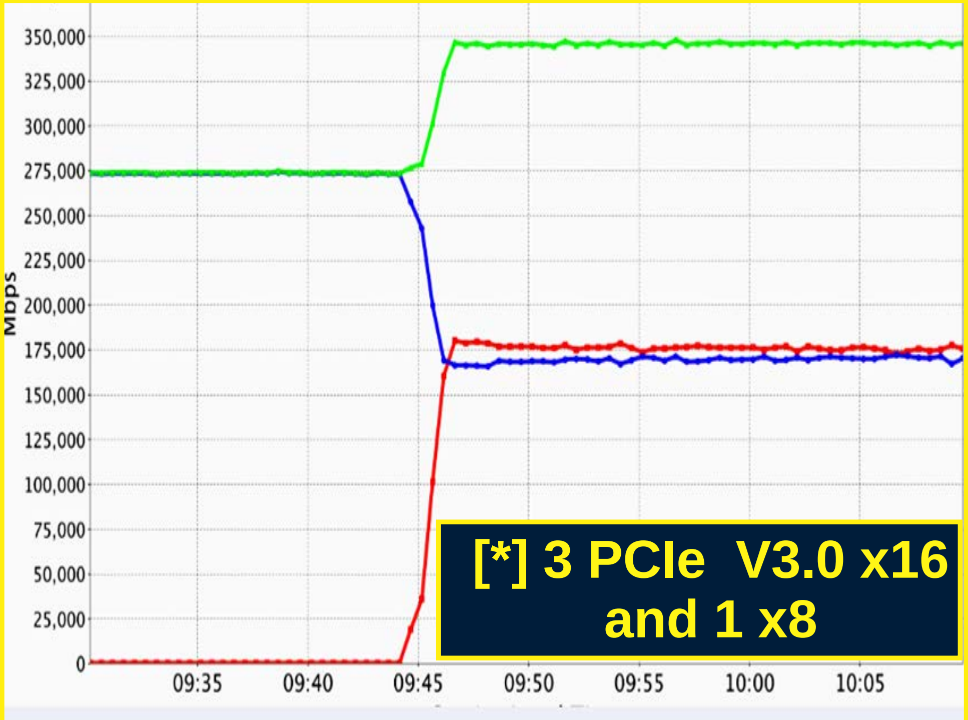
100G From FIU

80G+ to FIU

73G+ 47G to+from FIU

4 X 100G Server Pair in the Caltech Booth

275G out; 350G in+out [*]
Stable Throughput

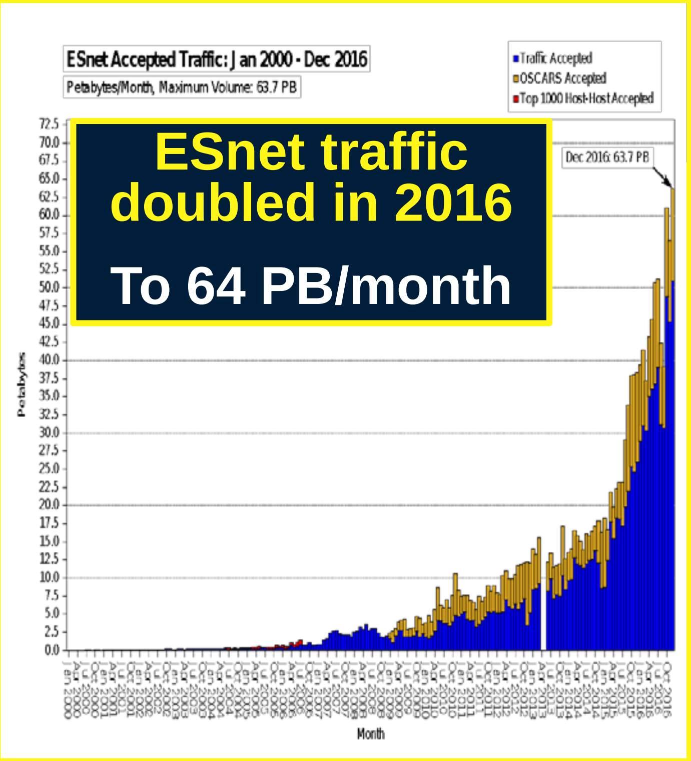


[*] 3 PCIe V3.0 x16 and 1 x8

Using Caltech's FDT Open Source TCP Application
http://monalisa.caltech.edu/FDT

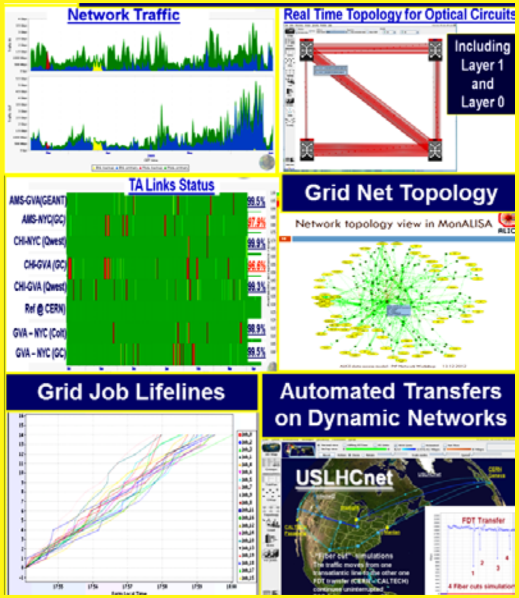
Entering a new Era of Exploration and Discovery in HEP and Other Data Intensive Sciences

- The resilient high capacity advanced network services provided by ESnet, US LHCNet, and partner networks around the world
  - Have been keys enabling the Higgs discovery
- New challenges in scale, complexity, global reach
- *A new class of intelligent software defined Systems encompassing N X 100G networks, computing and storage will be the new cornerstone*
- Enabling the next rounds of discovery, at LHC and in many other fields


ESnet Accepted Traffic: Jan 2000 - Dec 2016
Petabytes/Month, Maximum Volume: 63.7 PB
Traffic Accepted
OSCARS Accepted
Top 1000 Host-Host Accepted
Dec 2016: 63.7 PB
Petabytes
Month

ESnet traffic doubled in 2016

To 64 PB/month


Network Traffic
Real Time Topology for Optical Circuits
Including Layer 1 and Layer 0
TA Links Status
Grid Net Topology
Network topology view in MonALISA
Grid Job Lifelines
Automated Transfers on Dynamic Networks
USLHCnet
FDT Transfer

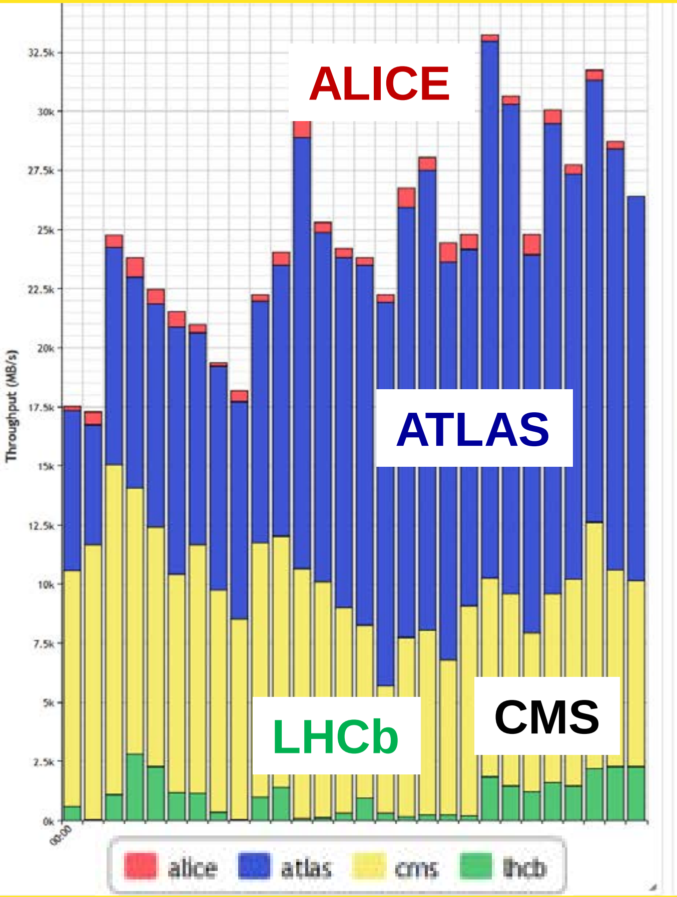
A New Era of Challenges and Opportunity

For Science, Networks and Society

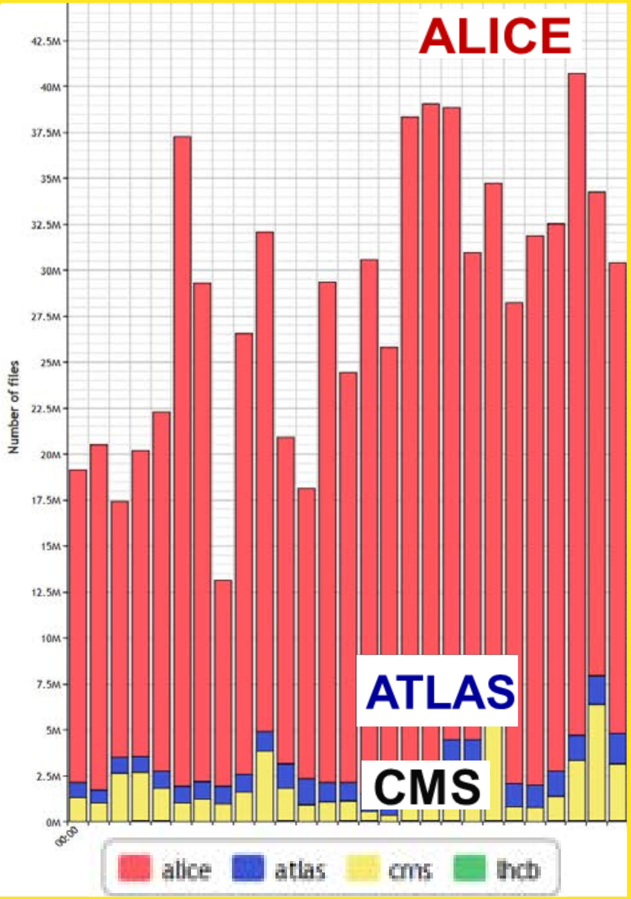
Complex Workflow: the Flow Patterns Have Increased in Scale and Complexity, even at the start of LHC Run2

WLCG Dashboard Snapshot April-May: Patterns Vary by Experiment

Transfer Throughput



Transfers Done/Day



28 GBytes/s yearly average; 40+ GBytes/s Peak Transfer Rates

Complex Workflow

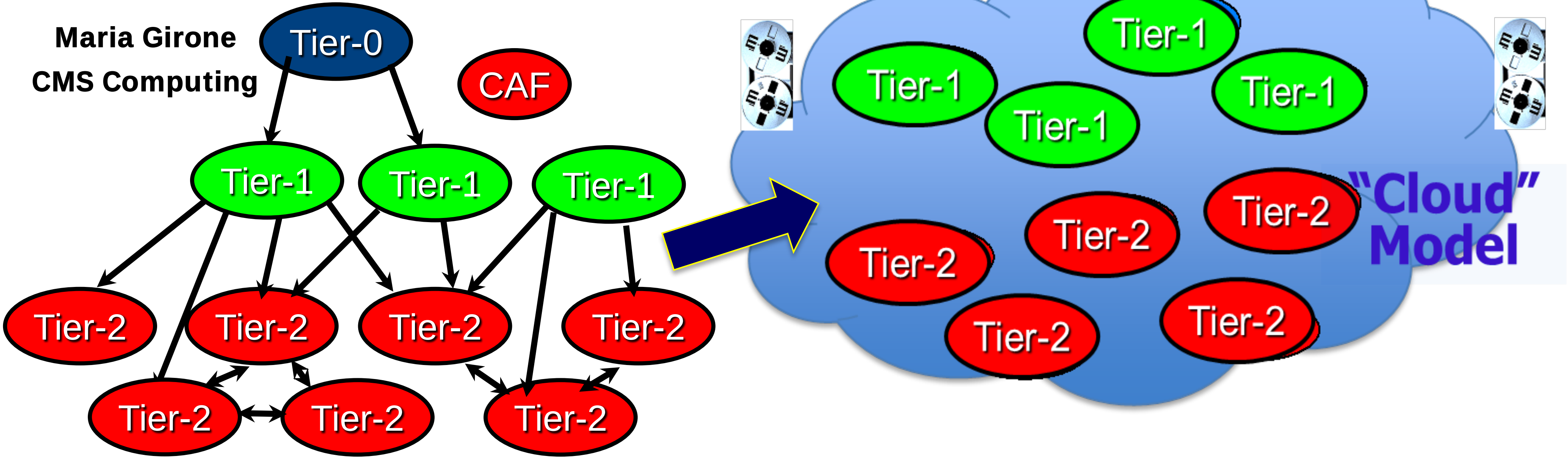
- Multi-TByte Dataset Transfers
- Transfers of up to 60 Million Files Daily
- Access to Tens of Millions of Object Collections/Day
- >100k of remote connections (e.g. AAA) simultaneously

2.7X Traffic Growth (+166%) in Last 12 Months; +60% in April Alone

Location Independent Access: Blurring the Boundaries Among Sites + Analysis vs Computing

- ❐ Once the archival functions are separated from the Tier-1 sites, the **functional difference between Tier-1 and Tier-2 sites becomes small [and the analysis/computing-ops boundary blurs]**
- ❐ **Connections and functions of sites are defined by their capability, including the network!!**

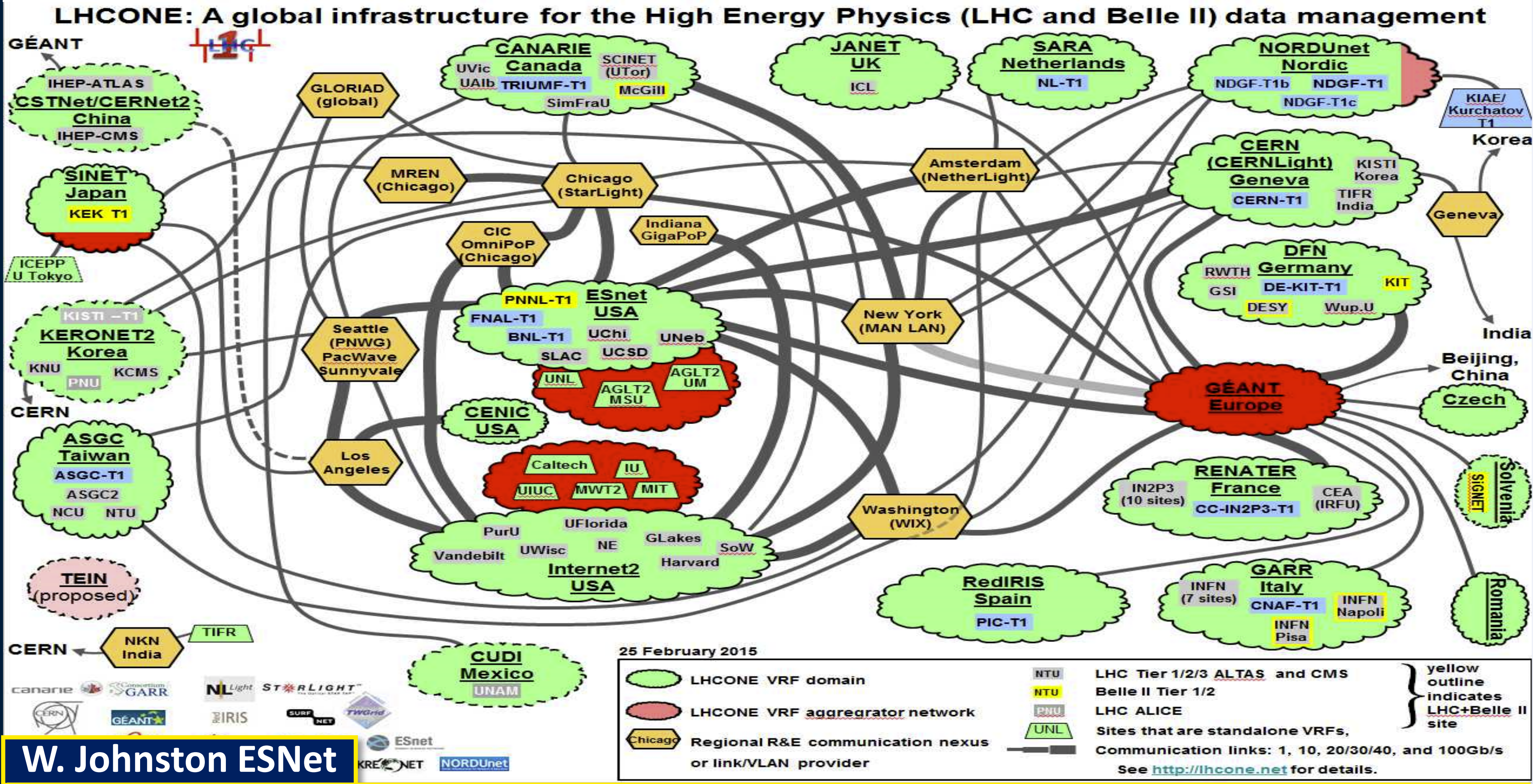
Maria Girone
CMS Computing



+Elastic Cloud-like access from some Tier1/Tier2/Tier3 sites

LHCONE: a Virtual Routing and Forwarding (VRF) Fabric

A global infrastructure for HEP (LHC, Belle II, NOvA, Xeon) data management



W. Johnston ESNet

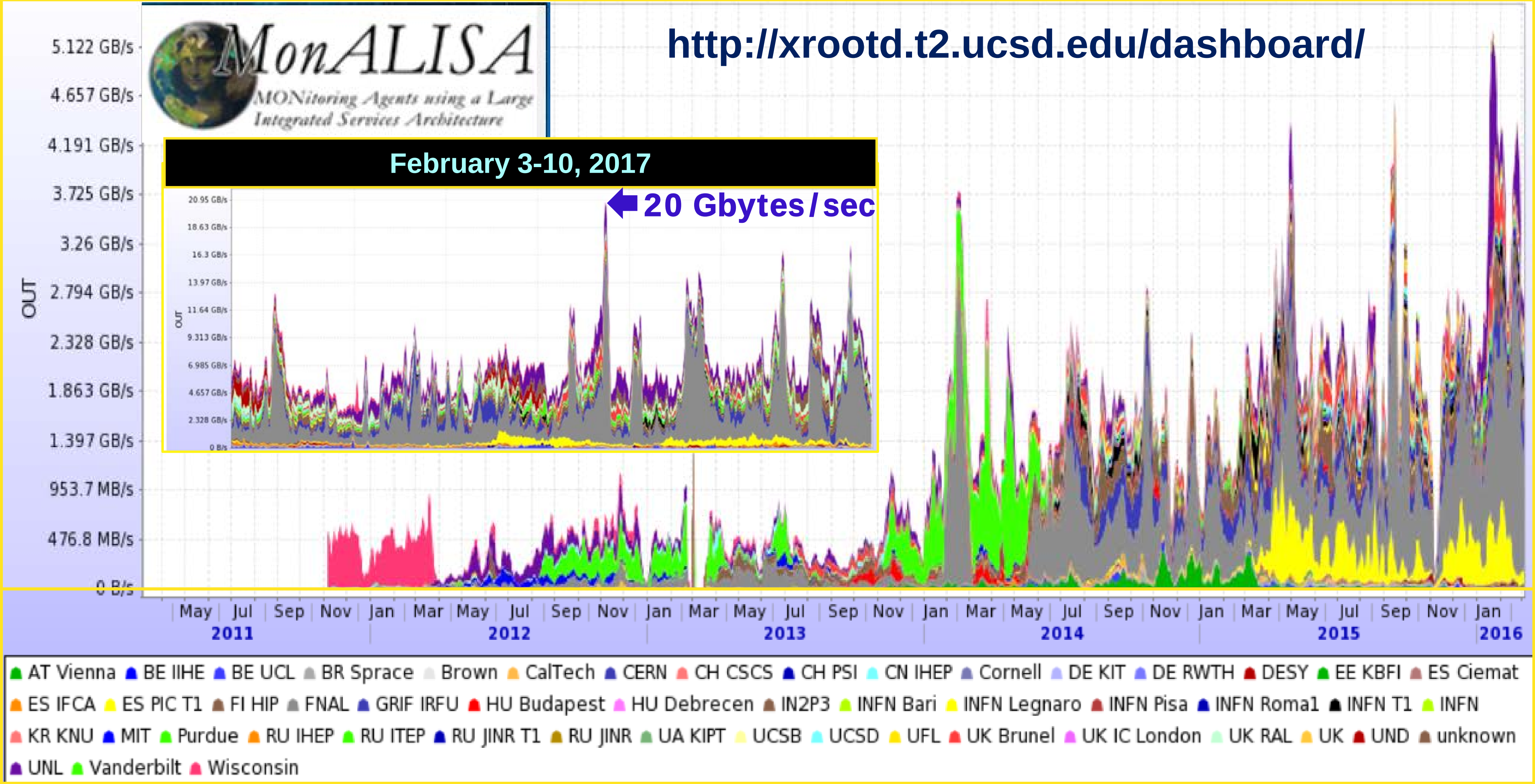
Good News: The Major R&E Networks Have Mobilized on behalf of HEP

Issue: A complex system with limited scaling properties.

LHCONE traffic grew by ~3-4X in 12 months: a challenge during Run2

Xrootd Traffic: Rapid Rise Since Fall 2013

US CMS XRootD Federation: *Any Data, Anytime, Anywhere*



http://xrootd.t2.ucsd.edu/dashboard/

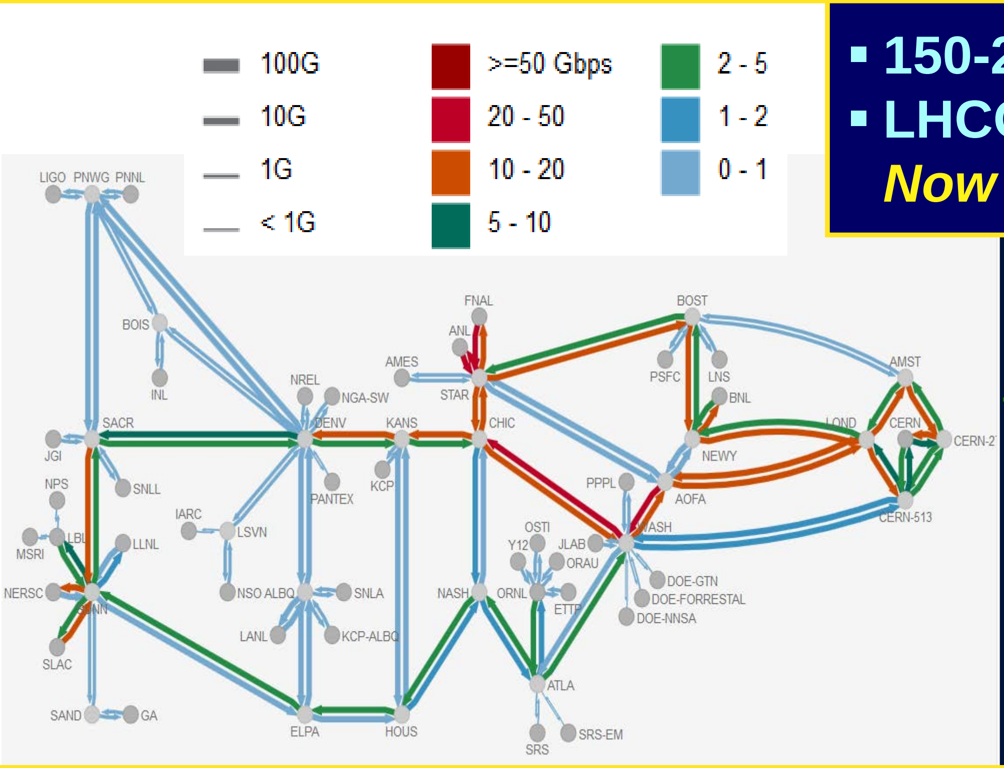
February 3-10, 2017

20 Gbytes/sec

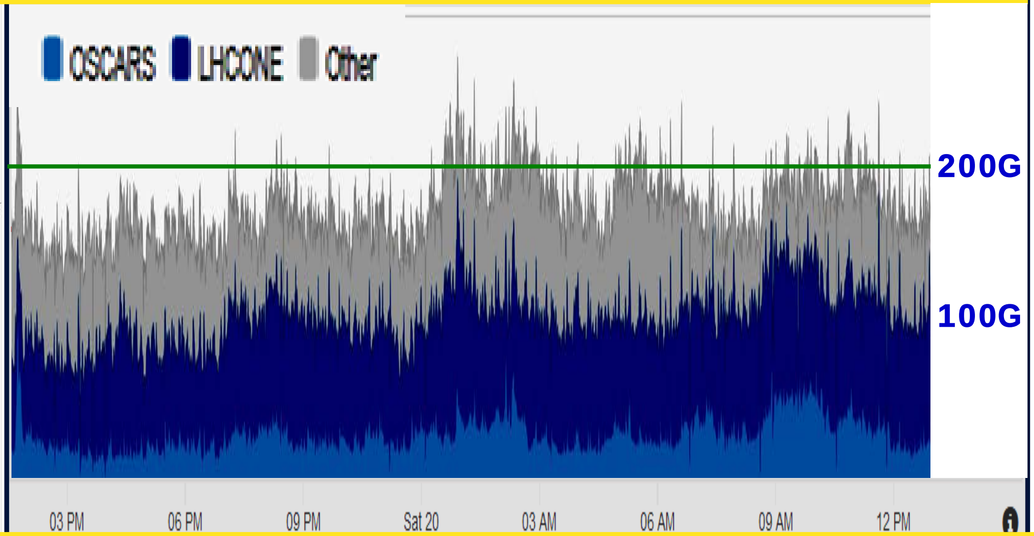
Several Gbytes/sec Sustained in 2016; Short term Peaks to 20 Gbytes/sec;
Flows will be greater if Tier2 throughput issues resolved

Energy Sciences Network
Updates and Outlook for 2016-18

- Long term traffic growth of 72%/year (10X per 4 Years) continues:
- 64 PB in Dec. 2016: 100% Growth in 2016; 100+ PB/mo by end 2017
- Stronger support for Universities, including through LHCONE
- PerfSONAR monitoring tools for users
- MyESnet (my.es.net) traffic portal: for both users and IT experts



- 150-200 Gbps Typical; Peaks to 300+ Gbps
- LHCONE *Rapid* Growth in 2015-16: *Now the largest class of ESnet traffic*

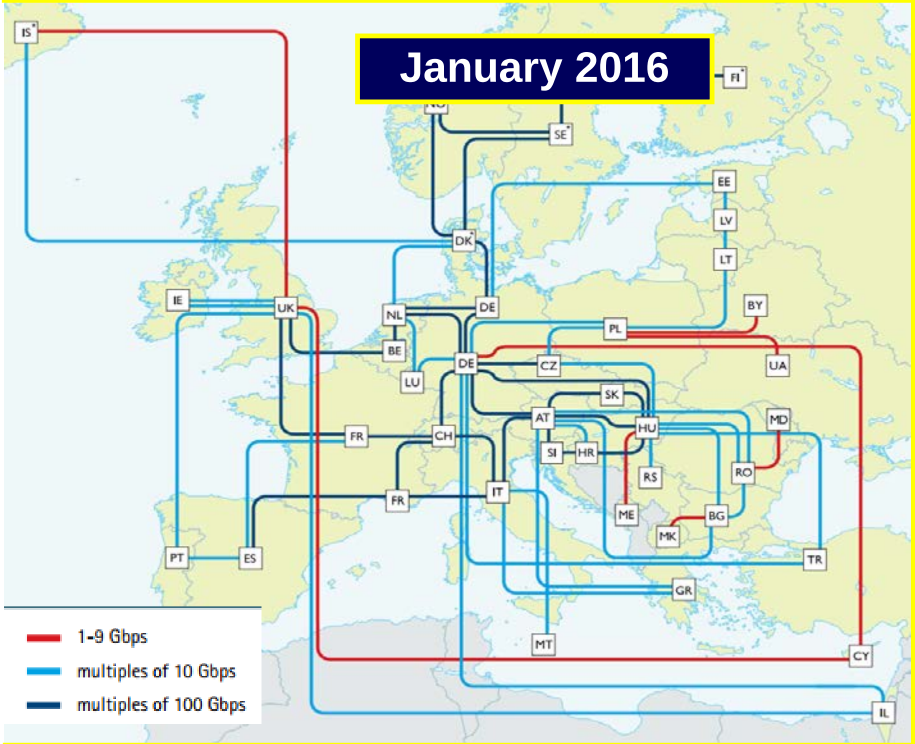


✷ ESnet6: the next SDN-enabled generation, is planned by 2019

GÉANT Pan-European Backbone
Highlights in 2016

www.geant.net

https://www.geant.org/Projects/GEANT_Project_GN4/Documents/GEANT%20Project%20Highlights%20GN4-1.pdf



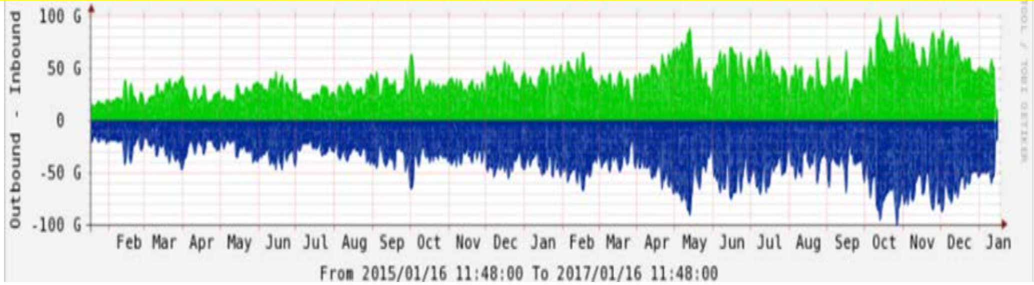
January 2016

- Traffic growth: core IP traffic increased by 64%. Combined with dedicated services for large users, total traffic volume was 1.425 Exabytes in 2016.
- 2nd iteration of the network evolution plan was developed. Great progress in such areas as fibre sharing, SDN and packet optical integration.
- Future evolution will include assisting in delivery of the European Open Science Cloud.
- GÉANT Testbed Service: 5 new GTS nodes deployed in Europe, supporting innovative uses of the network.
- Several new 10G circuits in Southern and Eastern Europe were added: improved connectivity to NRENs in those regions at lower cost.

Support to CERN/LHC

- Deployed 2nd 100G link between CERN and Wigner Center in Budapest
- LHCONE expansion to Asia-Pacific: ThaiREN 1st Asian NREN in TEIN to join
- Inclusion of Poland in LHCONE; discussing adding Portugal in 2017

LHCONE Traffic Grew 72% in 2016 With peaks above 100 Gbps

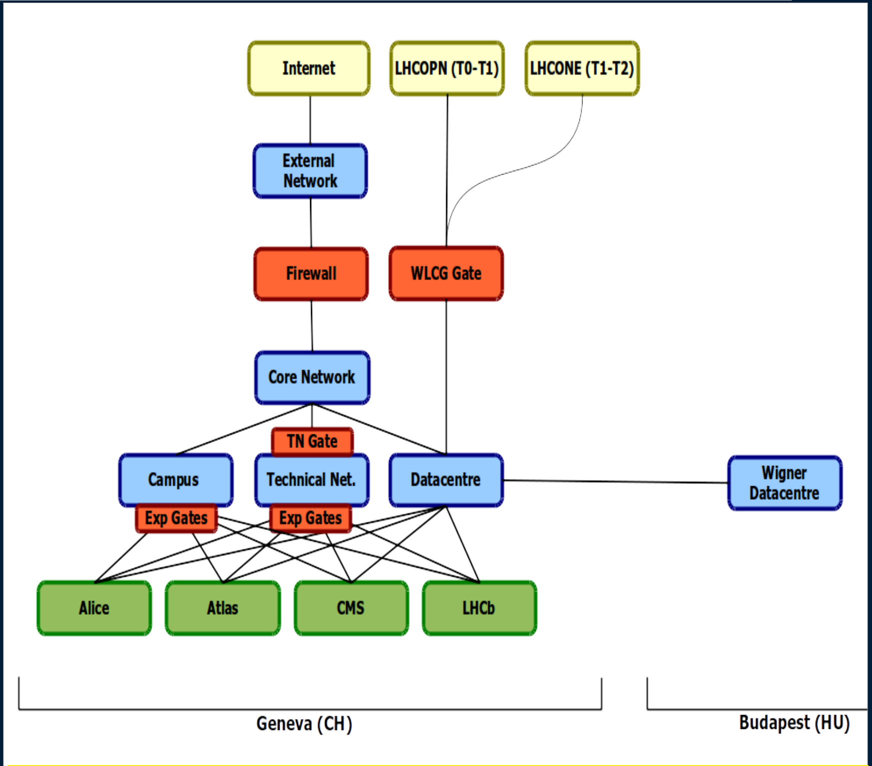


CERN Network Evolution Responding to the Demands

E. Martelli

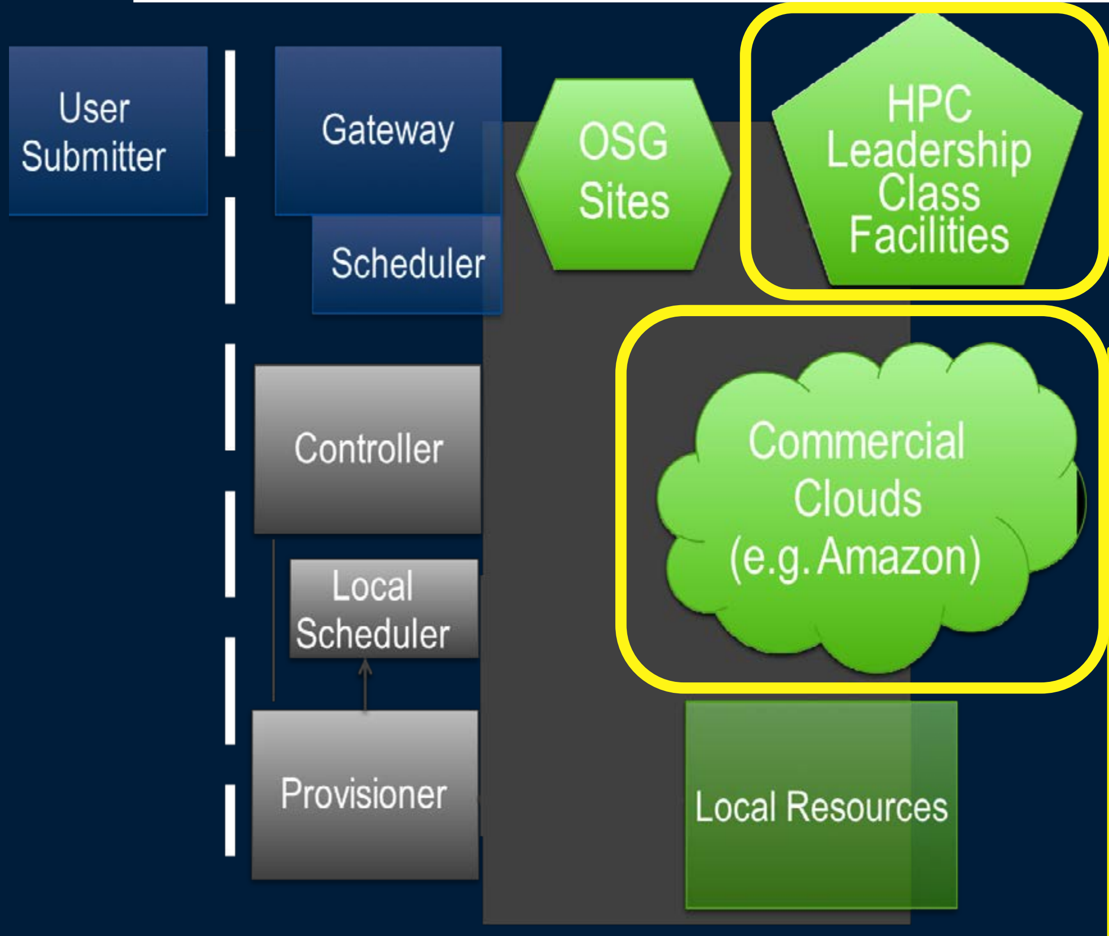
- In 2016 the traffic was unprecedented.
- **The LCG and GPN previously separate datacentre network infrastructures were merged.**
- With the delivery of the 3rd 100Gbps link Geneva-Budapest, the Wigner data centre also will be integrated in the unified infrastructure.

- The capacity of the datacentre backbone is 9.6 Tbps non-blocking
  - **It may double during 2017-2018, if the budget allows it**
- The LCG server farm is connected to the Tier1s **with an aggregate capacity of 300 Gbps**
- The connection to LHCONE is **200 Gbps**



- External connectivity to R&E networks has an aggregate capacity of 180Gbps and is provided by **ESnet, GEANT, NORDUnet, RENATER, SURFnet, and SWITCH.**

Fermilab has moved ahead with the HEPCloud Facility
To provision local, cloud and HPC Leadership Resources


User Submitter

Gateway

Scheduler

OSG Sites

HPC Leadership Class Facilities

Controller

Local Scheduler

Provisioner

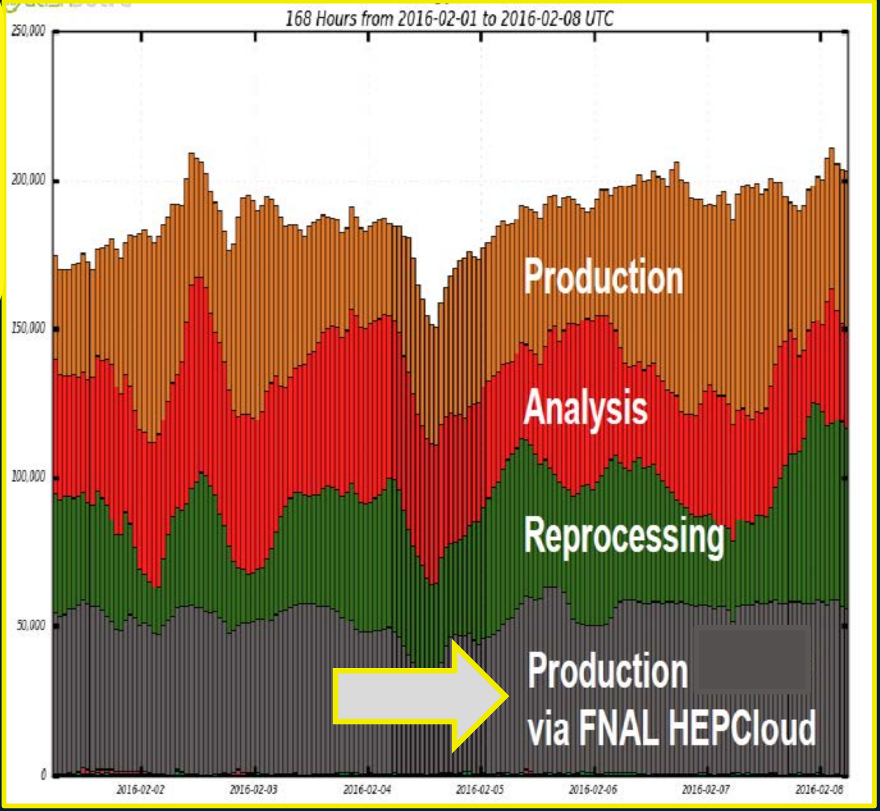
Commercial Clouds (e.g. Amazon)

Local Resources

LATBauerdick I DOE SC Exascale Requirements Reviews —High Energy Physics 06/10/15

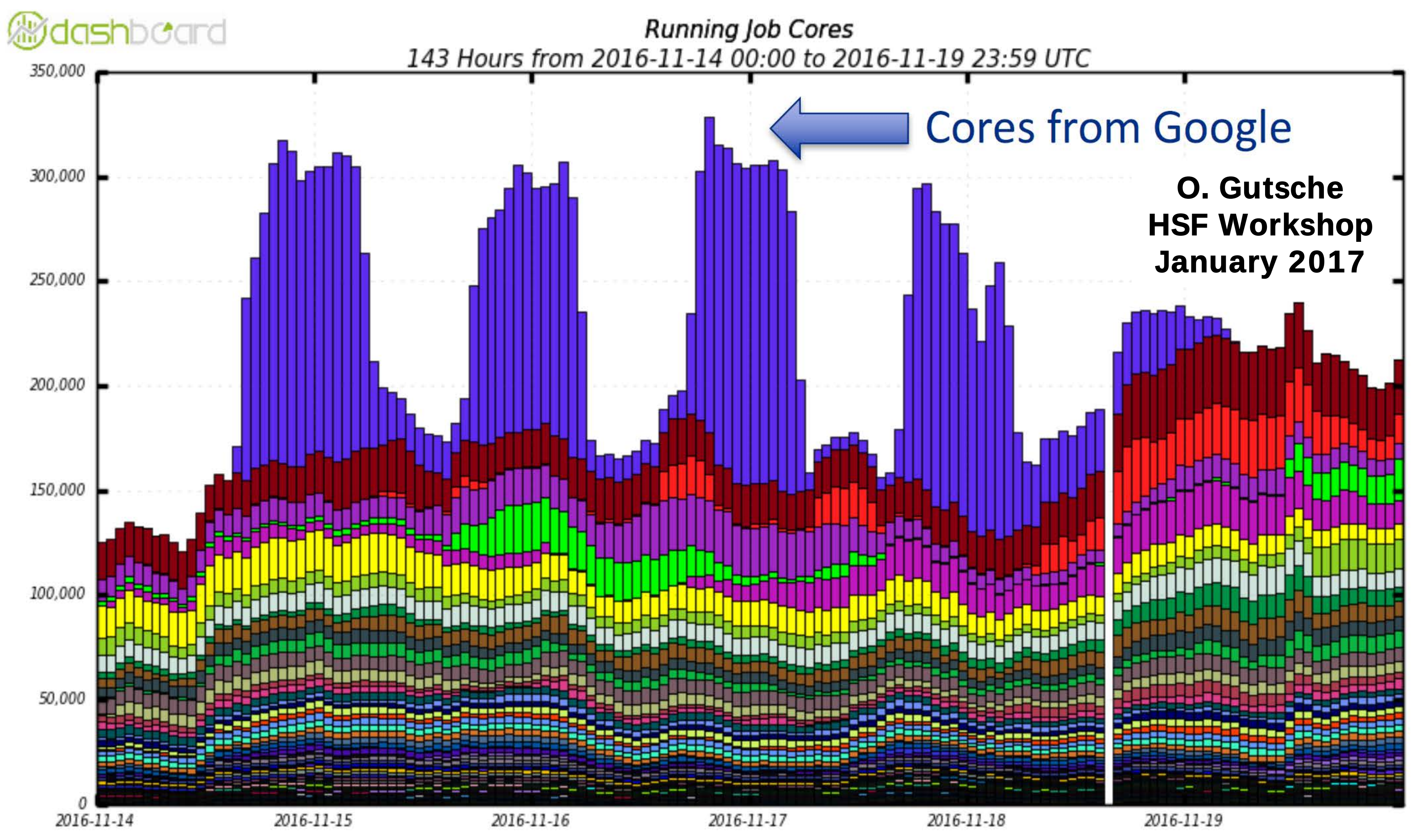
150k Core Demo at SC16: Double CMS Computing on Google Cloud

https://cloudplatform.googleblog.com/2016/11/Google-Cloud-HEPCloud-and-probing-the-nature-of-Nature.html



Issue beyond the 1st trials: Cost of extracting data from the Cloud

HEPCloud Facility: Doubling CMS Compute Capacity



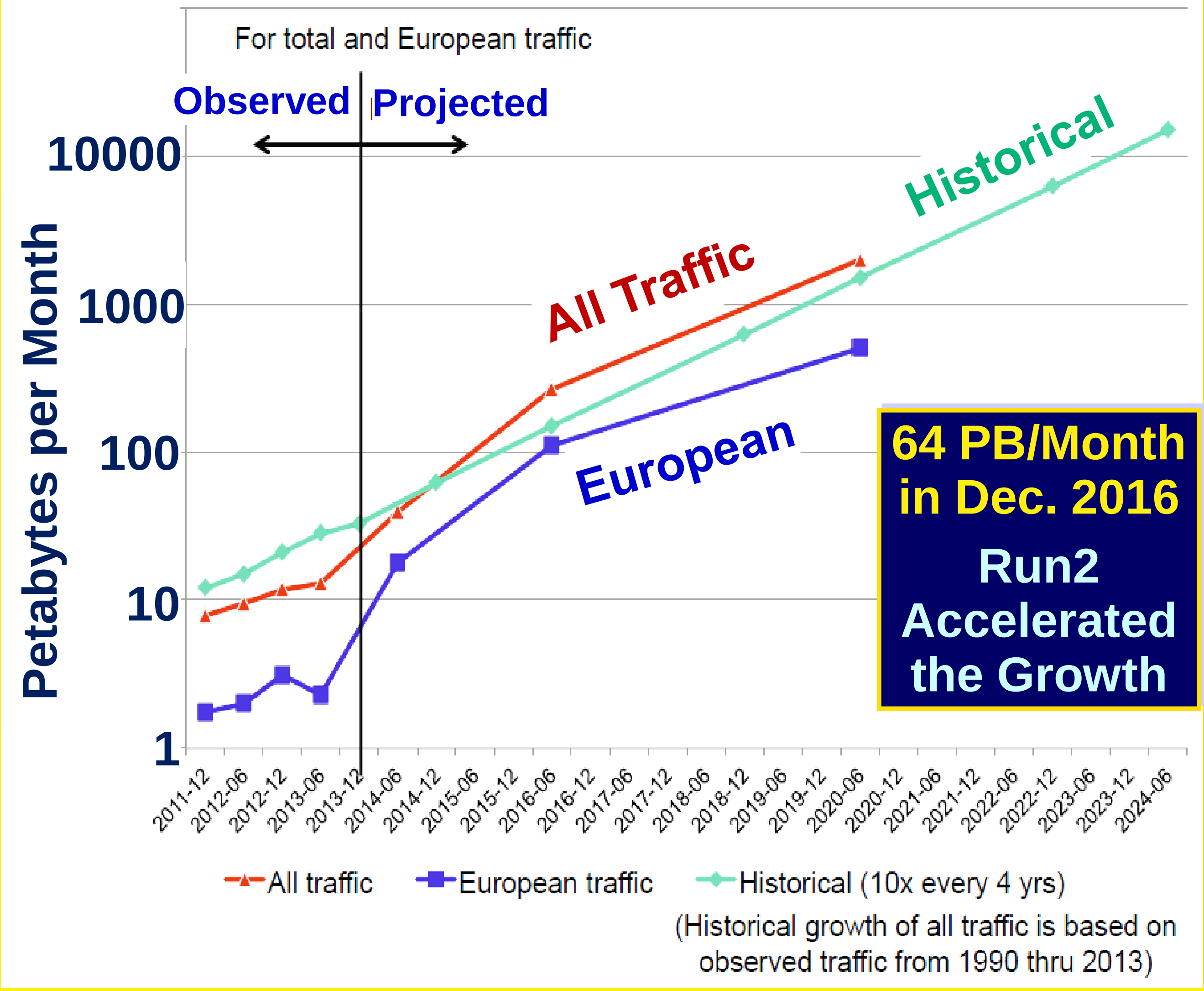
O. Gutsche
HSF Workshop
January 2017

Issue beyond the 1st trials is Cost: Cloud Provider Business Model

ESnet Science projection to 2024
Compared to historical traffic

E. Dart
W. Johnston

ESnet Total traffic handled in Petabytes per Month



(Historical growth of all traffic is based on observed traffic from 1990 thru 2013)

64 PB/Month in Dec. 2016

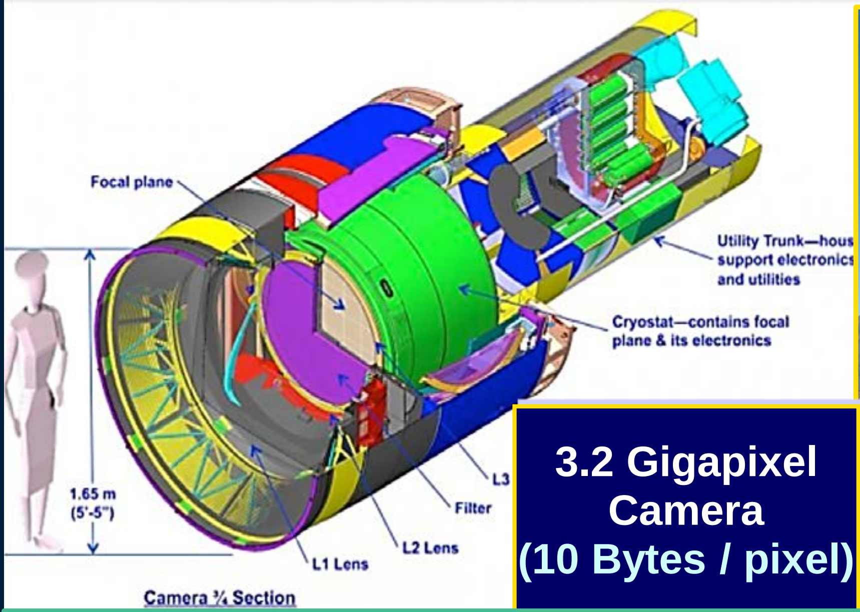
Run2 Accelerated the Growth

Projected Traffic Reaches 1 Exabyte Per Month. by ~2020
10 EB/Mo. by ~2024

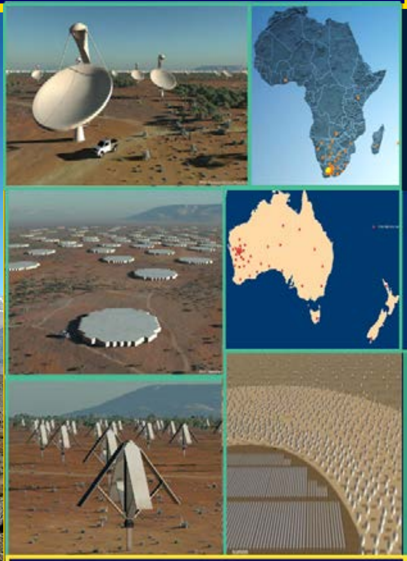
Rate of increase follows or exceeds Historical trend of 10X per 4 Years

HEP traffic will compete with BES, BER and ASCR

LSST + SKA Data Movement
Upcoming *Real-time* Challenges for Astronomy



3.2 Gigapixel Camera
(10 Bytes / pixel)



SKA

❐ Planned Networks: Dedicated 100G for image data, Second 100G for other traffic, and 40G for a diverse path

❐ Lossless compressed Image size = 2.7GB
(~5 images transferred in parallel over a 100 Gbps link)

- ❐ Custom transfer protocols for images (UDP Based)

❐ Real-time Challenge: delivery in seconds to catch cosmic “events”

❐ + SKA in Future: 3000 Antennae covering > 1 Million km2;
15,000 Terabits/sec to the correlators ➡ 1.5 Exabytes/yr Stored

The Future of Big Data Circa 2025:
Astronomical or Genomical ? By the Numbers

PLoS Biol 13(7): e1002195. doi:10.1371/journal.pbio.1002195

Domains of Big Data in 2025. In each, the projected annual and storage needs are presented, across the data lifecycle

Basis: 0.1 to 2B Humans with Genomes, replicated 30Xs;

+ Representative Samples of 2.5M Other Species' Genomes

Data Phase	SKA	Twitter	YOU TUBE	GENOMICS	HL LHC
Acquisition	25 ZB/Yr	0.5–15 billion tweets/year	500–900 million hours/year	1 Zetta-bases/Yr	
Storage	1.5 EB/Yr	1–17 PB/year	1–2 EB/year	2-40 EB/Yr	2-10 EB/Yr
Analysis	In situ data Reduction	Topic and sentiment mining	Limited requirements		
	Real-time processing	Metadata analysis		Variant Calling 2 X 10^{12} CPU-h	
	Massive Volumes			All-pairs genome alignment 10^{16}CPU-h	0.065 to 0.2 X 10^{12} CPU Hrs
Distribution	DAQ 600 TB/s	Small units of distribution	Major component of modern user's bandwidth (10 MB/s)	Many at 10 MBps Fewer at 10 TB/sec	DAQ to 10 TB/s Offline ~0.1 TB/s

Conclusion: Genomics Needs Realtime Filtering/Compression Before a Meaningful Comparison Can Be Made

The ICFA-SCIC
Network Monitoring WG

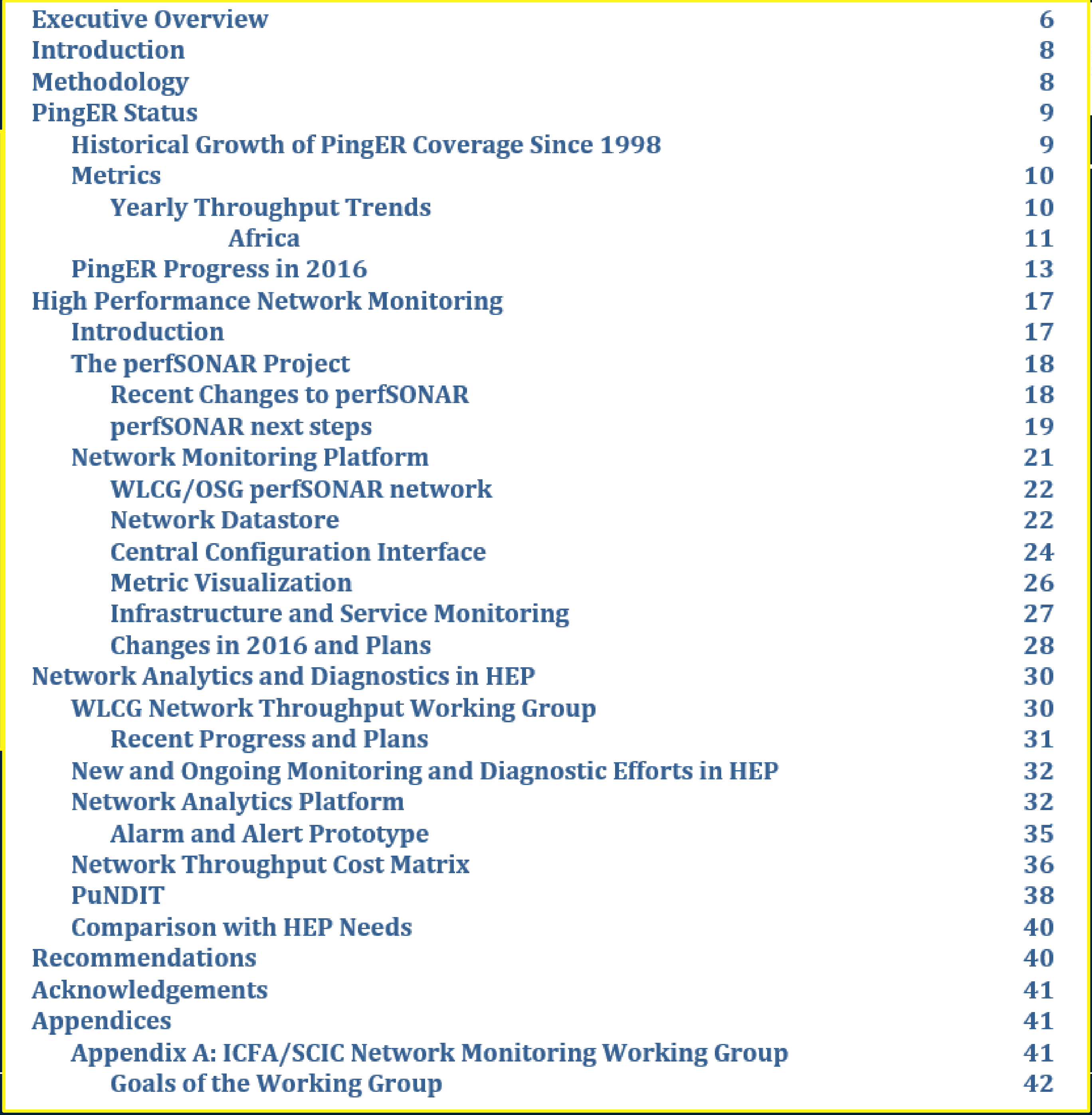
Shawn McKee/UM, Les Cottrell/SLAC,
Marian Babik/CERN, Ilija Vukotic/U Chicago

With contributions from Brian Tierney/LBNL,
Soichi Hayashi/IU, Mike O'Connor/ESnet

The 2016-2017 Monitoring WG Report is Available at:
https://docs.google.com/a/umich.edu/document/d/17odQd2C3CLKt7ZkOtLohP_MVY6A-jnnUZuOFB3st0r0/edit?usp=sharing

NOTE: "The PingER portion of the report is shortened relative to previous years due to support constraints"

The ICFA SCIC Network Monitoring WG Report Feb. 2017 ToC



The ICFA SCIC
Network Monitoring WG

- The ICFA-SCIC network monitoring group continues to organize and maintain global monitoring of the Research & Education networks relevant to high energy physics
  - Two methods are used to measure our networks: PingER and perfSONAR
  - PingER provides generic, low intrusiveness monitoring to track global trends
  - perfSONAR captures the state of our high-performing excellent networks
- The current report updates the January 2016 report. Some new areas related to network monitoring in HEP are included:
  - Updates and status on the perfSONAR efforts globally, and
  - WLCG Network and Transfer Metrics Working Group activities

ICTP The Abdus Salam International Centre for Theoretical Physics

ICFA

Universiti Malaysia Sarawak UNIMAS

NUST Defining futures

SLAC NATIONAL ACCELERATOR LABORATORY

UNIVERSITY OF MALAYA

UTM UNIVERSITI TEKNOLOGI MALAYSIA

UUM Universiti Utara Malaysia

Monitoring the World's Networks

PingeR Project (R. Cottrell et al.)

Mapping the Digital Divide

http://www-iepm.slac.stanford.edu/pinger/

IEPM Internet End-to-end Performance Monitoring

SCIC Monitoring WG
PingER (Also IEPM-BW)

IEPM
Internet End-to-end Performance Monitoring

R. Cottrell

- **Measurements from 1995 On**

  Reports link reliability & quality

- **Countries monitored**

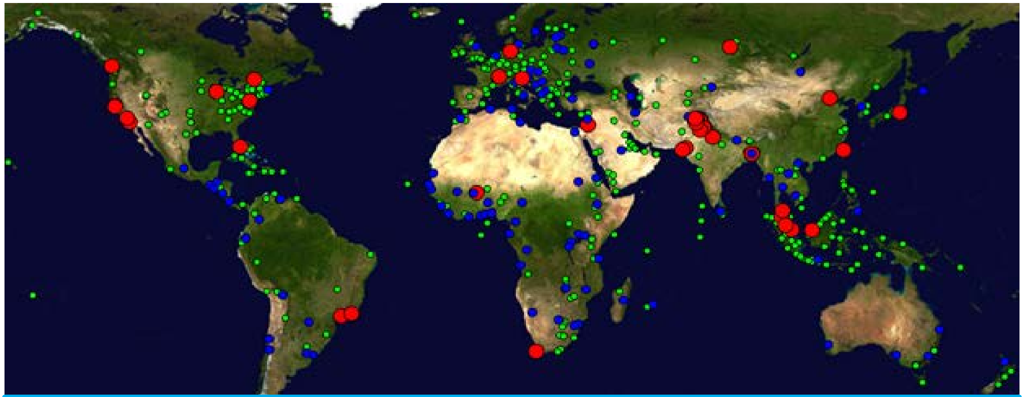
  ➔ **Contain >99% of world pop. and of World's Internet Users**

- **~800 remote sites monitored in 160-70 nations; from 97 (2011) down to 50 monitor nodes (2016)**

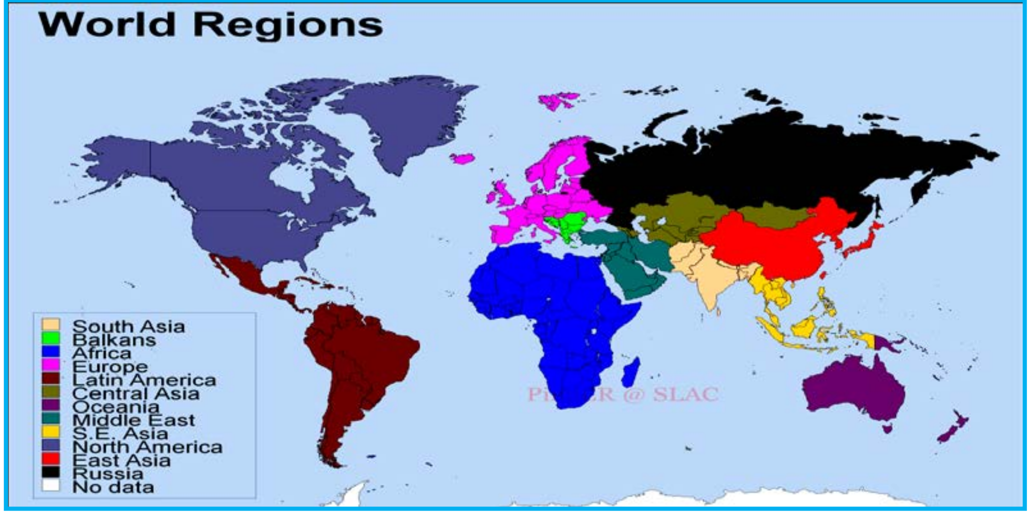
- **Excellent, Vital Work; Funding issue**

Countries (2016): N. America (3), Latin America (*25*), Europe (36), Balkans (*10*), Africa (47), Middle East (*16*), Central Asia (9), South Asia (*8*), East Asia (5), SE Asia (11), Russia (1), Oceania (5)

Monitoring & Remote Nodes Jan 2017)



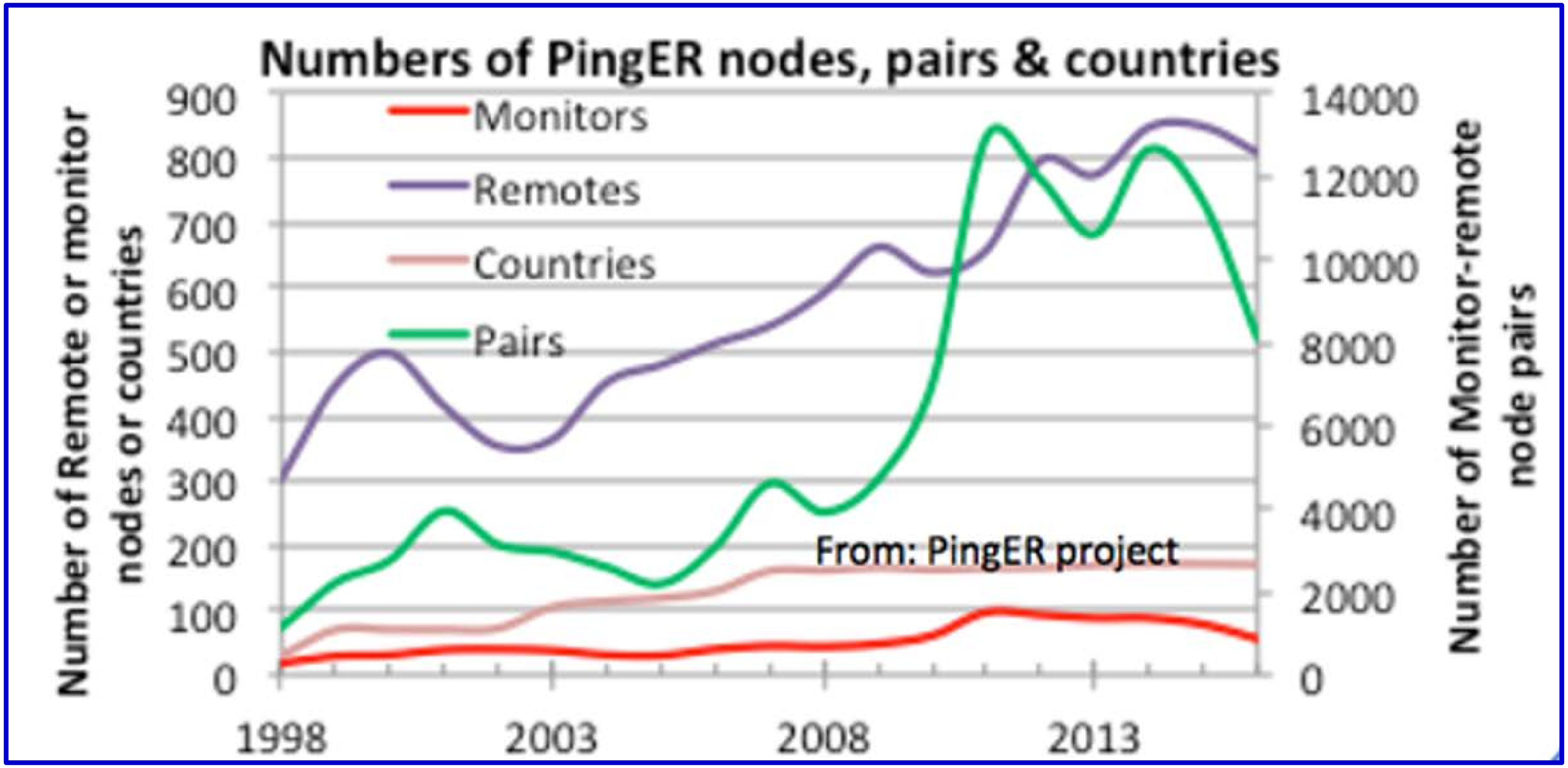
Locations of PingER monitoring and remote sites as of January 2017. Red sites are monitoring sites, blue sites are beacons that are monitored by most monitoring sites, and green sites are remote sites that are monitored by 1 or more monitoring sites.



PingER: Number of Nodes, Monitor – Remote Site Pairs and Countries

R. Cottrell

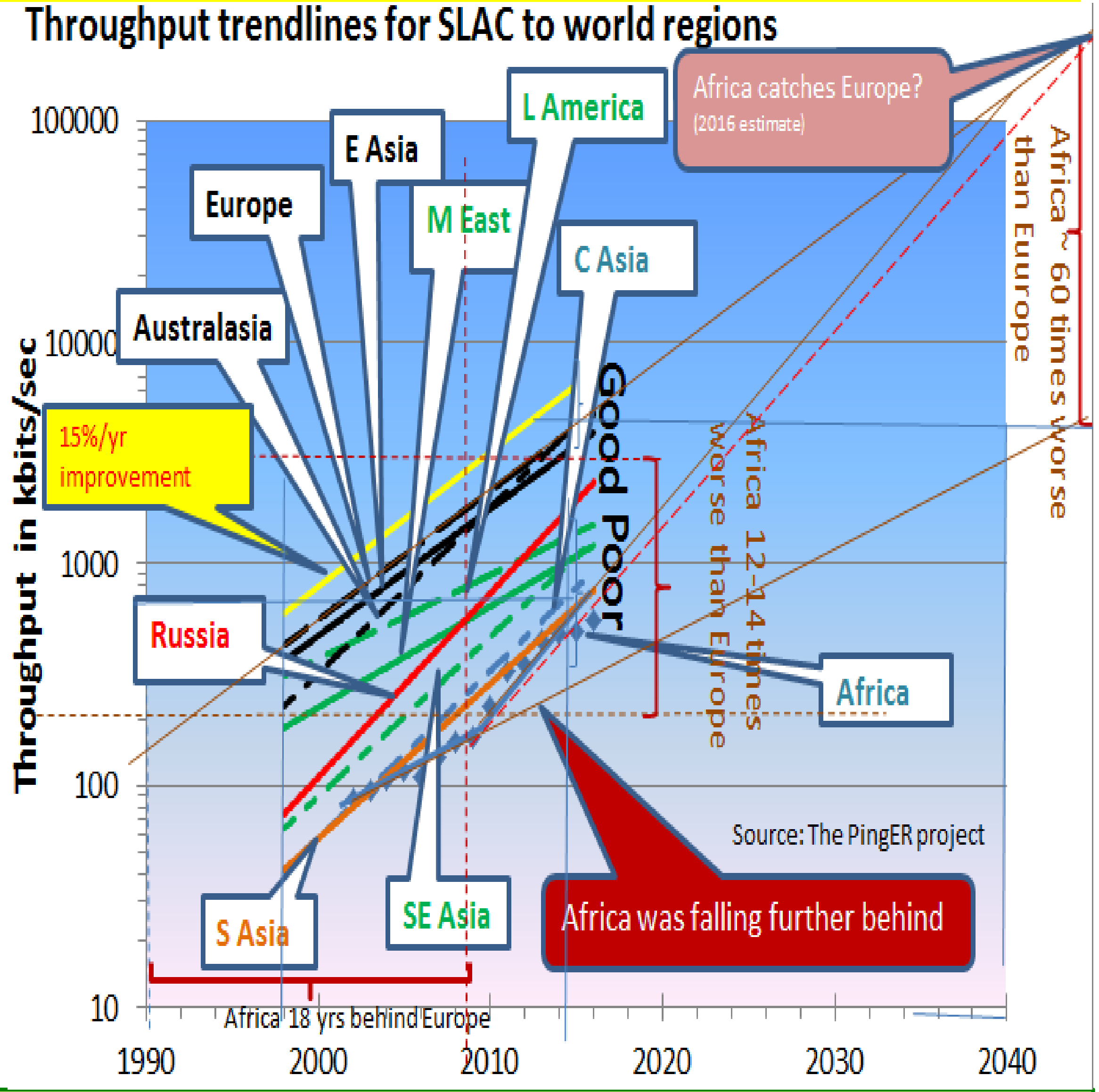
- **Number of Monitors has declined to half that in 2011(97 to 50)**
- **Number of Remote hosts has declined from 850 to 800 since 2013**
- **Number of Remote host – Monitor pairs has declined from 13k to 8k**
- **Number of Countries Monitored has plateaued**



Throughput Trendlines from SLAC 1990-2040+

- East Asia and Oceania are catching up to Europe
- Russia is 6 years behind Europe and catching up
- Latin America and the Middle East are 8 years behind and falling further behind
- S. E. Asia is also 8 years behind but is catching up
- S. Asia & Central Asia are 13 years behind & keeping up
- Africa has been catching up since 2010
  - But the rate has slowed
  - Africa now estimated to catch up in *2046*

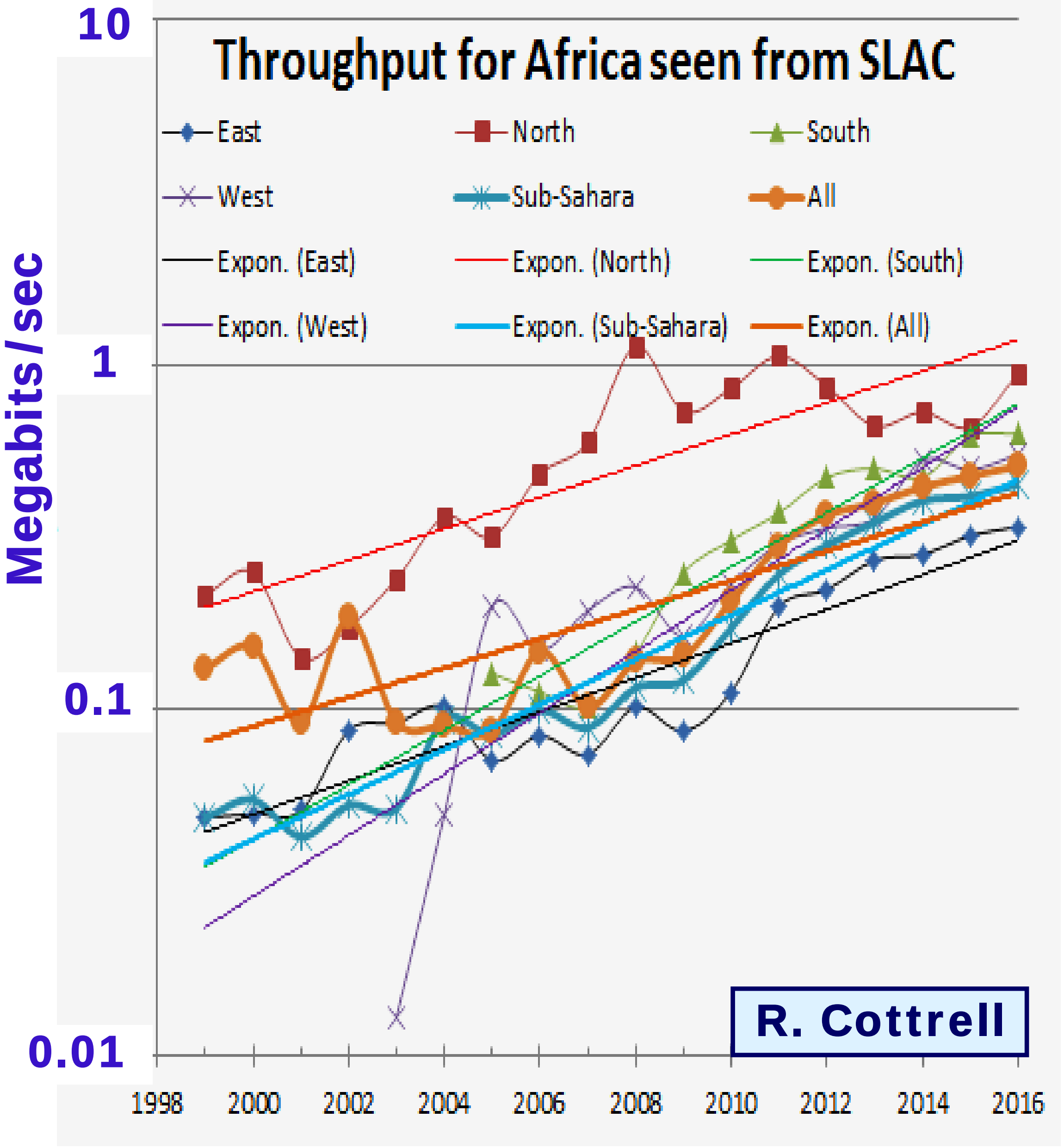
R. Cottrell



Derived TCP Throughput = 1460 Bytes*8bits/Byte/ (RTT * Sqrt(loss)); Matthis et al.

Trendlines from SLAC to Africa 1998-2016

- Instability in the All Africa data (up through 2003) is related to only monitoring <=3 sites in Africa
- North Africa, for long the leader, is being caught by the South and West African countries,
- The instability and lack of growth from 2009 on may be partially due to the "Arab Spring"
- Sub-Sahara is tracking all Africa but slightly lower.
- East Africa and West Africa saw a big improvement in 2010. They are still improving but much more slowly, possibly linearly rather than exponentially.



How to Reach the Rest of the World 2: O4B: "The Other 4 Billion"

R. Cottrell

- Refers to the population of the world without broadband:

✸ Medium Earth Orbit Satellites (MEOS)
Constellation of 8 at 8000km altitude launched in 2013-14

- Min RTTs factor of 4 less than GEOS: ~125ms, similar to intercontinental land lines
- Backed by SES World Skies, HSBC, Google…
- Needs steerable ground stations: Lifetime ~ 10 years:

✸ Low Earth Orbiting Satellites (LEOs)

- SpaceX asked FCC approval for 4425 LEO ($ 10-15B) fleet; 1st of 800 for US, Puerto Rico and Virgin Islands https://cdn.geekwire.com//wp-content/uploads/2016/11/Technical-Attachment.pdf
- Google plans to invest $ 1B in fleet of 180
- Virgin and Qualcomm have invested in launching 648 low orbit satellites

How to Reach the Rest of the World: Summary

- Satellites can last decades, balloons & drones must be constantly replenished, and many more are needed to cover the Earth.
- Google and SpaceX believe they have a real shot at connecting the 57% (4 billion) of the world's population still offline.
- Google’s Loon Project Balloons are deploying; its OneWeb LEO project is still in the formative stage
- SpaceX has a well developed plan for a huge LEO constellation and has applied to the US FCC for the necessary spectrum
- It’s likely we’ll end up in an “all of the above” world, in which distant, powerful satellites provide for streaming media while an assortment of balloons, and close-in satellites will provide a more responsive Internet.

R. Cottrell + HN

PingER Status and Progress

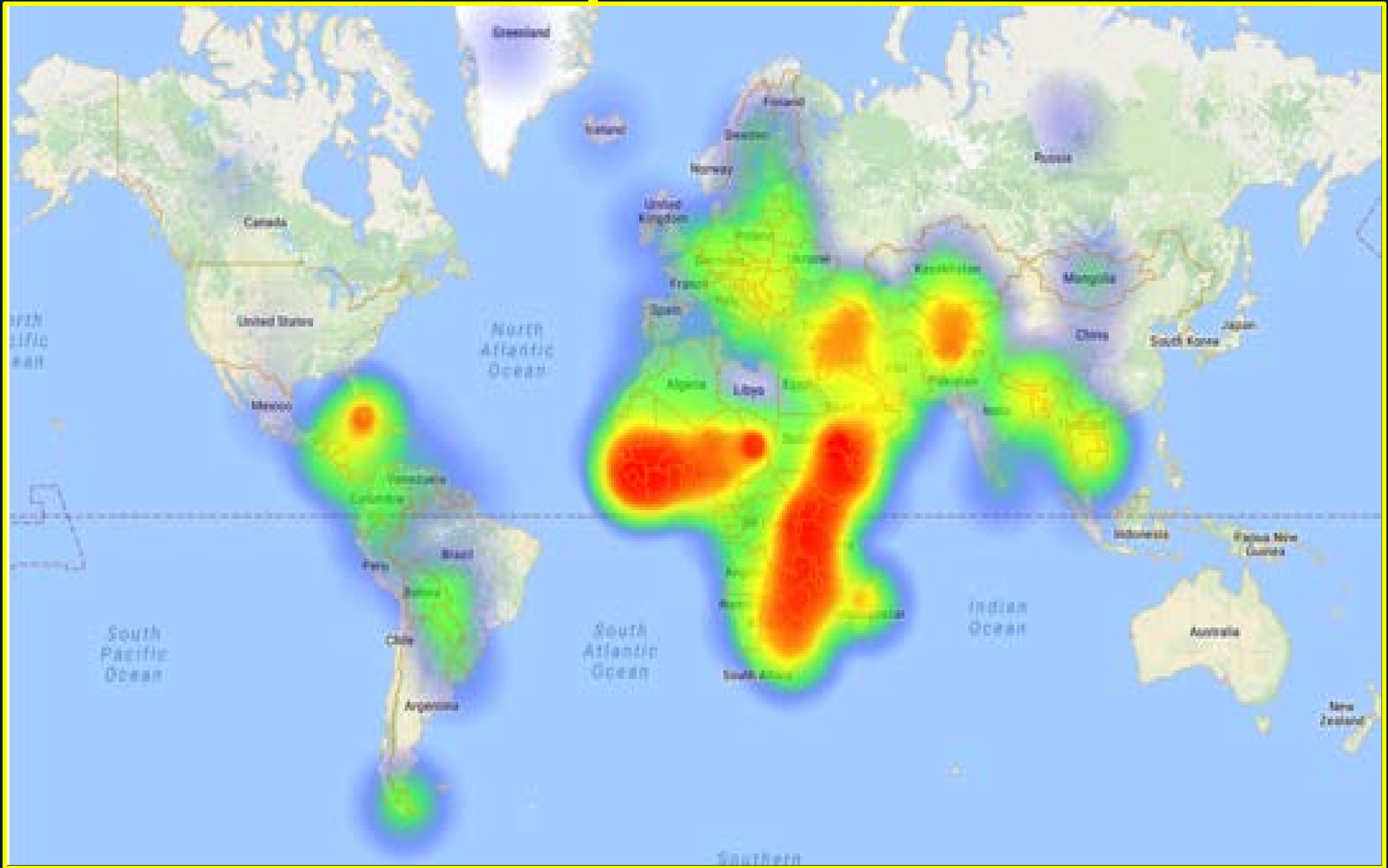
The PingER collaboration meets monthly by Skype:

- SLAC
- National University of Sciences and Technology (NUST), Islamabad, Pakistan
- University of Agriculture (Faisalabad) (UAF), Faisalabad, Pakistan
- University of Malaysia in Sarawak (UNIMAS), Kuching, Malaysia
- University Utara Malaysia (UUM), Sintok, Kedah, Malaysia
- Amity University, Noida, Uttar Pradesh, India

Progress in 2016

- Joao Rulff a student from Brazil spent 3 months at SLAC working on a PingER data warehouse
- New automatic updates of the FTP site with PingER data
- PingER Measurement Archive moved to a virtual machine for ease of backup/mgmt.
- PingER under active development in Brazil
- Data normalization and visualization pipeline using Python created.
- New “heat-map” data visualization created

PingER Heat Map: Showing the Round Trip Time from SLAC



PingER Project Issue
Funding for Managing the Effort

R. Cottrell

- The management and operation includes maintaining data collection and archiving, explaining needs, identifying and reopening broken connections, identifying and opening firewall blocks, finding replacement hosts, making limited special analyses and case studies, preparing and making presentations, responding to questions.
  - The equipment performing this in this country is currently in place at SLAC. There is also an archive/analysis site in Pakistan.
- Management, operation and supervision requires central funding at a level of about 20% of a Full Time Equivalent (FTE) person, plus travel. This had been provided by discretionary funding from the HEP budgets of SLAC and FNAL, But this ended at the beginning of 2008.

PingER: Uncertainty for Managing the Future

R. Cottrell

- Many agencies/organizations have expressed interest (e.g DoE, ESnet, NSF, ICFA, ICTP, IDRC, UNESCO, IHY) in this work, also Google is interested in the historical interest now and going forward, but none have so far stepped up to funding the management and operation.
- Without funding, for the operational side, the future of PingER and reports such as this one is unclear, and the level of effort sustained in previous years will not be possible.

PingER: What is Needed

R. Cottrell

- Moral support, legitimacy
  - Most work is on my spare time
  - Immediate management aware but limited interest
- Travel money for conferences (one or two/year, typically international), workshop
- Some % of an FTE to supervise etc. students
- Graduate student funding for visit to SLAC for up to a year
- Without funding, for the operational side, the future of PingER and reports such as this one is unclear, and the level of effort sustained in previous years will not be possible.

WLCG, Open Science Grid, Network Related Developments

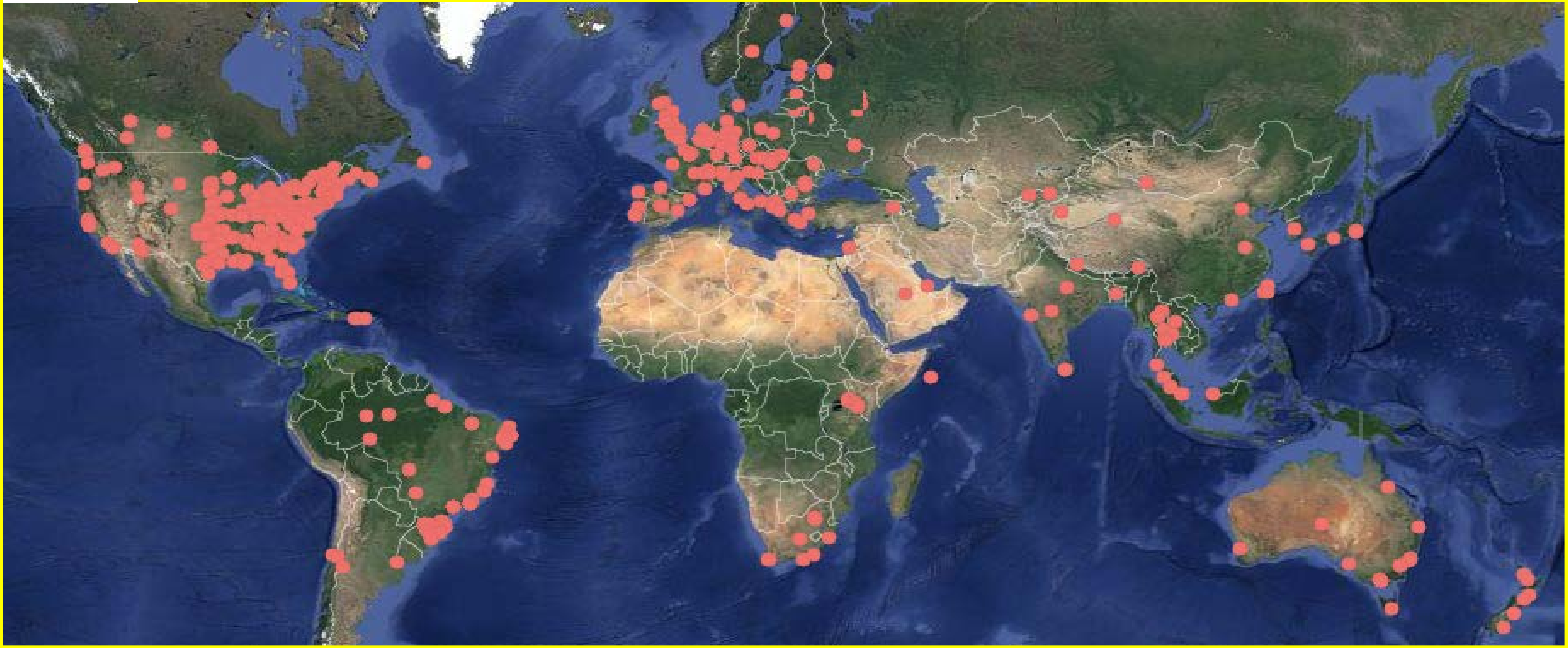
WLCG Network Throughput WG

- WLCG has a Network and Transfer Metrics WG with several tasks: https://twiki.cern.ch/twiki/bin/view/LCG/NetworkTransferMetrics
- A WLCG Network Throughput WG was formed in 2014 within the scope of WLCG operations with the objectives:
  - Ensure sites and experiments can better understand and fix networking issues
  - Measure end-to-end network performance and use the measurements to single out on complex data transfer issues
  - Help determine the current status of our networks to improve overall transfer efficiency
- Core activities:
  - Deployment and operations of perfSONAR infrastructure: to gain visibility into how our networks operate and perform
  - Network performance incidents response team: To provide support to help debug and fix network performance issues
  - Network Analytics: To Improve our ability to fully utilize the existing network capacity

Network Analytics Activities

- Ilija Vukotic (Univ. of Chicago) has developed an ELK/Jupyter stack for ATLAS Analytics and worked with Xinran Wang on *anomaly detection and advanced alerting/notifications* for network problems
- Jerrod Dixon and Brian Bockelman (Nebraska) are exploring network analytics in CMS
- Shawn McKee (Michigan) is working on real-time root cause analysis (PuNDIT) in collaboration with ESNet
- Henryk Giemza (NCBJ), Federico Stagni are integrating perfSONAR in DIRAC for LHCb
- Hendrik Borras (Heidelberg) and Marian Babik (CERN) are working on developing models for *network cost-matrices,* to determine the performance of network paths

WLCG perfSONAR Network



- ~2K perfSONAR instances deployed world-wide within ~1K domains
- ~ 50% on 10Gbps connectivity, > 60 instances at 40Gbps
- 8% of instances running on virtual machines, rest bare metal, mostly Centos6
- Current perfSONAR version: 4.0
- New features: web-based config interface; new test scheduler (pscheduler replaces bwctl), pluggable support, archive backends (RabbitMQ), REST API; improved graphing support and dashboards.

perfSONAR Developers

- The perfSONAR developers continue to focus on improving and supporting a robust network measurement toolkit
- The HEP community has been one of their most important customers and has provided feedback about bugs and needed features for many years
- The global HEP community has helped shape the current perfSONAR toolkit and continues to be an important partner in perfSONAR development
- The mutual goal is to provide a robust, standardized way to measure network metrics to better manage, maintain and upgrade our global networks

WLCG Monitoring WG Roadmap

Short Term

- **Focus on improving efficiency of current network links.** Continuing developments in network analytics, integration of more flow/router information and FTS data, alerting/notifications, etc.
- **Validation and deployment campaign for perfSONAR 4.0 and 4.1,** to be completed this year
- **Updates to central services** (configuration, monitoring, collectors, messaging, etc.)
- **Tracking the evolution in Software Defined Networks,** where new pilots/demonstrator projects will be proposed

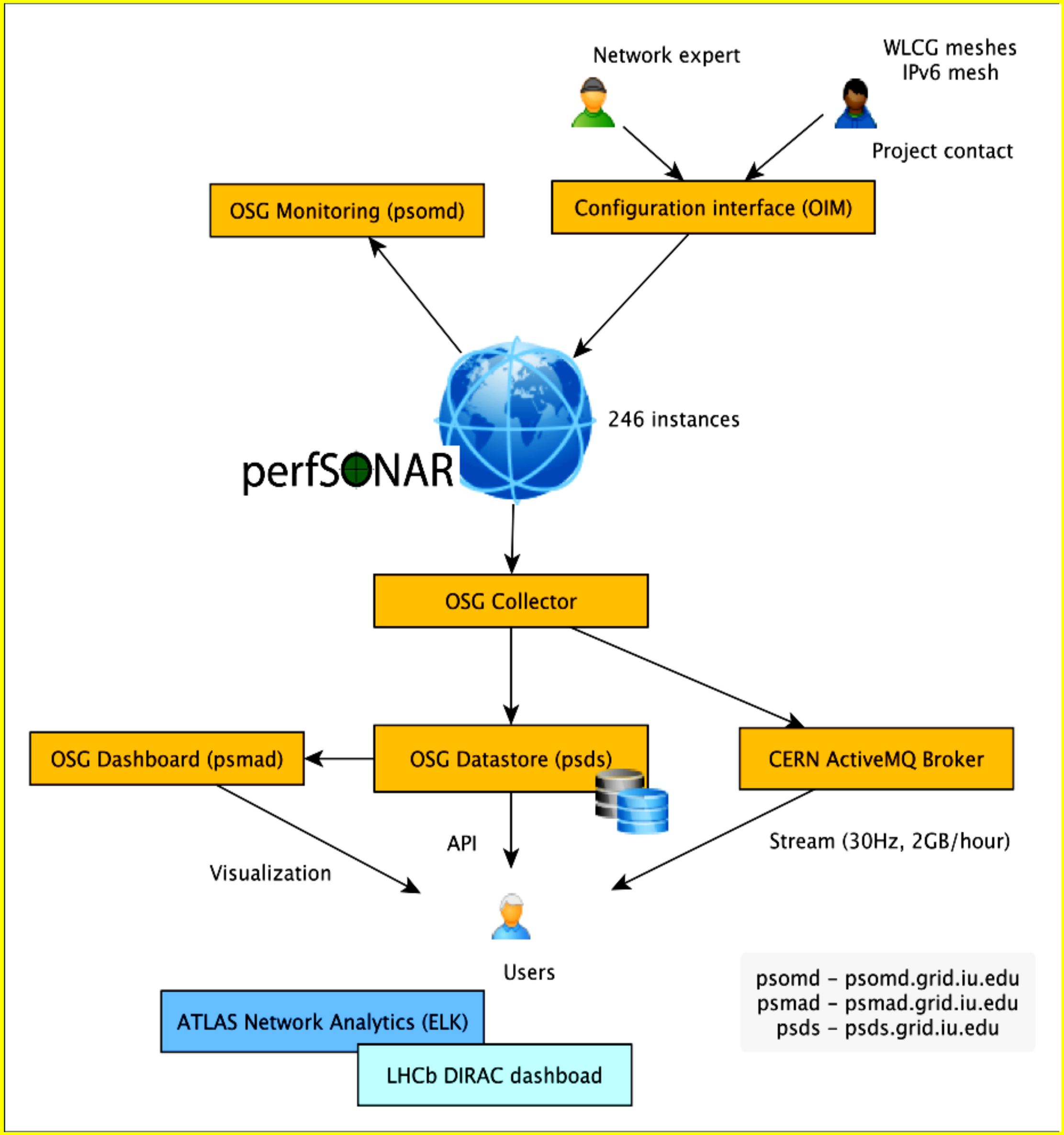
Long Term

- **Increased focus** on foreseeing network capacity needs
- **Sharing the future capacity needs projections;** will require greater interaction with R&E networks
- **Use of "containers":** could accelerate adoption of SDNs on campus
- **See LHC Network Evolution and pre-GDB on Networking** for further details

Open Science Grid (OSG)

- The Open Science Grid has added a network area as of 2012
  - The goal is to become a source of network metrics for its constituents, including HEP and WLCG
- The network service OSG now provides perfSONAR network metric collection from all OSG and WLCG perfSONAR instances
  - This data is continually collected globally
  - Will be made available to users, higher level services and users, indefinitely
- OSG additionally provides tools to allow HEP collaborations to organize and monitor their perfSONAR deployments

OSG Network Measurement Platform



- **OSG has developed an extensive network measurement platform using perfSONAR**
- **Tests can be centrally configured and are continuously gathered by the OSG collectors**
- **Long-term storage of network measurements is provided by the OSG Datastore with a public API**
- **All measurements are also available for subscriptions via ActiveMQ netmon brokers at CERN**

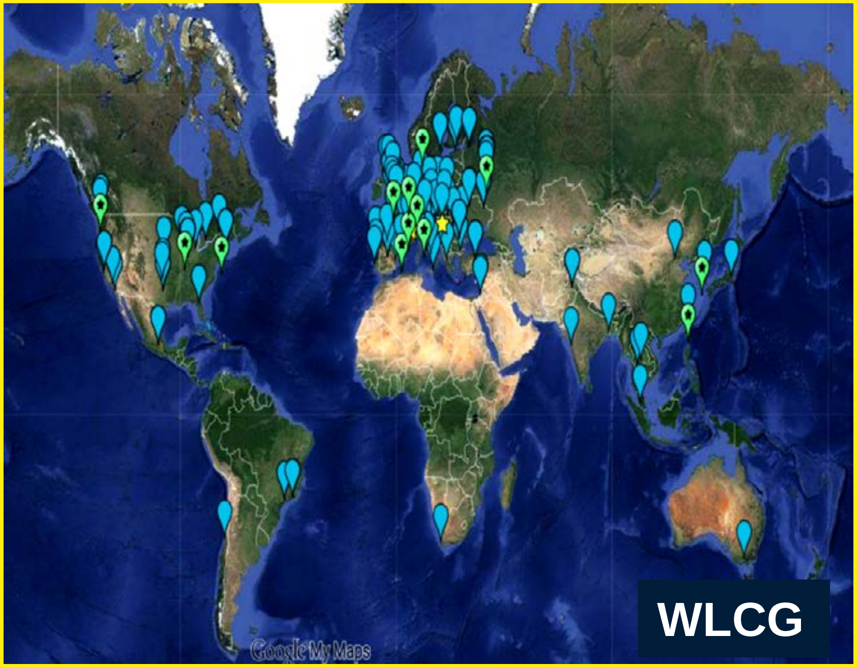
Towards a Next Generation Network-Integrated Systems for HEP and Other Data Intensive Science Programs

Vision: Next Gen Integrated Systems for Exascale Science: a Major Opportunity

✷ A new CPU/Storage/Network ecosystem + LCFs as focal points in the global workflow to meet otherwise daunting needs

Opportunity: Exploit the Synergy among

1. Global operations data and workflow management systems developed by HEP programs, *to respond to peak demands*
   - Evolving to work with *increasingly diverse and elastic resources*
   - Riding on high capacity (mostly still-passive) networks
   - Enabled by distributed operations and security infrastructures



2. Deeply programmable, agile software-defined networks (SDN), emerging as multi-domain network operating systems (e.g. SENSE project with ESnet):

3. Caltech, Esnet et al: New consistent ops methods with end-to-end control

3. Machine Learning, modeling and simulation, and game theory methods Extract key variables; optimize; move to real-time self-optimizing workflows

Next Gen SDN Systems for Exascale Science

Vision: Distributed environments where resources can be deployed flexibly to meet the demands

- SDN is a natural path to this vision:
  - Separating the functions that control the flow of traffic, from the switching infrastructure that forwards the traffic
  - Through open deeply programmable "controllers".

With many benefits:

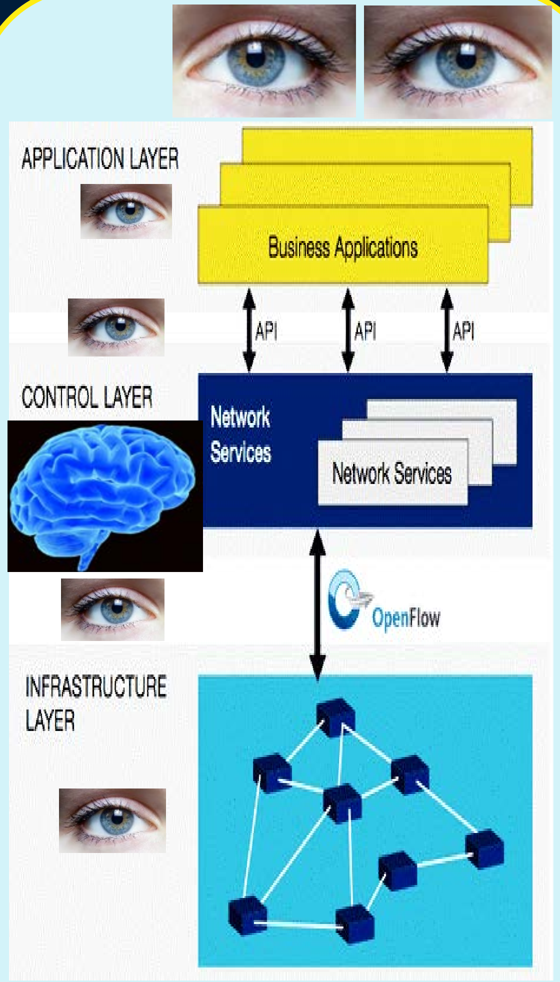
❒ Replacing stovepiped vendor HW/SW solutions by open platform-independent software services

❒ Virtualizing services and networks: lowering cost and energy, with greater simplicity

❒ Adding intelligent dynamics to system operations

A major direction of Research networks + Industry

❒ A Sea Change that is still emerging and maturing

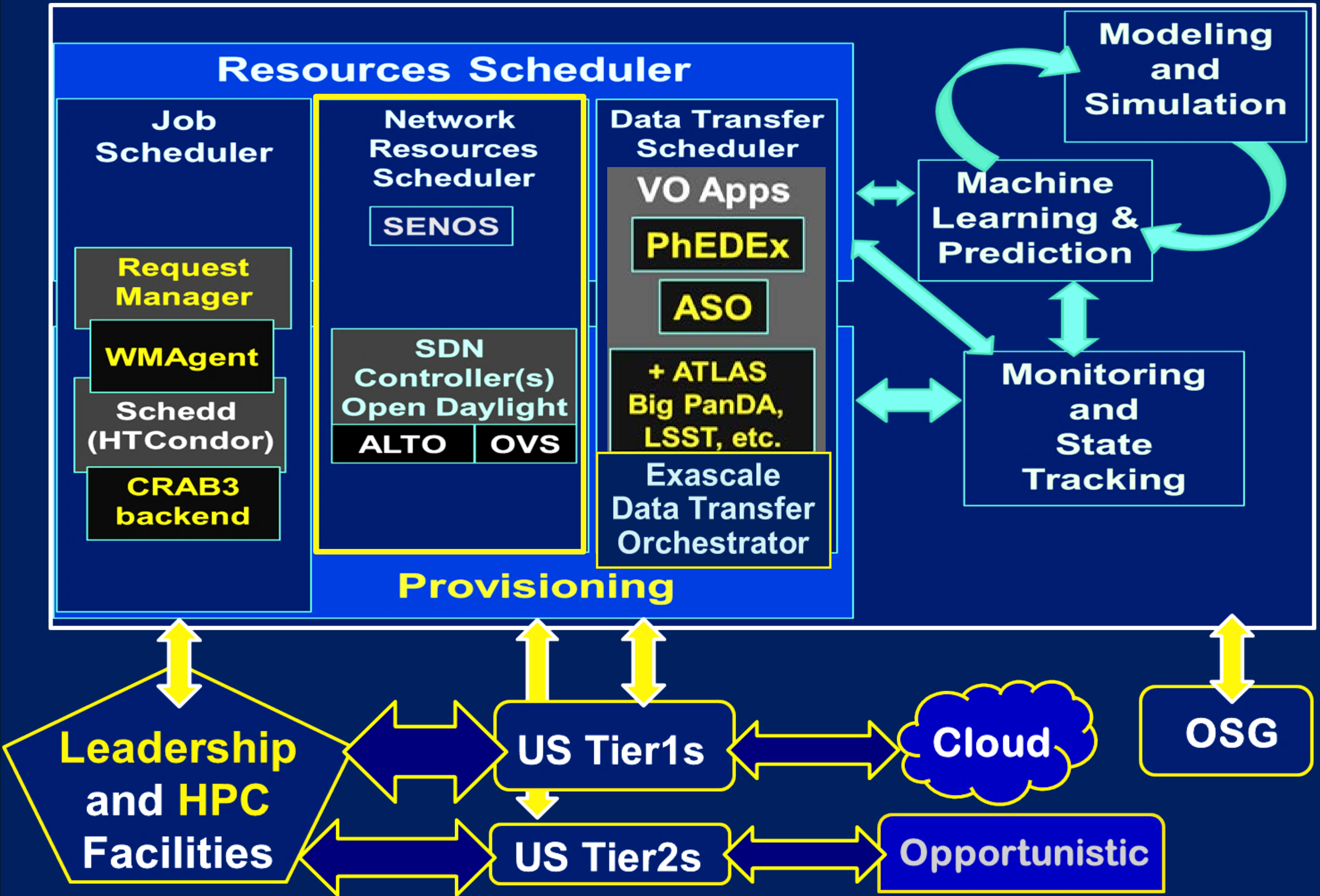


opennetworking.org

A system with built in intelligence

Requires excellent monitoring at all levels

NGenIA-ES Services and Data Flow Diagram


Resources Scheduler
Job Scheduler
Request Manager
WMAgent
Schedd (HTCondor)
CRAB3 backend
Network Resources Scheduler
SENOS
SDN Controller(s) Open Daylight
ALTO
OVS
Data Transfer Scheduler
VO Apps
PhEDEx
ASO
+ ATLAS Big PanDA, LSST, etc.
Exascale Data Transfer Orchestrator
Provisioning
Modeling and Simulation
Machine Learning & Prediction
Monitoring and State Tracking
Leadership and HPC Facilities
US Tier1s
US Tier2s
Cloud
Opportunistic
OSG

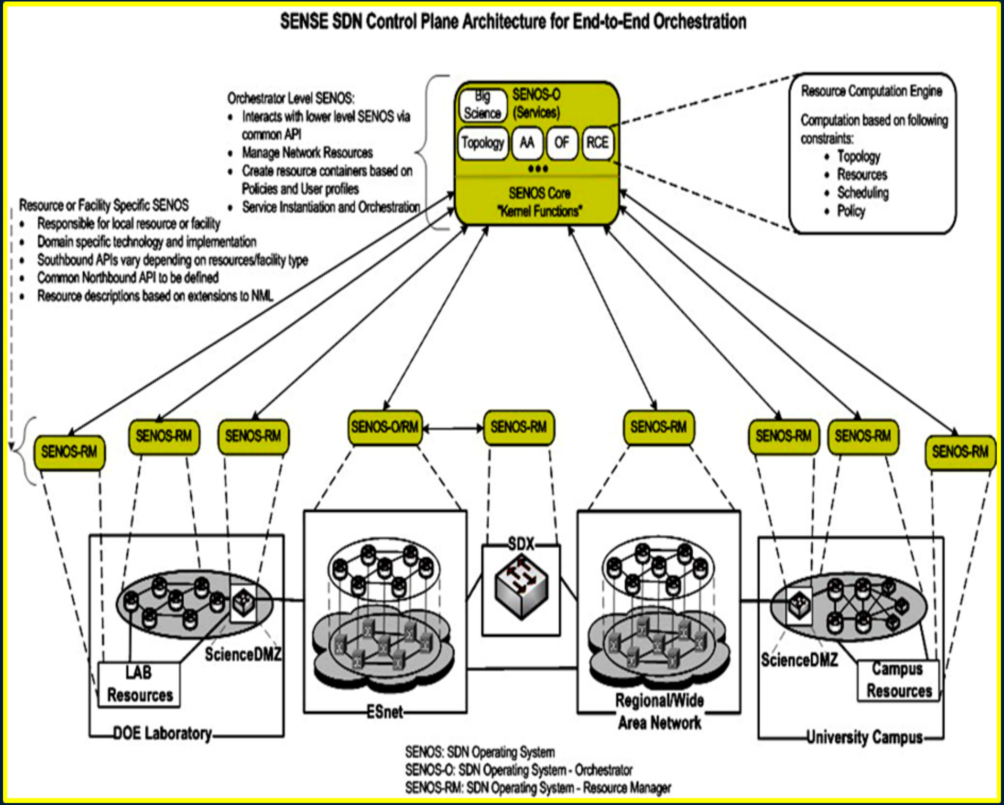
SENSE: SDN for End-to-end Networked Science at the Exascale

ESnet Caltech Fermilab Argonne Maryland LBNL

❒ **Mission Goals:**

- ❒ Improve end-to-end performance of science workflows
- ❒ Enabling new paradigms: e.g. creating dynamic distributed ‘Superfacilities’.

❒ **Comprehensive Approach:** An end-to-end SDN Operating System (SENOS), with:



SENSE SDN Control Plane Architecture for End-to-End Orchestration

Orchestrator Level SENOS:
- Interacts with lower level SENOS via common API
- Manage Network Resources
- Create resource containers based on Policies and User profiles
- Service Instantiation and Orchestration

Resource or Facility Specific SENOS
- Responsible for local resource or facility
- Domain specific technology and implementation
- Southbound APIs vary depending on resources/facility type
- Common Northbound API to be defined
- Resource descriptions based on extensions to NML

Resource Computation Engine

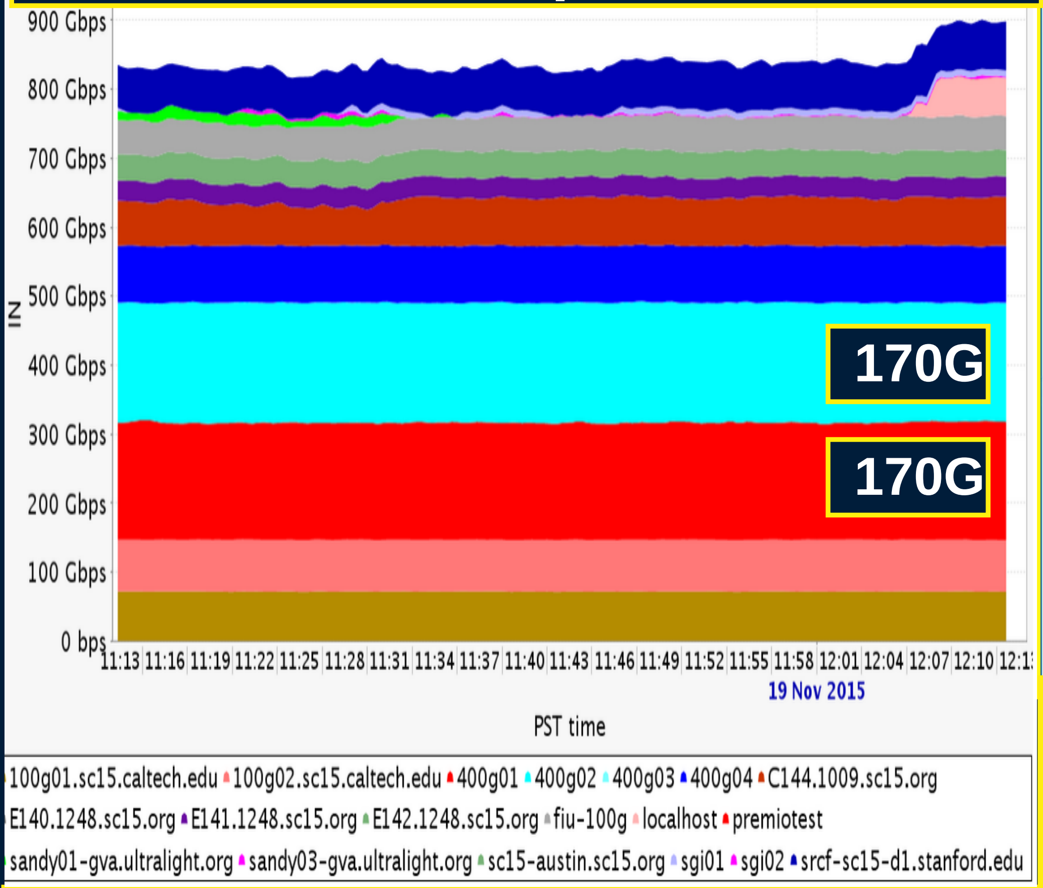
Computation based on following constraints:
- Topology
- Resources
- Scheduling
- Policy

SENOS: SDN Operating System
SENOS-O: SDN Operating System - Orchestrator
SENOS-RM: SDN Operating System - Resource Manager

- ❒ Intent-based interfaces, providing intuitive access to intelligent SDN services
- ❒ Policy-guided E2E orchestration of resources
- ❒ Auto-provisioning of network devices and Data Transfer Nodes
- ❒ Network measurement, analytics and feedback to build resilience

SC15: Caltech and Partners Terabit/sec SDN Driven Agile Network: Aggregate Results

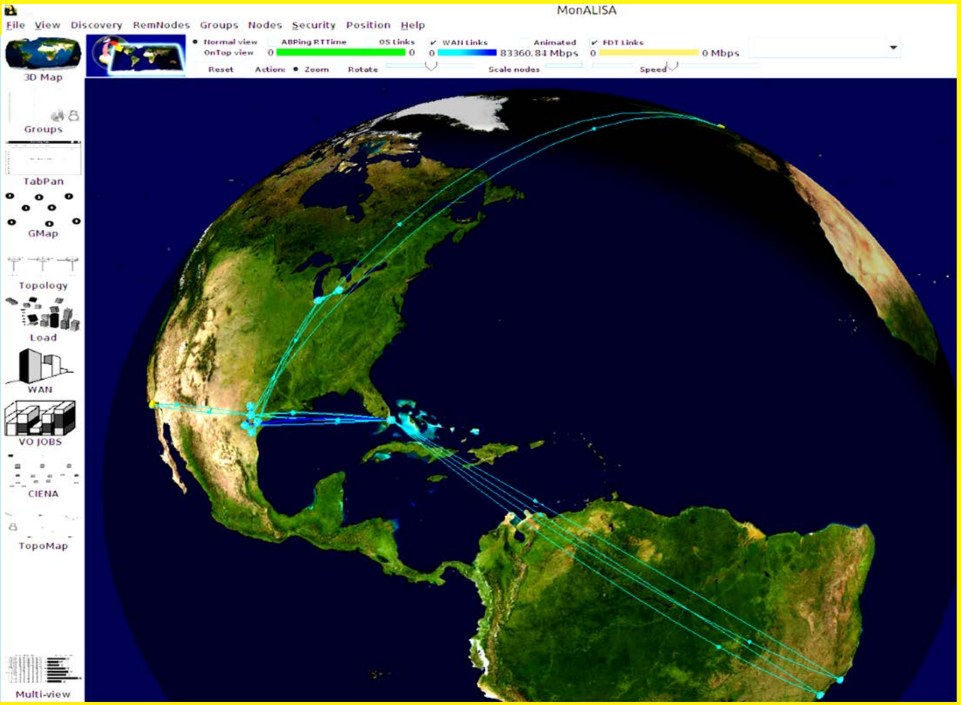
900 Gbps Total
Peak of 360 Gbps in the WAN



170G

170G

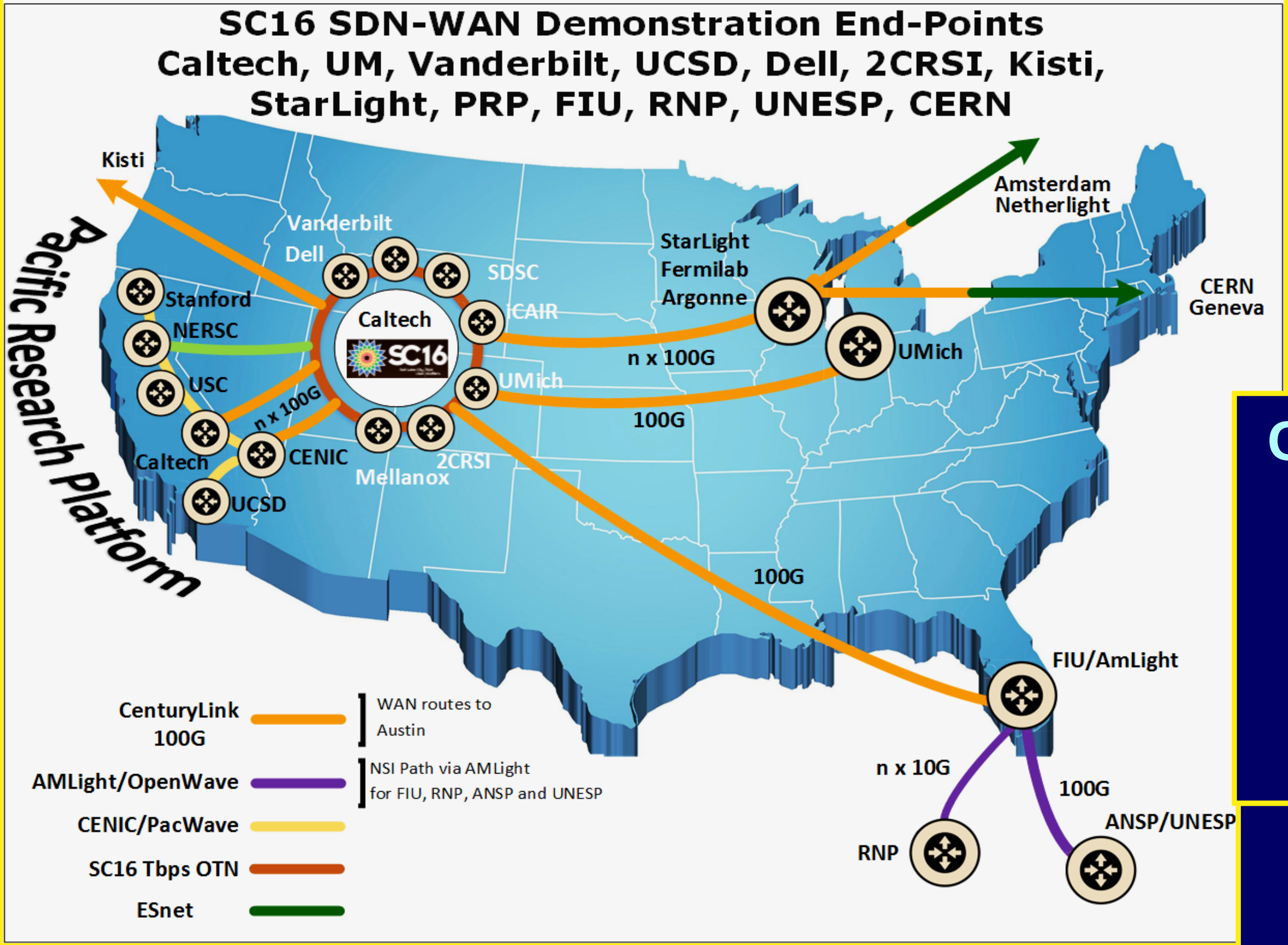
MonALISA Global Topology



29 100G NICs; Two 4 X 100G and Two 3 X 100G DTNs; 9 32 X100G Switches

Smooth Single Port Flows up to 170G; *120G over the WAN.* With Caltech’s FDT TCP Application http://monalisa.caltech.edu/FDT

SC16: SDN Next Generation Terabit/sec Integrated Network for Exascale Science



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

SDN-driven load balanced flow steering and site orchestration Over Terabit/sec Global Networks

Consistent Operations: Edge & Core Limits With Agile Feedback: Major Science Flow Classes Up to High Water Marks

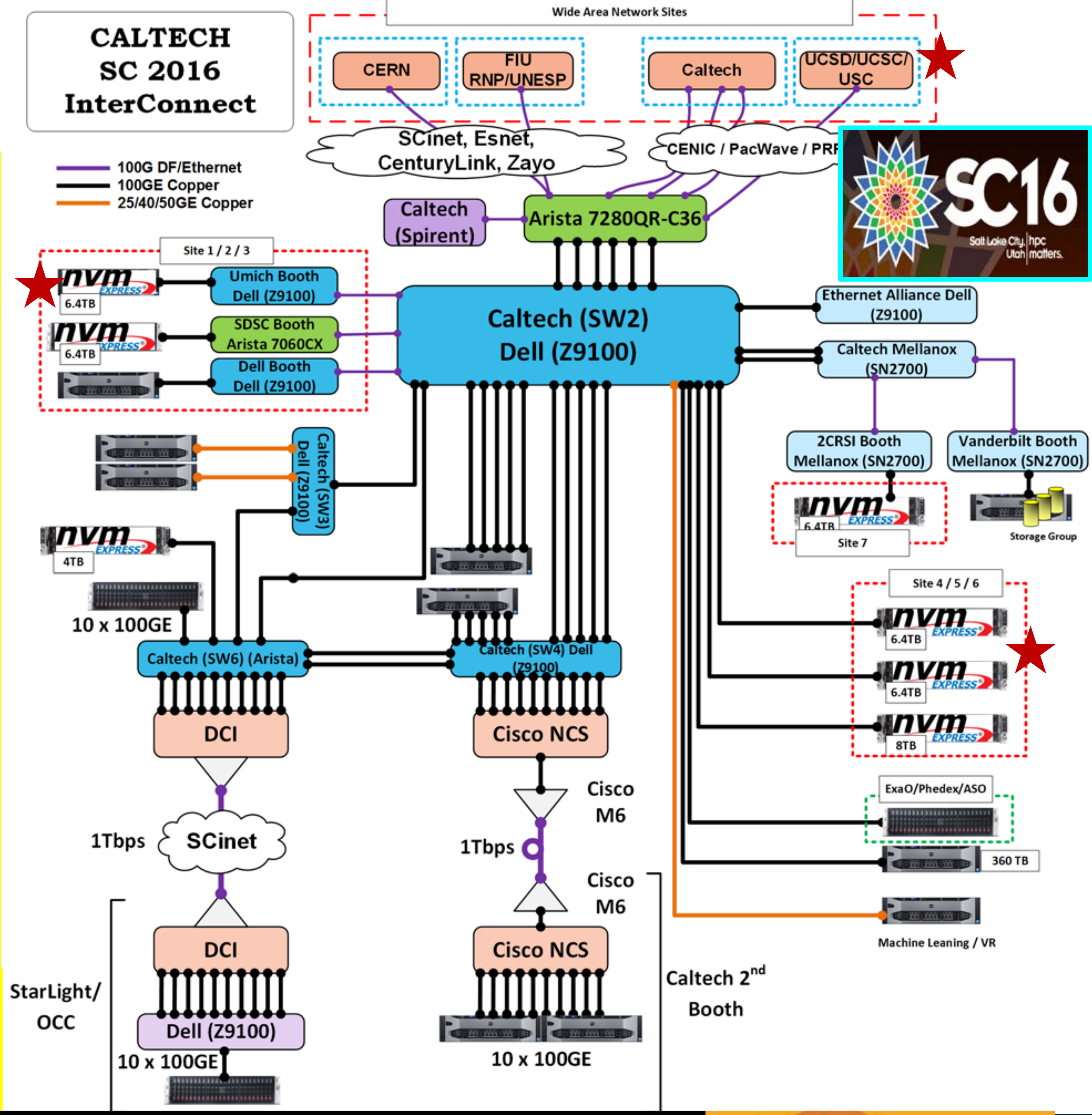
Preview PetaByte Transfers to/from Site Edges of Exascale Facilities With 400G+ DTNs

Caltech, Yale, UNESP & Partners: Open Daylight Controller, OVS and ALTO higher level services, New SDN Programming Framework

Caltech at SC16

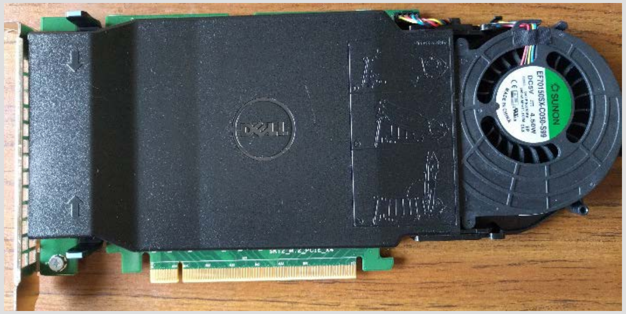
- ❑ Terabit/sec ring topology: Caltech – Starlight – SCInet; > 100 Active 100G Ports
- ❑ Interconnecting 9 Booths: Caltech 1 to 1 Tbps in booth, and to Starlight 1 Tbps; UCSD, UMich, Vanderbilt, Dell, Mellanox, HGST @100G
- ❑ WAN: Caltech, FIU +UNESP (Sao Paulo), PRP (UCSD, UCSC, USC), CERN, KISTI, etc.

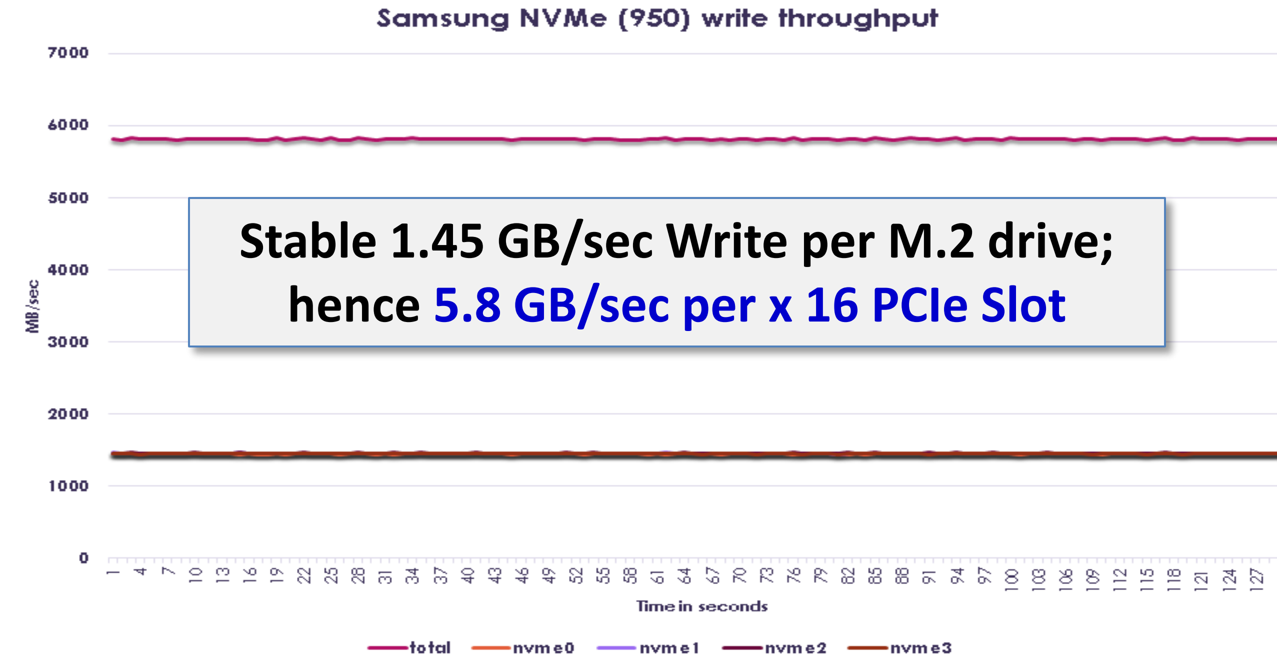
★ **ExaO + PhEDEx/ASO CMS Sites**



Looking Forward: We will start work on SC17 and will be looking for network and research site partners Soon

A low cost NVMe based Data Transfer Node TN Server





Stable 1.45 GB/sec Write per M.2 drive; hence 5.8 GB/sec per x 16 PCIe Slot

Ingredients:

- 2U SuperMicro Server (with 3 x16 slots)
- Dual Dell Quad-M.2 adapter card
- 8 Samsung 950 Pro M.2 drives

(We are now testing SM961 and SM 960 Pro)

4TB NVMe Storage

~90 Gbps disk I/O using NVMe over Fabrics or FDT

Also see http://www.anandtech.com/show
10754/samsung-960-pro-ssd-review

Further slides on DTNs designs and performance tests:
https://www.dropbox.com/s/y1ln4m68tdz2lhj/DTN_Design_Mughal.pptx?dl=0

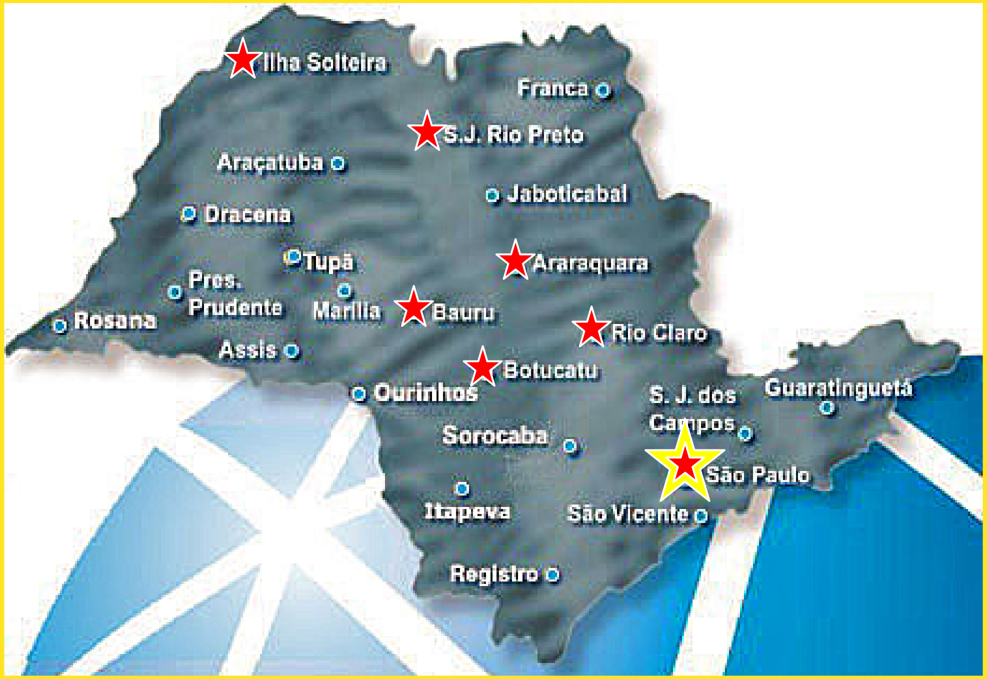
GridUNESP

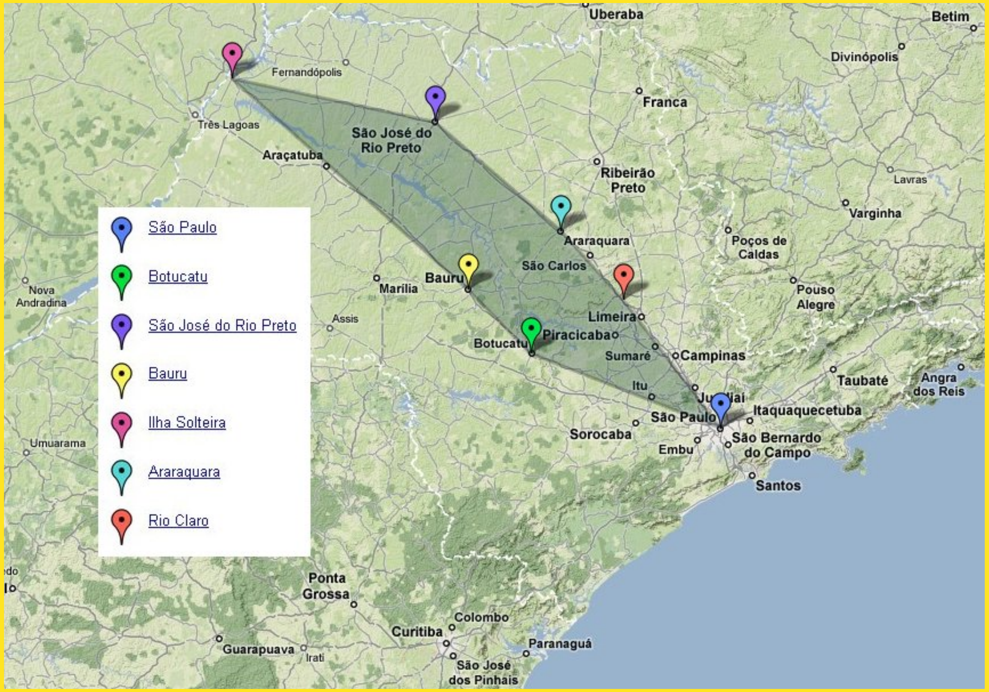
Spin-off of SPRACE project, the first Campus Grid in Latin-America

- Scientific Computing for UNESP
- Partnership with US OSG: the only OSG VO outside US
- Provides ANSP Grid Certificate Authority for State of São Paulo

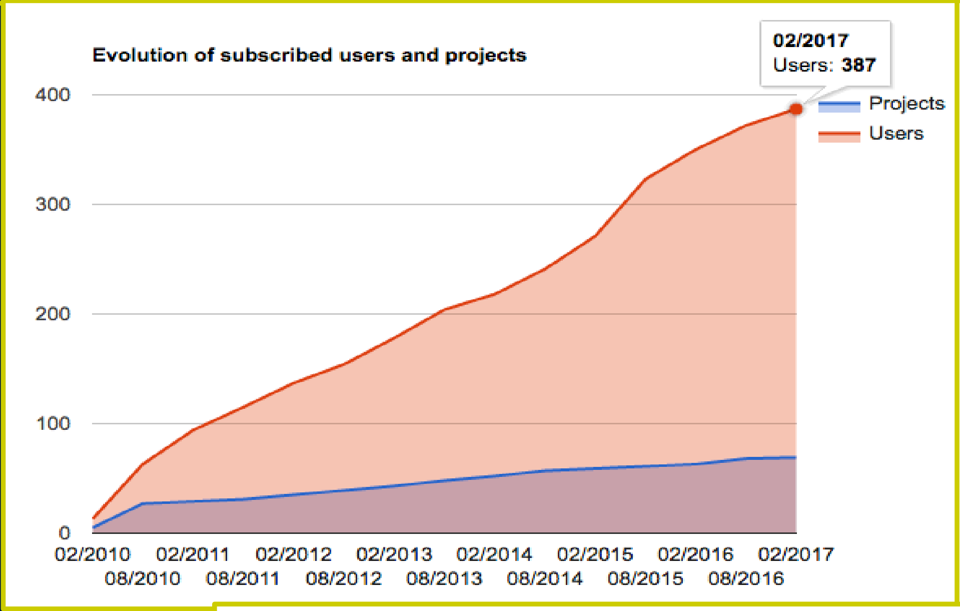
Distributed computational system with widely dispersed resources

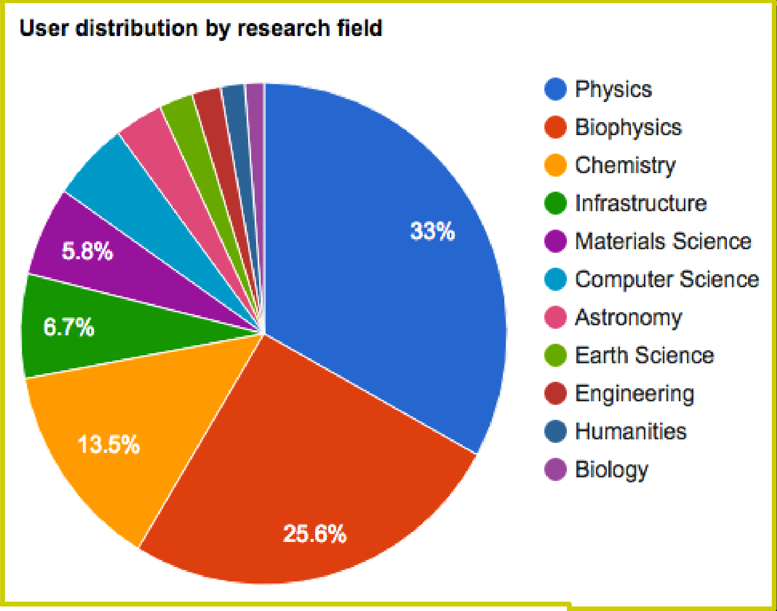
- Two-tiered architecture
- 1 central cluster in São Paulo capital, ~90 Tflops
- 6 secondary clusters on other campuses, with 2 headnodes, 16 worker nodes on each site

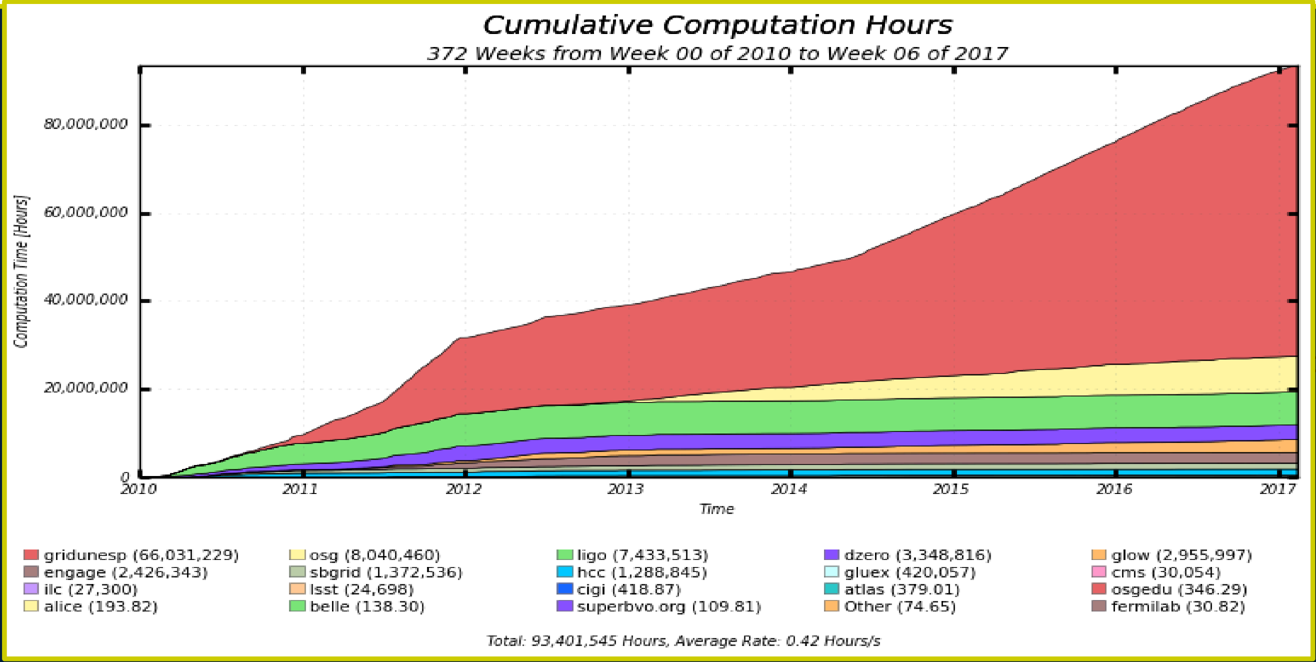




GridUnesp: Projects and Users

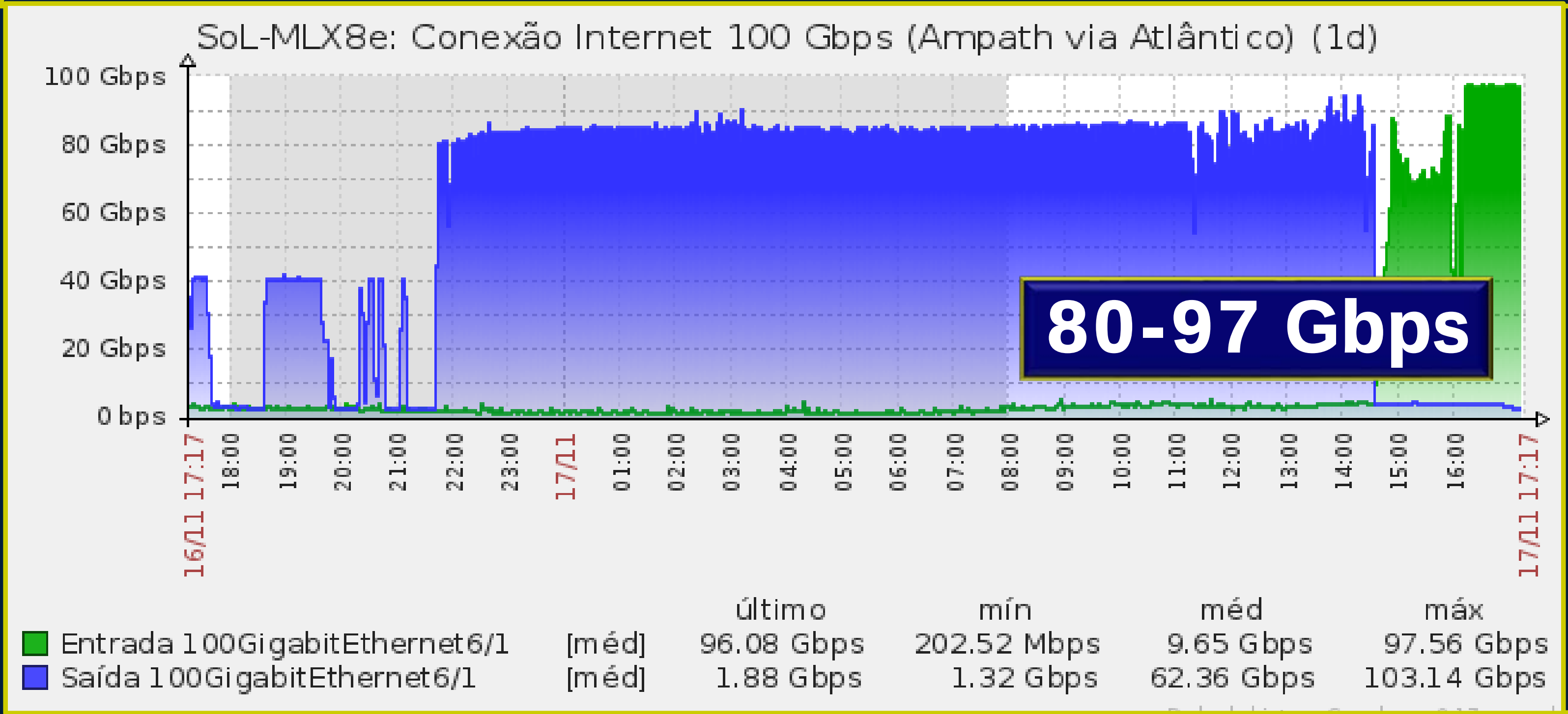






ennetworking.org

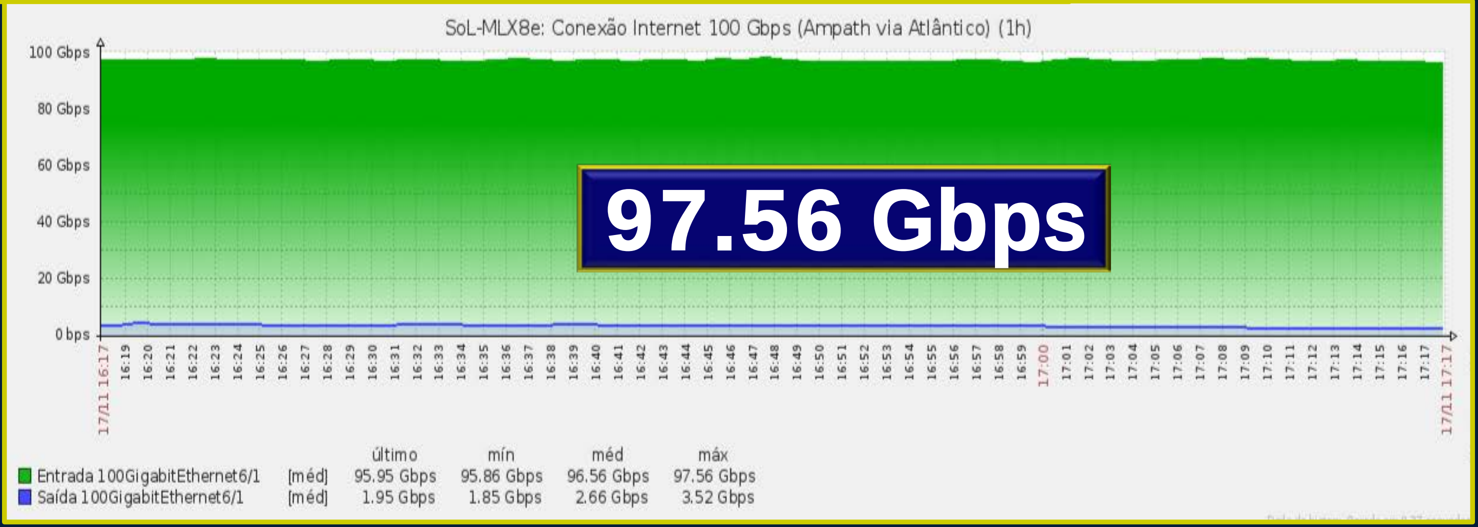
GridUnesp: Transfer Demo at SC16



80-97 Gbps

		último	mín	méd	máx
🟩 Entrada 100GigabitEthernet6/1	[méd]	96.08 Gbps	202.52 Mbps	9.65 Gbps	97.56 Gbps
🟦 Saída 100GigabitEthernet6/1	[méd]	1.88 Gbps	1.32 Gbps	62.36 Gbps	103.14 Gbps

17 Hour transfer overnight on Miami-Sao Paulo Atlantic link



97.56 Gbps

		último	mín	méd	máx
🟩 Entrada 100GigabitEthernet6/1	[méd]	95.95 Gbps	95.86 Gbps	96.56 Gbps	97.56 Gbps
🟦 Saída 100GigabitEthernet6/1	[méd]	1.95 Gbps	1.85 Gbps	2.66 Gbps	3.52 Gbps

1 Hour transfer on Miami-Sao Paulo Atlantic link

Exascale Ecosystems with Petabyte Transactions

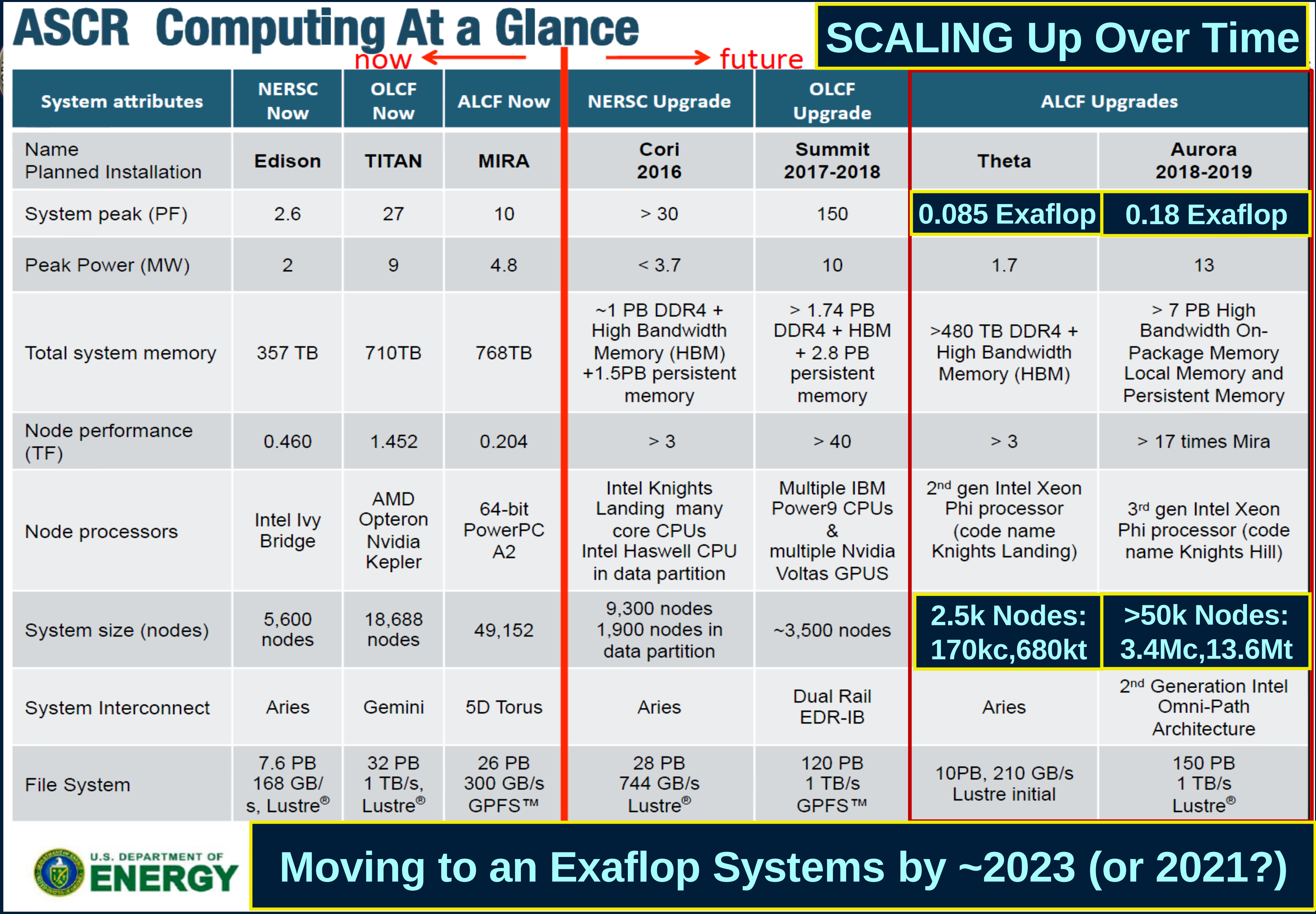
for Next-Generation Data Intensive Sciences

- Opportunity for HEP (CMS example):
  - CPU needs will grow 65 to 200X by HL LHC
  - Dedicated CPU that can be afforded will be an order of magnitude less; even after code improvements on the present trajectory
- Short term Goal: Making such systems a grid resource for CPU using data resident at Tier1s and US Tier2s
- Method: Petabyte transactions over 400G then Terabit/sec networks with Secure proxies at the site edge
- Important Long Term benefits
  - Folding LCFs into a global ecosystem for HEP and data intensive sciences
  - Building a modern coding workforce
  - Helping to Shape the future architecture and operational modes of Exascale Computing Facilities

Pilots Programs with Argonne, ORNL

1. MIRA as a grid resource
2. Precise NLO generators on Mira with new more efficient methods
3. DTN and process design for 100G+ data transfers

SCALING Up Over Time



ASCR Computing At a Glance

now ← | → future

System attributes	NERSC Now	OLCF Now	ALCF Now	NERSC Upgrade	OLCF Upgrade	ALCF Upgrades	
Name Planned Installation	**Edison**	**TITAN**	**MIRA**	**Cori** **2016**	**Summit** **2017-2018**	**Theta**	**Aurora** **2018-2019**
System peak (PF)	2.6	27	10	> 30	150	**0.085 Exaflop**	**0.18 Exaflop**
Peak Power (MW)	2	9	4.8	< 3.7	10	1.7	13
Total system memory	357 TB	710TB	768TB	~1 PB DDR4 + High Bandwidth Memory (HBM) +1.5PB persistent memory	> 1.74 PB DDR4 + HBM + 2.8 PB persistent memory	>480 TB DDR4 + High Bandwidth Memory (HBM)	> 7 PB High Bandwidth On-Package Memory Local Memory and Persistent Memory
Node performance (TF)	0.460	1.452	0.204	> 3	> 40	> 3	> 17 times Mira
Node processors	Intel Ivy Bridge	AMD Opteron Nvidia Kepler	64-bit PowerPC A2	Intel Knights Landing many core CPUs Intel Haswell CPU in data partition	Multiple IBM Power9 CPUs & multiple Nvidia Voltas GPUS	2nd gen Intel Xeon Phi processor (code name Knights Landing)	3rd gen Intel Xeon Phi processor (code name Knights Hill)
System size (nodes)	5,600 nodes	18,688 nodes	49,152	9,300 nodes 1,900 nodes in data partition	~3,500 nodes	**2.5k Nodes: 170kc,680kt**	**>50k Nodes: 3.4Mc,13.6Mt**
System Interconnect	Aries	Gemini	5D Torus	Aries	Dual Rail EDR-IB	Aries	2nd Generation Intel Omni-Path Architecture
File System	7.6 PB 168 GB/s, Lustre®	32 PB 1 TB/s, Lustre®	26 PB 300 GB/s GPFS™	28 PB 744 GB/s Lustre®	120 PB 1 TB/s GPFS™	10PB, 210 GB/s Lustre initial	150 PB 1 TB/s Lustre®

U.S. DEPARTMENT OF ENERGY

Moving to an Exaflop Systems by ~2023 (or 2021?)

NGenIA Summary

- *Advanced networks will continue to be a key* to the discoveries in HEP and other data intensive fields of science
- *Near Term and Decadal Challenges must be addressed: Greater scale, complexity and scope; challenging the available capacity*
- *New approaches: a new class of deeply programmable software driven networked systems to handle globally distributed Exabyte-scale data are required, and being developed*
- NGenIA: New paradigm - Consistent SDN-driven end-to-end ops with stable, load balanced, high throughput managed flows
  - *A new horizon in the way networks are operated and managed*

✷ Adapting Exascale Computing Facilities to meet the needs of data intensive science, with high energy physics as the first use case (followed by others) will have multiple benefits

- *Short Term:* Enable Rapid Responses, including full reprocessing
- *Medium Term: Paving the Way to the next LHC Computing Model, within the bounds of networking and storage*
- *Long Term:* Empower the HEP and other communities *to make the next rounds of discoveries in science*

Examples of Major Network Developments

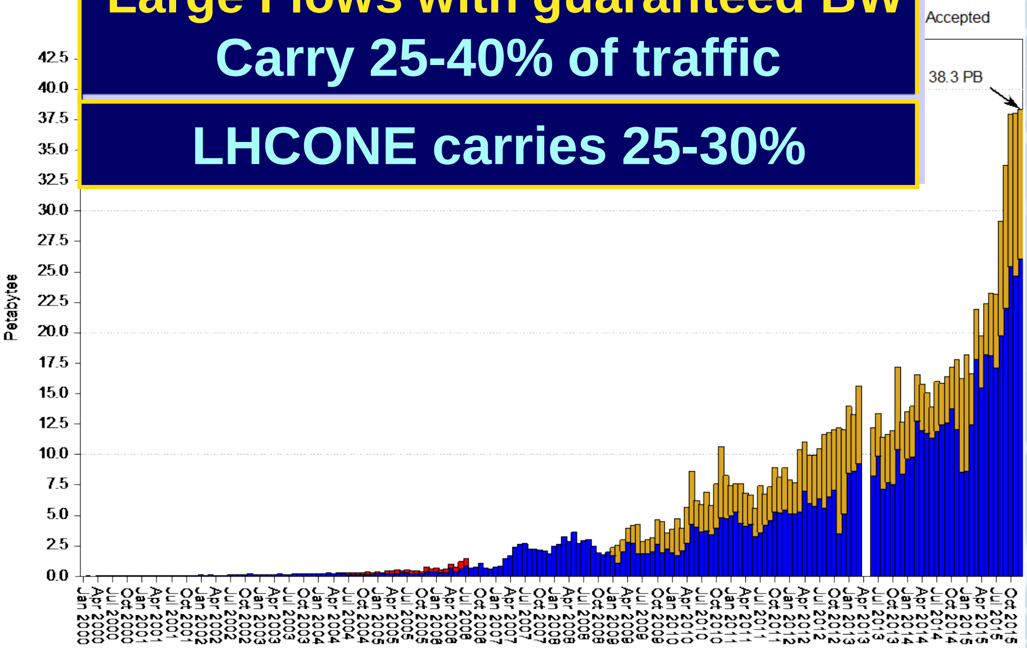
Energy Sciences Network

Inder Monga
J. Metzger

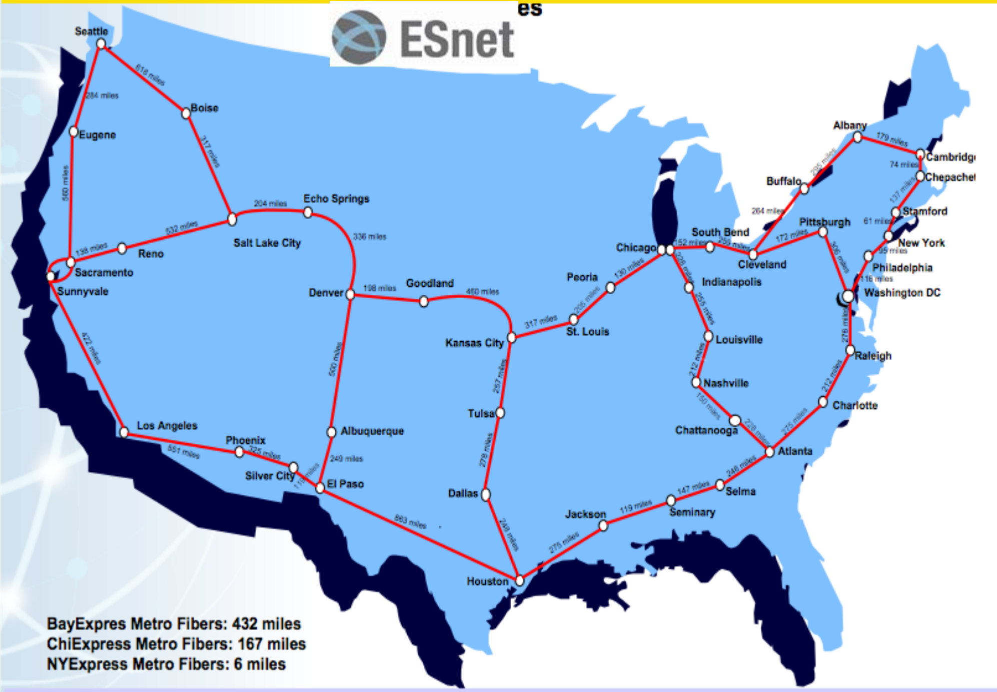
A Wide Range of Advanced Network Services

OSCARS Dynamic Circuits for Large Flows with guaranteed BW Carry 25-40% of traffic

LHCONE carries 25-30%



13000 km long haul dark fiber network



BayExpres Metro Fibers: 432 miles
ChiExpress Metro Fibers: 167 miles
NYExpress Metro Fibers: 6 miles

Leading the NSI Emerging dynamic circuit standard effort in OGF

Enables 100G testbed to test new network technologies and architectures

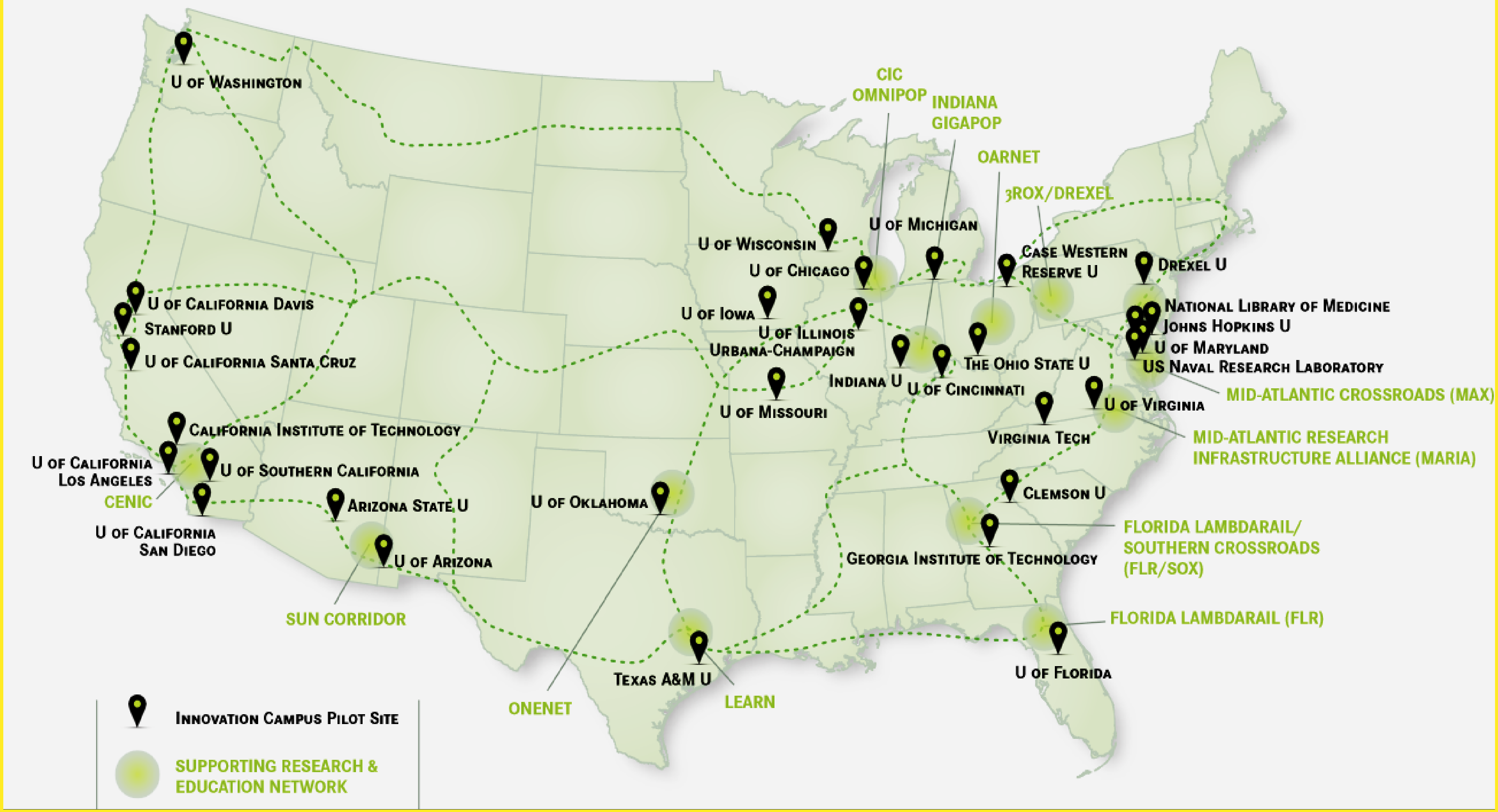
- Dedicated Science Engagement Team: consulting support in data transfer, network architecture, performance measurement, and visualization tools
- SDN Development, including SENOS a Network Operating System
- High Throughput Trials with HEP, NASA, Livermore et al. Including bringing 4 X 100G (1/3 of total) to SC2015: 1 Tbps trials; RDMA over Ethernet

Innovation Campus Pilot Program

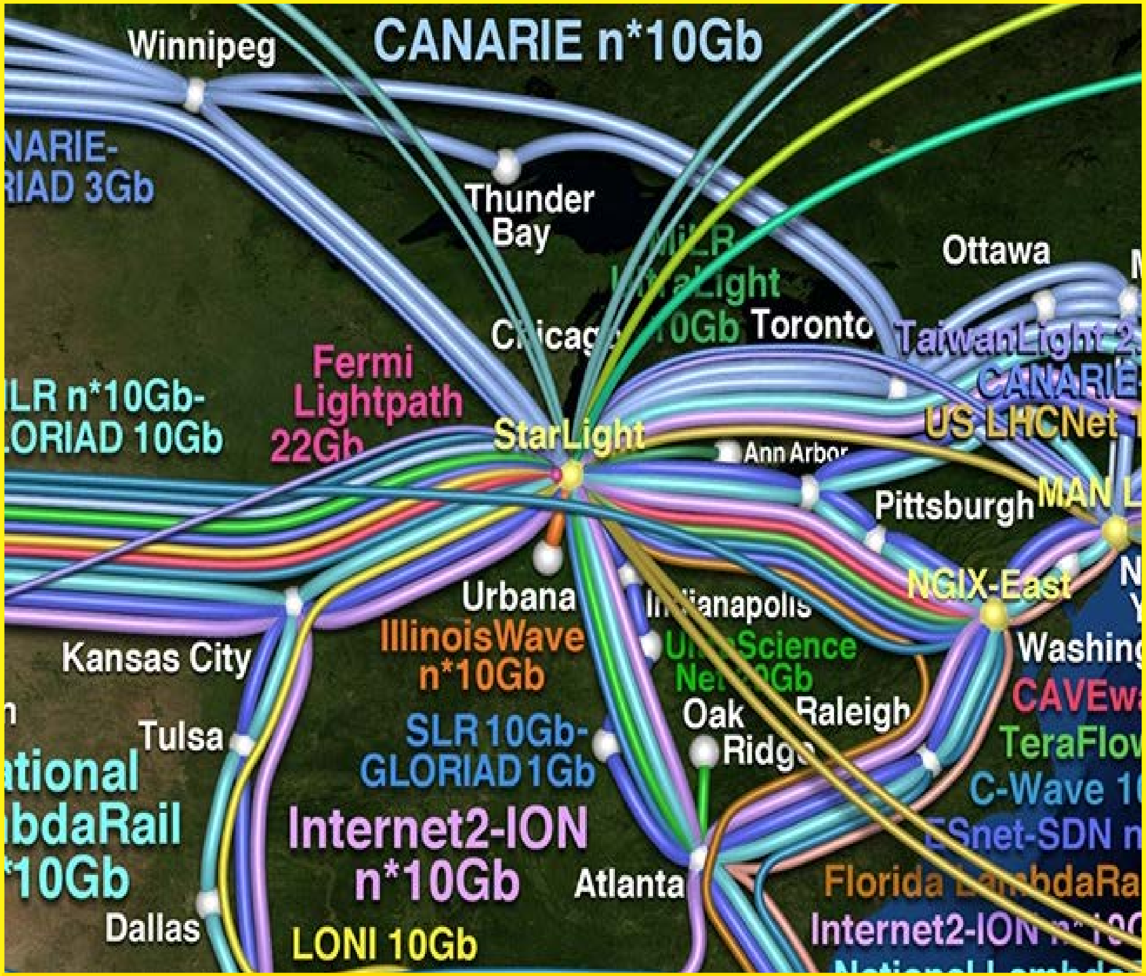
1. 100GE Now at 21 Campuses, 9 Regional Nets

2. “Science DMZs” to Separate, Support Large Flows

3. SDN at 20 Campuses, 4 Regional Nets


U of Washington
CIC OMNIPOP
Indiana GigaPOP
OARnet
3ROX/Drexel
U of Michigan
U of Wisconsin
U of Chicago
Case Western Reserve U
Drexel U
National Library of Medicine
Johns Hopkins U
U of Maryland
US Naval Research Laboratory
Mid-Atlantic Crossroads (MAX)
U of California Davis
Stanford U
U of California Santa Cruz
U of Iowa
U of Illinois Urbana-Champaign
Indiana U
U of Cincinnati
The Ohio State U
U of Missouri
U of Virginia
Virginia Tech
Mid-Atlantic Research Infrastructure Alliance (MARIA)
California Institute of Technology
U of California Los Angeles
U of Southern California
CENIC
U of California San Diego
Arizona State U
U of Oklahoma
U of Arizona
Clemson U
Florida LambdaRail/ Southern Crossroads (FLR/SoX)
Georgia Institute of Technology
Sun Corridor
Florida LambdaRail (FLR)
Texas A&M U
U of Florida
OneNet
LEARN
Innovation Campus Pilot Site
Supporting Research & Education Network

StarLight: Major Scientific R&E Hub in Chicago



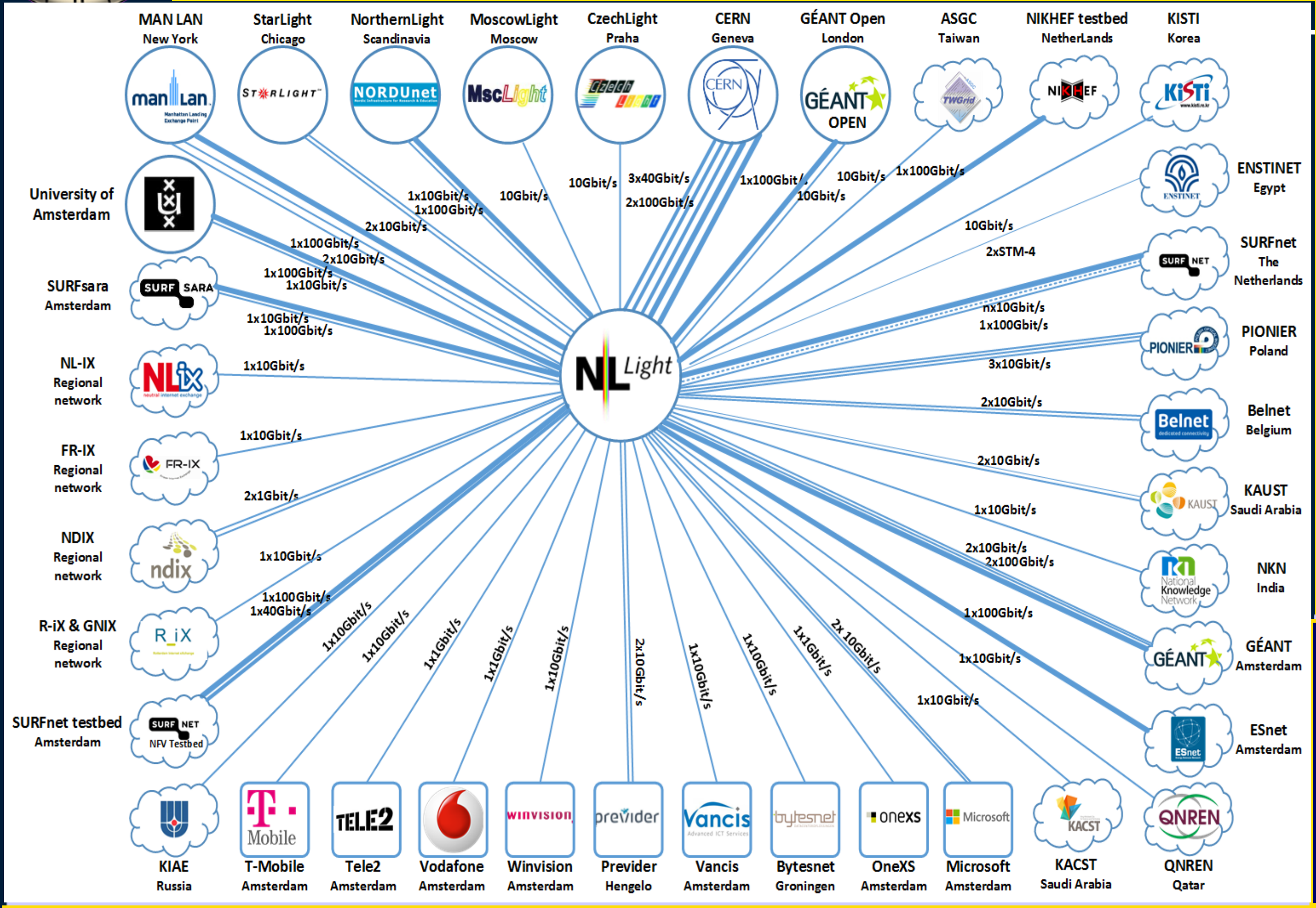
2015-16 Highlights

1. 34 Individual 100GE WAN paths
2. At SC15, iCAIR and OSDC conducted 15 100G demos
3. Active work on Software Defined Exchanges (part of GENI project)
4. Recently connected at 100GE with the Pacific Research Platform (PRP)

Fiber and circuits from many vendors, including:
AboveNet, AT&T, Cogent, Global Crossing, Level3, CenturyLink, RCN, Lightower, Zayo Group, and Sunesys

Open Exchange Points: NetherLight Example

12 X 100G, 4 x 40G, 37+ 10G Lambdas, Use of Dark Fiber

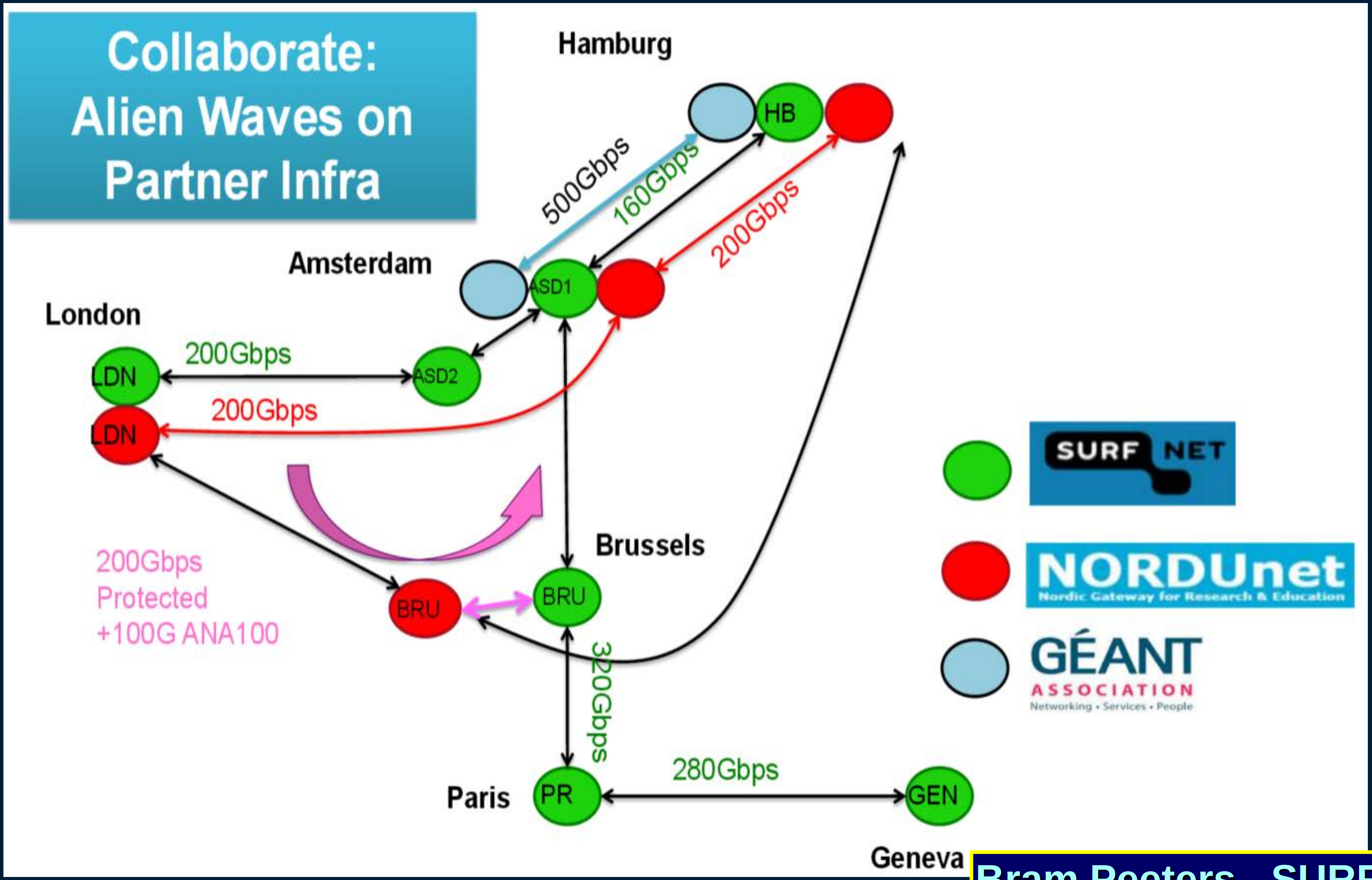


Inspired Other Open Lightpath Exchanges
Daejon (Kr)
Hong Kong
Tokyo
Praha (Cz)
Seattle
Chicago
Miami
New York

2013-16:
Dynamic Lightpaths + IP Services at and Above 100G

Convergence of Many Partners on Common Lightpath Concepts
ESnet, Internet2, GEANT, USLHCNet; nl, cz, ru, be, pl, es, tw, kr, hk, in, nordic

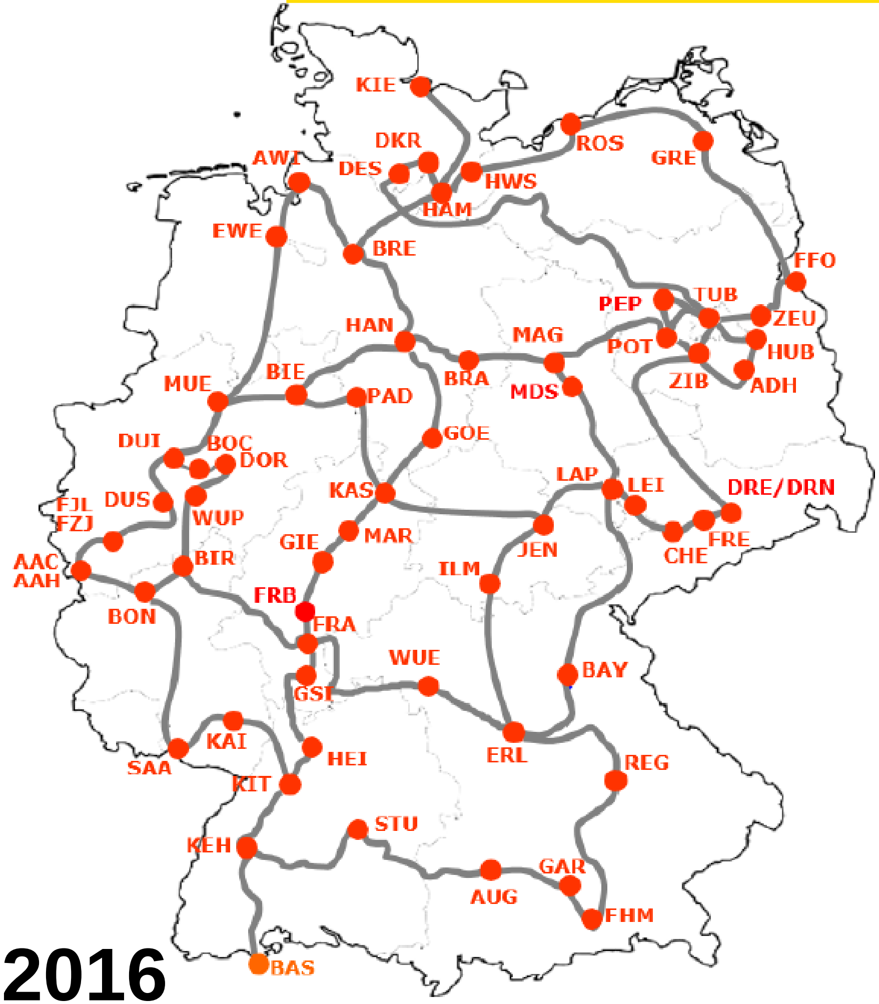
Cross Border Alien Waves
SURFnet, NORDUnet and GEANT



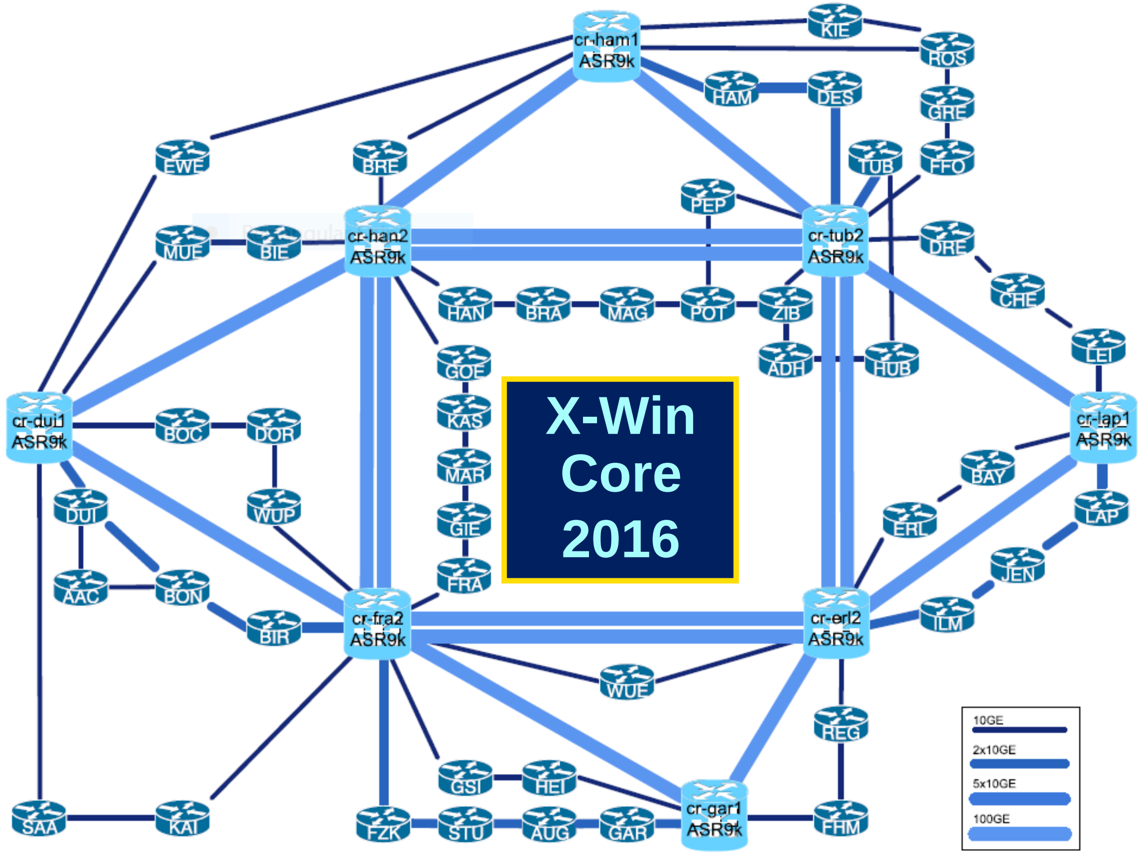
Bram Peeters - SURFnet

DFN X-Win Network

100G Optical Waves Supported Across the Network



2016



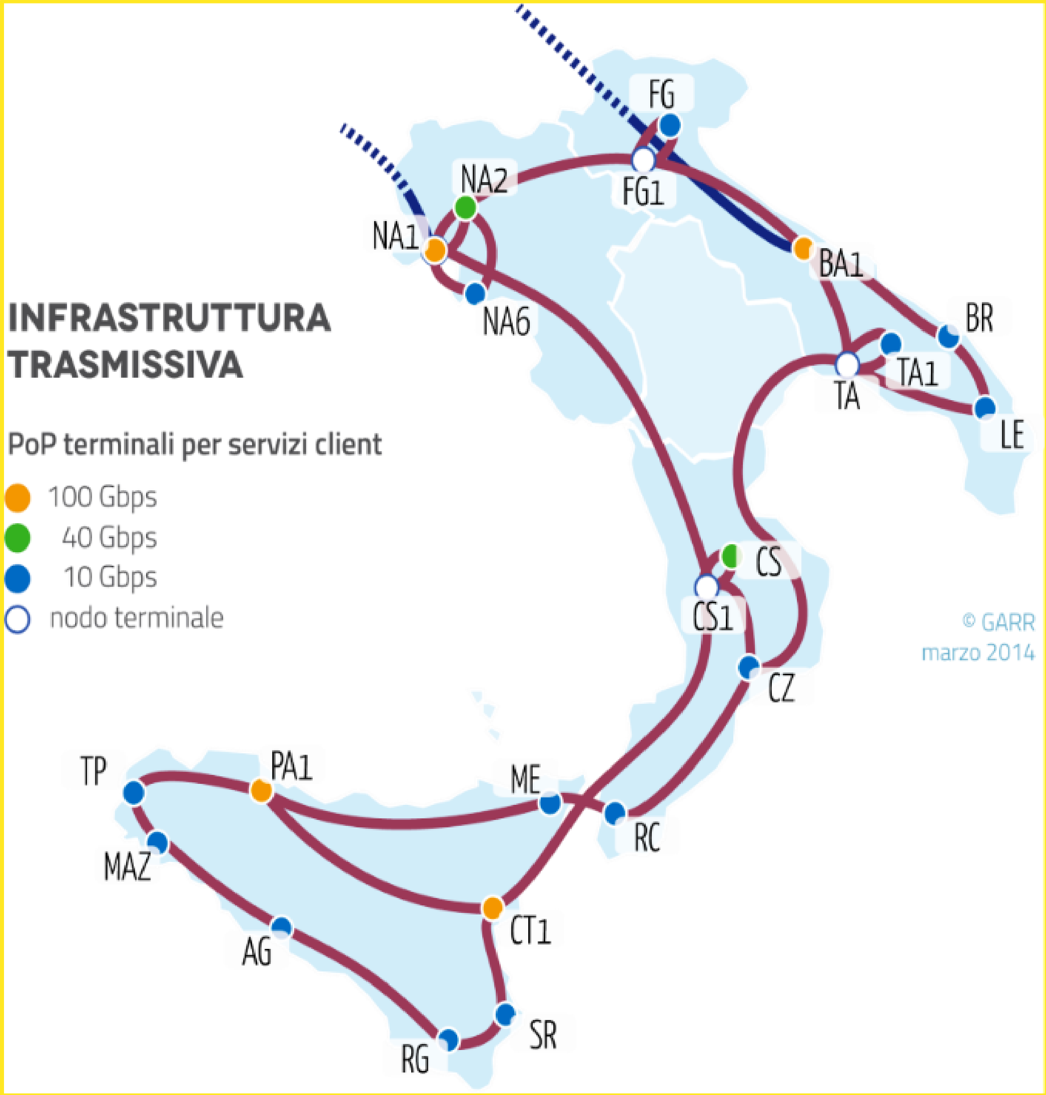
11000 km of Cross Dark Fiber
Total capacity 640 Gbps

During 2017 All Cisco 76XX routers (35)
will be replaced with CISCO ASR900

GARR-X Progress
Closing the Digital Divide in Italy

M. Marletta
E. Valente

GARR-X Progress Client Services



GARR-X

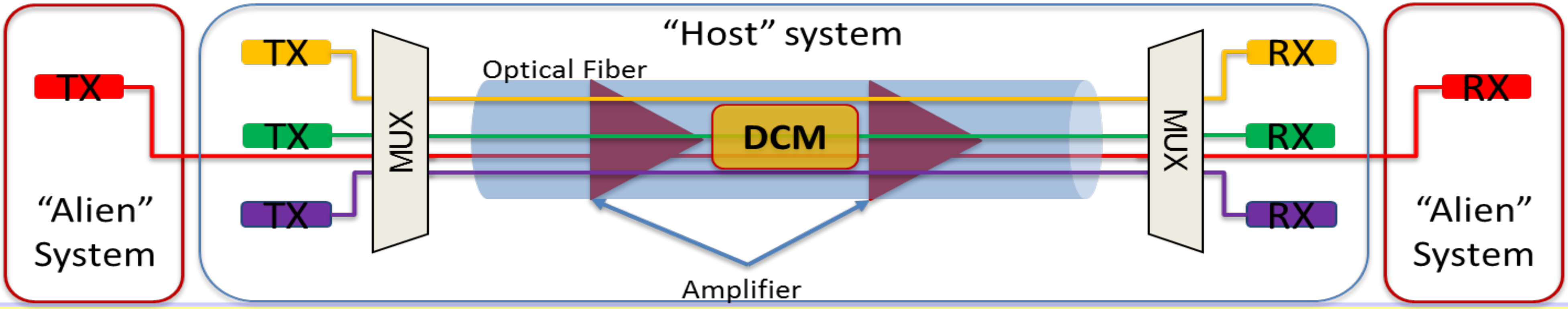
46.5 M€ Program to Reduce the Digital Divide in 4 regions of Southern Italy

- 100G Optical Fiber Ring Core
- 2500 Km new backbone fiber
- Connecting 100 Schools
- Allows 40G or 100G Tier2 connections now:
  - Catania (ALICE)
  - Naples (ATLAS)
  - Bari (CMS)
- Includes computing and storage for internally developed cloud services distributed among 5 sites (>8000 Vcores and 10PB)

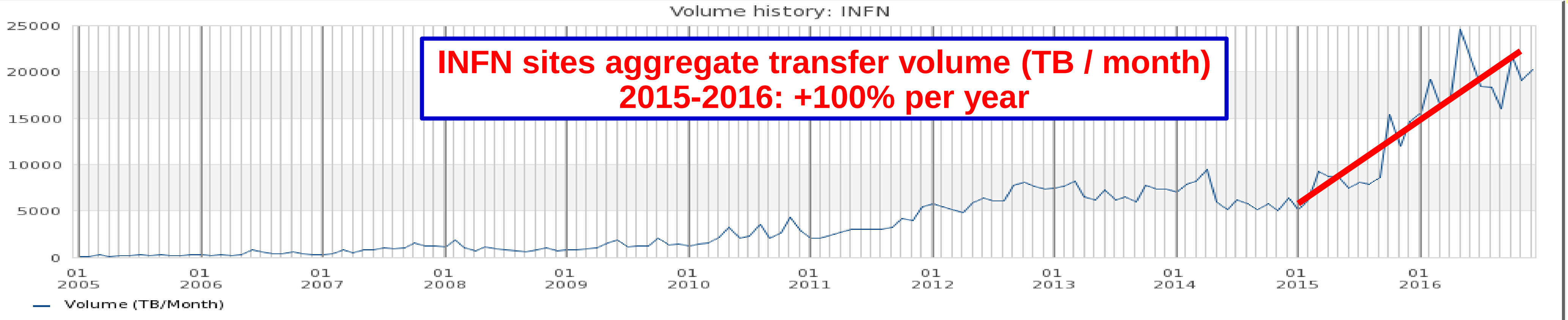
ICFA SCIC

GARR-X Progress
Alien Wavelength Technique (AWT)

M. Marletta
E. Valente



- ❑ Hybrid solution based on transmission and reception of optical signals generated by infrastructure different from the one providing transport and regeneration
- ❑ GARR plans to use AWT to provide 100G through **Infinera equipment** on the main backbone nodes of the **Huawei infrastructure**, in northern and central part of Italy



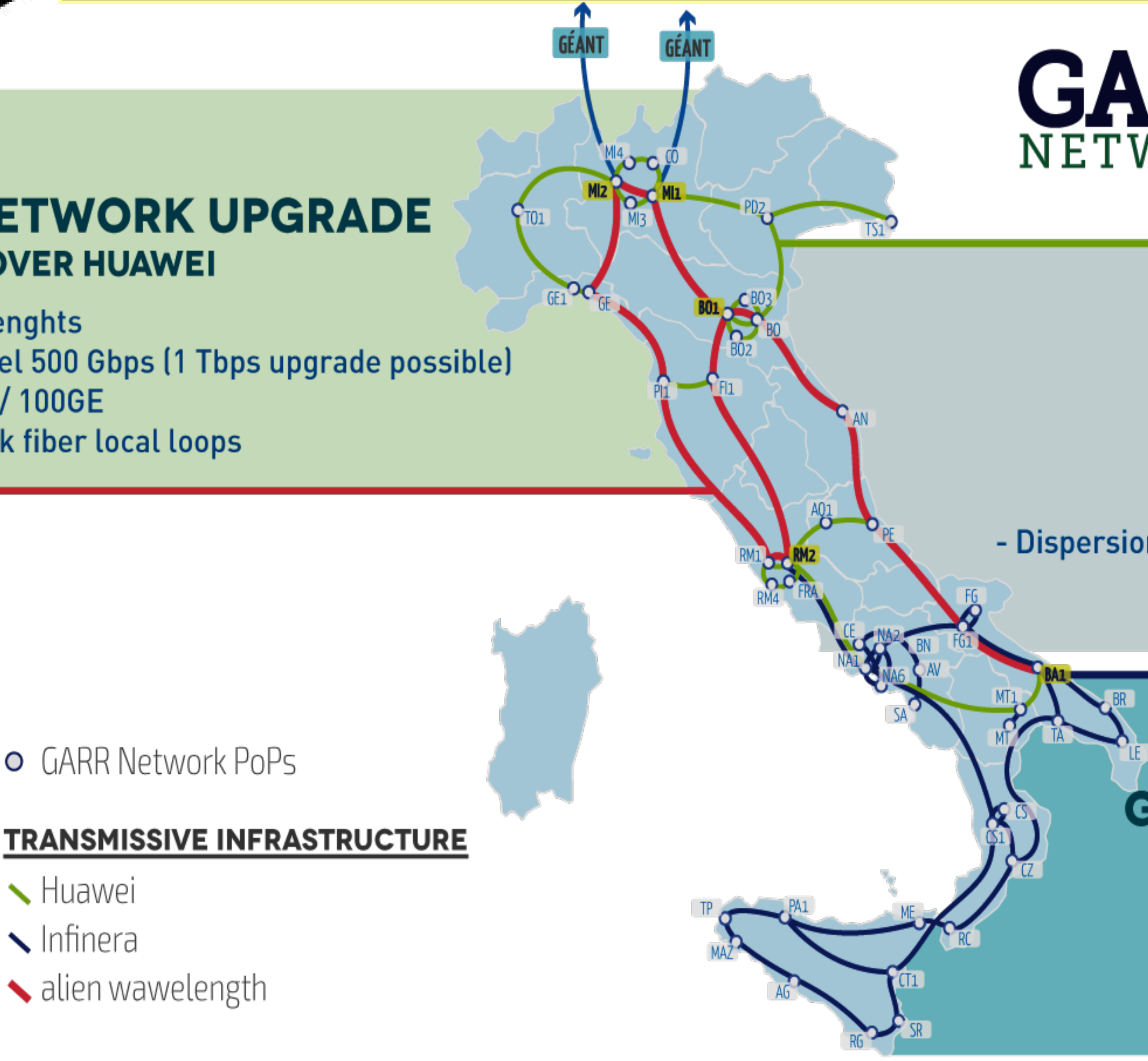
INFN sites aggregate transfer volume (TB / month)
2015-2016: +100% per year

ICFA SCIC

GARR-X Progress + Upgrade
Nationwide Advanced Optical Network

M. Marletta
E. Valente

GARR
NETWORK



2017
GARR NETWORK UPGRADE
INFINERA OVER HUAWEI

- Alien wavelenghts
- Superchannel 500 Gbps (1 Tbps upgrade possible)
- Client 10GE / 100GE
- 120 new dark fiber local loops

2011
GARR-X
HUAWEI

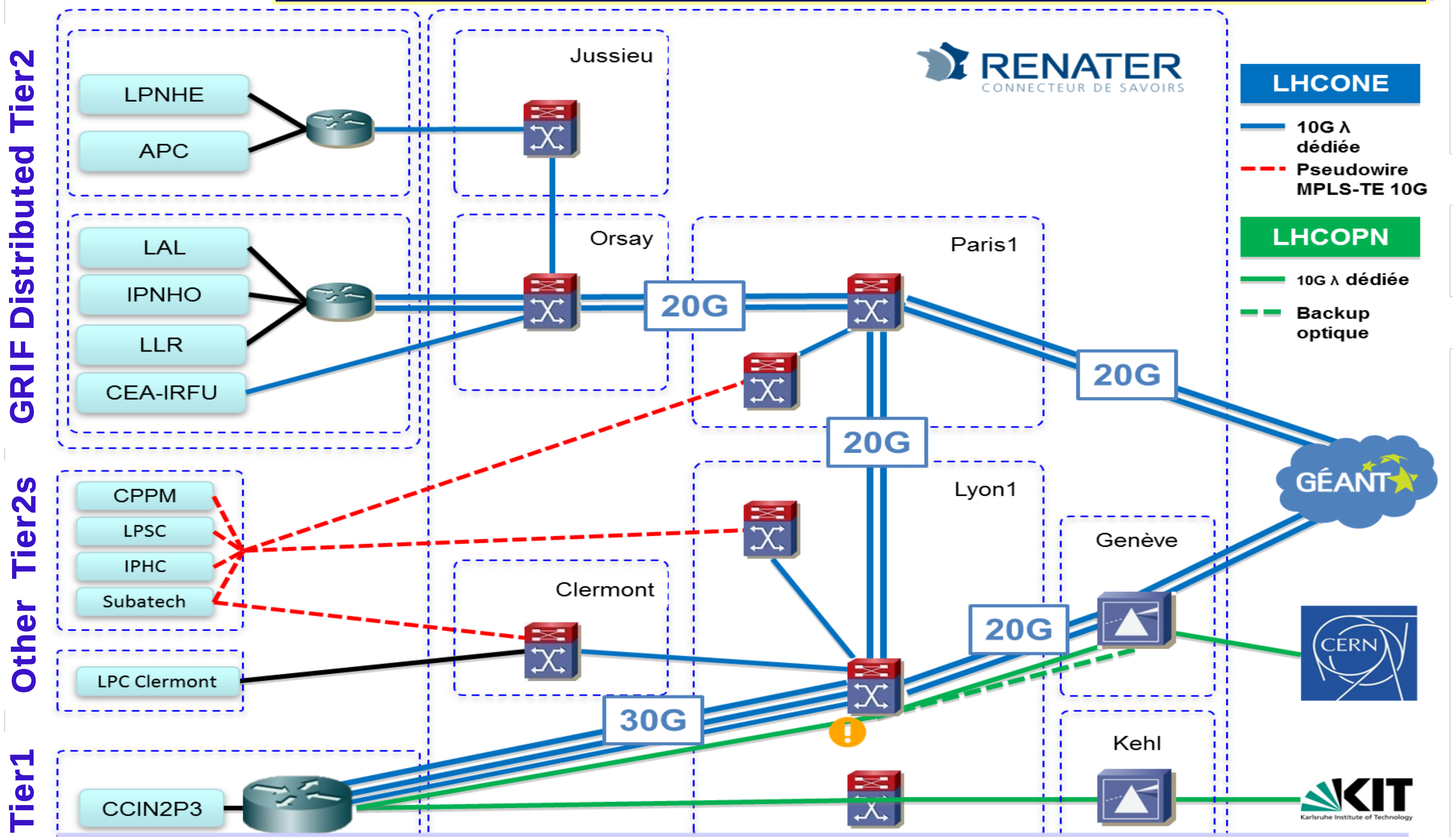
- IM-DD (OOK) network
- 10 Gbps / 40 Gbps Channels
- Dispersion Compensation Module (DCM) based infrastructure
- Client 1GE / 10GE

2014
GARR-X PROGRESS
INFINERA

- Coherent network
- Superchannel 500 Gbps
- DCM-Free Infrastructure
- Client 10GE / 40GE / 100GE

RENATER
CONNECTEUR DE SAVOIRS

France: 19 10G Links Dedicated to HEP for the LHCOPN and LHCONE



2015: 100G Core Paris-Lyon-GEANT Planned
Some 100G Ports to the Tier1 and Tier2s Possible

CESNET, Czech Republic

National Research and Education Network Operator

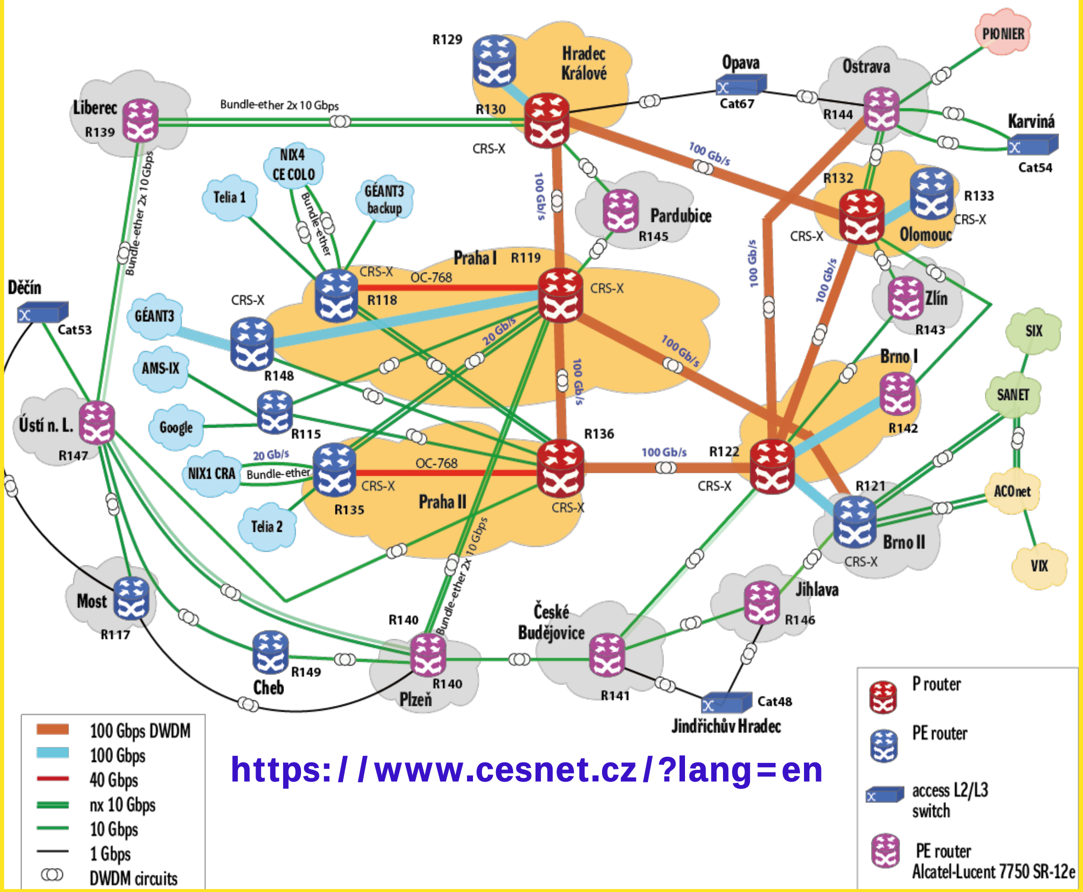
- ❒ Completed 100G Network Core by 2015
- ❒ 2016: CESNET2, 6000km of leased optical fibers + DWDM

External:

- ❒ 100 Gbps to Géant
- ❒ 20 Gbps to LHCONE
- ❒ 10 Gbps commodity traffic
- ❒ 10 Gbps to NetherLight for GLIF
- ❒ 10 Gbps to AMS-IX

Crossborder connections:

- ❒ 20 Gbps to SANET (NREN of the Slovak Republic) and SIX
- ❒ 20 Gbps to ACONet (NREN of Austria) and VIX, including precise time transmission
- ❒ 10 Gbps to PIONIER (NREN Poland)
- ❒ 2x20 Gbps to the Czech Neutral Internet Exchange (NIX.CZ)

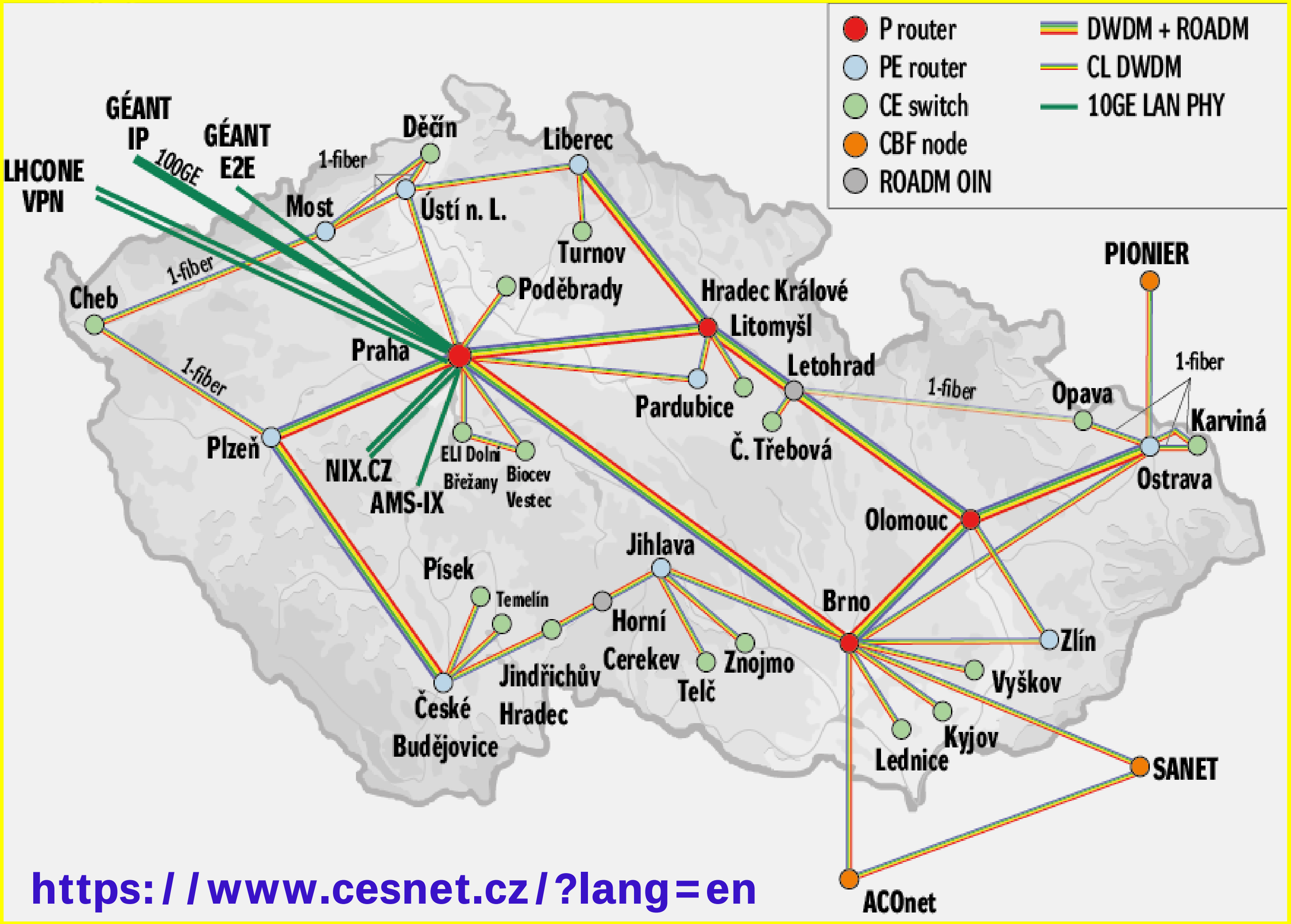


https://www.cesnet.cz/?lang=en

Supports the national computing grid infrastructure

CESNET2 DWDM Optical Topology

Hybrid Communication Network



https://www.cesnet.cz/?lang=en

Based on 6000 km of leased optical fiber

100GE Connection to GEANT IP Services

Diverse photonic technology offers increased availability and reliability for R&E collaboration

Offers 10G and 100G Wavelengths; IPv4/v6 multicast, MPLS, QoS Services

Canada: Pioneered “Light Paths”
Participation in LHCONE

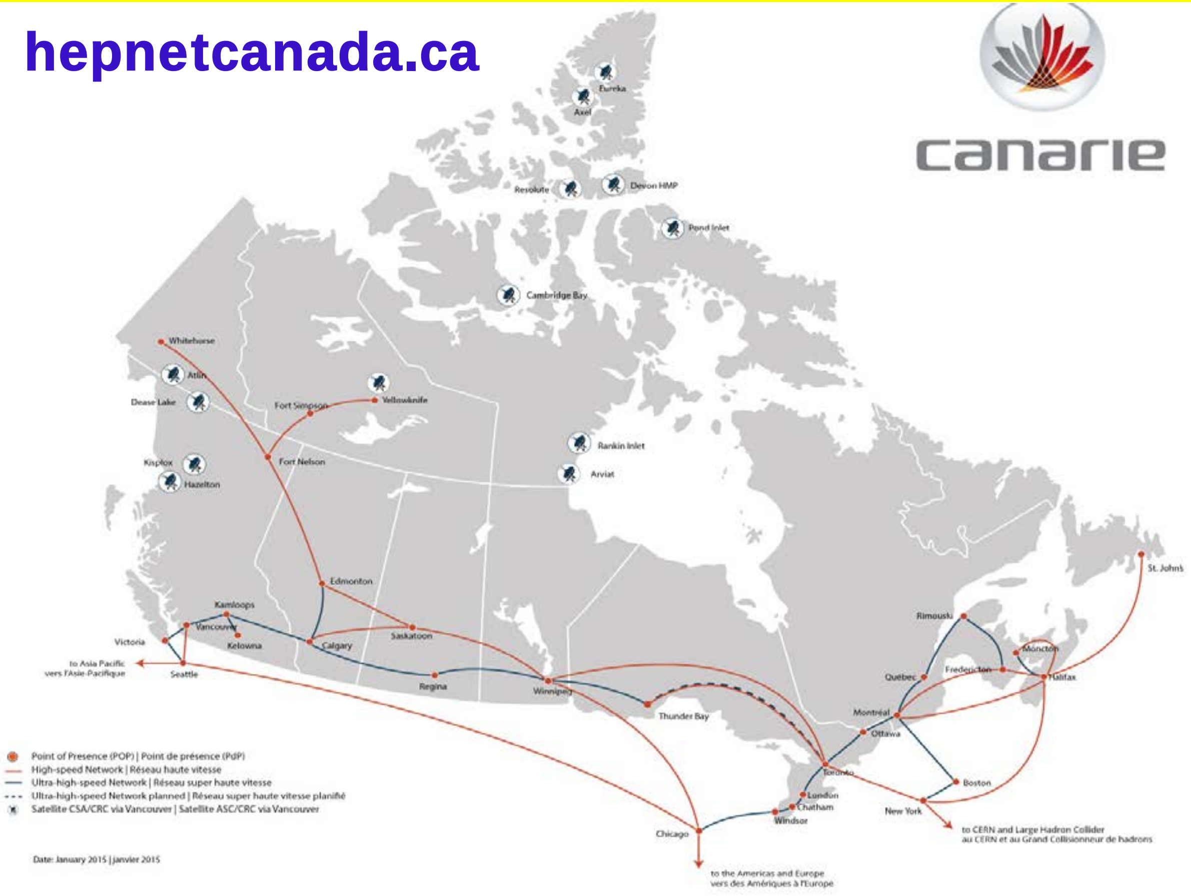
- ❒ 1M users at 1100 Institutions
- ❒ 88 wavelengths up to 100Gbps per wavelength
- ❒ Lightpaths in CANARIE available to researchers
- ❒ LHCONE VRF
- ❒ Tier1 & all Tier2s connected
- ❒ 2015: 100G IP link from Victoria to NYC

T. Tam, R. Sobie, I. Gable

2016

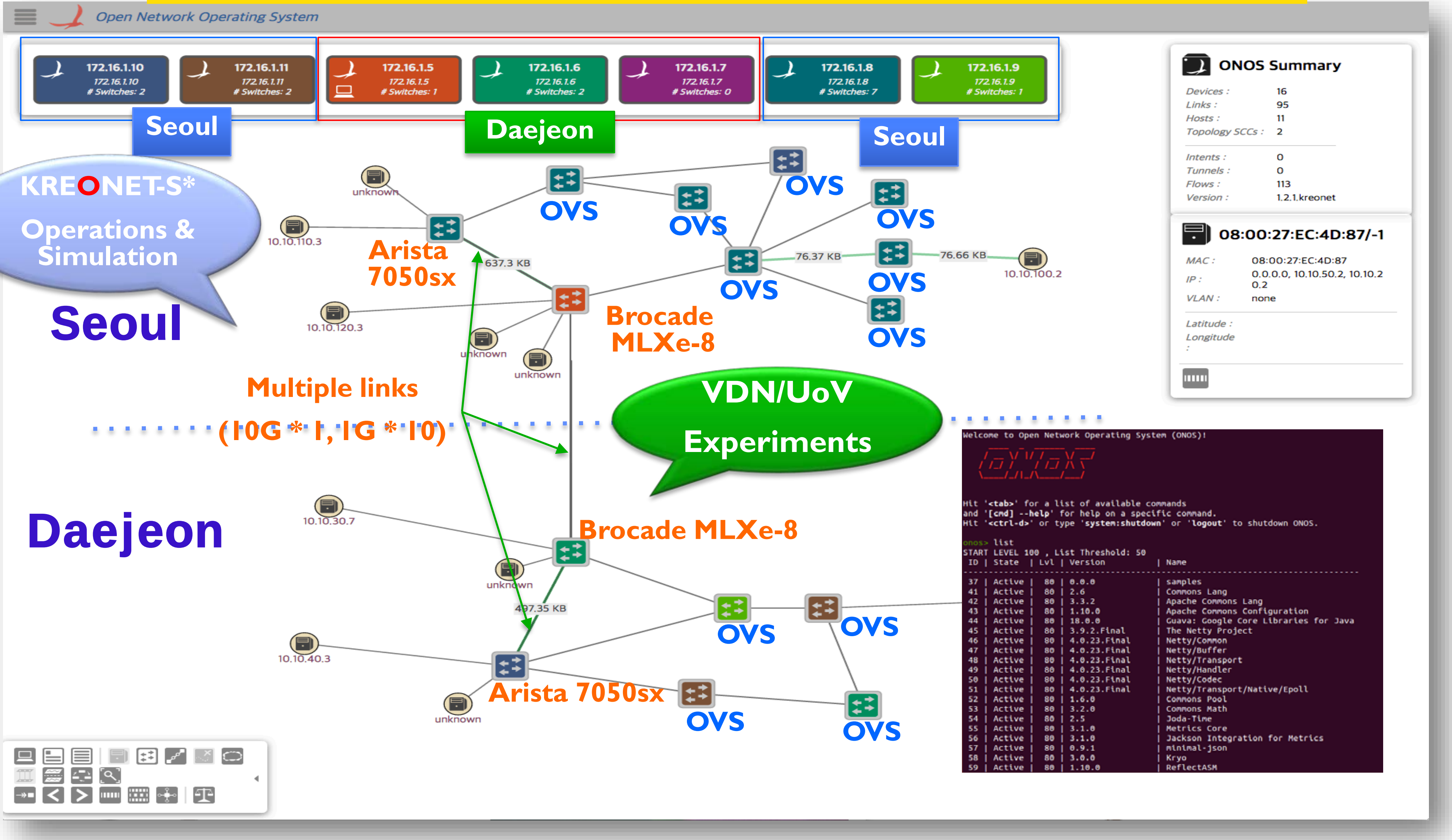
- ❒ 100G redundant core IP network
- ❒ International ANA-300G: 3x100 Gbps across the Atlantic

hepnetcanada.ca


canarie
Eureka
Axel
Resolute
Devon HMP
Pond Inlet
Cambridge Bay
Whitehorse
Atlin
Dease Lake
Fort Simpson
Yellowknife
Rankin Inlet
Arviat
Kisplox
Fort Nelson
Hazelton
Edmonton
Kamloops
Vancouver
Victoria
Kelowna
Calgary
Saskatoon
to Asia Pacific
vers l'Asie-Pacifique
Seattle
Regina
Winnipeg
Thunder Bay
St. Johns
Rimouski
Moncton
Québec
Fredericton
Halifax
Montréal
Ottawa
Toronto
Boston
London
Chatham
Windsor
New York
Chicago
to CERN and Large Hadron Collider
au CERN et au Grand Collisionneur de hadrons
to the Americas and Europe
vers des Amériques à l'Europe
● Point of Presence (POP) | Point de présence (PdP)
High-speed Network | Réseau haute vitesse
Ultra-high-speed Network | Réseau super haute vitesse
Ultra-high-speed Network planned | Réseau super haute vitesse planifié
Satellite CSA/CRC via Vancouver | Satellite ASC/CRC via Vancouver
Date: January 2015 | janvier 2015

KREONET SDN Deployment: KREONET-S

ONOS SDN Controller: Managing multi-vendor OpenFlow Switches including OVS on servers



ONOS Summary

Devices : 16
Links : 95
Hosts : 11
Topology SCCs : 2

Intents : 0
Tunnels : 0
Flows : 113
Version : 1.2.1.kreonet

08:00:27:EC:4D:87/-1

MAC : 08:00:27:EC:4D:87
IP : 0.0.0.0, 10.10.50.2, 10.10.2 0.2
VLAN : none

Latitude :
Longitude :

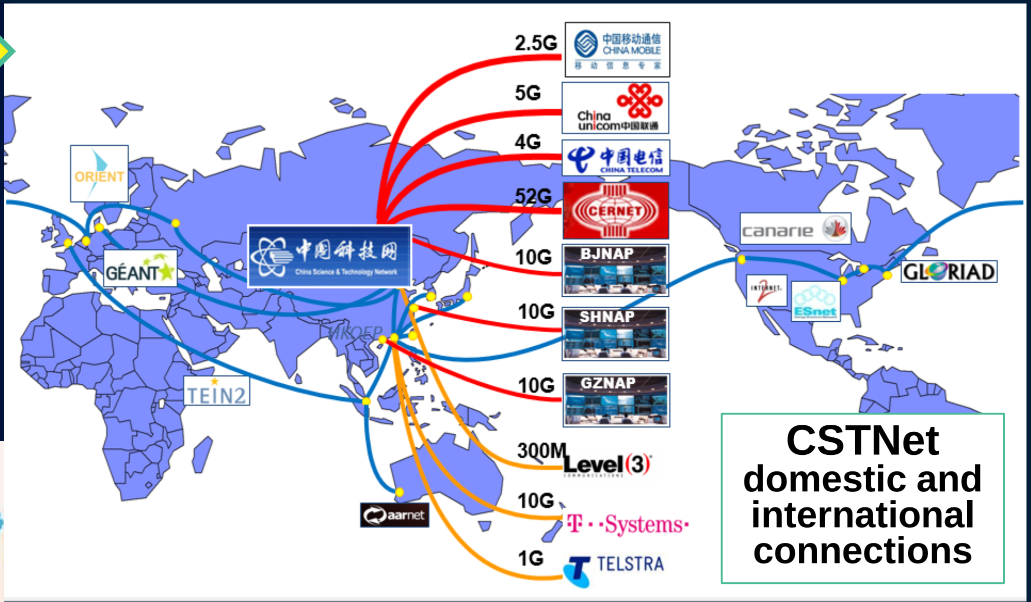
```
Welcome to Open Network Operating System (ONOS)!

Hit '<tab>' for a list of available commands
and '[cmd] --help' for help on a specific command.
Hit '<ctrl-d>' or type 'system:shutdown' or 'logout' to shutdown ONOS.

onos> list
START LEVEL 100 , List Threshold: 50
 ID | State  | Lvl | Version             | Name
----------------------------------------------------------------------------
 37 | Active |  80 | 0.0.0               | samples
 41 | Active |  80 | 2.6                 | Commons Lang
 42 | Active |  80 | 3.3.2               | Apache Commons Lang
 43 | Active |  80 | 1.10.0              | Apache Commons Configuration
 44 | Active |  80 | 18.0.0              | Guava: Google Core Libraries for Java
 45 | Active |  80 | 3.9.2.Final         | The Netty Project
 46 | Active |  80 | 4.0.23.Final        | Netty/Common
 47 | Active |  80 | 4.0.23.Final        | Netty/Buffer
 48 | Active |  80 | 4.0.23.Final        | Netty/Transport
 49 | Active |  80 | 4.0.23.Final        | Netty/Handler
 50 | Active |  80 | 4.0.23.Final        | Netty/Codec
 51 | Active |  80 | 4.0.23.Final        | Netty/Transport/Native/Epoll
 52 | Active |  80 | 1.6.0               | Commons Pool
 53 | Active |  80 | 3.2.0               | Commons Math
 54 | Active |  80 | 2.5                 | Joda-Time
 55 | Active |  80 | 3.1.0               | Metrics Core
 56 | Active |  80 | 3.1.0               | Jackson Integration for Metrics
 57 | Active |  80 | 0.9.1               | minimal-json
 58 | Active |  80 | 3.0.0               | Kryo
 59 | Active |  80 | 1.10.0              | ReflectASM
```

R&E Networking in China

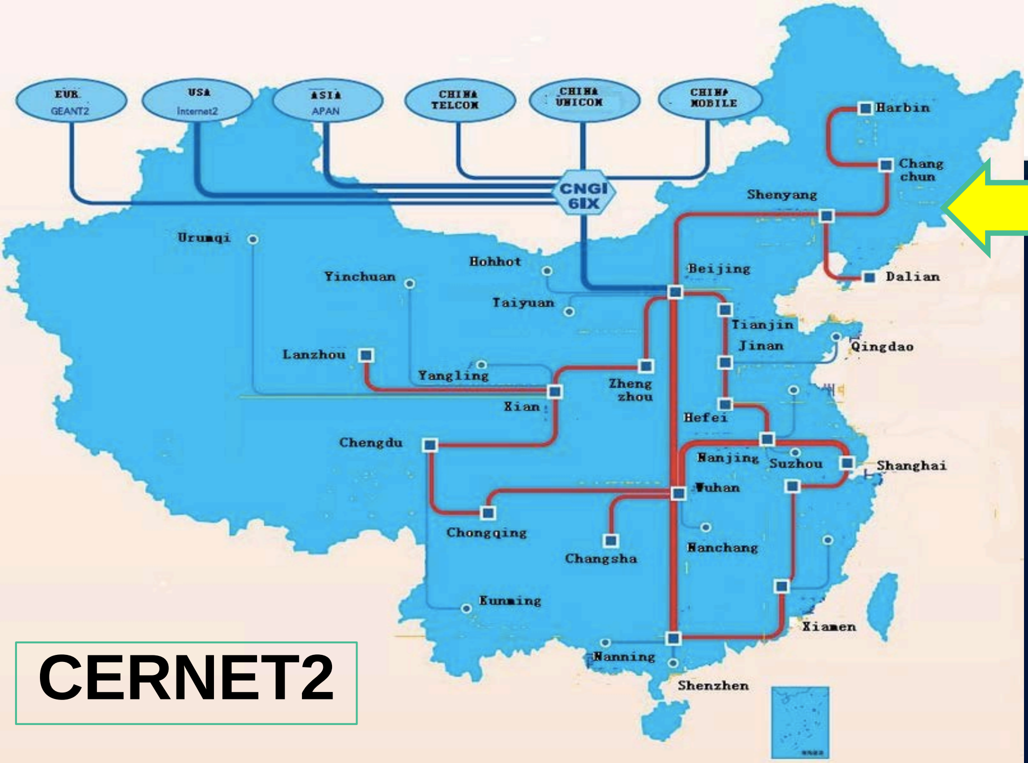
Gang Chen, IHEP Beijing

CSTNet

- China Science and Technology Network,
- academic network system, Chinese Academy of Sciences
- 12 regional centers, 370 institutes, 1M users
- 2.5 Gbps between major cities
- 10Gbps between Beijing and Europe



CSTNet domestic and international connections

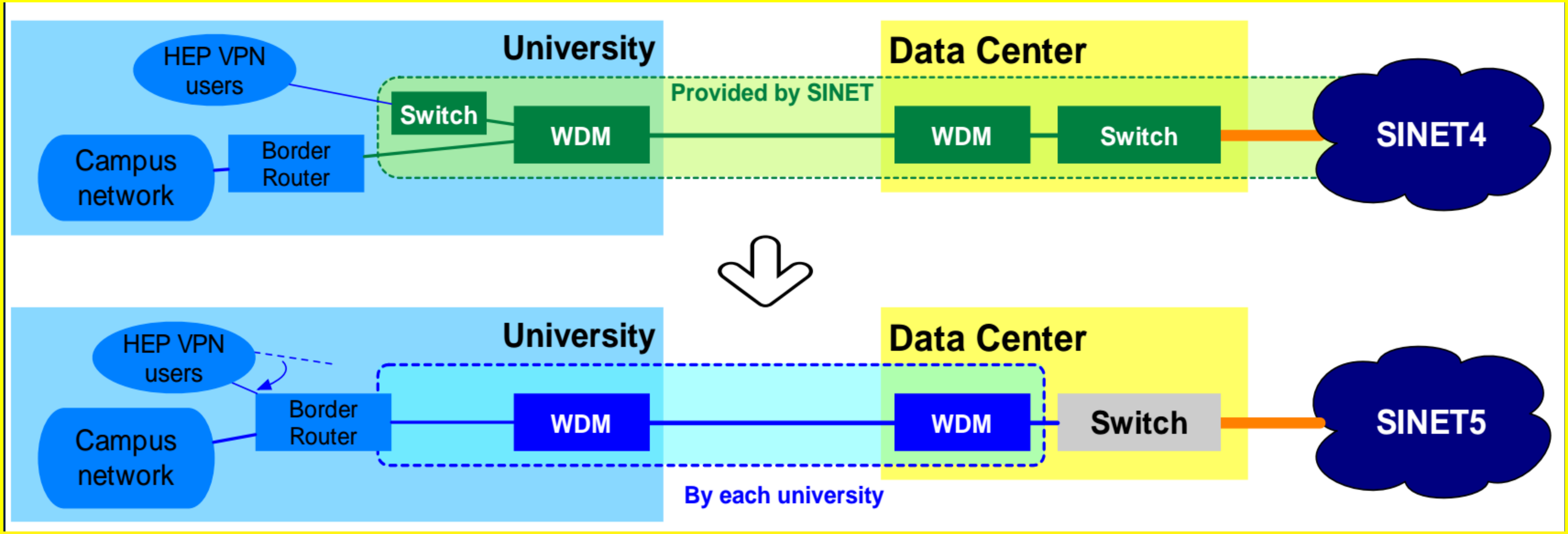


CERNET2

CERNET

- China Education and Research Network
- largest academic network in the country
- backbone: 10~100Gbps with 38 PoPs in 36 cities and over 2600 institutes
- total number of CERNET users > 25M
- CERNET2: 2nd generation 2.5~10Gbps with 25 PoPs in 20 cities and over 600 institutes

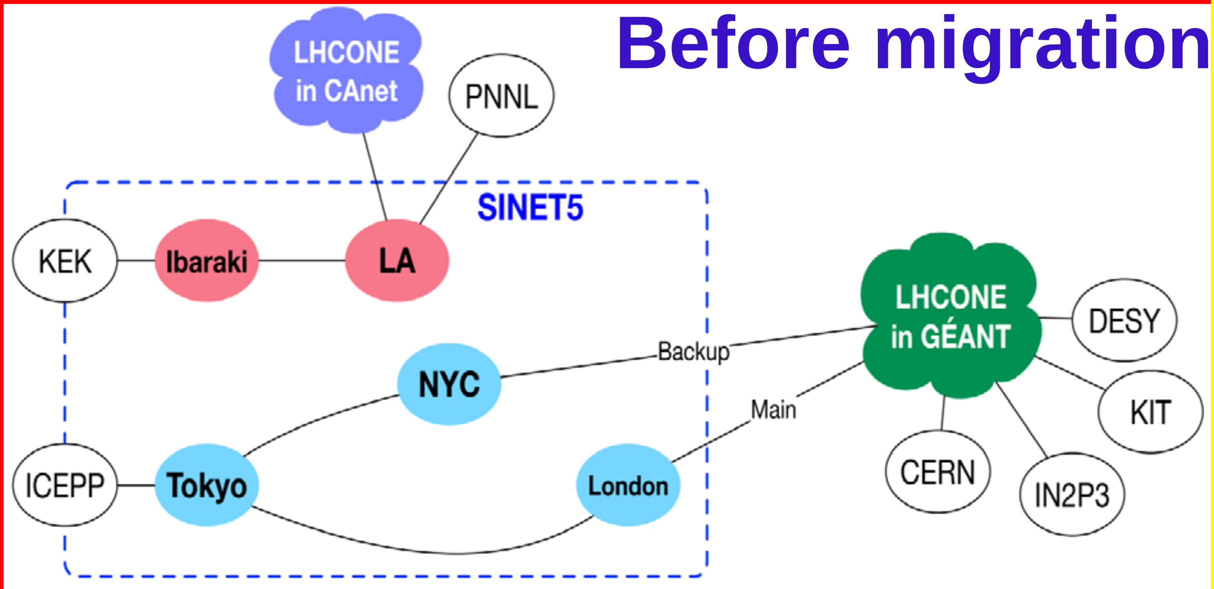
SINET4, SINET5 and HEPNet-J (Japan) Update



- Re-arranged physical connections at all HEPnet-J sites. Some got new 10G links on border switches
- SINET4 had four international links: 3 x 10G to US exchanges (LAX, MANLAN, WIX) and 1 x 10G to Singapore.
- SINET5 upgraded links to 1 x 100G to LAX and replaced WIX connection with 2 x 10G to London

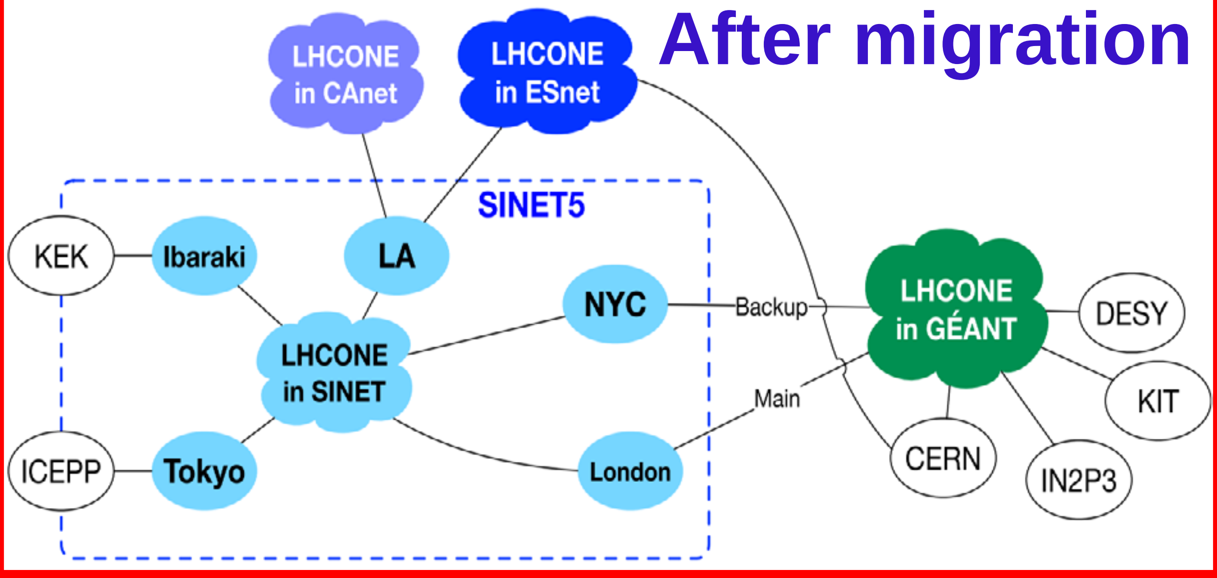
SINET4, SINET5 and HEPNet-J (Japan) Update

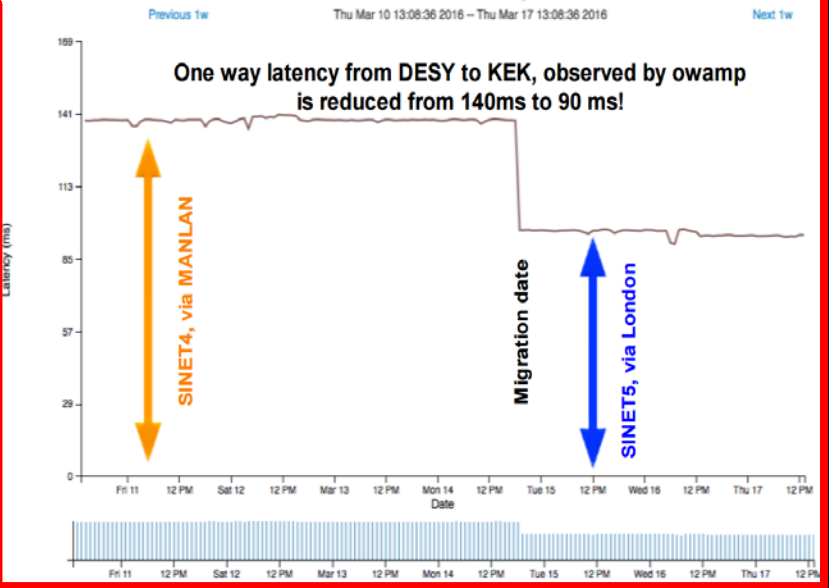
Before migration



- **Reduced RTT, Improved transfer speed** by factor ~3x: from 3 Gpbs to 8.8 Gbps
- **Both ICEPP and KEK** are now accessible from LHCONE

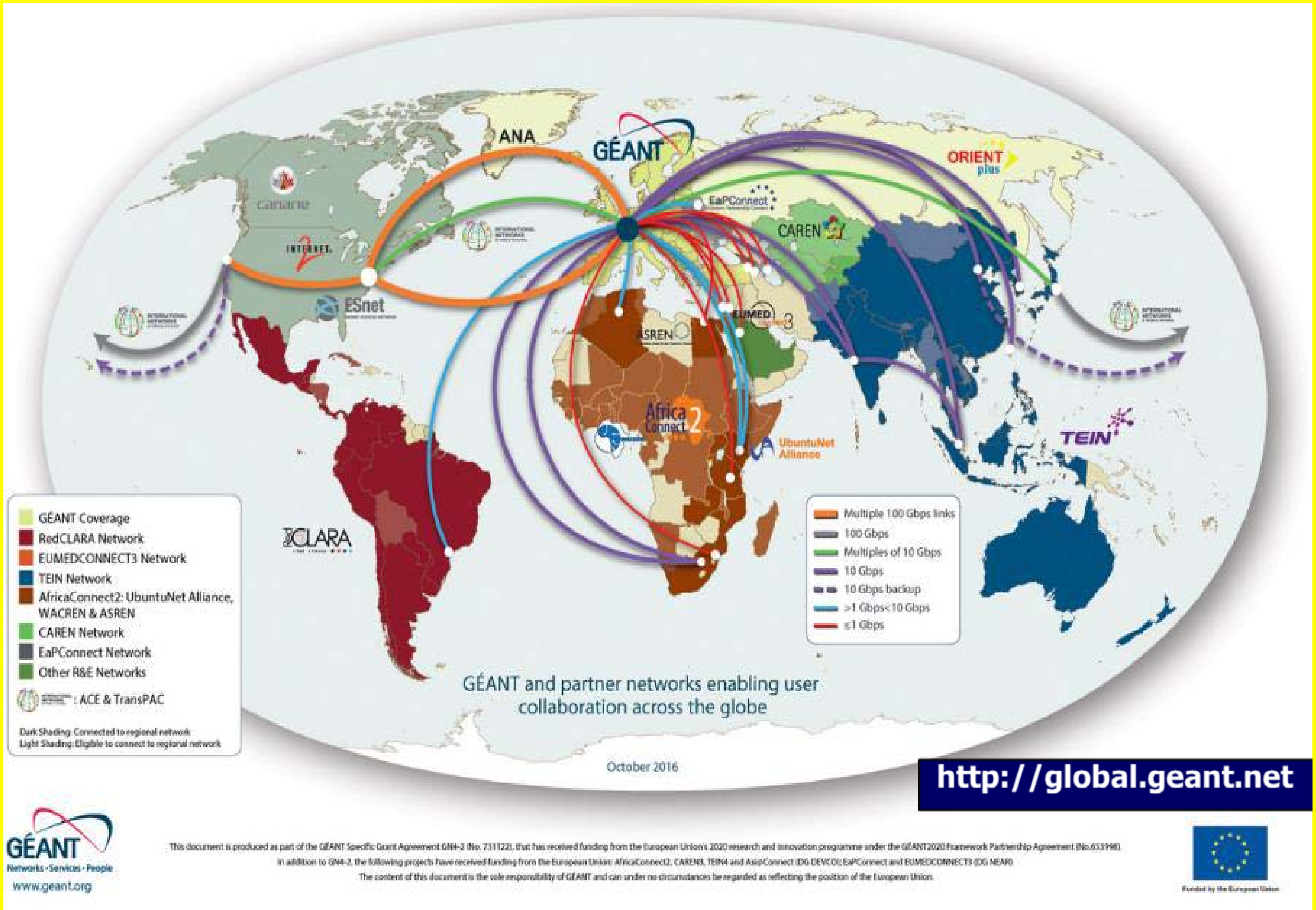
After migration




One way latency from DESY to KEK, observed by owamp is reduced from 140ms to 90 ms!

Closing the Digital Divide

GÉANT Global Connectivity



GÉANT and partner networks enabling user collaboration across the globe

October 2016

http://global.geant.net

www.geant.org

This document is produced as part of the GÉANT Specific Grant Agreement GN4-2 (No. 731122), that has received funding from the European Union's 2020 research and innovation programme under the GÉANT2020 Framework Partnership Agreement (No.653998).
In addition to GN4-2, the following projects have received funding from the European Union: AfricaConnect2, CAREN3, TEIN4 and AsiaConnect (DG DEVCO); EaPConnect and EUMEDCONNECT3 (DG NEAR).
The content of this document is the sole responsibility of GÉANT and can under no circumstances be regarded as reflecting the position of the European Union.

AfricaConnect: London – S. Africa 10G Links to **Ubuntunet Alliance;**
EUMEDCONNECT3 to Eastern and Southern Mediterranean
C@ribNet to Caribbean; **CAREN** to Central Asian FSU Republics;
OrientPlus to CSTNet and CERNET in China; **RedCLARA** to Latin America

African NRENS: UbuntuNet www.ubuntunet.net and WACREN Alliance formed in 2014

UbuntuNet Alliance

16 **Eastern and Southern Africa NRENs**

BERNET *(Burundi)*

Eb@le (Dem. Rep. of Congo)

EthERNet (Ethiopia)

iRENALA (Madagascar)

KENET (Kenya)

MAREN (Malawi)

MoRENet (Mozambique)

RENU (Uganda)

RwEdNet (Rwanda)

SomaliREN (Somalia)

SudREN (Sudan)

TENET (South Africa)

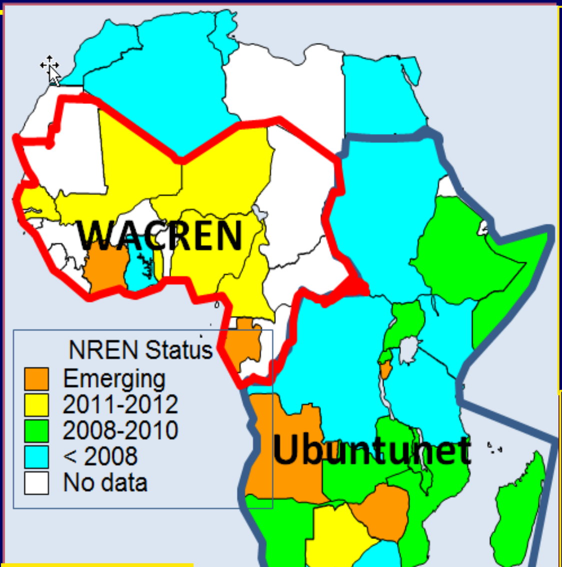
TERNET (Tanzania)

Xnet (Namibia)

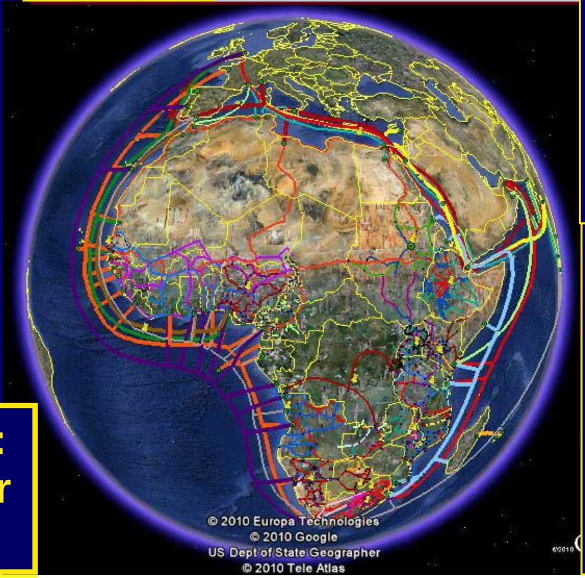
ZAMREN (Zambia)

ZARNet (Zimbabwe)

More Information on Ubuntunet: Nov. 2016 "NUANCE" Newsletter
http://www.ubuntunet.net/november2016



R. Cottrell
SLAC


© 2010 Europa Technologies
© 2010 Google
US Dept of State Geographer
© 2010 Tele Atlas

NRENs provide

- Leadership + coordination
- Training
- Leverage in contract negotiations: $4000 to $135 Per Mbps/Mo. in 4 Yrs

- *WACREN* West and Central Africa Research & Education Network http://wacren.net: *Alliance formed with Ubuntunet*

- N. Africa connected via EUMED to Europe
- ASREN: Arab States R&E Net formed in 2011

- With connection to GÉANT UbuntuNet provides sub-Saharan Africa with infrastructure for global and regional research collaboration and e-learning

ICFA SCIC

Africa: Setting up International eXchange Points for Better Connectivity

Connections between African countries are no longer via Europe or USA

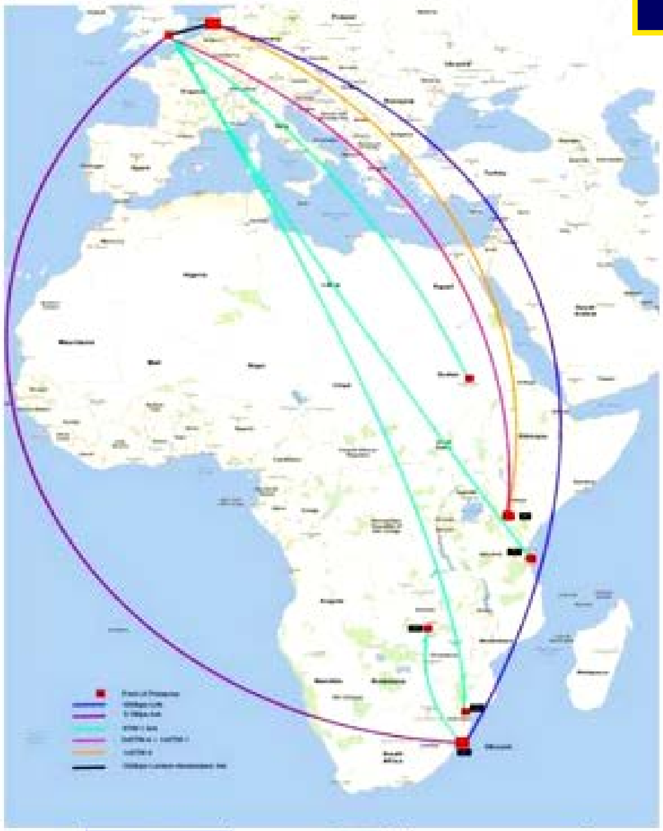
Much reduced Round Trip Times

Better reliability and performance

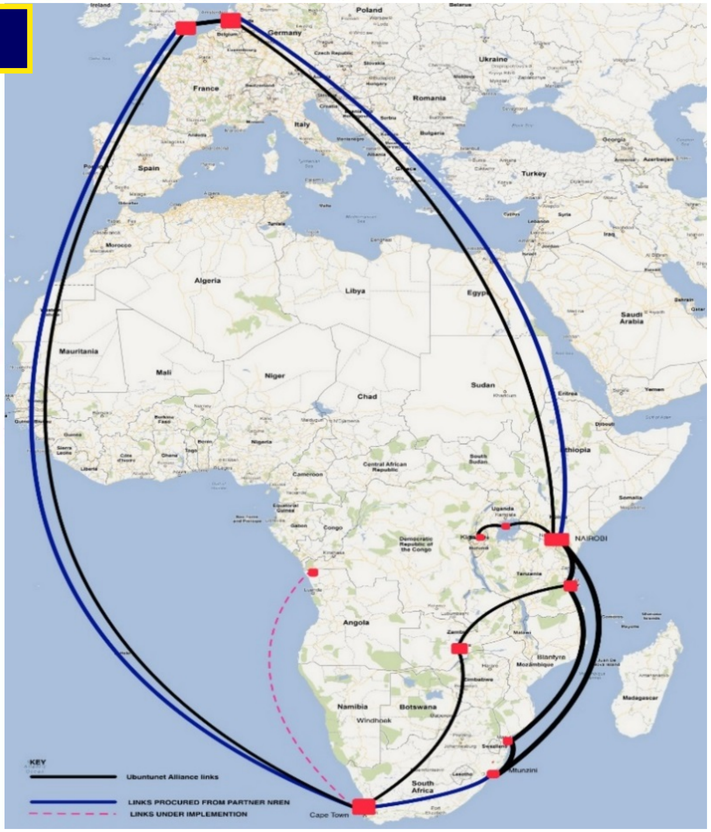
R. Cottrell
SLAC

AfricaConnect: Filling part of the regional connectivity gap

Africa Connect



2015



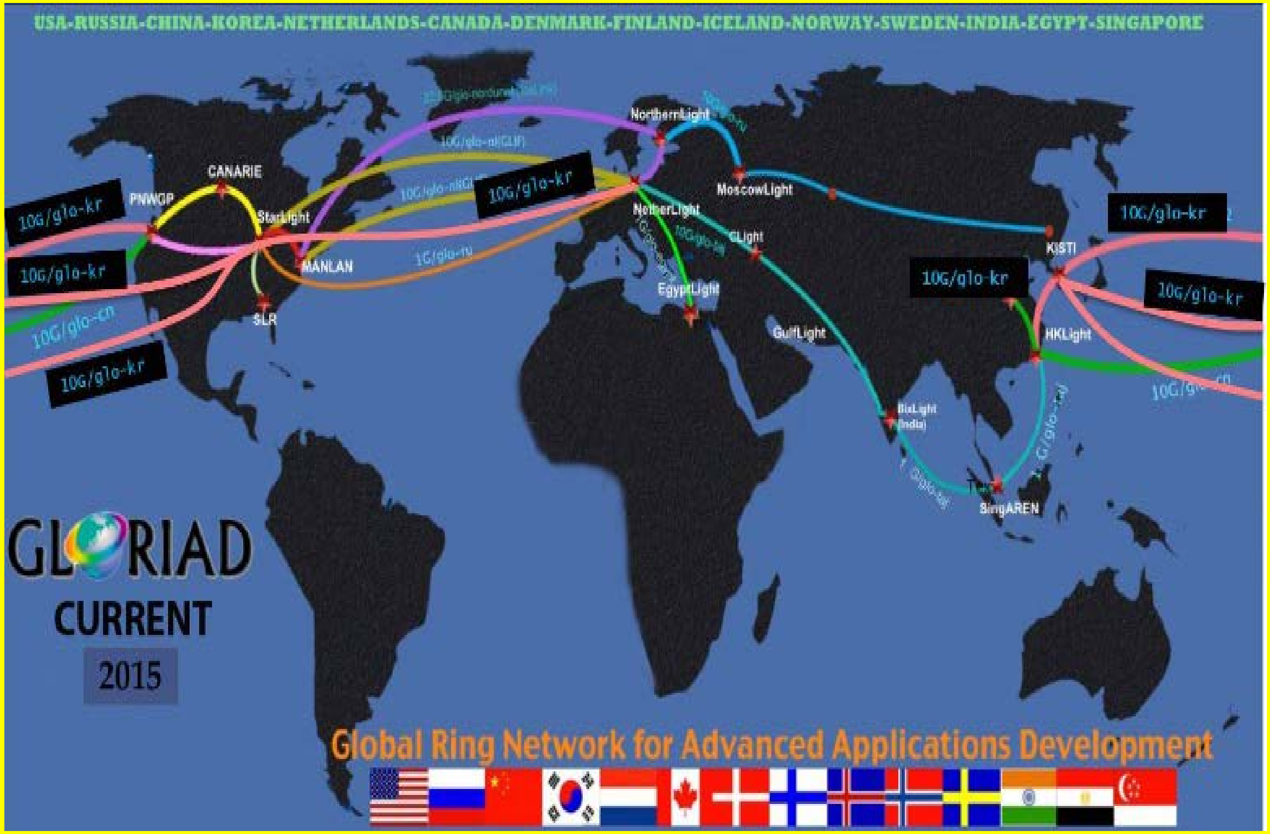
https://www.ubuntunet.net/network-topology

DANTE

UbuntuNet Alliance

The Research and Education Network for sub-Saharan Africa

GLORIAD: An Optical Ring Encircling the Globe



A Collaborative Hybrid Network with multiple 10G circuits to support Layer 1-3 Optical lightpath, switched and routed network services

2014-16 Highlights

Korea now 100G to US

GLORIAD exchange point in Amsterdam; 10GE to the NetherLight Exchange

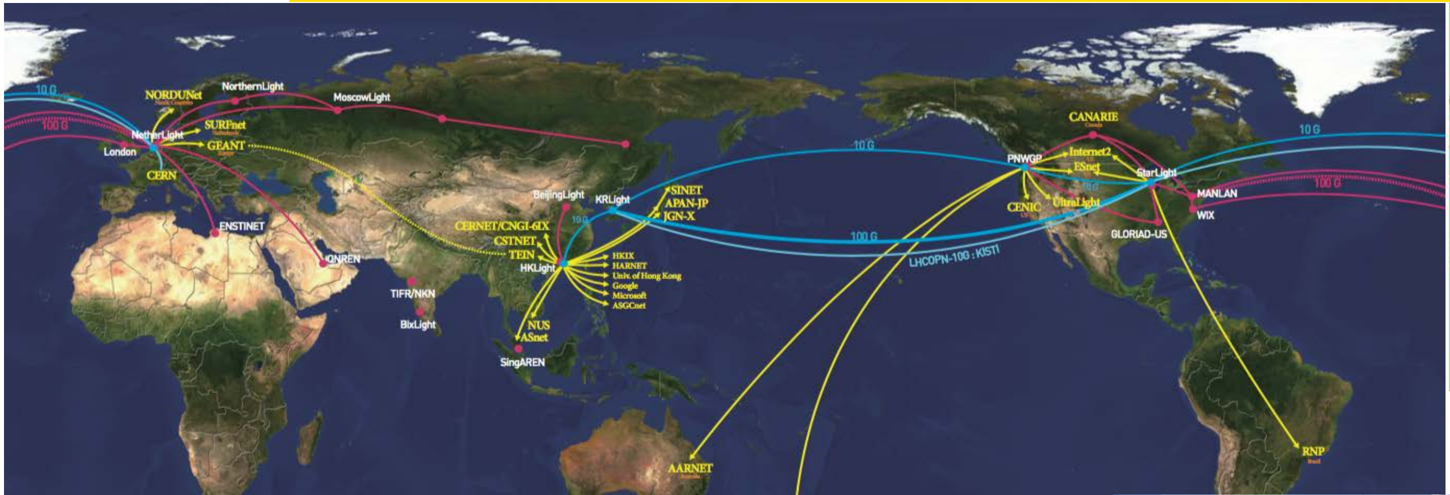
US – Russia link upgraded to 10GE

US-Qatar/Kuwait 10GE Link

China CSTnet connected at 10GE in Chicago; 4 X 1GE peerings with US Networks

Partners: US, Canada, Netherlands, Russia, China, Korea, Denmark, Finland, Iceland, Norway, Sweden, India, Egypt, Singapore, Malaysia

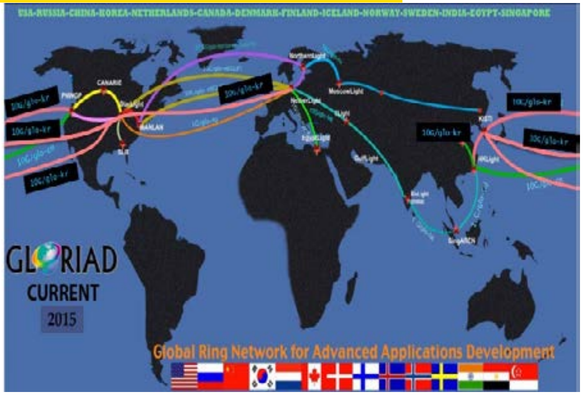
KREONet2 and GLORIAD-KR And SDN Deployment (KREONET-S)



AARNET

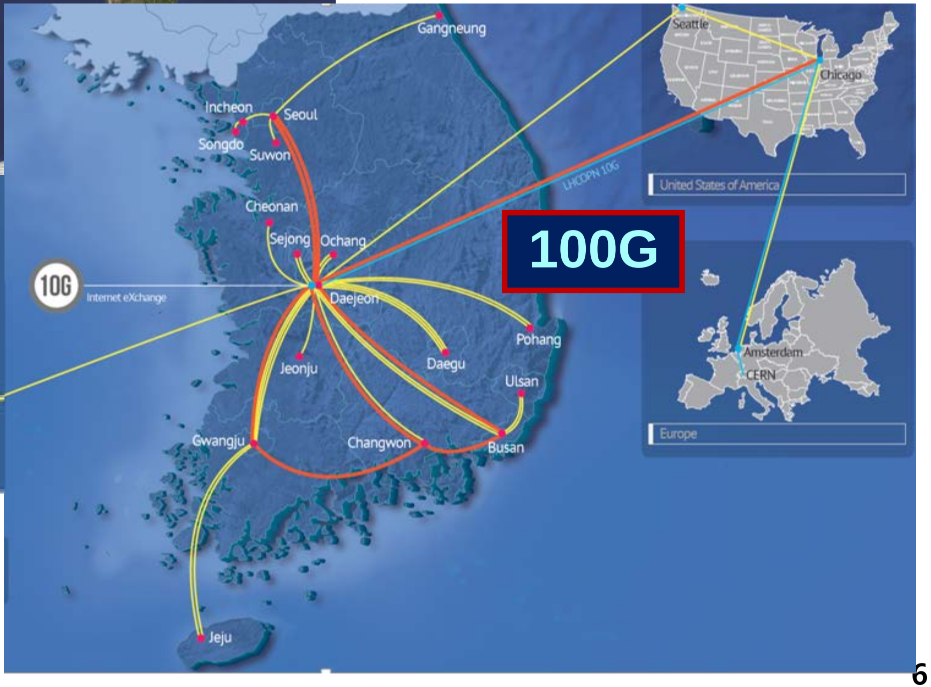
REANNZ

- GLORIAD-KR
- GLORIAD
- KRLight



2015-2016 Highlights

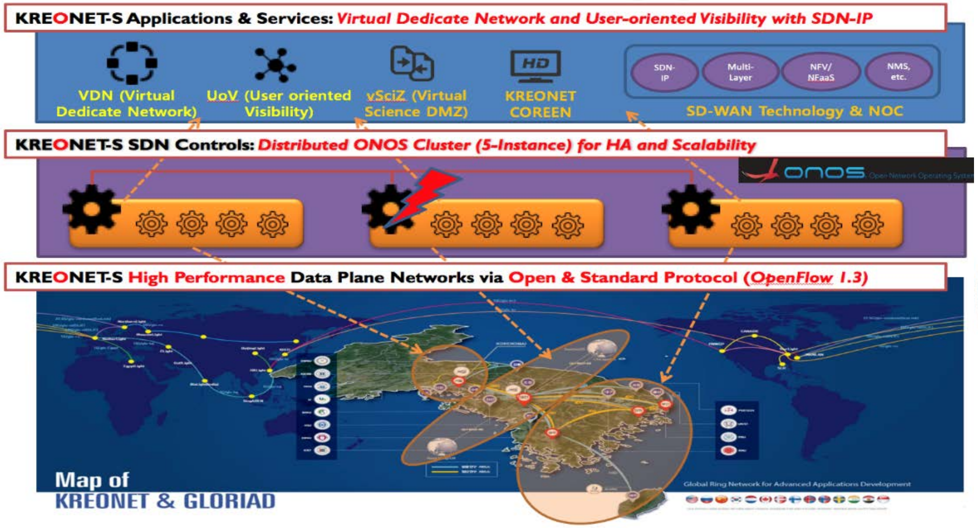
1. 100G from Daejon to Chicago/StarLight
2. 100G Ring linking major cities
3. 17 GigaPoPs with 1G, 10G or 40G



100G

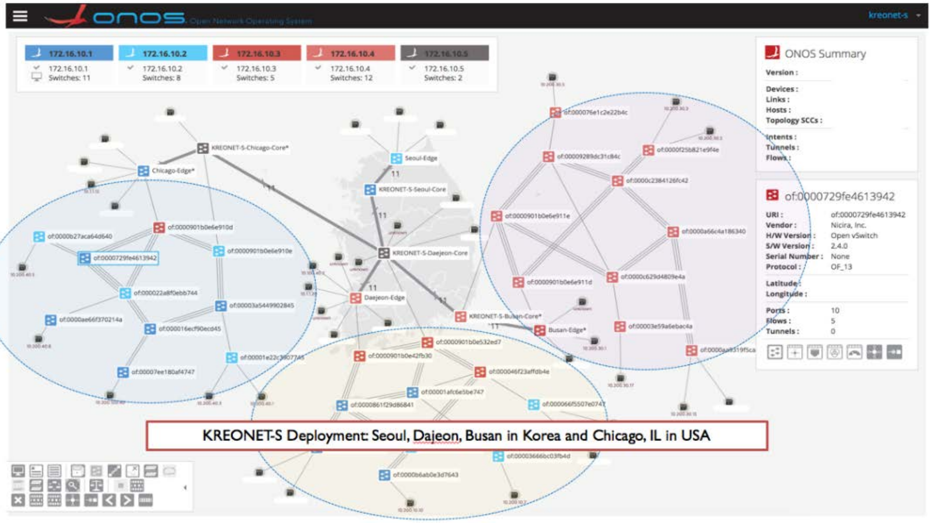
KREONET SDN Deployment: KREONET-S

User-oriented & On-Demand Virtual Dedicated Network (VDN) Provisioning based on ONOS

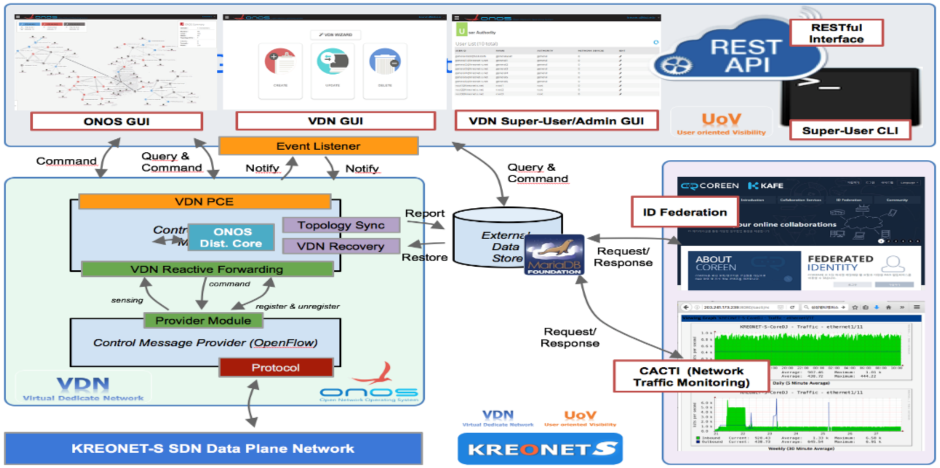


KREONET-S Applications & Services: Virtual Dedicate Network and User-oriented Visibility with SDN-IP
VDN (Virtual Dedicate Network)
UoV (User oriented Visibility)
vSciZ (Virtual Science DMZ)
HD
KREONET COREEN
SDN-IP
Multi-Layer
NFV/NFaaS
NMS, etc.
SD-WAN Technology & NOC
KREONET-S SDN Controls: Distributed ONOS Cluster (5-Instance) for HA and Scalability
onos Open Network Operating System
KREONET-S High Performance Data Plane Networks via Open & Standard Protocol (OpenFlow 1.3)
Map of KREONET & GLORIAD
Global Ring Network for Advanced Applications Development

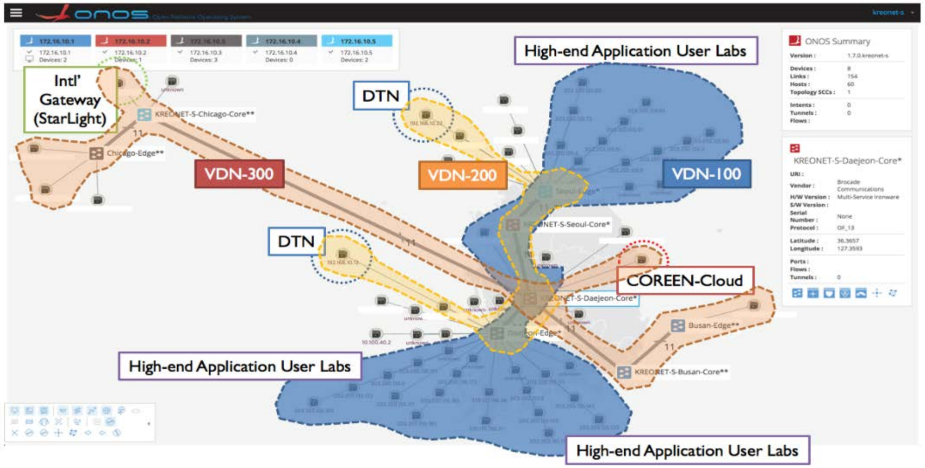
KREONET-S Primary Building Blocks



KREONET-S (International) SD-WAN Deployment



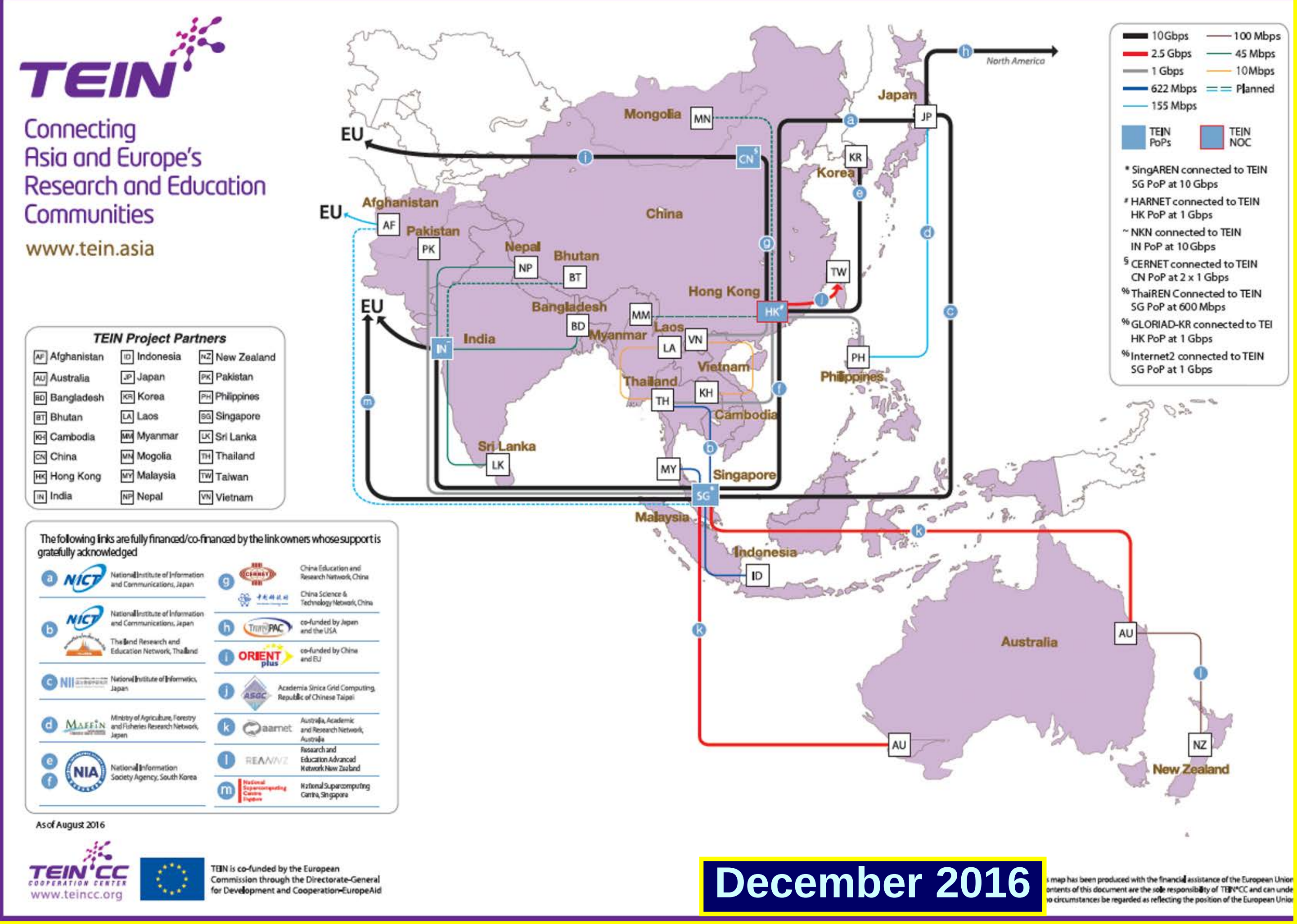
KREONET-S VDN Application Architecture



KREONET-S VDN Use Cases

TEIN4 Network

Enabling research communities in 20 Asian Countries


TEIN
Connecting Asia and Europe's Research and Education Communities
www.tein.asia
TEIN Project Partners
AF Afghanistan
AU Australia
BD Bangladesh
BT Bhutan
KH Cambodia
CN China
HK Hong Kong
IN India
ID Indonesia
JP Japan
KR Korea
LA Laos
MM Myanmar
MN Mogolia
MY Malaysia
NP Nepal
NZ New Zealand
PK Pakistan
PH Philippines
SG Singapore
LK Sri Lanka
TH Thailand
TW Taiwan
VN Vietnam
The following links are fully financed/co-financed by the link owners whose support is gratefully acknowledged
a NICT National Institute of Information and Communications, Japan
b NICT National Institute of Information and Communications, Japan
Thailand Research and Education Network, Thailand
c NII National Institute of Informatics, Japan
d MAFFIN Ministry of Agriculture, Forestry and Fisheries Research Network, Japan
e f NIA National Information Society Agency, South Korea
g CERNET China Education and Research Network, China
China Science & Technology Network, China
h TransPAC co-funded by Japan and the USA
i ORIENT plus co-funded by China and EU
j ASGC Academia Sinica Grid Computing, Republic of Chinese Taipei
k aarnet Australia, Academic and Research Network, Australia
l REANNZ Research and Education Advanced Network New Zealand
m National Supercomputing Centre Singapore National Supercomputing Centre, Singapore
As of August 2016
TEIN*CC COOPERATION CENTER www.teincc.org
TEIN is co-funded by the European Commission through the Directorate-General for Development and Cooperation-EuropeAid
10Gbps
2.5 Gbps
1 Gbps
622 Mbps
155 Mbps
100 Mbps
45 Mbps
10Mbps
Planned
TEIN PoPs
TEIN NOC
* SingAREN connected to TEIN SG PoP at 10 Gbps
# HARNET connected to TEIN HK PoP at 1 Gbps
~ NKN connected to TEIN IN PoP at 10 Gbps
§ CERNET connected to TEIN CN PoP at 2 x 1 Gbps
% ThaiREN Connected to TEIN SG PoP at 600 Mbps
% GLORIAD-KR connected to TEI HK PoP at 1 Gbps
% Internet2 connected to TEIN SG PoP at 1 Gbps
North America
EU
EU
EU
Mongolia
MN
China
CN
Japan
JP
Korea
KR
Afghanistan
AF
Pakistan
PK
Nepal
NP
Bhutan
BT
Bangladesh
BD
India
IN
Hong Kong
HK
TW
Myanmar
MM
Laos
LA
VN
Vietnam
Thailand
TH
KH
Cambodia
Philippines
PH
Sri Lanka
LK
MY
Malaysia
Singapore
SG
Indonesia
ID
Australia
AU
AU
New Zealand
NZ
a
b
c
d
e
f
g
h
i
j
k
k
l
m
s map has been produced with the financial assistance of the European Union
ontents of this document are the sole responsibility of TEIN*CC and can unde
o circumstances be regarded as reflecting the position of the European Unio

December 2016

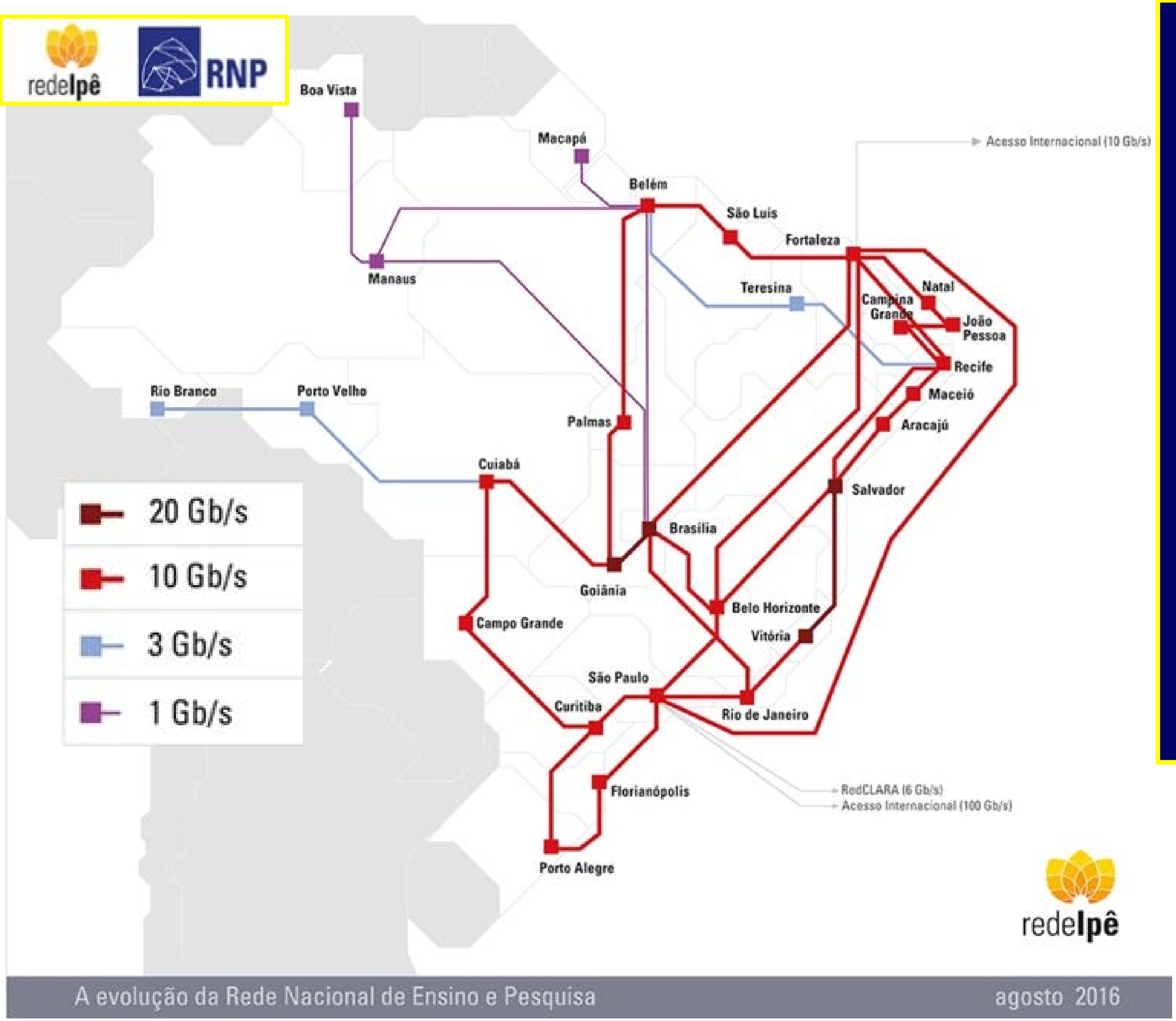
Managed by TEIN NOC in Hong Kong

Provides engineering, ops & research services to TEIN NRENs and partners for R&E collaboration

Offers: IPv4/v6, multicast, MPLS, QoS

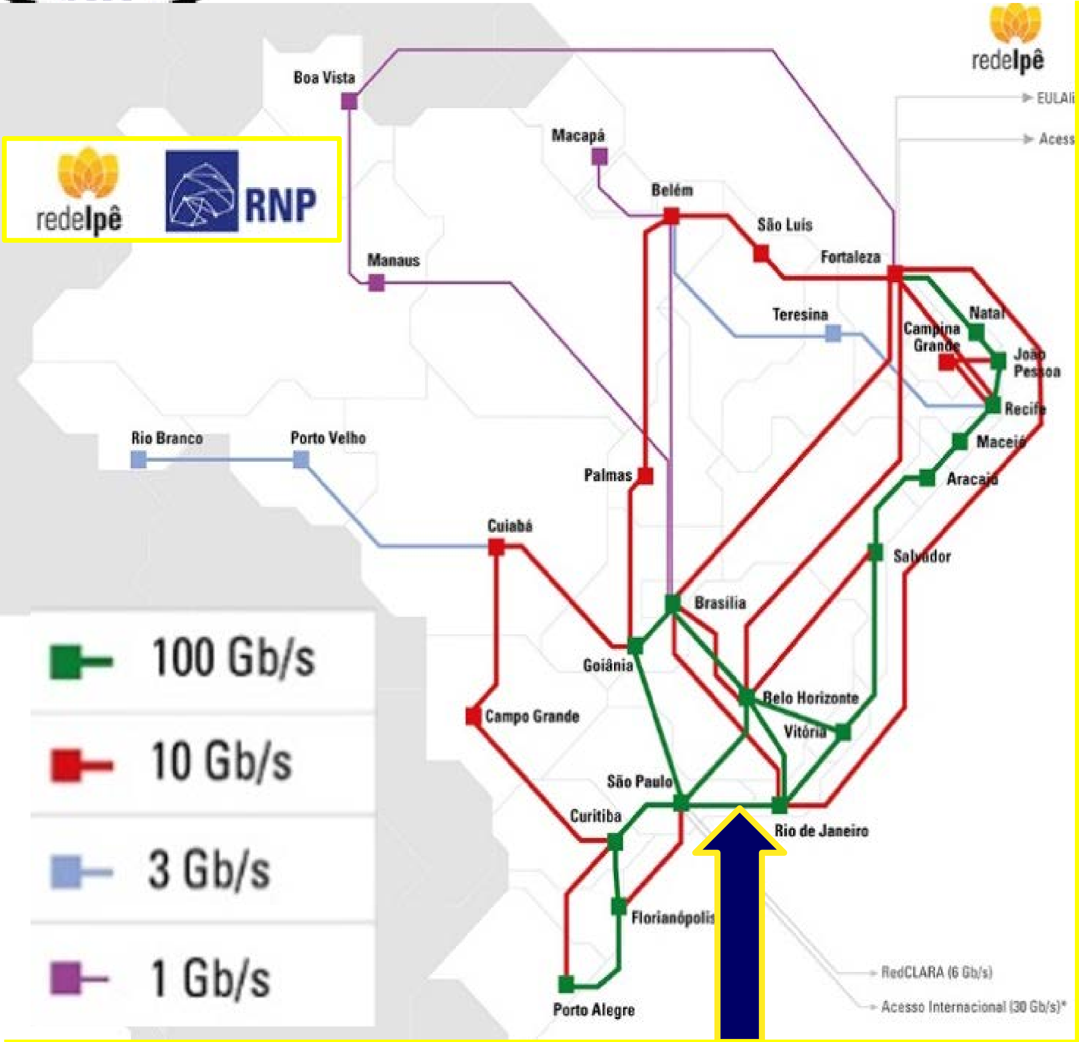
NOTE Some intra-regional links are still at 10 – 622 Mbps

Brazil: RNP Phase 6 Backbone:
347 Gbps Aggregate Capacity; 116 Gbps Int’l



- ❒ 10G + some 2 X 10G links in the Core
- ❒ Connecting all the State Capitals
- ❒ 1G Links across the Amazon to Manuas to the NW capitals
- ❒ 3G Links to the West capitals
- ❒ First 100G Int’l “OpenWave” link arrived in 2016

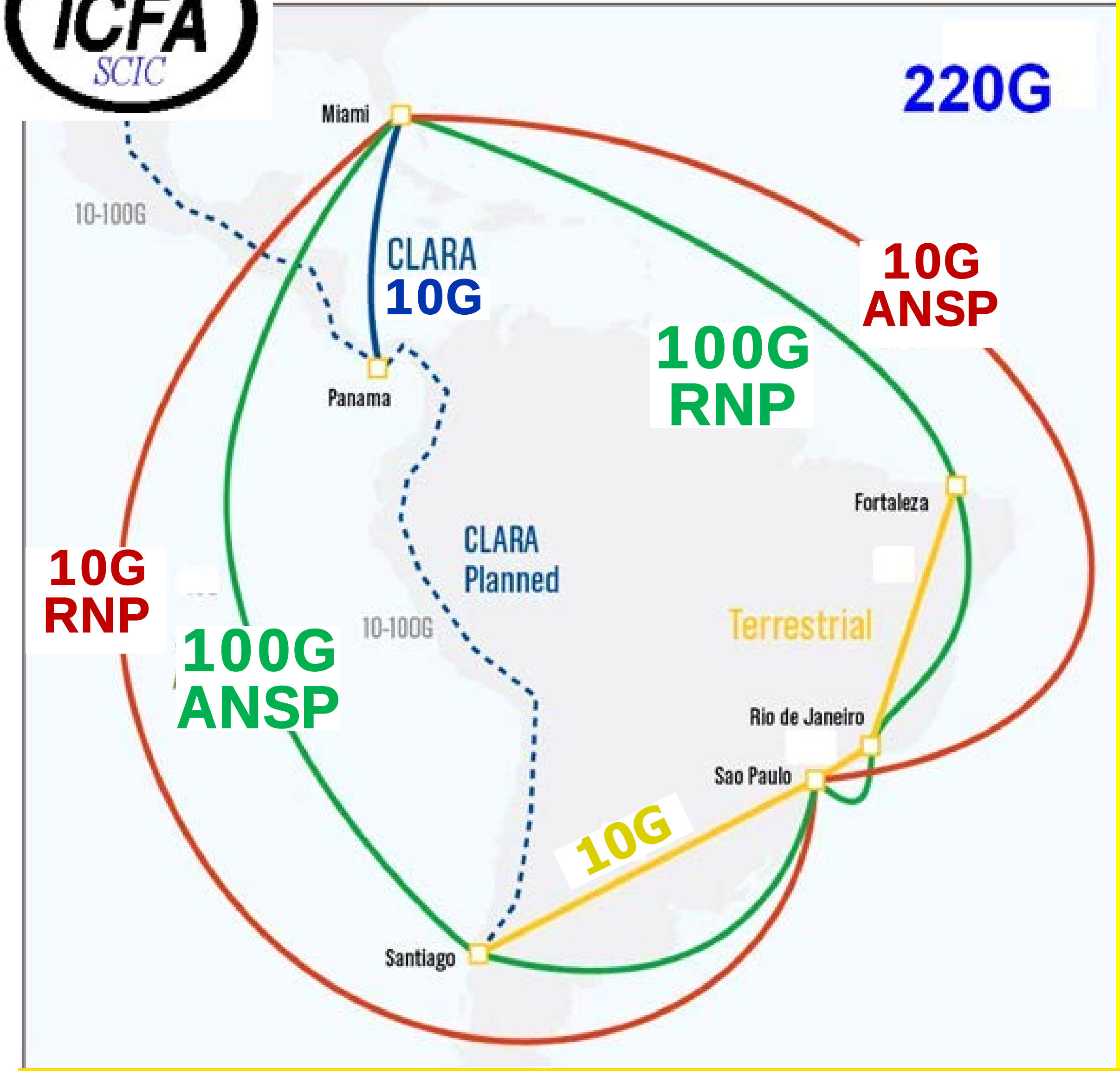
RNP Phase 7 Backbone with 100G Core Planned by 2019



Phase 7 RNP Backbone with 100G Core planned by 2019

- ❑ Requirement to support 100G waves starts in 2017
- ❑ By 2019 100G central rings and a 4000 km 100G backbone along the eastern coast are planned
- ❑ RNP is acquiring long-term rights to an extensive optical fiber infrastructure for the 100G transition

Brazil: Major Upgrade of Int'l Connectivity

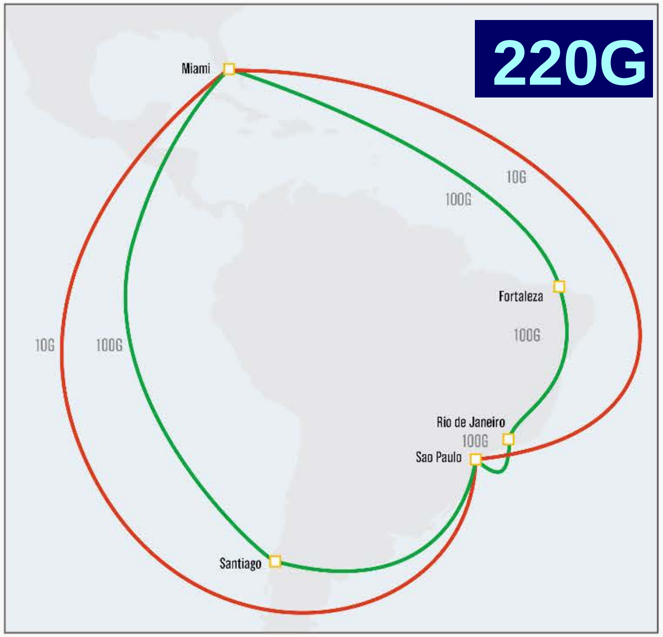


- ❒ RNP, ANSP with AmLight (US NSF): 220G Capacity from 2016
- ❒ Further expansion (N X 100G) foreseen from 2018
- ❒ Precedent-setting access to frequency spectrum by the academic community
- ❒ Backbone Sao Paulo-Rio-Fortaleza -St. Croix-Miami
- ❒ Will be extended to Chile at 100G then N X 100G
- ❒ To be heavily used by LSST into the 2030s

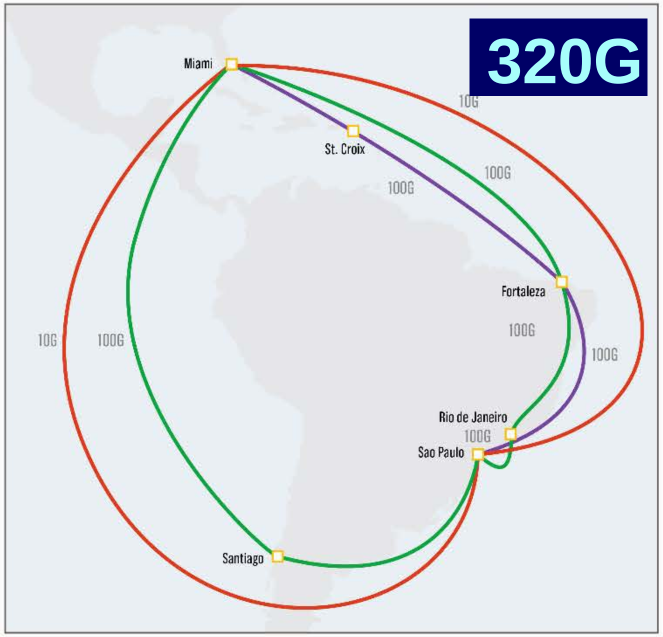
➡ Using Padtec (BR) 100G equipment. Demonstrations with the HEP team (Caltech, FIU, RNP, ANSP et al) at SC2013-14

Americas Lightpaths (AmLight)

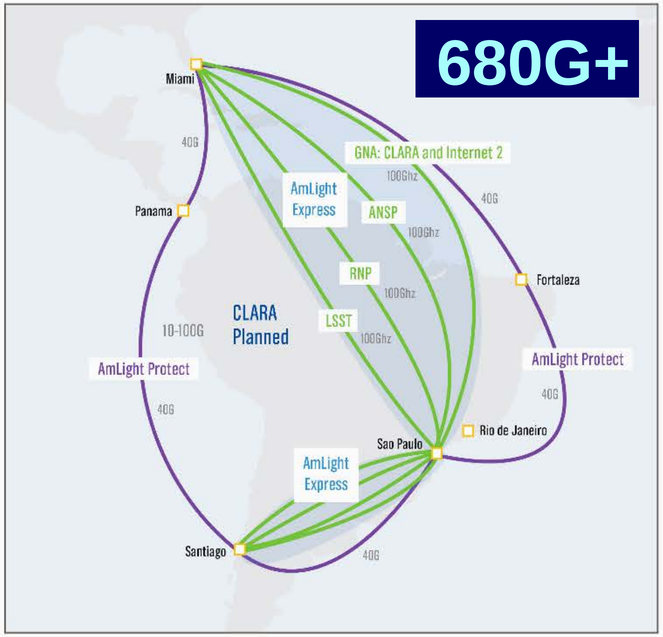
US-Latin America Amlight Backbone Plan: 220G to 680G+



2016



2017



2018-31

NSF support for *AmLight* is part of a scalable rational architecture, designed to support the needs of the U.S.-Western Hemisphere research and education community that supports the evolving nature of discovery and scholarship.

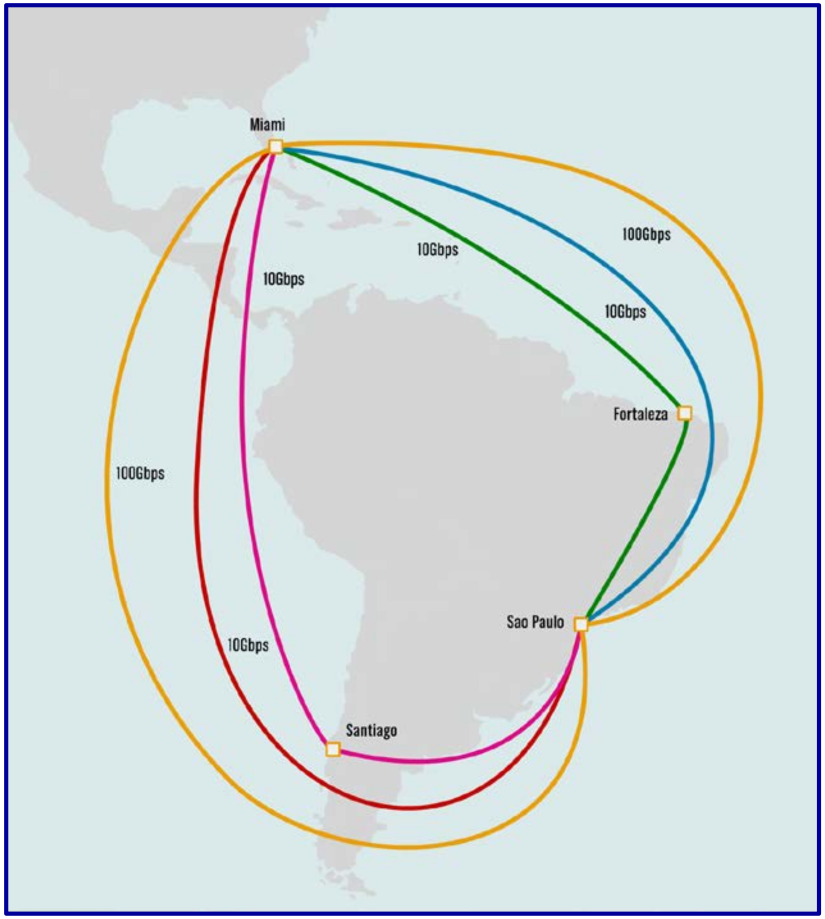
NSF Award # ACI-0963053 , 2010-2017

Americas Lightpaths Express and Protect
(AmLight Exp): US – Latin America

AmLight ExP implements a hybrid network strategy that combines the use of optical spectrum (Express) and leased capacity (Protect), in order to build a reliable, leading-edge network infrastructure for research and education

Links:

- **100G Miami-São Paulo, Atlantic**
- **100G Miami-São Paulo, Pacific**
- **4x10G links, landings in São Paulo, Fortaleza, Santiago**
- **240G of aggregate bandwidth capacity**
- **100G ring including Santiago and Fortaleza**



(NSF Award# ACI-1451018 2015-2020)

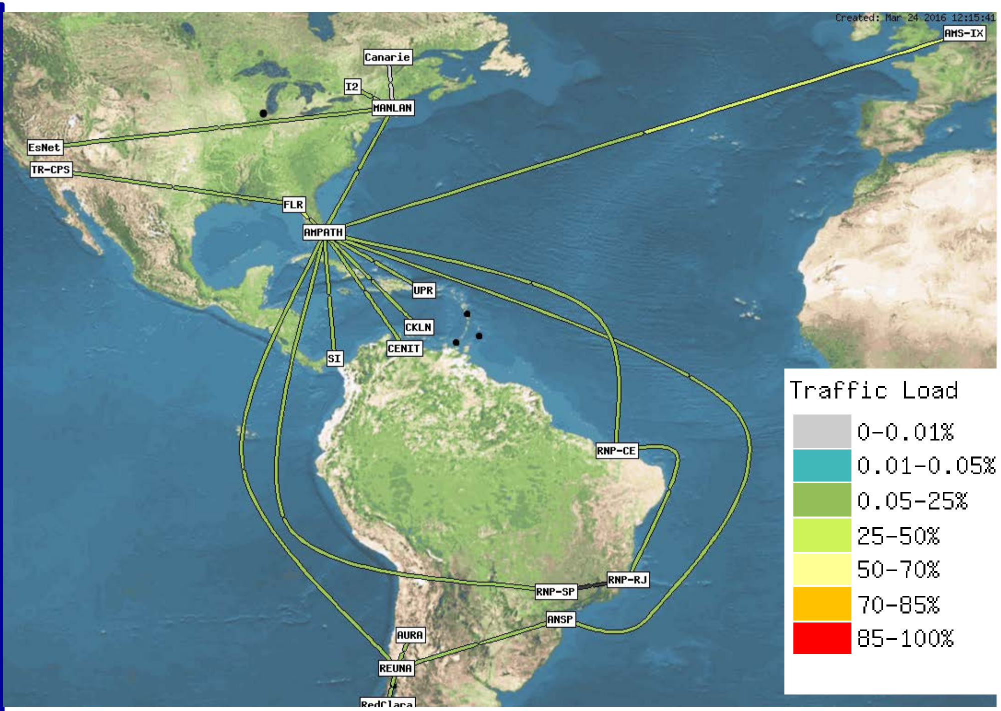
Ampath Intenational Exchange Point (IXP)

- **AMPATH is an Open R&E eXchange Point (RXP) led by Florida International University (FIU)**

- **Serves as the premiere interconnection point for network-enabled U.S.- Latin America and Caribbean science research and education.**

- **Supports science research and education programs of the NSF, DOE, etc.**

- **Operates 3x100G and multiple 10G circuits in collaboration with FLR, ANSP and RNP**



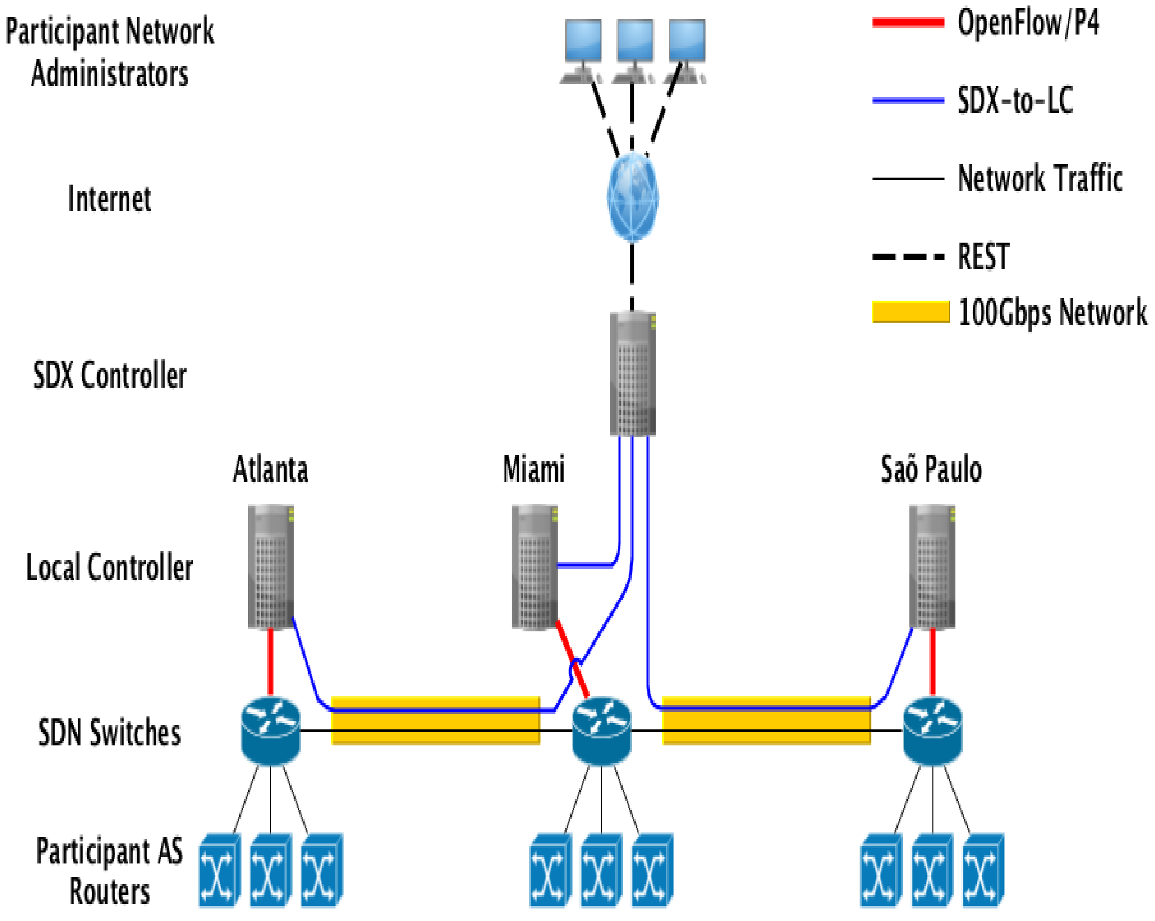
http://measurements2.ampath.net/

www.ampath.net

AtlanticWave-SDX: A Distributed Intercontinental Experimental Software Defined Exchange (SDX)

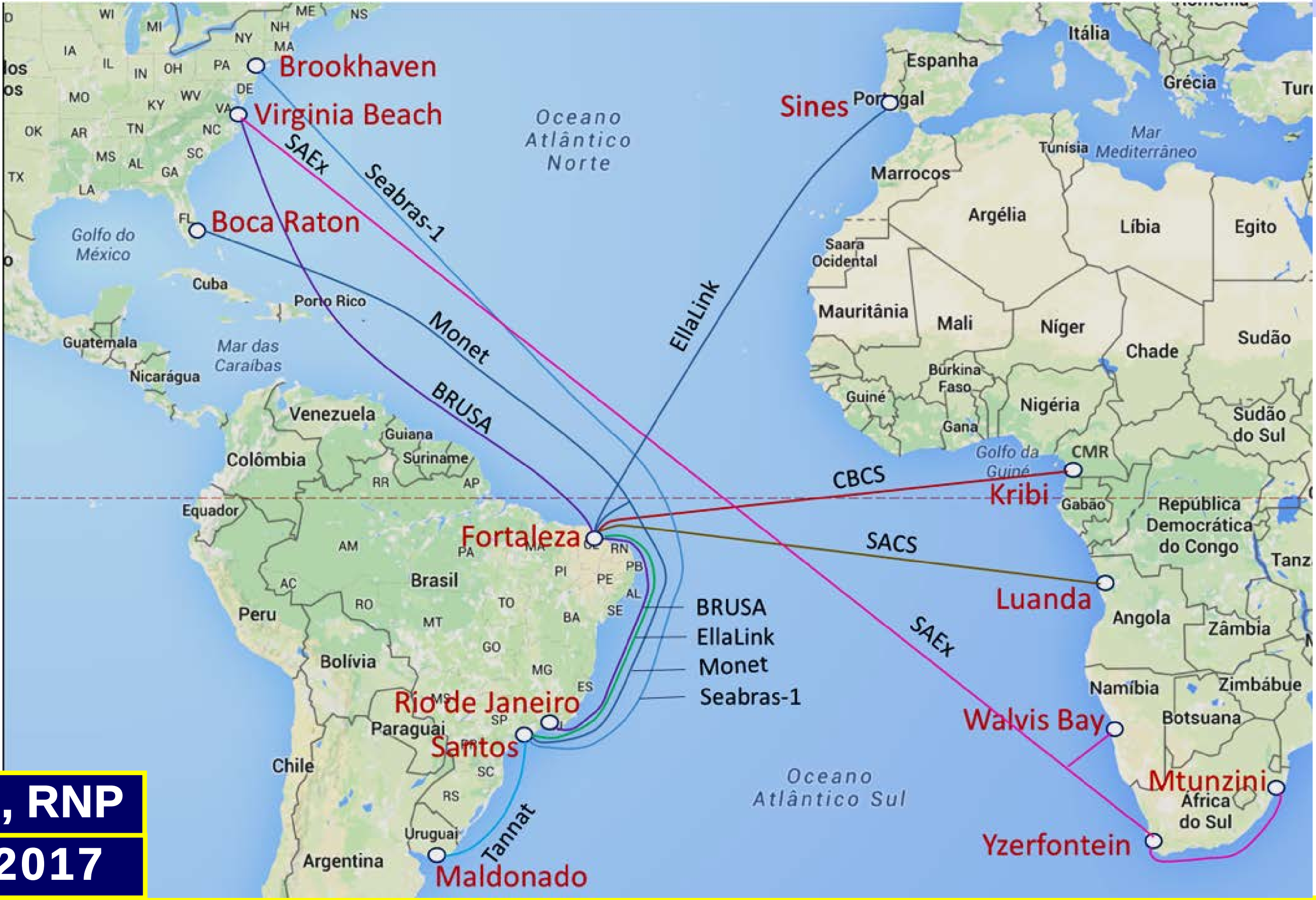
- **AtlanticWave-SDX is responding to the demands of big data scientific instruments through the development of an international software defined exchange point (SDX)**

- **Collaborators: Open exchange point resources at SoX (Atlanta), AMPATH (Miami), and Southern Light (Sao Paulo, Brazil)**



(NSF Award # ACI-1451024 2015-2020)

A New Generation of Cables with 100G Channels to Brazil in 2016-18



M. Stanton, RNP

February 2017

Potential great use for data intensive science programs, including the ALMA (Chile) and SKA (South Africa to the Americas) telescope arrays

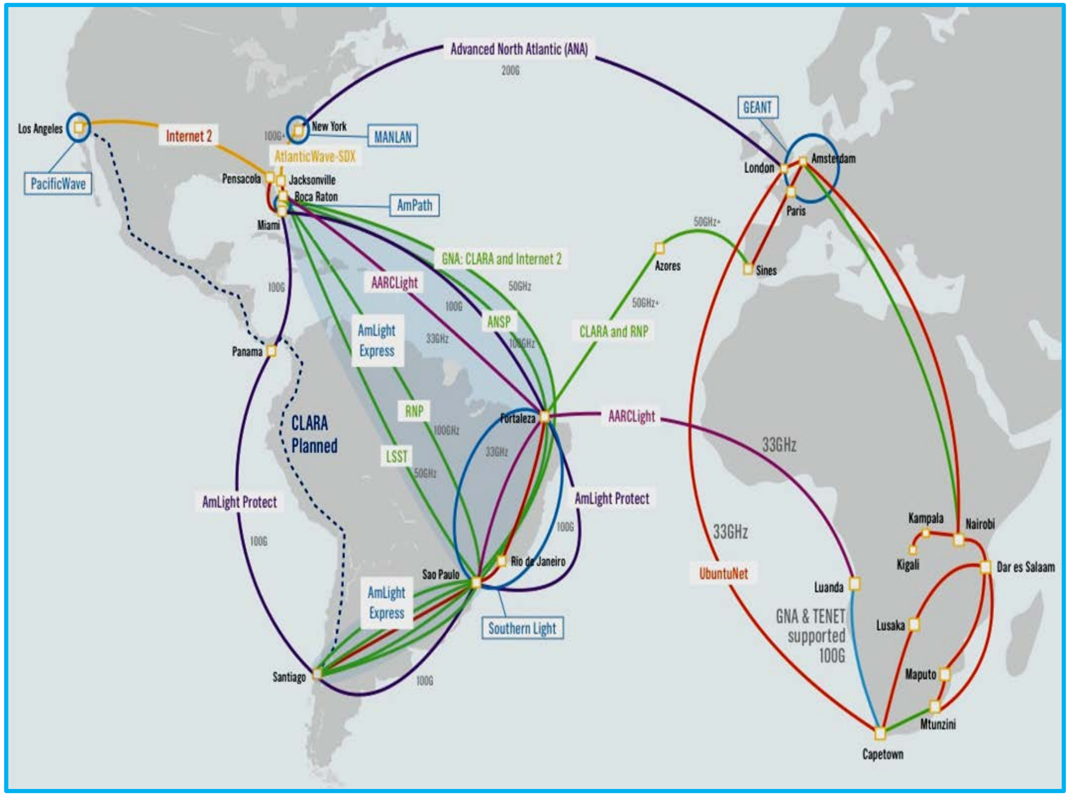
A New Generation of Cables to Brazil with 100G Channels in 2017-18

Cable	Owners	Ready for service	Capacity	Length (km)	Landing points in Brazil	Other countries served
Monet	Google, Antel, Angola Cables, Algar Telecom	2017	64 Tb/s	10,556	Fortaleza (branch) Santos	USA (Boca Ratón, FL)
South Atlantic Cable System (SACS)	Angola Cables	2018	40 Tb/s	6,165	Fortaleza	Angola (Luanda)
Ellalink	Telebras, IslaLink	2019	48 Tb/s	9,501	Fortaleza (branch) Santos	Portugal (Sines)
Tannat	Google, Antel	2018	90 Tb/s	2,000	Santos	Uruguay (Maldonado)
Seabras-1	Seaborn Networks	2017	72 Tb/s	10,500	Fortaleza (branch) Santos	USA (New York)
South Atlantic Interlink (SAIL)	Camtel, China Unicom	2018	32 Tb/s	5,900	Fortaleza	Cameroon (Kribi)
BRUSA	Telefonica	2018		11,000	Fortaleza Rio de Janeiro	USA (Virginia Beach)

M. Stanton, RNP

Americas Africa Research and education Lightpaths (AARCLight)

- **AARCLight** aims to enhance science research and education in the Americas
  - Planning, designing and defining a strategy for high-capacity connectivity
  - Engaging U.S., Brazil, Angola and all African science and engineering research and education communities
  - Serving the broadest communities of interest in research and education

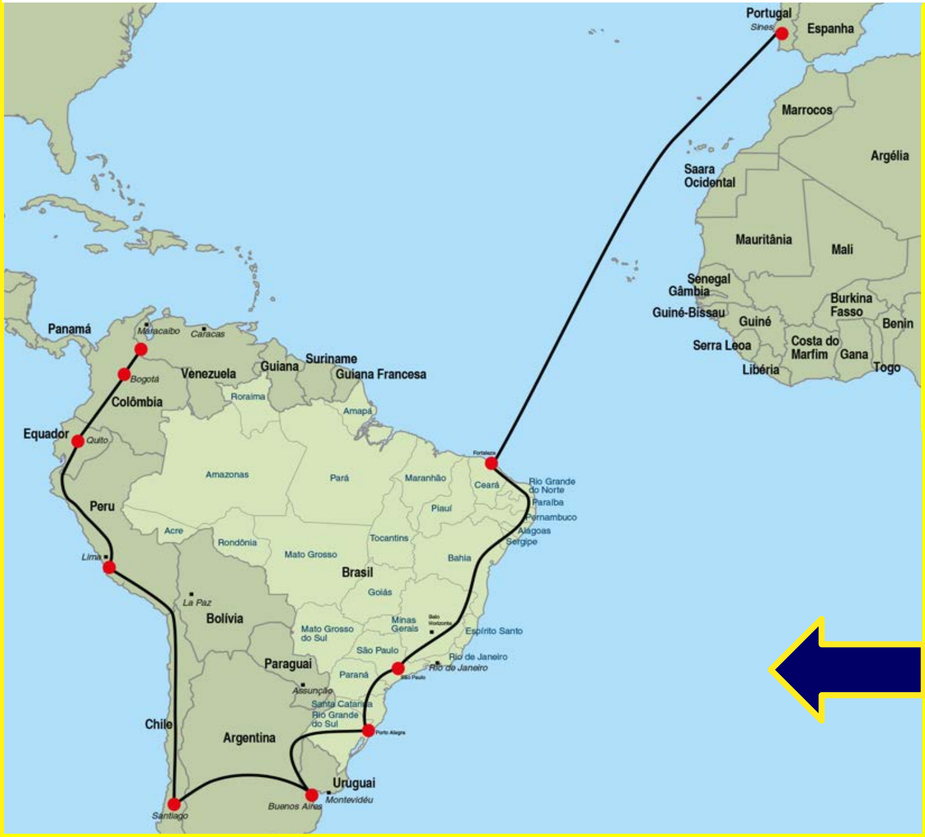


- **Collaborators:**
  - **USA:** FIU, FLR, Internet2
  - **Latin America:** RNP, CLARA, FAPESP
  - **Africa:** Angola Cable, UbuntuNet, and Wasace

BELLA-T: RedCLARA and GEANT Project

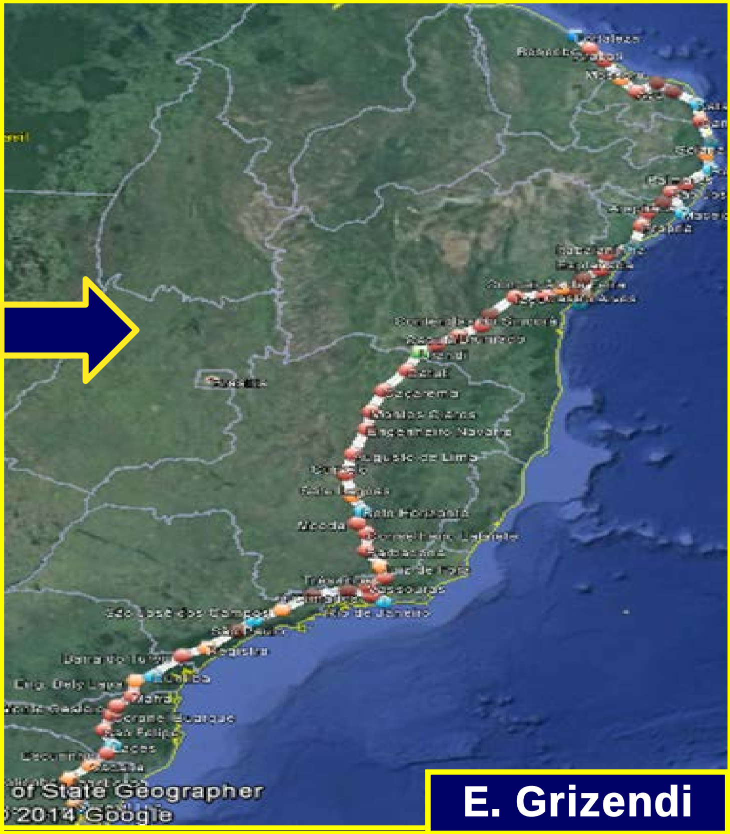
Linking Latin American NRENs to a BR-EU Cable

Country	Brazil	Argentina	Chile	Peru	Ecuador	Colombia	TOTAL
Route length (km)	6223	2500	2000	2594	1330	1803	16450



Planned Access Network to South America

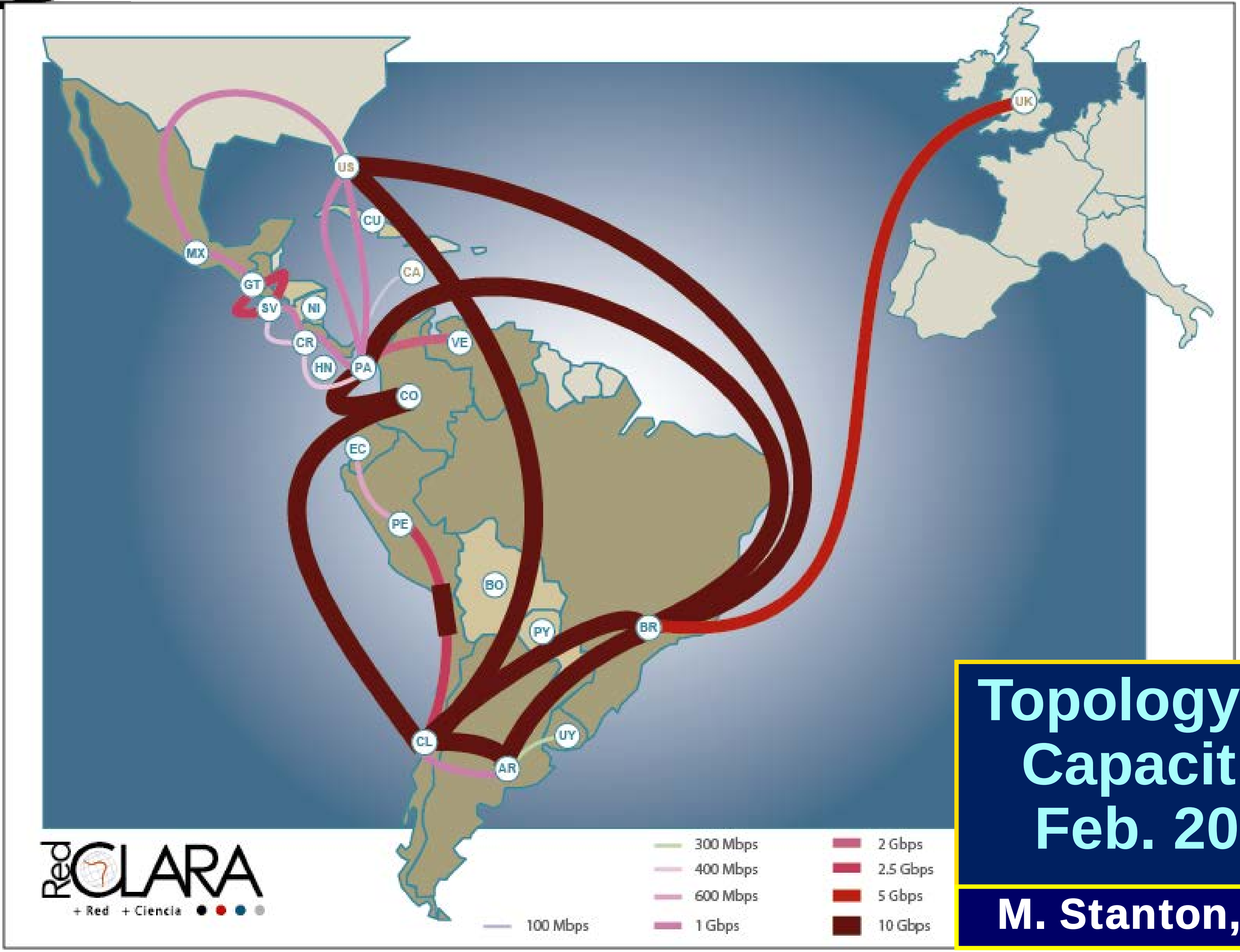
Projected Landing points in Brazil



E. Grizendi

EU and S. American NRENs Plan to Acquire ~45 100G Lambdas on the submarine cable

RedCLARA: Interconnecting Latin American NRENs

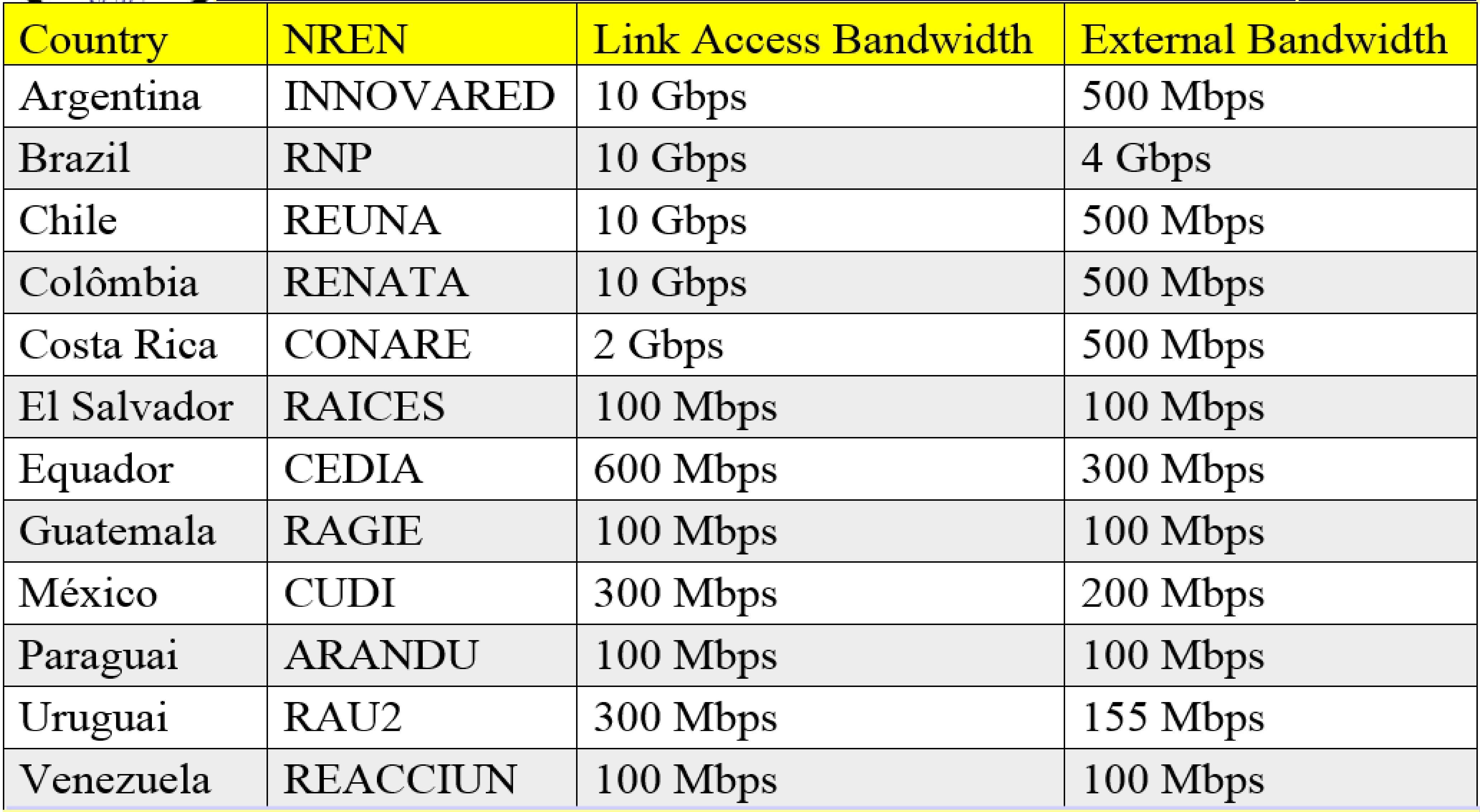


Topology and Capacities Feb. 2017

M. Stanton, RNP

RedCLARA: Extra-regional Connectivity to Participating Latin American Networks

Marco Teixeira (RNP)

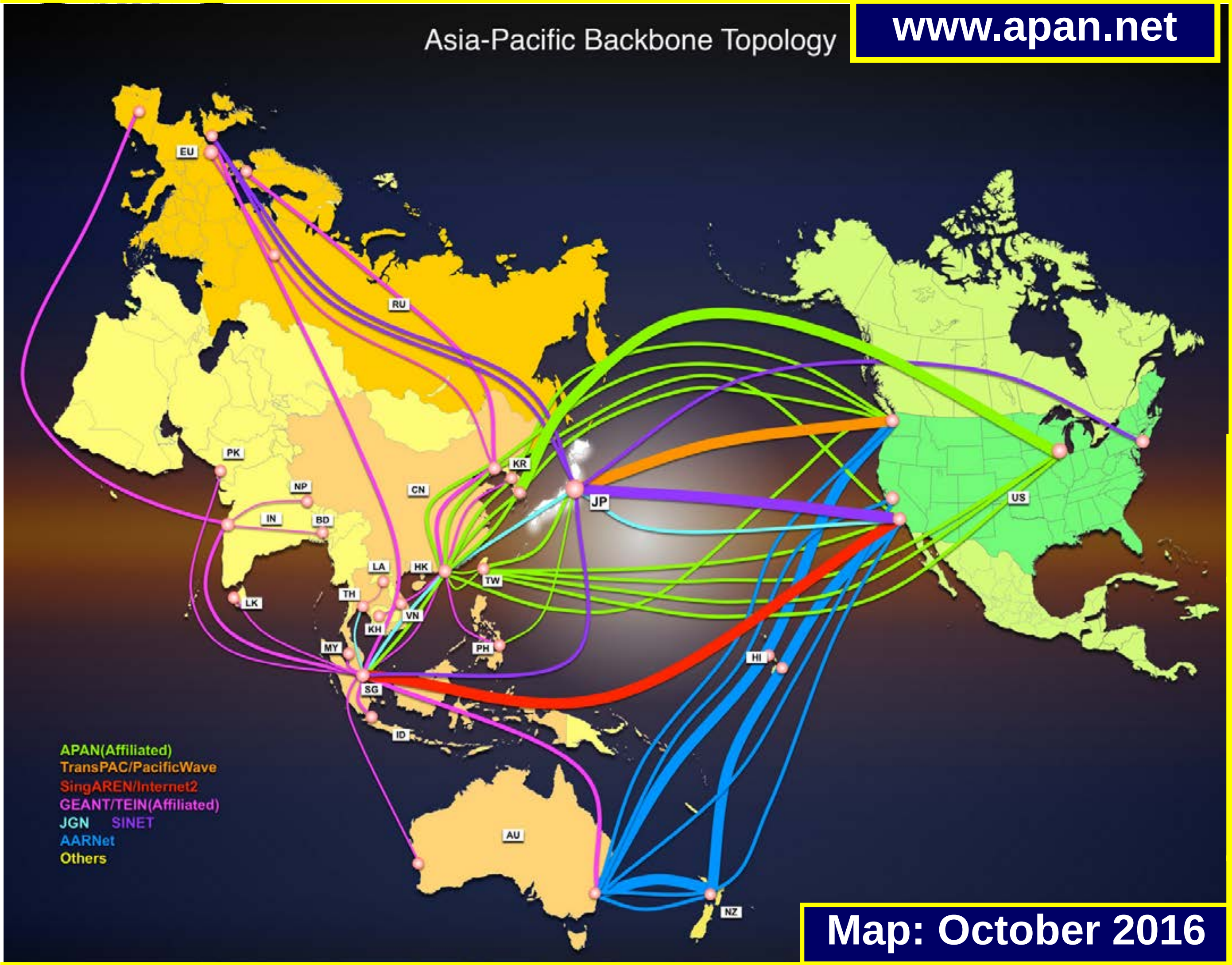


Country	NREN	Link Access Bandwidth	External Bandwidth
Argentina	INNOVARED	10 Gbps	500 Mbps
Brazil	RNP	10 Gbps	4 Gbps
Chile	REUNA	10 Gbps	500 Mbps
Colômbia	RENATA	10 Gbps	500 Mbps
Costa Rica	CONARE	2 Gbps	500 Mbps
El Salvador	RAICES	100 Mbps	100 Mbps
Equador	CEDIA	600 Mbps	300 Mbps
Guatemala	RAGIE	100 Mbps	100 Mbps
México	CUDI	300 Mbps	200 Mbps
Paraguai	ARANDU	100 Mbps	100 Mbps
Uruguai	RAU2	300 Mbps	155 Mbps
Venezuela	REACCIUN	100 Mbps	100 Mbps

RedCLARA: Low External Bandwidth Issue

Asia Pacific Advanced Network (APAN)

Global Partnership of R&E Networks and Advanced R&D Projects



www.apan.net

Map: October 2016

Some NRENs focus on the high end, others on breadth of access first

Working Groups
Cloud WG exploring possible federated Cloud development and access

Internet of Things WG

6 X 100G + 16 X 10G TransPacific Links

More developed NRENs also have moved some domestic links from 10G to 100G; some at N*10G or 40G

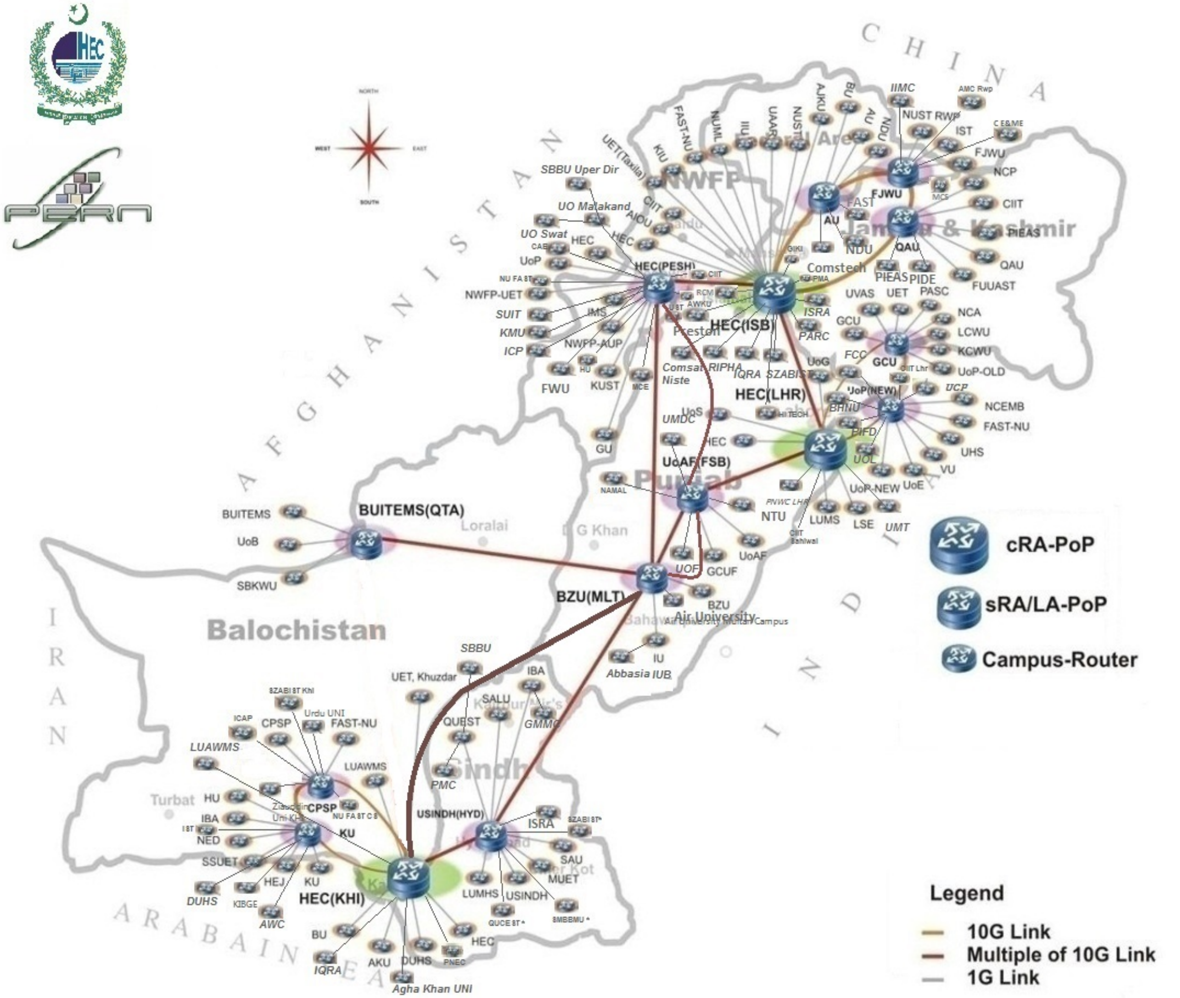
Contrast: some intra-regional NREN Links still in the 1G range or less

~50 Transpacific + Regional Links: < 1G to 100G

NREN Network Connectivity within APAN

	Domestic	International
Australia	n * 100G + 10G	2x2.5G to Asia, 2x40G (R&E) to North America
Bangladesh	1 - 10G	45M
Afghanistan		155M to EU, 155M planned to Tein4
China	Multiple 10G	Multiple 1G and 10G links
Hong Kong	1 - 10G	Multiple 155M - 10G
India	1G - 10G	2.5G
Japan	Multiple <1G - 10G	1.5M (satellite) to multiple 10G
Korea	Multiple 10G	Multiple 10G, 100G to US
Sri Lanka	1M - 500M	45M → 1G
Malaysia	1G	100M - 622M
Nepal		45M
New Zealand	1G – 10G	1G → 40G
Philippines	1G – 10G	Multiple 155M – 1G
Pakistan	1G – Multiple 10G	1G to TEIN4
Singapore	1G - 10G	Multiple 155M - 10G
Thailand	1G	310M – 1G
Taiwan	10G -> 100G	Multiple 2.5 - 10G
Vietnam	30M – 1G	622M

Pakistan Educational Research Network PERN



10GE Metro Ring

23 PoPs sites
connecting
208 Universities/
Institutes in 49 cities

853km Metro Fiber

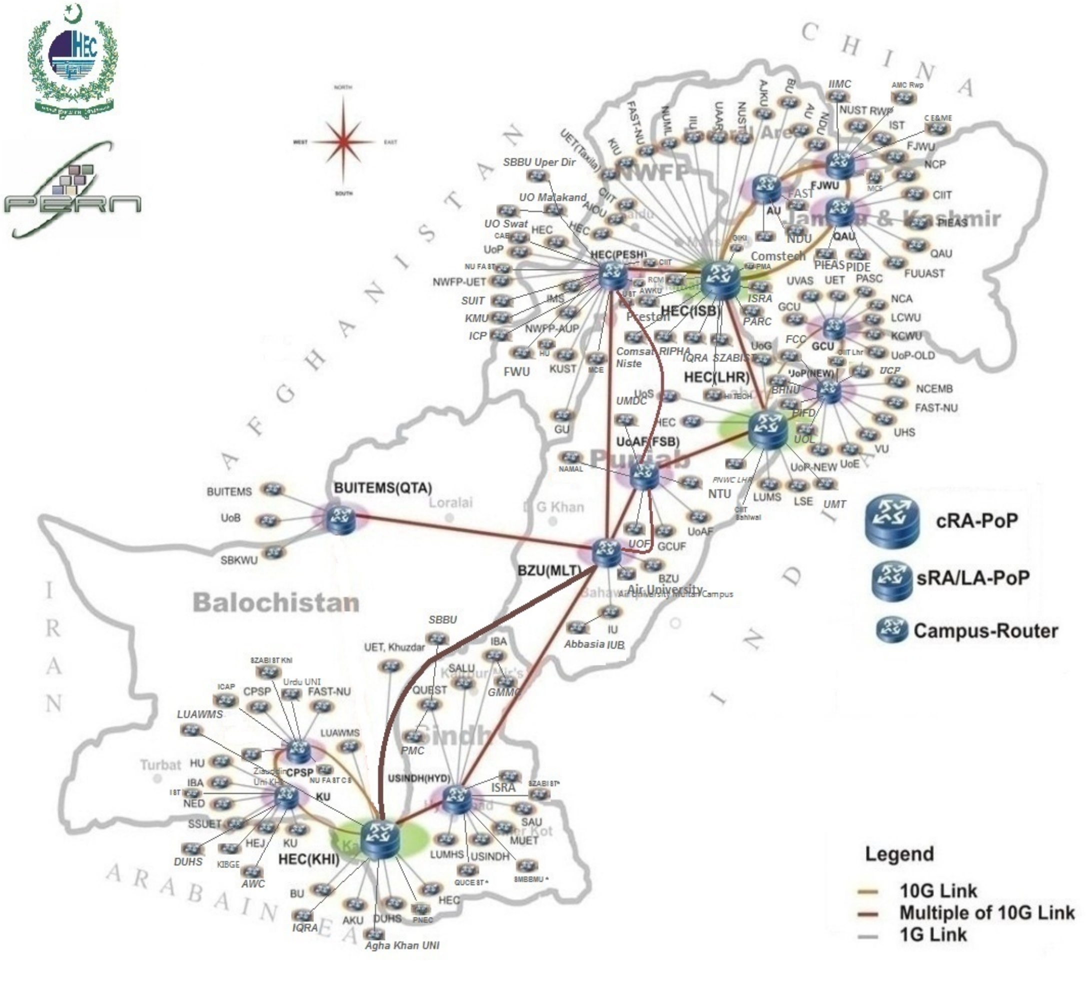
17 Gbps *aggregate*
Internet bandwidth

1GE Int'l R&E link
through TEIN4 over
the Transworld1 cable

8000 Km Dark Fiber Among 8 Cities of the country

Expansion Plans: PERN3

Ready to transition to a 100GE Core by 2017



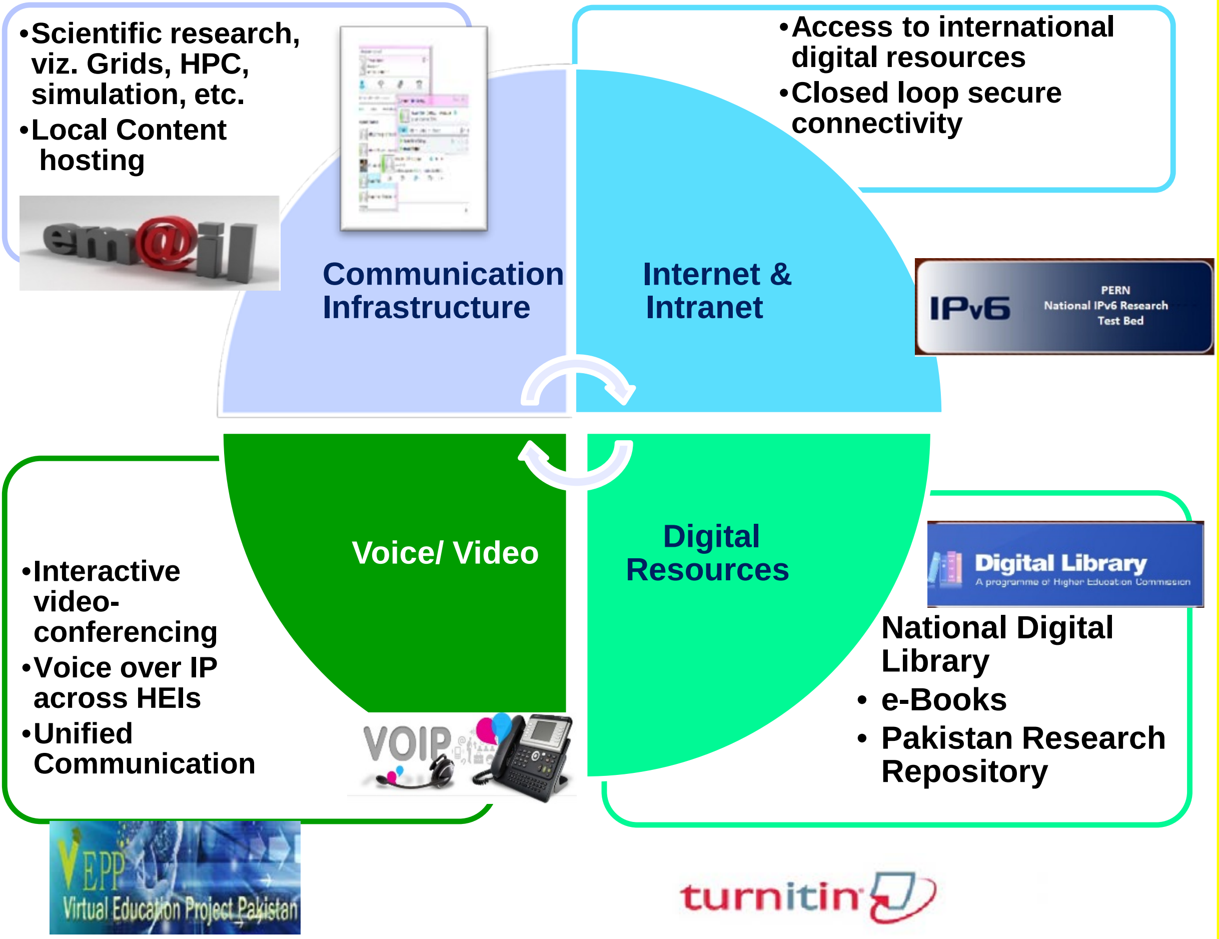
- ❒ Plan to transition to 100GE Core interlinking 6 major cities
  - ❒ Deployed in parallel to existing 10GE core
  - ❒ 10GE last mile access from the sites
- ❒ Establishment of PERN connectivity in 15 more cities
- ❒ Bifurcation of the PERN fiber rings to create more rings
  - ❒ To achieve 100% up time and resiliency of the network

Projects Under PERN

- **PERN to Eduroam network:** Eduroam (education roaming) is the secure, world-wide roaming access service developed for the international research and education community. PERN has deployed Eduroam in 15 Universities and planned to rollout on 100 Universities in 2017
- **IPv6 Implementation:** *As per APNIC Report , PERN is the Second largest IPv6 deployer in the country. It was started through* establishment of a research testbed for IPv6 among 12 higher education institutes (HEIs) using an existing infrastructure, and connecting to an international IPv6 backbone. IPv6 is now being extended to 26 Sites.
- **Telemedicine:** Highest quality content delivery for telemedicine sessions, and strengthening of the mutually beneficial relationship among doctors and medical students.
- **Smart Universities (WI-FI):** Blanket WiFi coverage across the campus to provide/extend wireless services while augmenting a highly conducive, technologically advanced, and cost effective learning environment at the HEIs of Pakistan.
- **IP Surveillance:** Under the Smart Universities project, a Safe Campus project has been initiated to provide HD cameras and intelligent video analysis technologies. This will be implemented with monitoring equipment at the campus main entry/exit, perimeter and building entries/exits etc.

PERN Service Area

Allied, Focused and Multidomain Services



- Scientific research, viz. Grids, HPC, simulation, etc.
- Local Content hosting

Communication Infrastructure

- Access to international digital resources
- Closed loop secure connectivity

Internet & Intranet

Voice/ Video

- Interactive video-conferencing
- Voice over IP across HEIs
- Unified Communication

Digital Resources

National Digital Library
- e-Books
- Pakistan Research Repository

em@il
IPv6
PERN
National IPv6 Research
Test Bed
Digital Library
A programme of Higher Education Commission
VOIP
VEPP
Virtual Education Project Pakistan
turnitin

- ❒ **Deploying videoconferencing services** at selected TEI (colleges and educational organizations) under a World Bank program
- ❒ **The National Digital Library accessible via PERN2 has launched an ebrary** & McGraw Hill Collections providing around 50,000 online books
- ❒ **Local Content Hosting through national data centers** is provided in three major cities
- ❒ **An IP video recording facility** for surveillance has been deployed

The ICFA-SCIC Network Monitoring WG Further Results and Studies

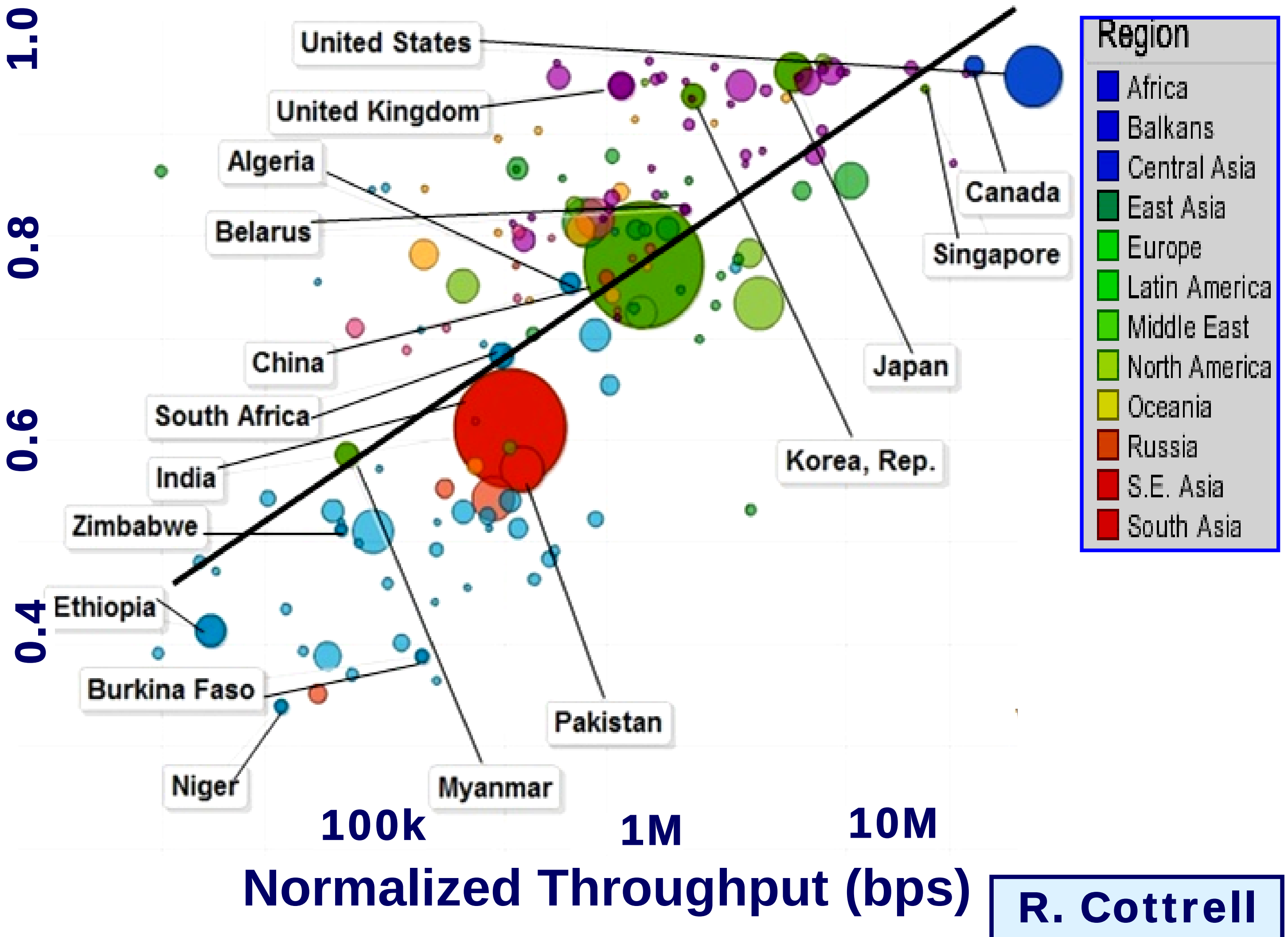
Shawn McKee/UM, Les Cottrell/SLAC,
Marian Babik/CERN, Ilija Vukotic/U Chicago
Brian Tierney/LBNL, Soichi Hayashi/IU,
Mike O’Connor/ESnet

IEPM
Internet End-to-end Performance Monitoring

TCP Throughput in 2015
vs. UN Human Development Index (HDI)

UNDP HDI:

- A long and healthy life, as measured by life expectancy at birth
- Knowledge as measured by the adult literacy rate (with 2/3 weight) and the combined primary, secondary and tertiary growth enrollment ratio (with 1/3 weight)
- A decent standard of living, measured by GDP per capita



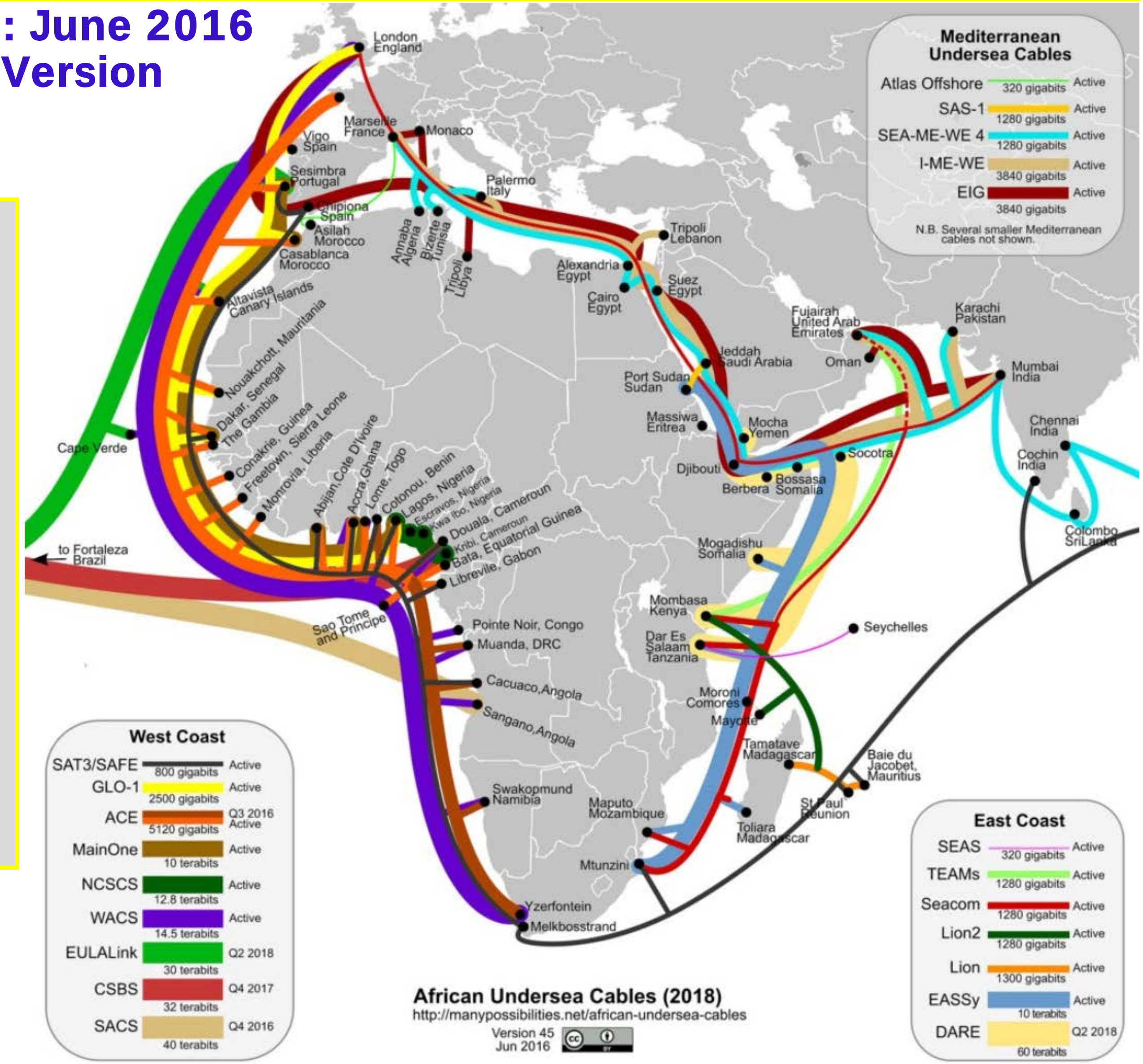
R. Cottrell

Clear Correlation Between the UNDP HDI and the Throughput

Throughput in Africa by Region

Map: June 2016 Version

- East & West Africa saw big improvements in 2010, following the World Cup
- East Africa growth rate slowed down
- West Africa now better than East
- Due to more cable capacity on the West



https://manypossibilities.net/african-undersea-cables/

New East African Undersea Cable

- Liquid Telecom (liquidtelecom.com) started the Liquid Sea project, for a new 10,000 km cable from South Africa to Middle East, and onward to Europe
  - Fully funded
  - 2 years to complete: by 2018
  - Up to ten times the capacity (20-30 Tbps) of existing undersea cables in the regionn
  - Adds new landing stations
  - Leverages extensive terrestrial fibre network

http://www.huffingtonpost.com/david-tereshchuk/a-giant-leap-in-2016-africa_b_8901556.html

How to Reach the Rest of the World
Geostationary Satellites (GEOS)

R. Cottrell

ViaSat: High bandwidth geosynchronous satellites

- Long delays (~0.5sec) avoided by aggregating multiple request/response for web objects in a page
  - Not good for real time
- Focus:
  - Aviation (Jet Blue & United), military, business, in the Americas, Europe, E. Asia
- 2016 launch ViaSat-2 250-300 Gbps
- 2020-2021 ViaSat-3 (3 satellites) in the Terabit range
  See http://investors.viasat.com/releasedetail.cfm?ReleaseID=954123

How to Reach the Rest of the World 3

Google plans on sending up 300 balloons Around the World at the 40th South Parallel

R. Cottrell

- Google balloons are active: early adopters Sri Lanka, Indonesia
- Stay aloft at 12 miles for up to 150 days
- Sept. 2016: Trial over Peru *steered by AI* https://www.technologyreview.com/s/602457/ai-is-taking-control-of-project-loon/
- Google hopes to eventually have thousands of balloons aloft

Google
BALLOON-POWERED INTERNET FOR EVERYONE

Towards a Next Generation Network-Integrated System for HEP and Other Data Intensive Science Programs

Additional Slides

NGenIA: Addressing a New Era of Challenges as we Move to Exascale Data and Computing

- The largest science datasets under management today, from the LHC program, are ~500 petabytes (PB)
  - Exabyte datasets are on the horizon, by the end of Run2 in 2018
  - 850 PB flowed Across the WLCG, 350 PB over Esnet in last 12 months
  - Data volumes could grow by to the ~50-100 Exabyte range, during the HL LHC era from 2026
- Reliance on high performance networks will continue to grow as many Exabytes are distributed, processed & analyzed at 100s of sites
- As needs of other fields continue to grow, HEP will face stiff competition for use of limited network resources.

1 EB = 2 milligrams of DNA

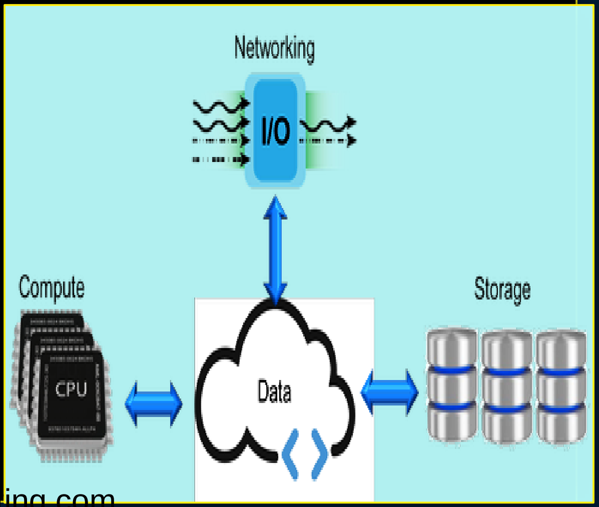
Earth Observation

Next Generation Integrated Architectures for HEP and Exascale Science

- **METHOD: Construct autonomous network-resident services that dynamically interact with site-resident services, and with the experiments' principal data distribution and management tools**
- **To coordinate use of network, storage + compute resources, using:**

1. **Smart middleware to interface to SDN-orchestrated data flows over network paths with allocated bandwidth levels all the way to a set of high performance end-host data transfer nodes (DTNs),**
2. **Protocol agnostic traffic shaping services at the site edges and the network core, coupled to high throughput data transfer applications that provide stable, predictable data transfer rates**
3. **Machine learning + system modeling and Pervasive end-to-end monitoring**

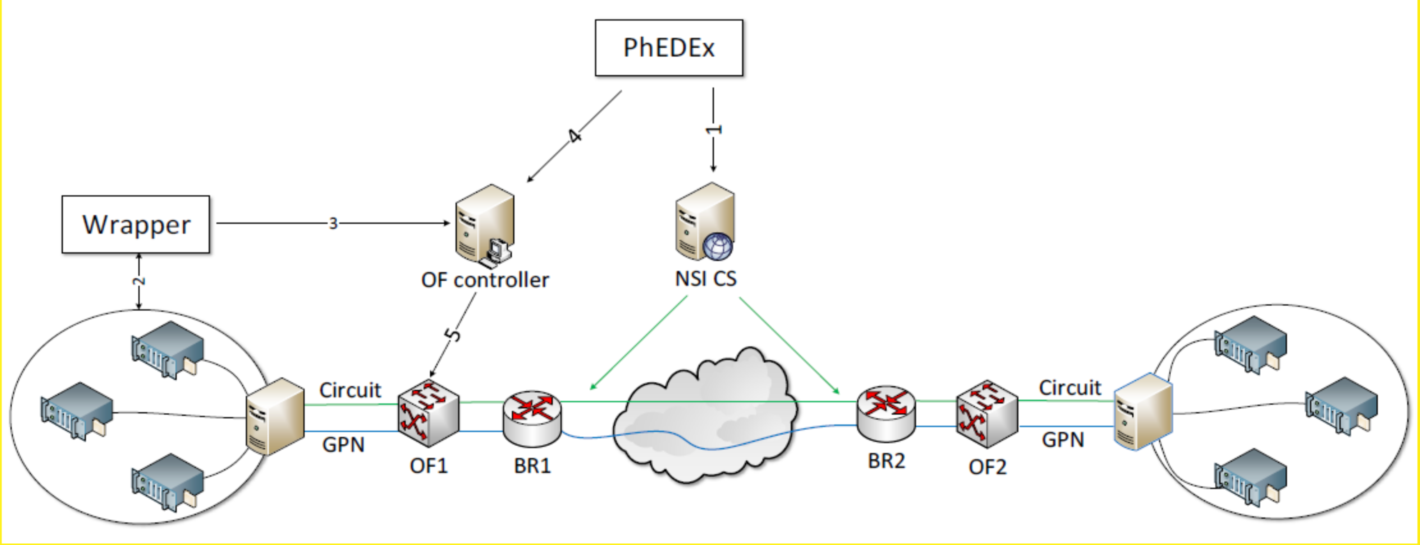
   - **To track, diagnose and optimize system operations on the fly**



networkcomputing.com

Prerequisites: Dynamic Circuits

- ❑ The team’s earlier work, in the DYNES and ANSE NSF projects with dynamic circuits, integrated with the CMS PhEDEx and ASO applications used a so-called “FDTAgent” to couple the data transfer nodes (DTNs) at the end-sites running Caltech’s FDT as the high throughput data transfer application
- ❑ The agent (1) requests the circuit, (2) waits for an answer, (3) configures both end-hosts if the circuit provisioning succeeds, and (4) modifies the local end-host routing including creating VLAN interfaces to use the new circuit.



Generalized to: Multicircuit, multisite, SDN driven systems:

- ✸ In LHCONE
- ✸ For LSST in the future

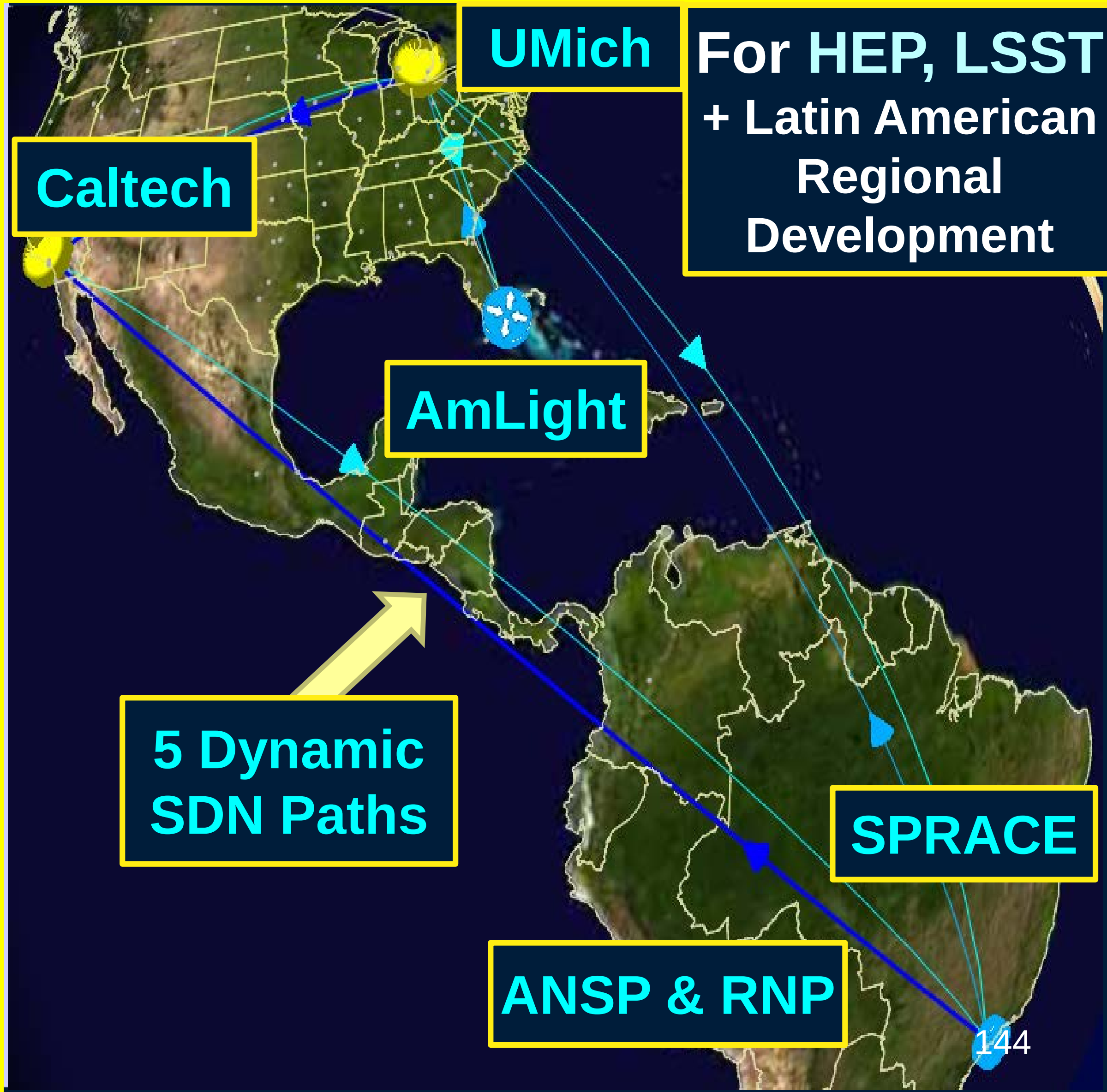
SDN Demonstration at the FTW Workshop. Partners:
Caltech, Amlight/FIU, ESnet, Internet2, Michigan, Sao Paolo

5 Dynamic Path creation:
Caltech – Umich
Caltech/Sprace
Umich/Sprace
Caltech – RNP
UMich – AmLight

Path initiation by the DYNES FDT Agent using OSCARS API Calls

OESS for OpenFlow data plane provisioning over Internet2/AL2S

MonALISA agents at the end-sites provide detailed monitoring information

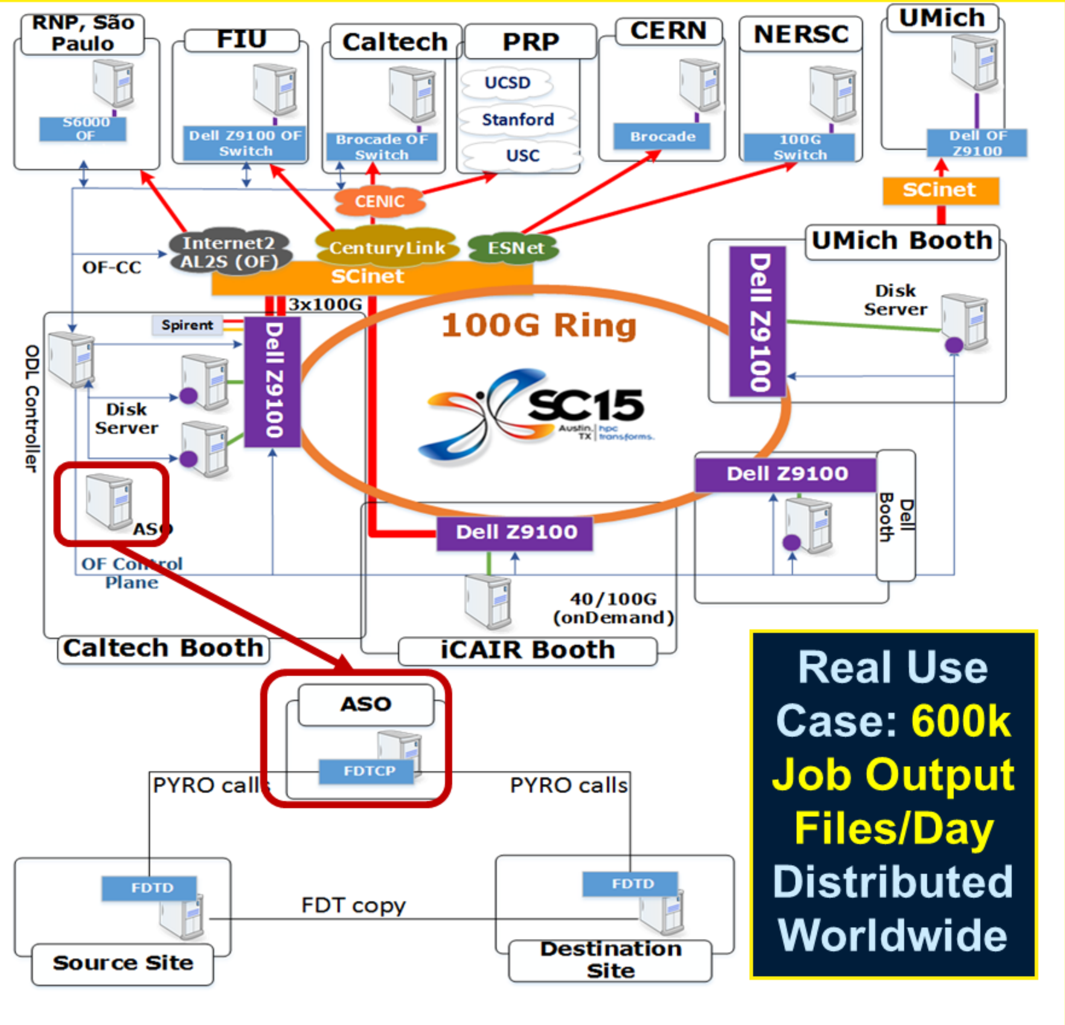


For HEP, LSST + Latin American Regional Development

CMS at SC15: Asynchronous Stage Out 3rd Party Copy Demonstration

SDN-driven Large Flow Steering, Load balancing, Site orchestration Over Terabit/sec Global Networks

- ❒ ASO: Stageout of out files from CMS Analysis Jobs
  - ❒ Groups multiple transfers per link; controls number of parallel transfers
- ❒ Tests among: Caltech, UMich, Dell booths and outside: FIU, Caltech, CERN, UMich
- ❒ PetaByte transfers from multiple sources to multiple destinations



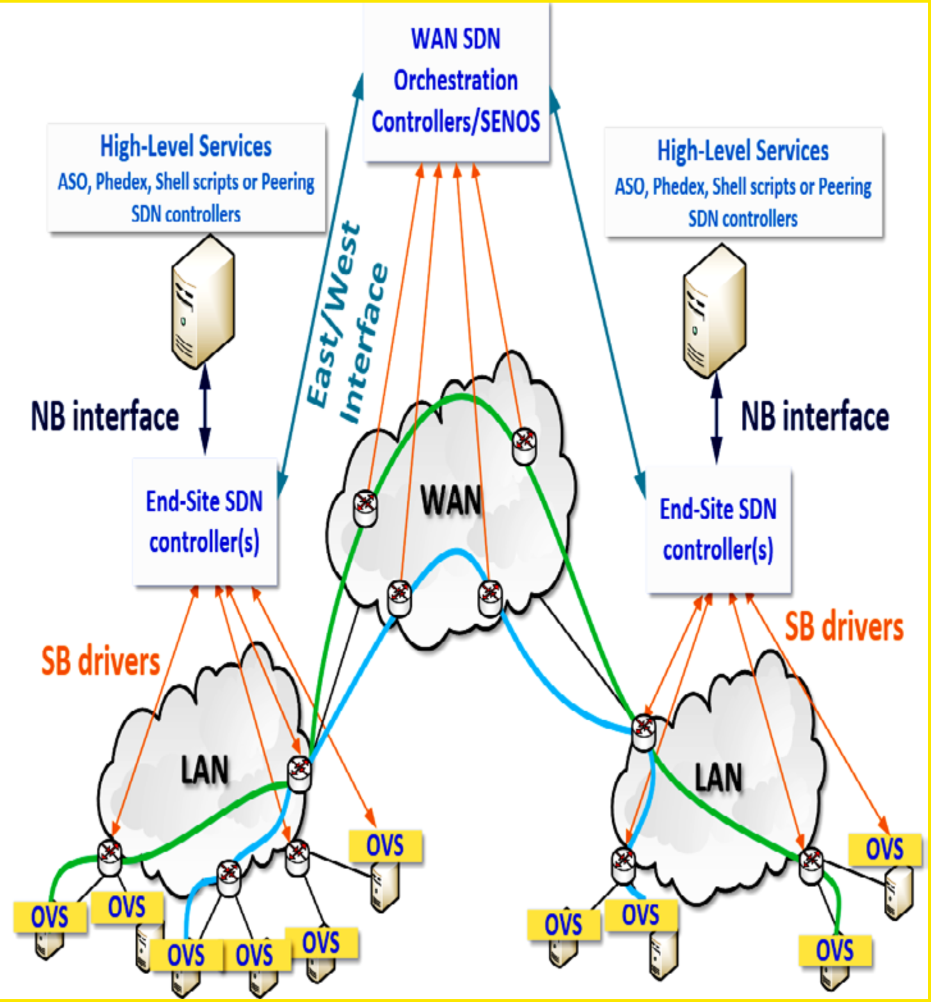
Real Use Case: 600k Job Output Files/Day Distributed Worldwide

Partners: UMich, StarLight, PRP, UNESP, Vanderbilt, NERSC/LBL, Stanford, CERN; ESnet, Internet2, CENIC, MiLR, AmLight, RNP, ANSP

End- and InterSite Orchestration with OVS

Among Multiple Host Groups with Different Paths & Policies

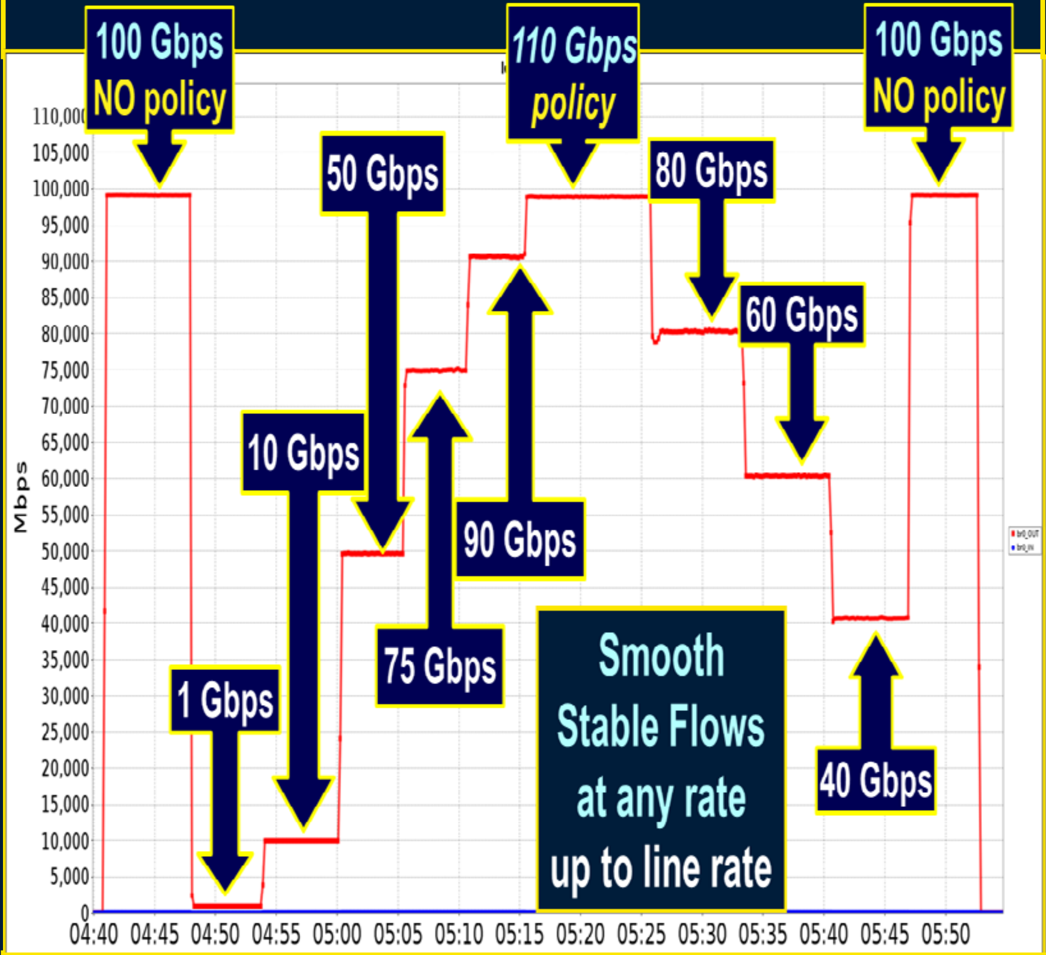
- ❑ Automatic discovery of end hosts in a priority dataset transfer: SDN controlling infrastructure becomes a distributed Lookup Service
- ❑ Automatic identification of data flows between pairs of hosts (IPs) which helps with flow steering
- ❑ The high level services/applications manage the OVS instances via "standard" RESTful NB APIs.
- ❑ SB protocols + drivers: handled by the SDN controller
- ❑ Coupled to Strategic Regional, National and Transoceanic workflow services
- ❑ Pervasive monitoring throughout



Northbound Interaction with SDN Controller(s)

OVS Dynamic BW: 100G Rate Limit Tests

RATES



100 Gbps NO policy

1 Gbps

10 Gbps

50 Gbps

75 Gbps

90 Gbps

110 Gbps policy

80 Gbps

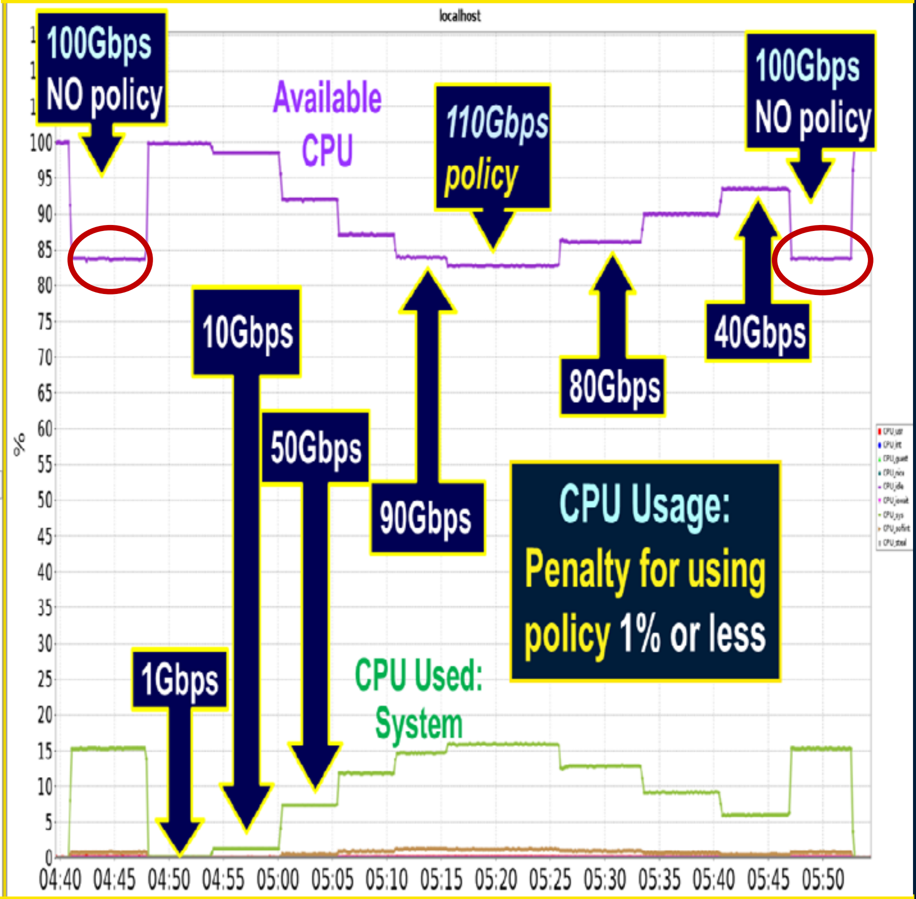
60 Gbps

40 Gbps

100 Gbps NO policy

Smooth Stable Flows at any rate up to line rate

CPU Utilization:
1 Core 16% at full 100G



100Gbps NO policy

Available CPU

1Gbps

10Gbps

50Gbps

90Gbps

110Gbps policy

80Gbps

40Gbps

100Gbps NO policy

CPU Used: System

CPU Usage: Penalty for using policy 1% or less

CPU Usage: Penalty for exerting policy: 1% or less

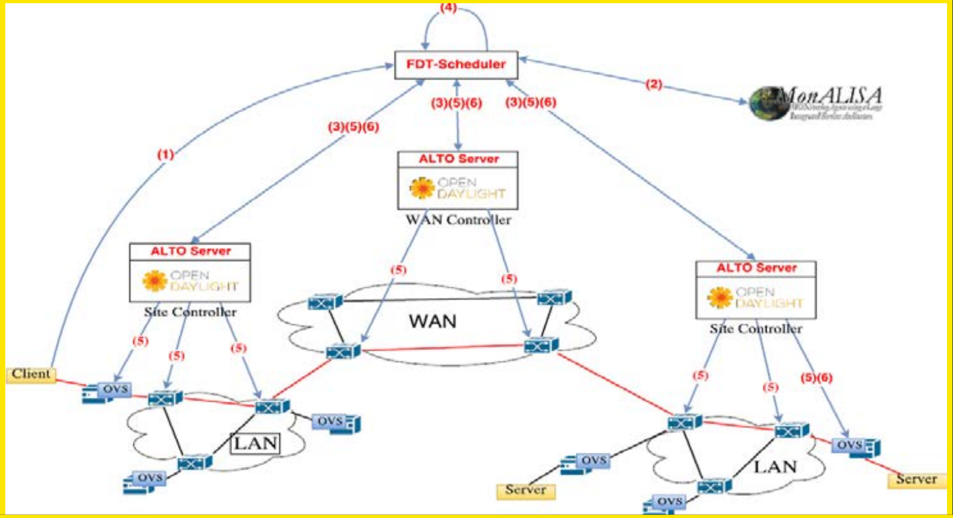
Next Generation "Consistent Operations"

Site-Core Interactions for Efficient, Predictable Workflow

- ❑ Key Components: (1) Open vSwitch (OVS) at edges to stably limit flows, (2) Application Level Traffic Optimization (ALTO) in Open Daylight for end-to-end optimal path creation, + flow metering and high watermarks set in the core
- ❑ Flow metering in network fed back to OVS edge instances: to ensure smooth progress of end-to-end flows
- ❑ Real-time flow adjustments triggered as below
- ❑ Optimization using "Min-Max Fair Resource Allocation" (MFRA) algorithms on prioritized flows

Demos: Internet2 Global Summit in May
SC16 in November

Consistent Ops with ALTO, OVS and MonALISA FDT Schedulers



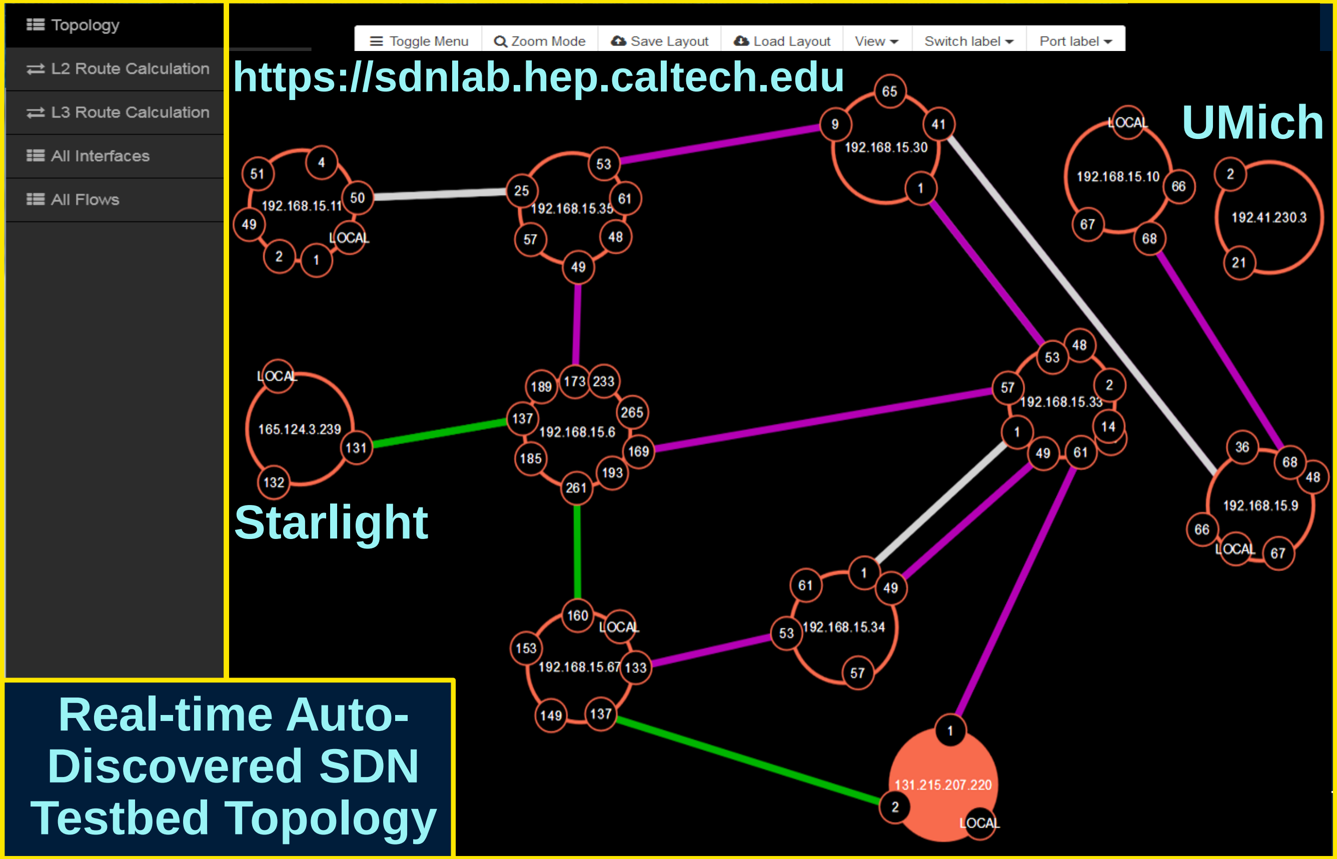
- ❑ Real-time adjustment of allocations triggered by: (1) new requests, (2) real-time feedback on progress of transfers, (3) network state changes or error conditions, (4) proactive load-balancing operations, or (5) rate-limiting operations imposed by controllers or emerging network operating systems (e.g. SENOS)

With Yale CS Team: Y. Yang, Q. Xiang et al.

SDN State of the Art Development Testbed

Caltech, Fermilab, StarLight, Michigan, UNESP; + CERN, Amsterdam, Korea

- ❑ 13+ Openflow switches: Dell, Pica8, Inventec, Brocade, Arista; Huawei
- ❑ Many 40G, N X 40G, 100G Servers: Dell, Supermicro, 2CRSI, Echostreams; and 40G and 100G Network Interfaces: Mellanox, QLogic
- ❑ Caltech Equipment funded through the NSF DYNES, ANSE, CHOPIN projects, and vendor donations

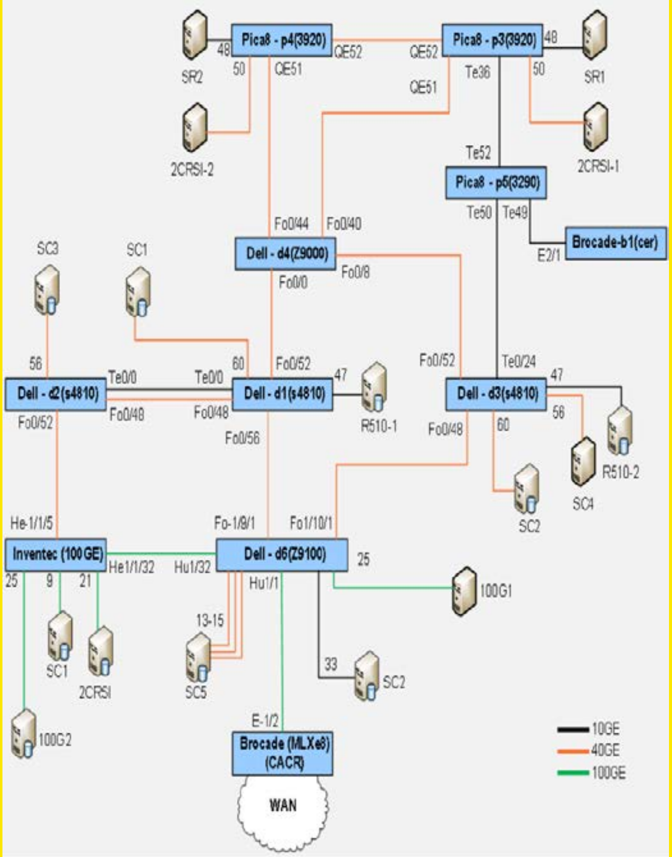


https://sdnlab.hep.caltech.edu

UMich

Starlight

Real-time Auto-Discovered SDN Testbed Topology



NGenIA

New SDN Paradigm

ExaO LHC Orchestrator

Tbps Complex Flows

Machine Learning

LHC Data Traversal

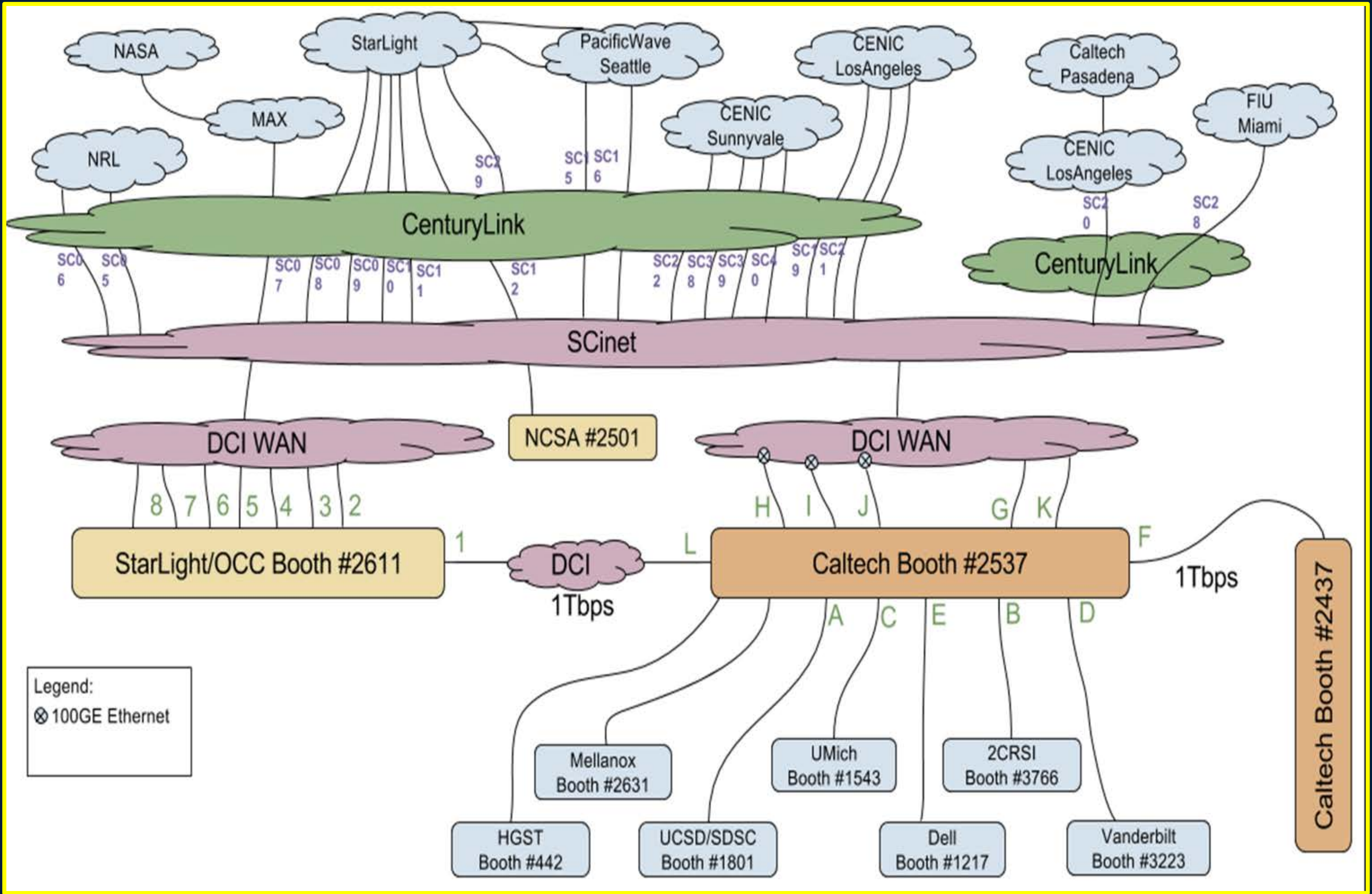
Immersive VR

Thanks to Ecostreams,
Orange Labs Silcon Valley

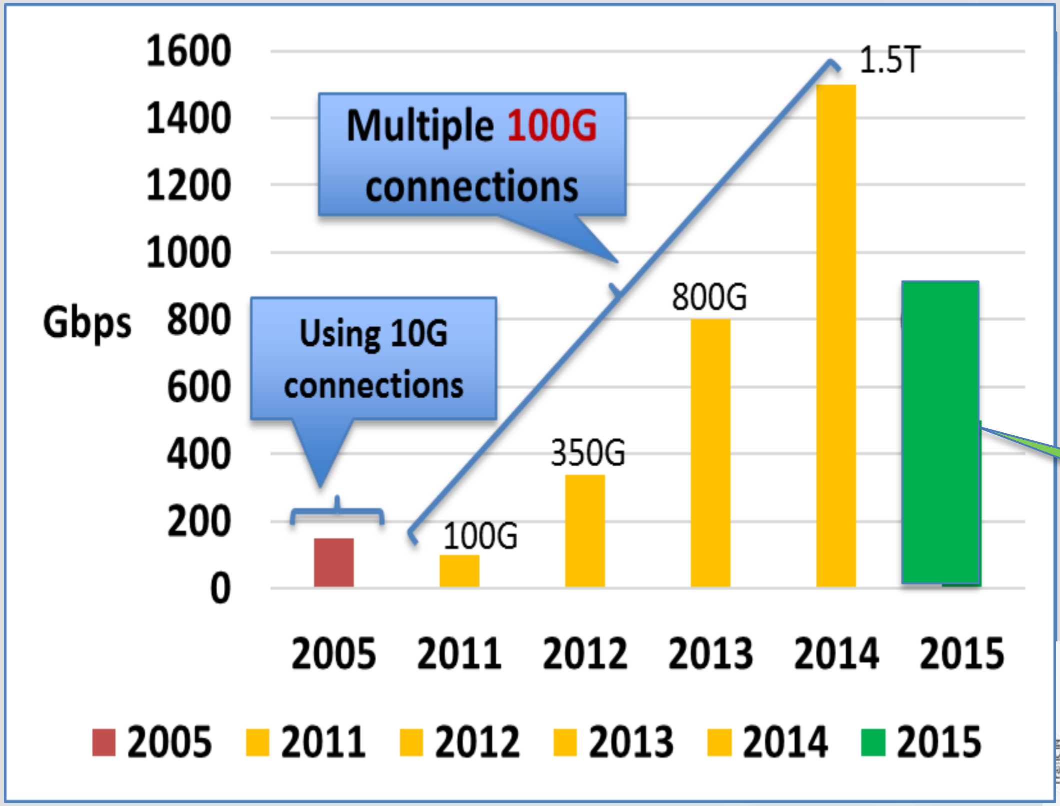
Caltech Booths 2437, 2537

+ the Starlight Booth 2611

SC16: Caltech and StarLight Interbooth and Wide Area Connections

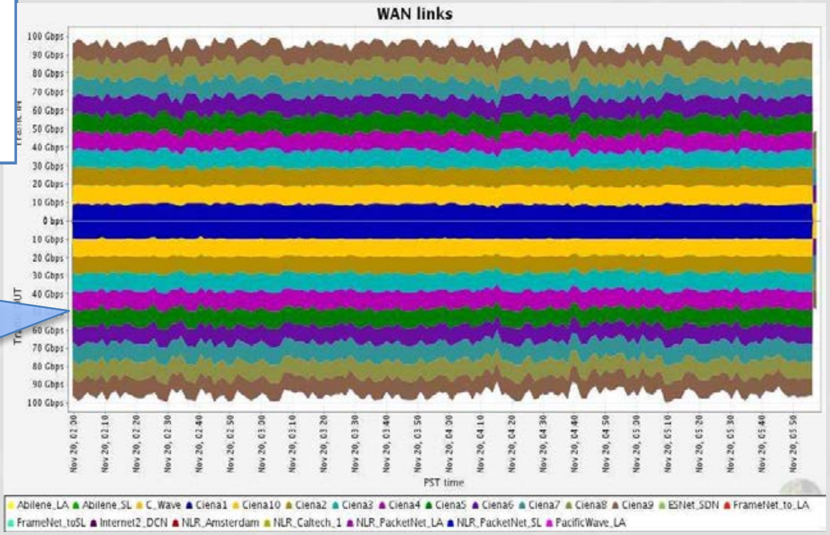


Bandwidth "explosions" by Caltech et al at SC



SC02: FAST
SC05 (Seattle): 155Gbps (15 racks)
SC06: FDT
SC11 (Seattle): 100Gbps
SC12 (Salt Lake): 350Gbps
SC13 (Denver): 800Gbps
SC14 (Louisiana): 1.5Tbps
SC15 (Austin): ~ 750 – 900 Gbps
SC16 (Salt Lake): ~ *2.5Tbps (est.)*

Fully SDN enabled



2008: First ever 100G OTU-4 trials using Ciena laid over multiple 10GE connections on the SC08 floor
191 Gbps bidirectional average:
1 Petabyte in 12 hours

Design options for High Throughput DTN Server

1U SuperMicro Server (Single CPU)

Single 40/100GE NIC

Dual NVME Storage Units (LIQID 3.2TB each)

~90 Gbps disk I/O using NVME over Fabrics

2U SuperMicro Server (Dual CPU)

Single 40/100GE NIC

Three NVME Storage Units (LIQID 3.2TB each)

~100 Gbps disk I/O using FDT/NVME over Fabrics

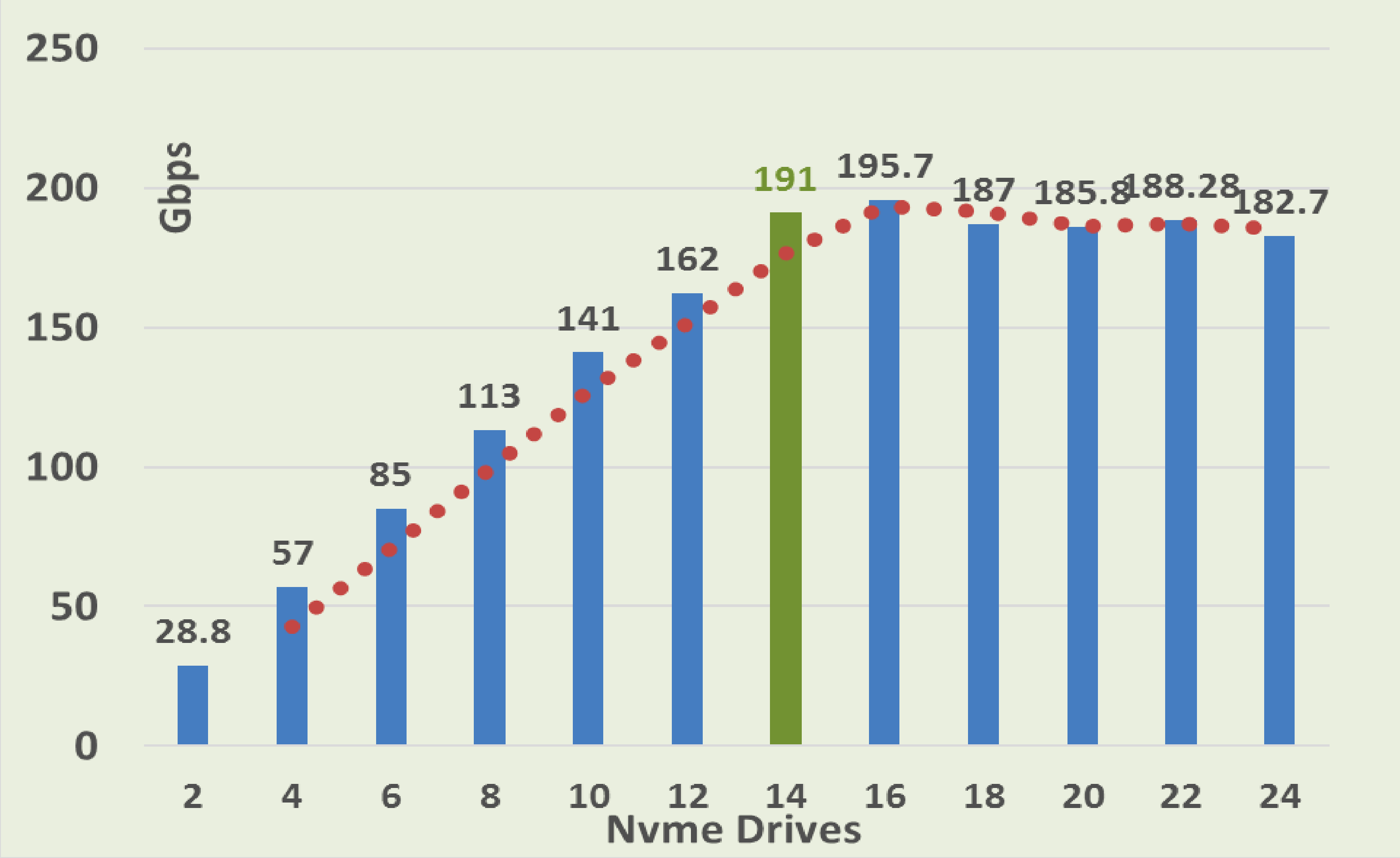
2U SuperMicro (Dual CPU)

Single/Dual 40/100GE NICs

24 NVME front loaded 2.5” drives

~200Gbps of disk I/O using FDT/NVME over Fabrics

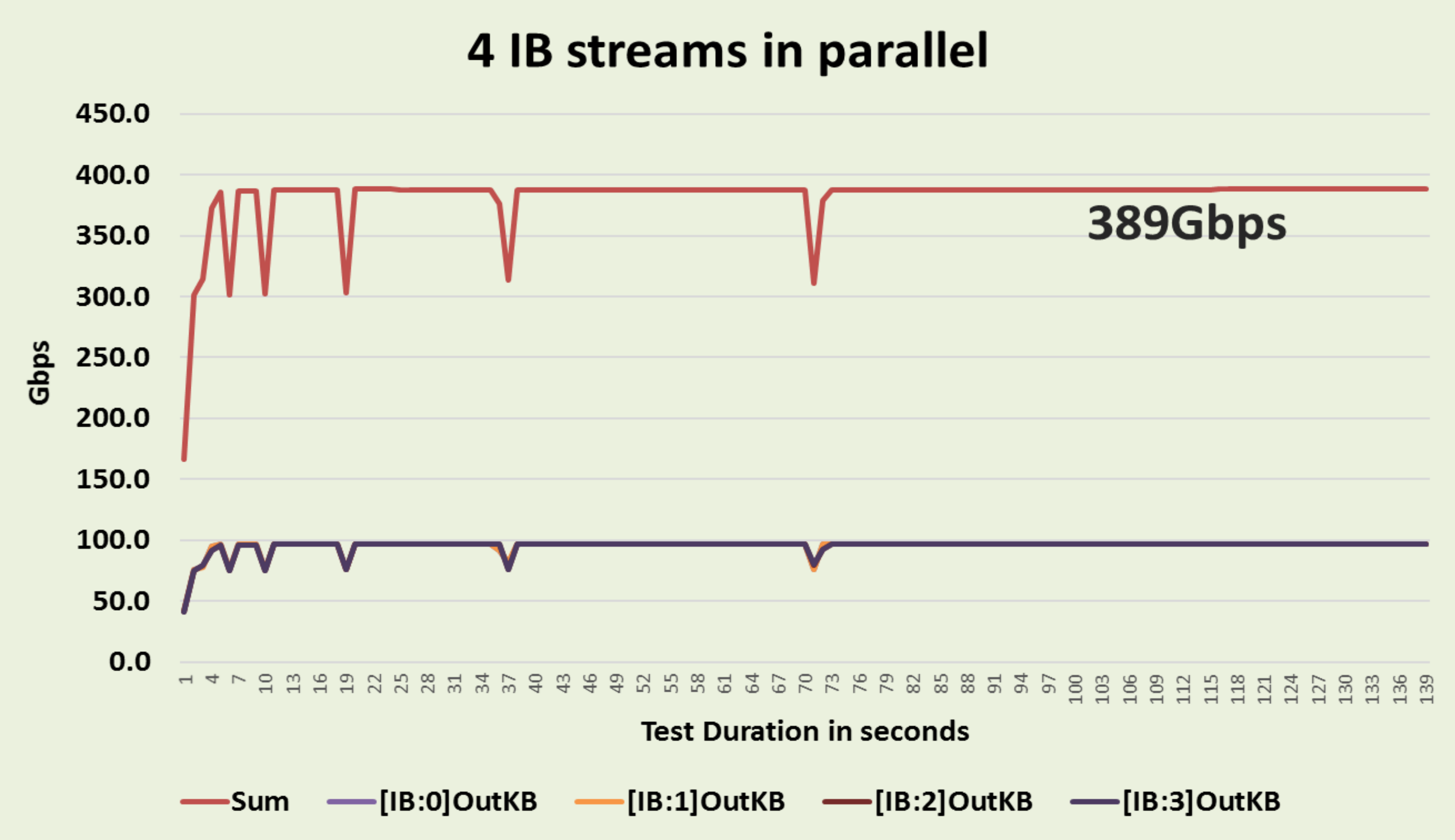
2CRSI + Supermicro Servers with 24 NVMe drives



Max throughput reached at 14 drives (7 drives per processor)

A limitation due to combination of single PCIe x16 bus (128Gbps), processor utilization and application overheads.

400GE Network Testing: Infiniband Reduces Overheads

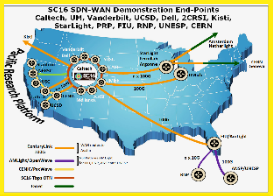


Transmission across 4 Mellnox VPI NICs.

Only 4 CPU cores are used out of 24 cores.

Yale and Caltech at SC16

State of the Art SDN Controller + Framework


SC16 SDN-WAN Demonstration End-Points
Caltech, UM, Vanderbilt, UCSD, Dell, 2CRSI, Kisti, StarLight, PRP, FIU, RNP, UNESP, CERN
Pacific Research Platform
Caltech
StarLight
Fermilab
Argonne
USC
UCSD
n x 100G
100G
n x 10G
RNP
CenturyLink
AmLight/OpenWave
CENIC/PacWave
SC16 Tbps OTN

Driving large load balanced smooth flows over optimally selected paths

See "Traffic Optimization for ExaScale Science Applications", Q. Xiang et al. IETF Internet Draft https://tools.ietf.org/pdf/draft-xiang-alto-exascale-network-optimization-00.pdf

- ❒ We are demonstrating and conducting tutorials at Booths 2437+2537 on our (evolving) state of the art OpenDaylight controller
- ❒ Based on a unified control plane programming framework, and novel components and developments, that include:
  - ❒ The Application Level Traffic Optimization (ALTO) Protocol
  - ❒ A Max-Min fair resource allocation algorithm-set providing flow control and load balancing in the network core
  - ❒ A data-driven function store for high-level, change-oblivious SDN programming
  - ❒ A data-path oblivious high-level programming framework.
- ❒ Smart middleware to interface to SDN-orchestrated data flows over network paths with guaranteed (flow-controlled) bandwidth to a set of DTNs
- ❒ Coupled to protocol agnostic (Open vSwitch-based) traffic shaping services at the site edges
- ❒ Will be used with Machine Learning to identify key variables controlling the system's throughput and stability, and for overall system optimization

Yale and Caltech at SC16: State of the Art SuperSDN Framework + Controller

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

New SDN Framework and Tools : Yale Team

Powerful state of the art, generic tools to substantially simplify SDN programming

Before (manual programming)

•Complex, manual maven programming

Web IDE

- Web-based automatic generation of projects
- Programmer focuses only on key aspects

Before (low level programming)

•Low-level, complex OpenFlow rule programming
•Programmer can define only at flow level
•Specific access control allowing only hosts partition

Maple programming (high-level programming)

- High-level, completely south-bound agnostic, cross-layer programming
- Programmer sees (logically) each and every packet
- Integrated access control supporting per-user or role based programming

Before (raw data store)

- Complex, manual tracking of execution dependency
- Manual cleanup, re-execute
- Designed directly on raw data store

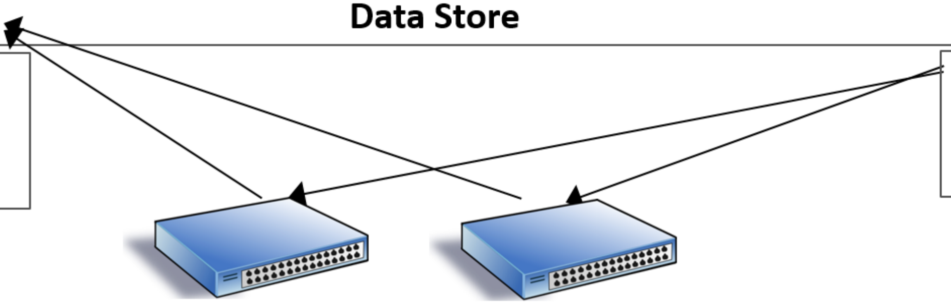
FAST (automated function store)

- Automatic execution dependency tracking
- Automatic cleanup, re-execution (intent ++)
- Can host generic network functions

Data Store

Before

- Ad hoc flow rule installation

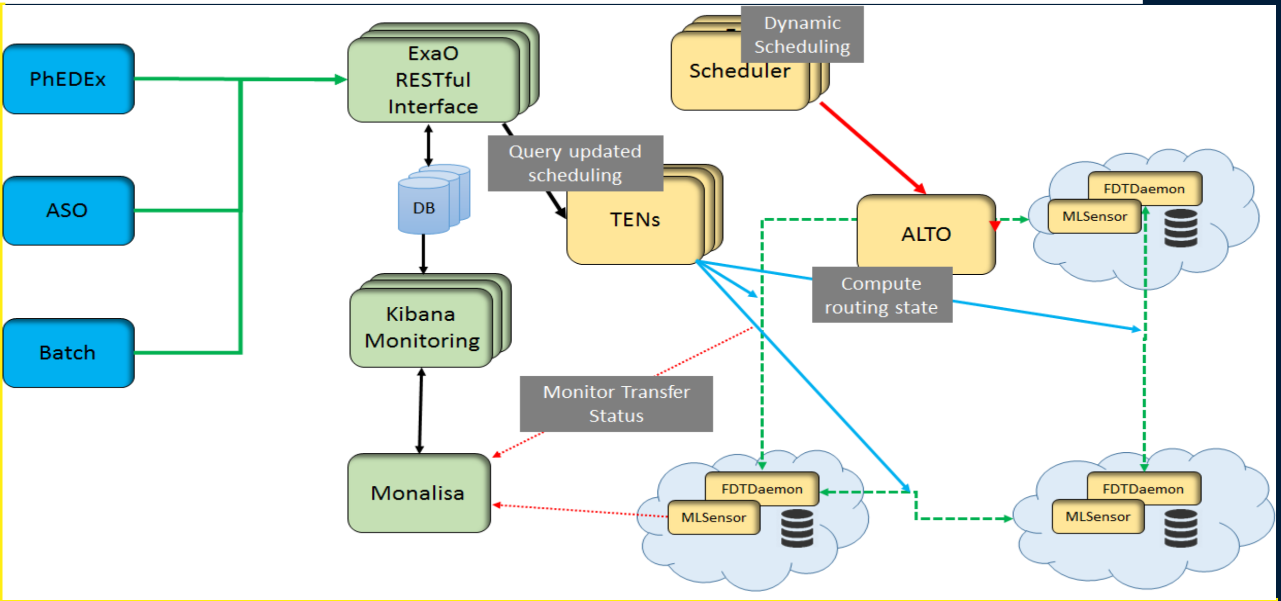


FAST Schedule

- Consistent, optimized flow-mod scheduling

CMS at SC16: *ExaO* - Software Defined Data Transfer Orchestrator with Phedex and ASO

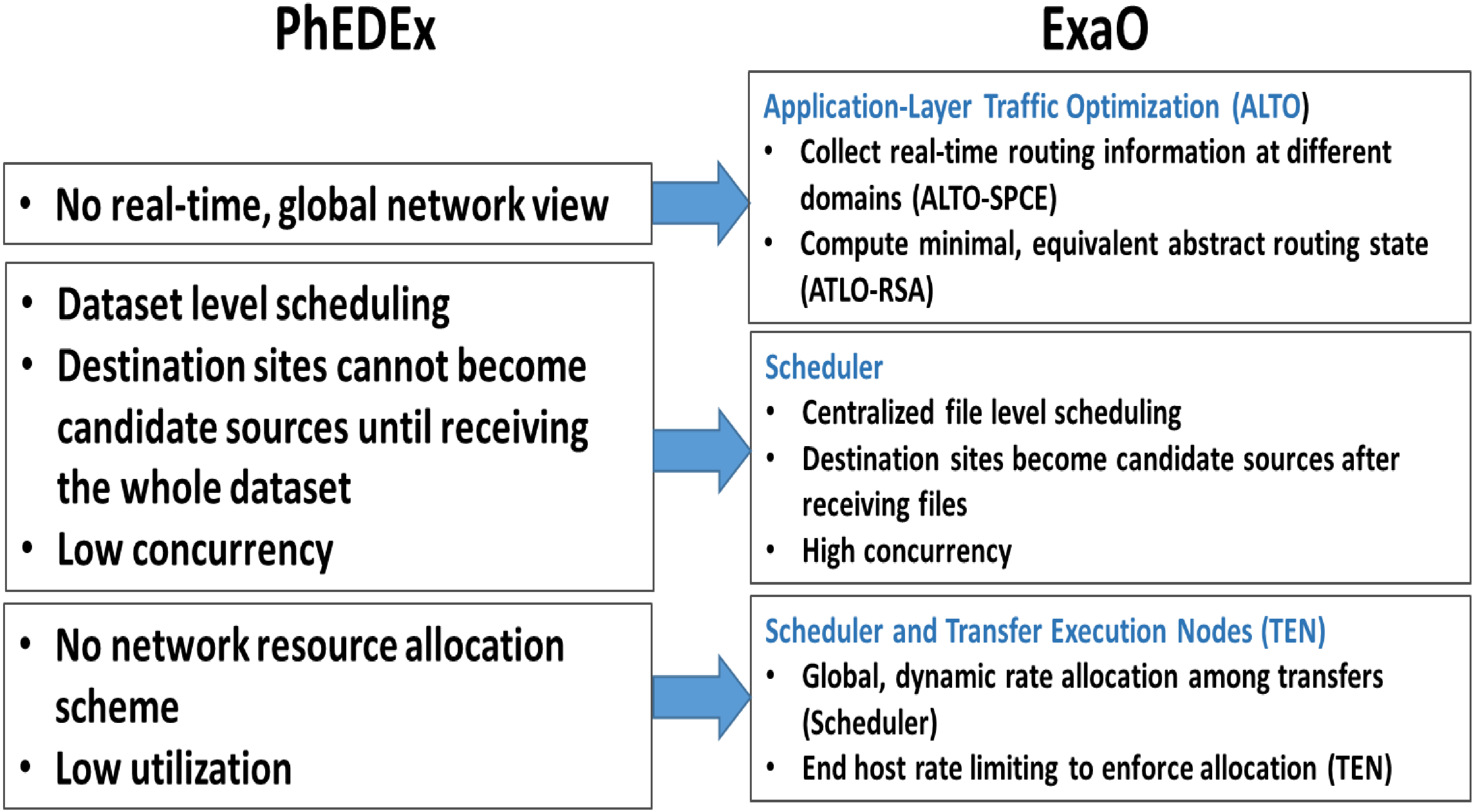
Leverage emerging SDN techniques to realize end-to-end orchestration of data flows involving multiple host groups in different domains



- ❒ Maximal link utilization with ExaO:
  - ▪ PhEDEx: CMS data placement tool for datasets
  - ▪ ASO: Stageout of output files from CMS Analysis Jobs
- ❒ Tests across the SC16 Floor: Caltech, UMich, Dell booths and Out Over the Wide Area: FIU, Caltech, CERN, UMich
- ❒ Dynamic scheduling of PetaByte transfers to multiple destinations

Partners: UMich, StarLight, PRP, UNESP, Vanderbilt, NERSC/LBL, Stanford, CERN; ESnet, Internet2, CENIC, MiLR, AmLight, RNP, ANSP

ExaO: Software Defined Data Transfer Orchestrator



PhEDEx

- No real-time, global network view

- Dataset level scheduling
- Destination sites cannot become candidate sources until receiving the whole dataset
- Low concurrency

- No network resource allocation scheme
- Low utilization

ExaO

Application-Layer Traffic Optimization (ALTO)

- Collect real-time routing information at different domains (ALTO-SPCE)
- Compute minimal, equivalent abstract routing state (ATLO-RSA)

Scheduler

- Centralized file level scheduling
- Destination sites become candidate sources after receiving files
- High concurrency

Scheduler and Transfer Execution Nodes (TEN)

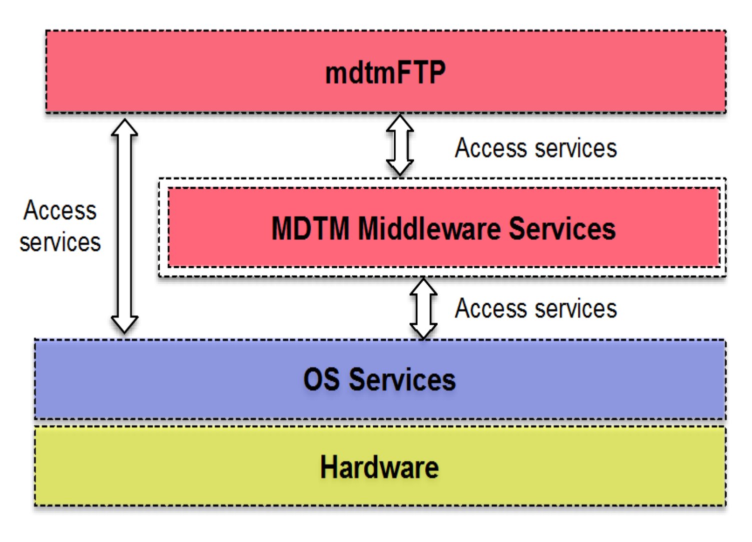
- Global, dynamic rate allocation among transfers (Scheduler)
- End host rate limiting to enforce allocation (TEN)

A Major Application of the New SDN Maple+Fast Framework
By the Yale Team and Caltech, towards CMS Data Operations

Multicore-Aware Data Transfer Middleware (mdtmFTP) Key Features

W. Wu, F. Demar et al. (Fermilab)

- **Key features**
  - **Pipelined I/O-centric design to streamline data transfer**
  - **Multicore-aware data transfer middleware (MDTM) optimizes use of underlying multicore system**
  - **Extremely efficient in transferring Lots Of Small Files**
  - **Various optimization mechanisms**
    - **Zero copy**
    - **Asynchronous I/O**
    - **Batch processing**



Note: mdtmFTP uses some basic Globus modules for rapid prototyping

http://mdtm.fnal.gov/

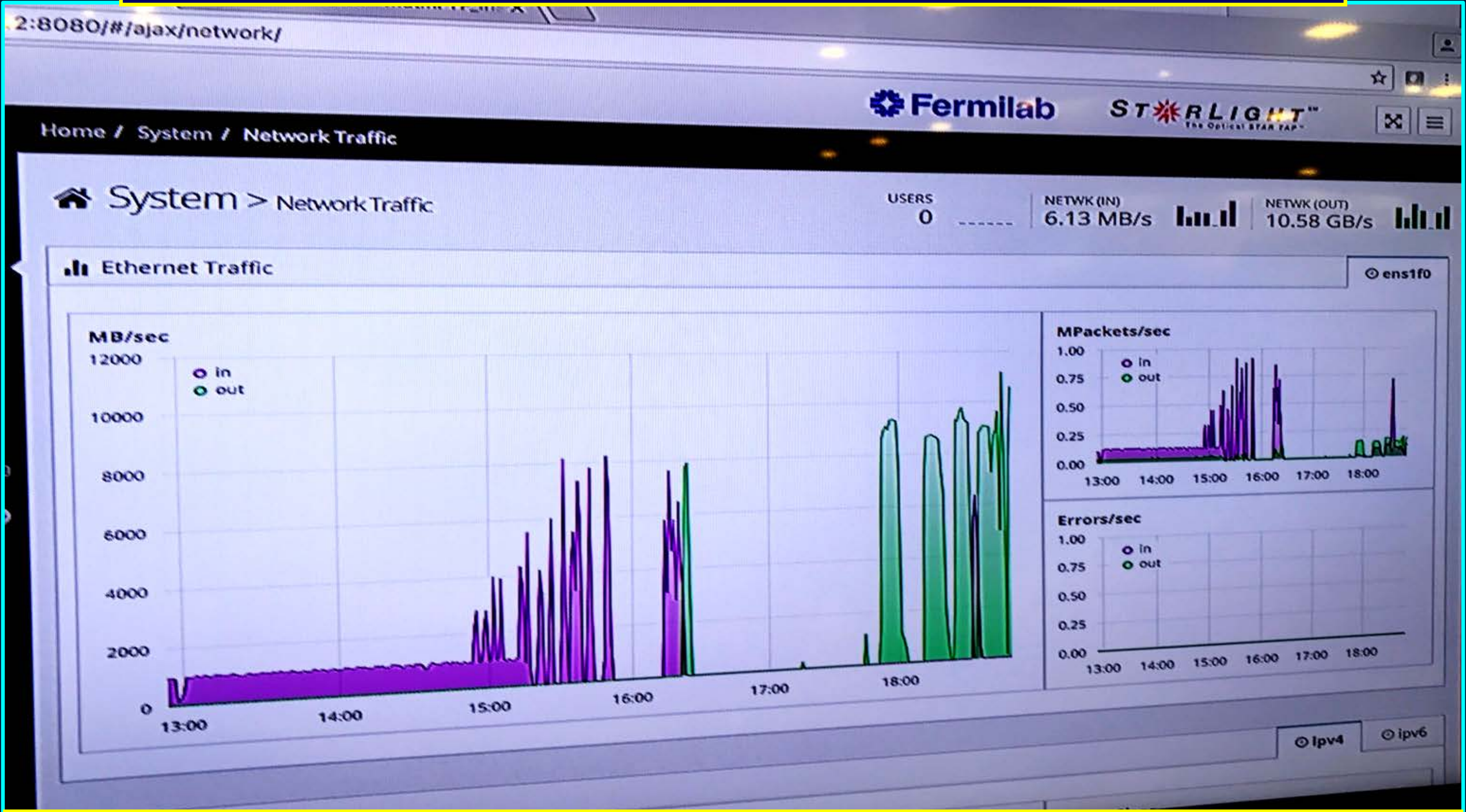
Multicore-Aware Data Transfer Middleware (mdtmFTP) Design (1)

- Dedicated I/O threads to perform network & disk I/O operations in parallel
- MDTM middleware to schedule cores for I/O threads
  - Each I/O thread pinned to a core
    near the I/O device the thread uses
    - I/O locality
    - Core affinity for I/O operations
  - An I/O thread is typically dedicated with a single core
  - System zoning to avoid interference with other applications
    - MDTM-zone for mdtmFTP
    - Non-MDTM-zone for other applications

Multicore-Aware Data Transfer Middleware (mdtmFTP) Design (2)

- Advanced data buffer mechanism to improve I/O performance
  - Pre-allocated data buffers to avoid costly memory allocation/deallocation in the critical I/O path of data transfer
  - Data buffers are pinned and locked to avoid memory swap and migration

"mdtmFTP @ 100GE Networks"
Demo At SC16, November 2016



2:8080/#/ajax/network/

Home / System / Network Traffic

System > Network Traffic

USERS 0

NETWK (IN) 6.13 MB/s

NETWK (OUT) 10.58 GB/s

mdtmFTP achieved ~85Gbs disk-to-disk

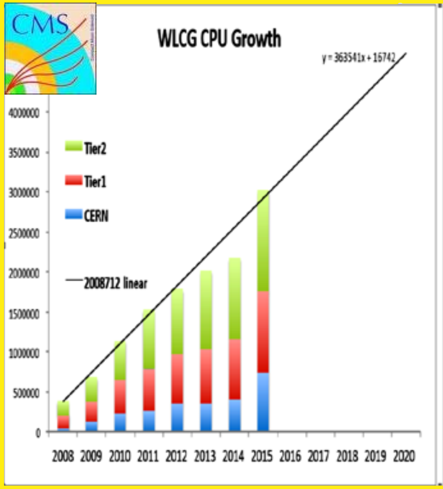
Bringing the Leadership HPC Facilities

Into the Data Intensive Echosystems of the LHC and Other Major Science Programs

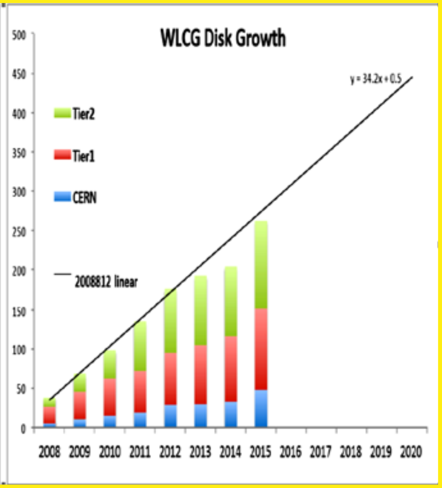
CMS Offline Computing Requirements

HL LHC versus Run2 and Run1 [*]

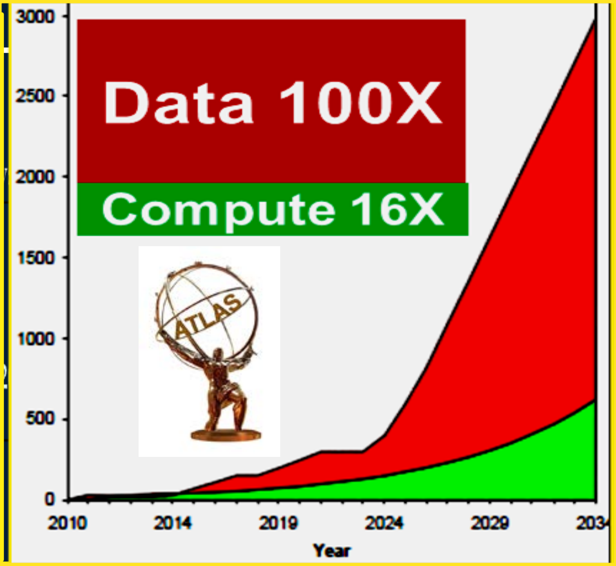
+ ~36k cores/Yr



+ ~34 PBytes/Yr



- Ratios in Computing and Storage for Run 2/Run1 are ≈ 2X.
- Hence HL-LHC to Run1 CPU: 130X to 400X



CPU Requirements Projections

- Projected CPU Needs:
  HL LHC/Run2 = 65 to 200X
- Anticipated increase in CPU resources at fixed cost/year: 8X
- Anticipated code efficiency improvements: 2X
- *Projected shortfall* at HL LHC
  4X to 12X

Storage Requirements Projections

- Projected Events:
  HL LHC / Run2 = 5 to 7.5X
- Event Size:
  HL LHC / Run2 = 4 to 6X
- Anticipated growth in Storage HL-LHC / Run2: 20-45X
- Projected shortfall at HL LHC
  3X or More

[*] CMS Phase2 Technical Proposal: https://cds.cern.ch/record/202088

Operational Pilot for Exascale and other HPC Facilities with Petabyte Transactions

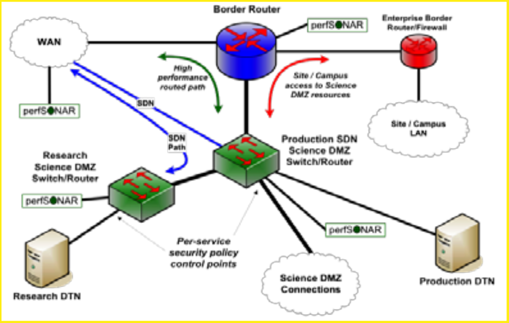
➡ Targeting the CPU Needs at LHC Run3 and HL LHC

❒ Develop system architectures in HW + software for petabyte transactions (to/from exabyte stores)

- ✸ Edge clusters with petabyte caches
  - ✸ Input + output pools: ~10 to 100 Pbytes
- ✸ A handful of proxies at the edge
  - ✸ To manage and focus security efforts
- ✸ Extending Science DMZ concepts
  - ✸ Enabling 100G to Tbps SDNs with Edge/WAN Coordination
- ✸ Identifying + matching HEP units of work to specific sub-facilities adapted to the task
- ✸ Site-Network End-to-End Orchestration
  - ✸ Efficient, smooth petabyte flows over 100G then 400G (2018) then ~1 Tbps (2021) networks

Leadership

Next Gen Science DMZ


Border Router
perfSONAR
Enterprise Border Router/Firewall
WAN
High performance routed path
Site / Campus access to Science DMZ resources
perfSONAR
SDN
SDN Path
Site / Campus LAN
Production SDN Science DMZ Switch/Router
Research Science DMZ Switch/Router
perfSONAR
perfSONAR
Per-service security policy control points
Science DMZ Connections
Research DTN
Production DTN

Pilots at Argonne (and ORNL) HPC Facilities

(1A) CMS HPC Prod: a pilot on Mira as a major resource on the CMS Grid

- ❑ Adapting and Interfacing CMS' job submission system based on HTCondor, to MPI and Cobalt
  - ❑ + Providing a generally useful interface

(1B) Moving to THETA by this Fall: Intel Knights Landing Architecture: 72 core X 4 threads

- ❑ Porting multi-threaded CMSSW reco + simulation

(2) HPC Sherpack: advanced multiparton generators with NLO accuracy (Sherpa,MC@NLO)

- ❑ Building on and advancing the work of Tom LeCompte et al.
- ❑ New boosting methods for multidimensional integration and space sampling: with order of magnitude advances in speed &/or accuracy
- ❑ CPU intensive integration step results will be retrieved for further CMS event generation on existing resources elsewhere

(3) HPC Data Transfer Nodes (DTNs)

Deployed in the Argonne JLSE subnet

Pilot bidirectional high throughput transfers of large data blocks

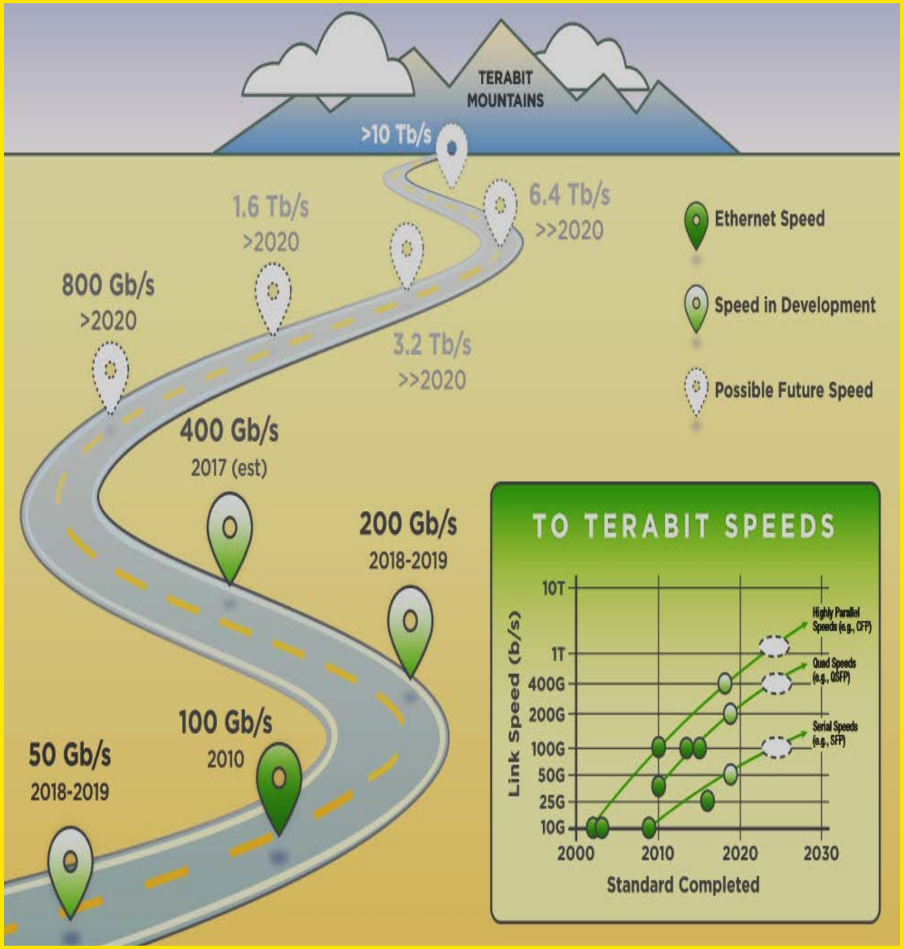
- ▪ ANL↔ Fermilab Tier1
- ▪ ANL↔ Caltech Tier2

Riding the Ethernet Wave: Petabyte Transactions
To Create the NG and NNG Ecosystems: 2016 – 2026+

- We are midway in the current 7-8 year generational cycle of 100G network links
- A petabyte transfer would occupy a 100G link for 24 hrs at wire speed now
- With Aurora circa 2019, a PB transfer would take 6 hours on a 400G link
- At the dawn of the exascale era, circa 2023 a PB would take 90 minutes on a 1.6 Tbps link
  - Providing some agility
  - Beginning to allow Multiple transactions
- Through the HL LHC era we can foresee Next-to-Next Generation Systems with
  - Increasing agility
  - Larger and multiple transactions

Ethernet Alliance Roadmap

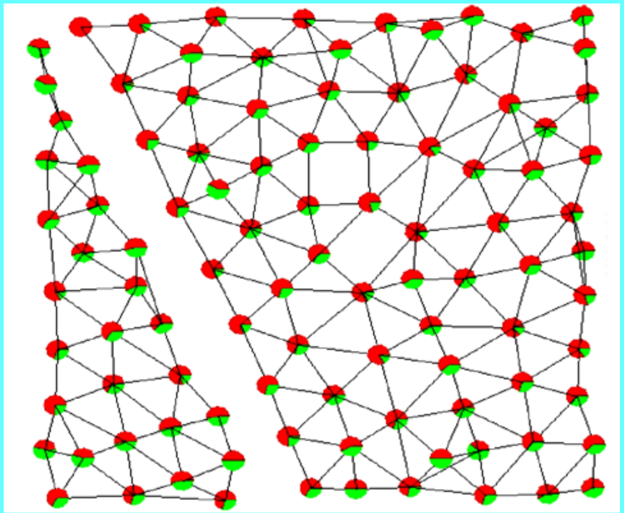
http://www.ethernetalliance.org/wp-content/uploads/2015/03/Ethernet-Roadmap-2sides-29Feb.pdf

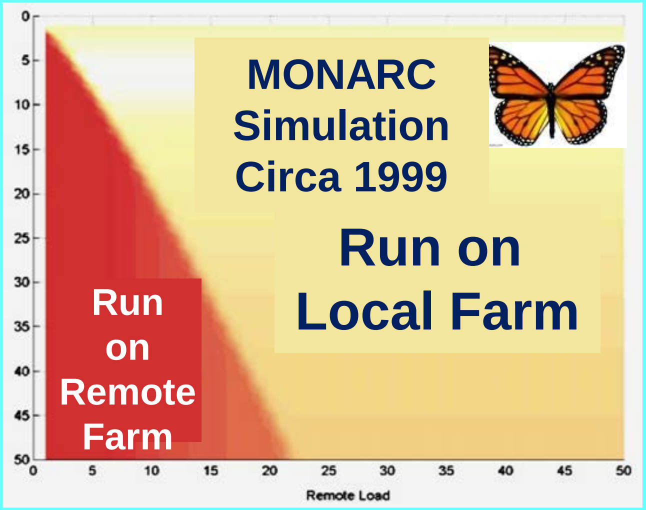


20 Years Forward: To the 10 Terabit/sec Mountains, and Beyond

Key Developments from the HEP Side: Machine Learning, Modeling, Game Theory

- Applying Deep Learning + Self-Organizing systems methods to optimize LHC workflow
  - Unsupervised: to extract the key variables and functions
  - Supervised: to derive optima
  - Iterative and model based: to find effective metrics and stable solutions [*]
- Complemented by modeling and simulation; game theory methods [*]
- Progressing to real-time agent-based dynamic systems
- With application to LHC Workflow





Self-organizing neural network for job scheduling in distributed systems

[*] T. Roughgarden (2005). *Selfish routing and the price of anarchy*

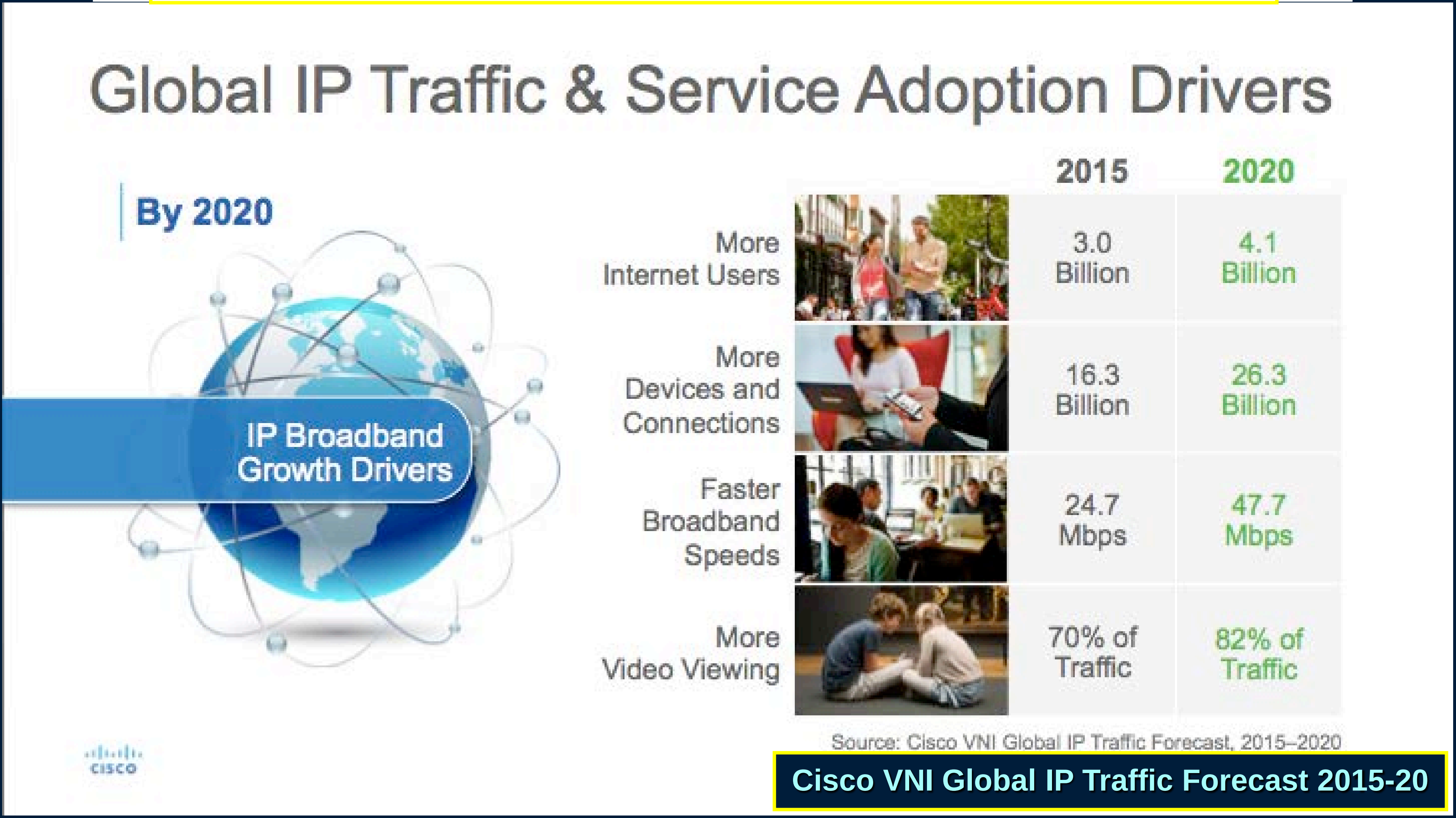
More On
Global Trends
The Internet and
International Networks

Future of International Networks
Summary

- Demand growth will remain strong at ~30-40% per year
- A small group of companies will likely account for a larger share of the international capacity
- Prices per unit bandwidth continue to decline
  - Disparity among regions decreasing but still striking
- Investment in new cables and technological advantages will delay the risk of capacity exhaustion: a potential issue by 2021-3
- Most damage to submarine cables from fishing and anchors (not sharks and Russian submarines . . .)

Cisco Network 2016 Update
Global Trends in 2015-2020

Global IP Traffic Growth and Service Adoption Drivers



Global IP Traffic & Service Adoption Drivers

By 2020

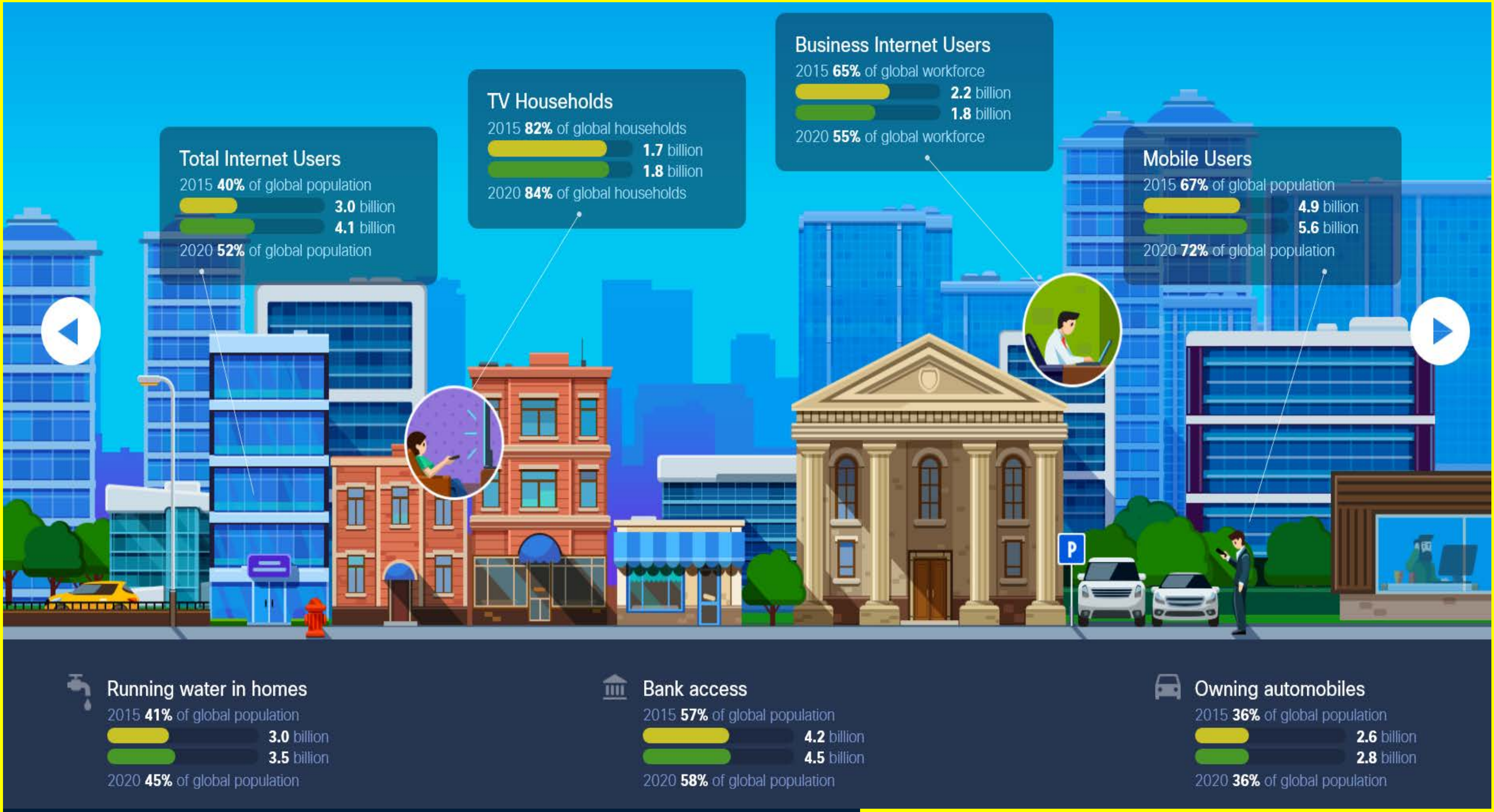
IP Broadband Growth Drivers

	2015	2020
More Internet Users	3.0 Billion	4.1 Billion
More Devices and Connections	16.3 Billion	26.3 Billion
Faster Broadband Speeds	24.7 Mbps	47.7 Mbps
More Video Viewing	70% of Traffic	82% of Traffic

Source: Cisco VNI Global IP Traffic Forecast, 2015–2020

Cisco VNI Global IP Traffic Forecast 2015-20

Internet vs. TV, Running Water and Other “Necessities”



Total Internet Users
2015 **40%** of global population
3.0 billion
4.1 billion
2020 **52%** of global population

TV Households
2015 **82%** of global households
1.7 billion
1.8 billion
2020 **84%** of global households

Business Internet Users
2015 **65%** of global workforce
2.2 billion
1.8 billion
2020 **55%** of global workforce

Mobile Users
2015 **67%** of global population
4.9 billion
5.6 billion
2020 **72%** of global population

Running water in homes
2015 **41%** of global population
3.0 billion
3.5 billion
2020 **45%** of global population

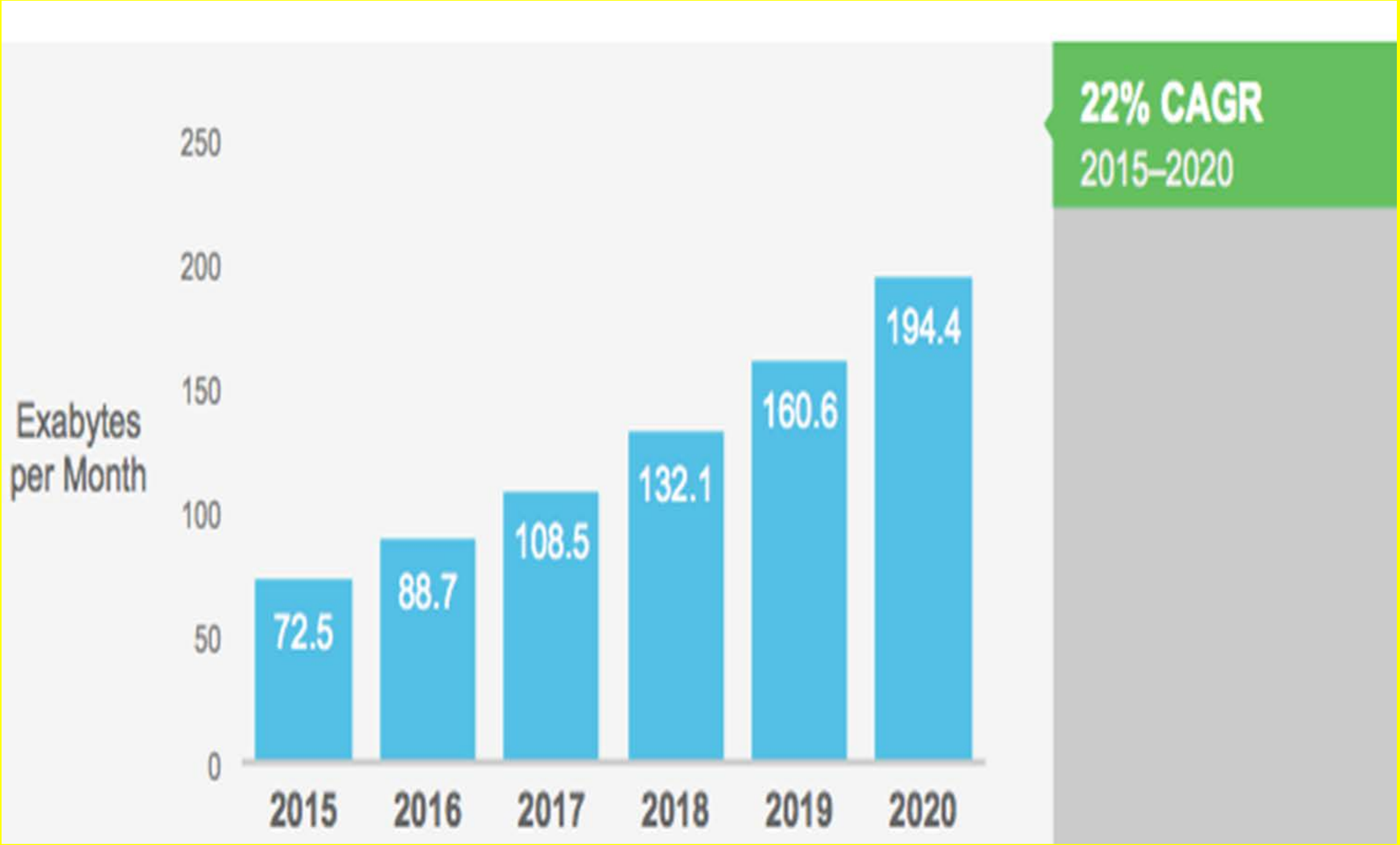
Bank access
2015 **57%** of global population
4.2 billion
4.5 billion
2020 **58%** of global population

Owning automobiles
2015 **36%** of global population
2.6 billion
2.8 billion
2020 **36%** of global population

http://www.cisco.com/c/m/en_us/solutions/service-provider/vni-complete-forecast/infographic.html

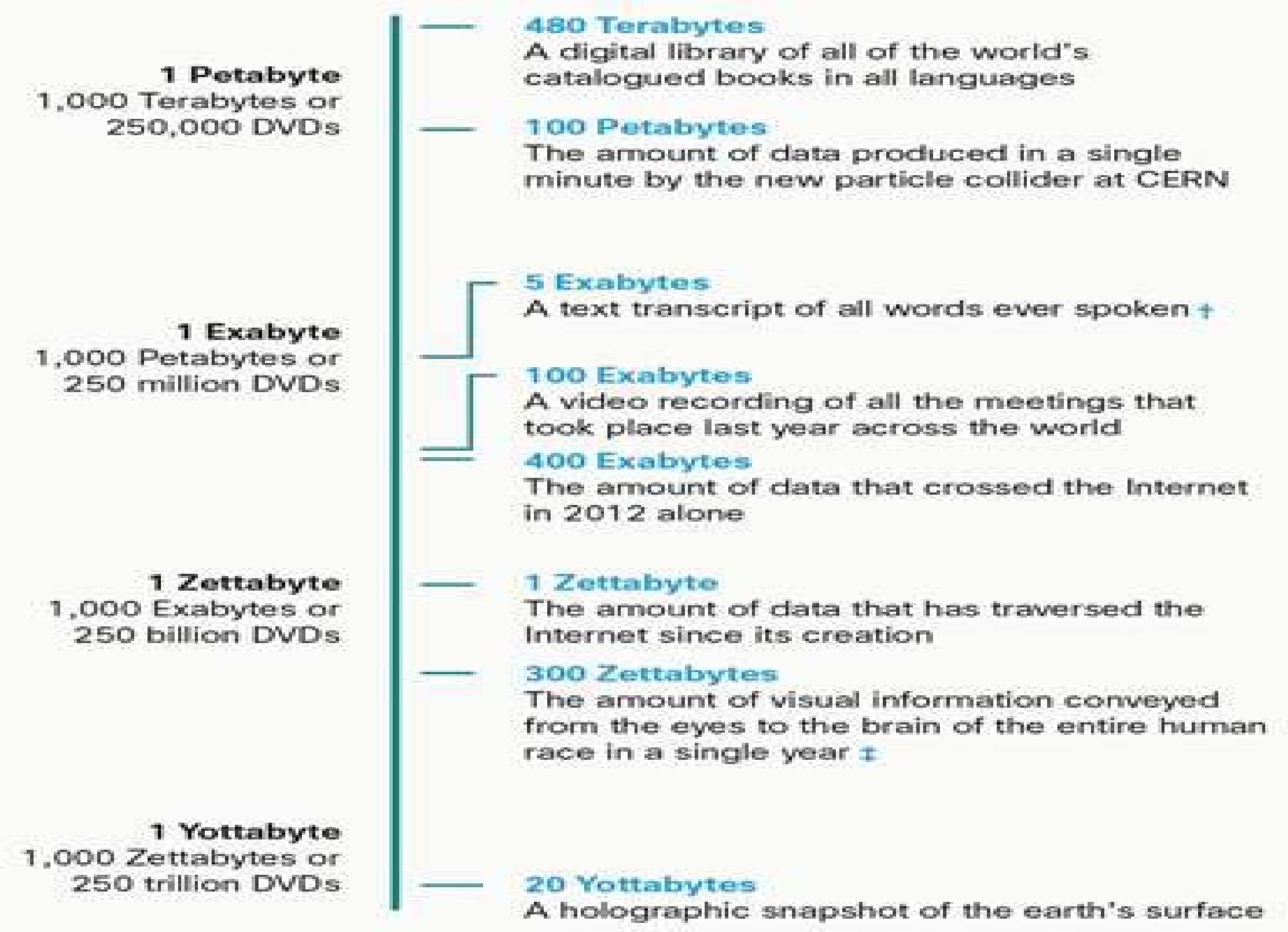
Cisco 2016 VNI Complete Forecast

Global IP Traffic will Grow 3X During 2014-19 (22% CAGR)



Cisco VNI Global IP Traffic Forecast 2015-20

Cisco VNI: Byte Scale and Equivalences



1 Petabyte
1,000 Terabytes or
250,000 DVDs

480 Terabytes
A digital library of all of the world's catalogued books in all languages

100 Petabytes
The amount of data produced in a single minute by the new particle collider at CERN

1 Exabyte
1,000 Petabytes or
250 million DVDs

5 Exabytes
A text transcript of all words ever spoken †

100 Exabytes
A video recording of all the meetings that took place last year across the world

400 Exabytes
The amount of data that crossed the Internet in 2012 alone

1 Zettabyte
1,000 Exabytes or
250 billion DVDs

1 Zettabyte
The amount of data that has traversed the Internet since its creation

300 Zettabytes
The amount of visual information conveyed from the eyes to the brain of the entire human race in a single year ‡

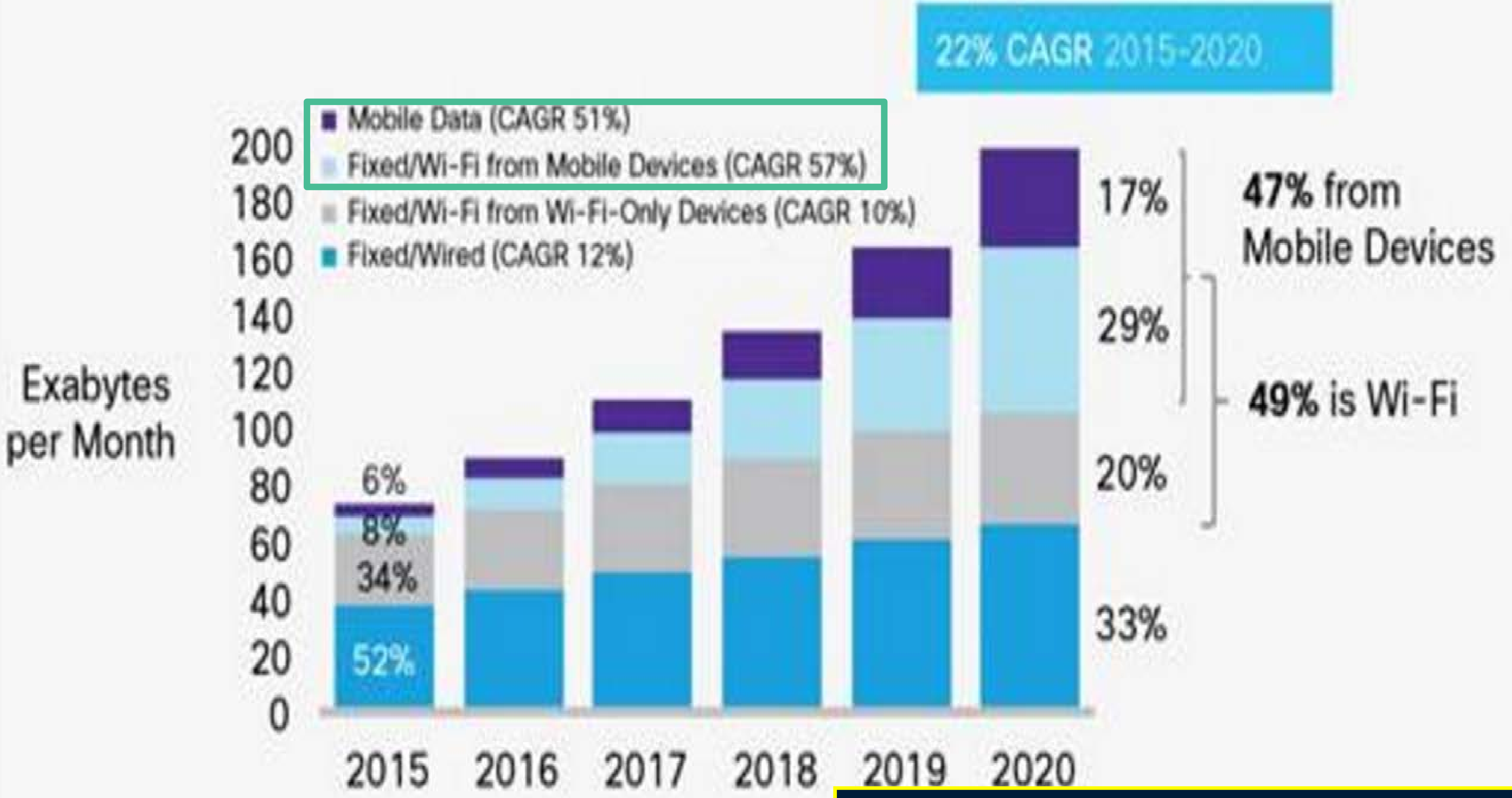
1 Yottabyte
1,000 Zettabytes or
250 trillion DVDs

20 Yottabytes
A holographic snapshot of the earth's surface

A digital holographic snapshot of the Earth' surface is estimated at 20 Yottabytes

http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-forecast-qa.html

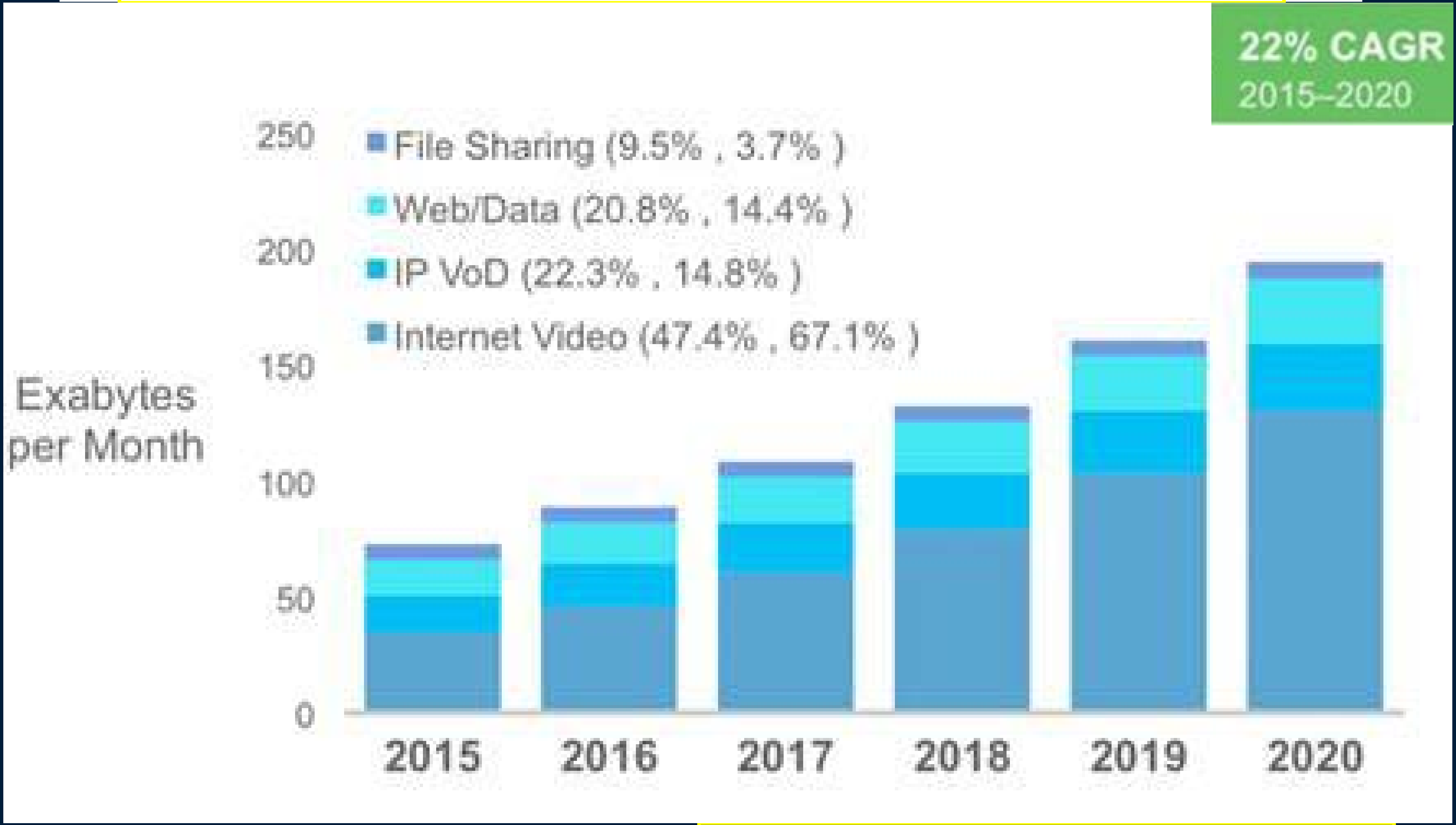
Global Internet Traffic
by Local Access Technology



Cisco VNI Global IP Traffic Forecast 2015-20

Global IP Video Traffic

Will Account for 80% of IP Traffic by 2019

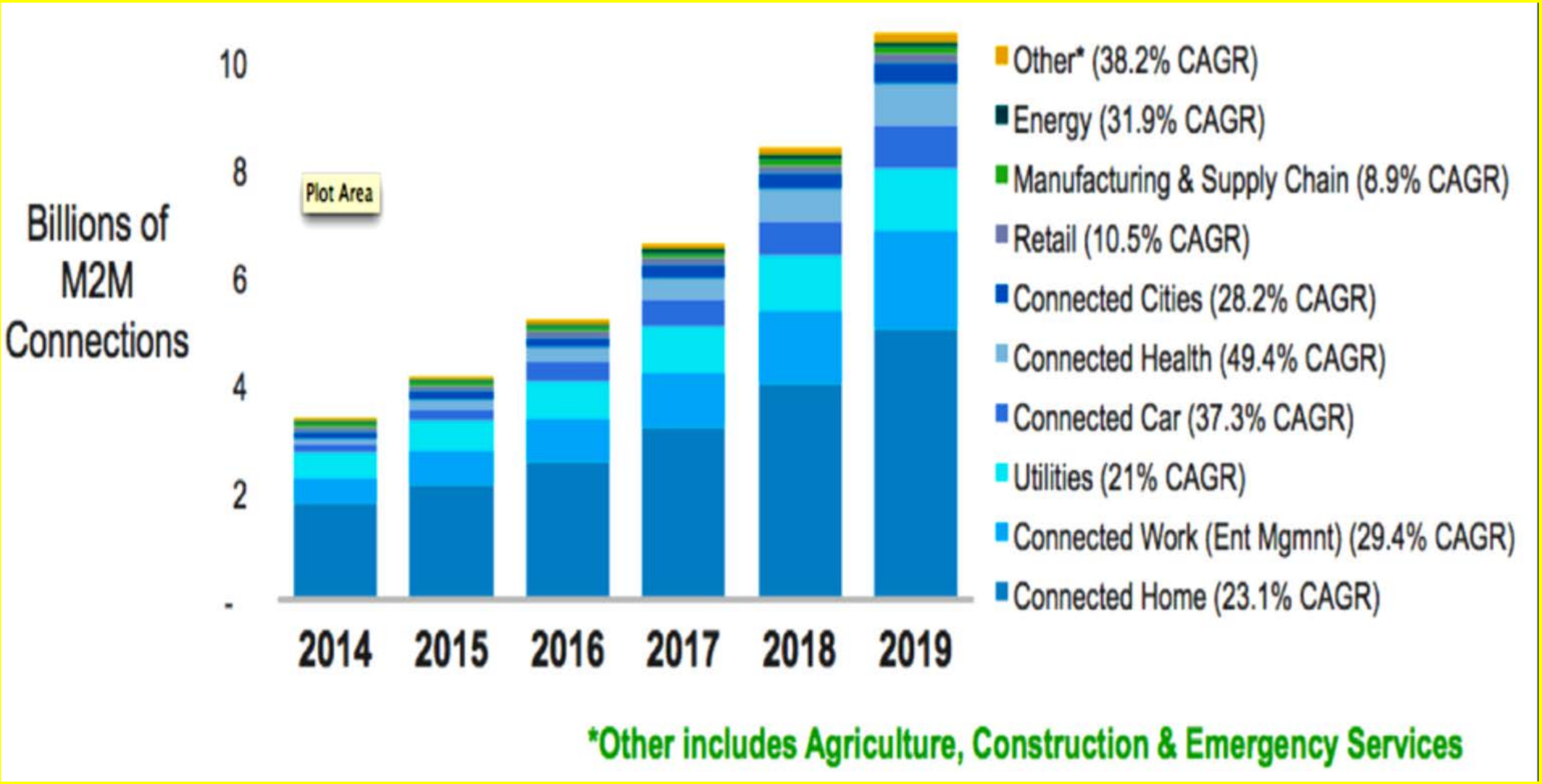


Cisco VNI Global IP Traffic Forecast 2015-20

Cisco VNI Global IP Traffic Outlook
The Zettabyte Era: Trends and Analysis

- Annual global IP traffic will reach 2.3 Zettabytes (ZB) in 2020; 1.6 ZB by 2018
  - ❐ Global IP traffic has increased 5X over the past 5 years, and will increase 3X over the next 5 years, equivalent to a CAGR of 21% [slowing growth]
- Busy-hour Internet traffic will increase 3.4X between 2013 & 2018, to 1.0 petabit/s while average Internet traffic will increase 2.8X to 0.3 Pbps.
- Metro traffic will surpass long-haul traffic in 2015, and account for 62% of total IP traffic by 2018.
  - ❐ Due in part to the increasing role of content delivery networks, which bypass long-haul links and deliver traffic to metro & regional backbones.
  - ❐ 55% of all Internet traffic will cross CDNs by 2018 globally, up from 36% in 2013.
- The Non-PC share of total IP traffic will grow to 57% by 2018.
  - ❐ CAGR of Traffic Sources: PC-originated 10%; TVs 35%; *Tablets 74%; Smartphones 64%; M2M (machine to Machine) 84%*
- Traffic from wireless and mobile devices will exceed traffic from wired devices by 2016.

Global Machine to Machine Connections
Internet of Everything (IoE) Growth

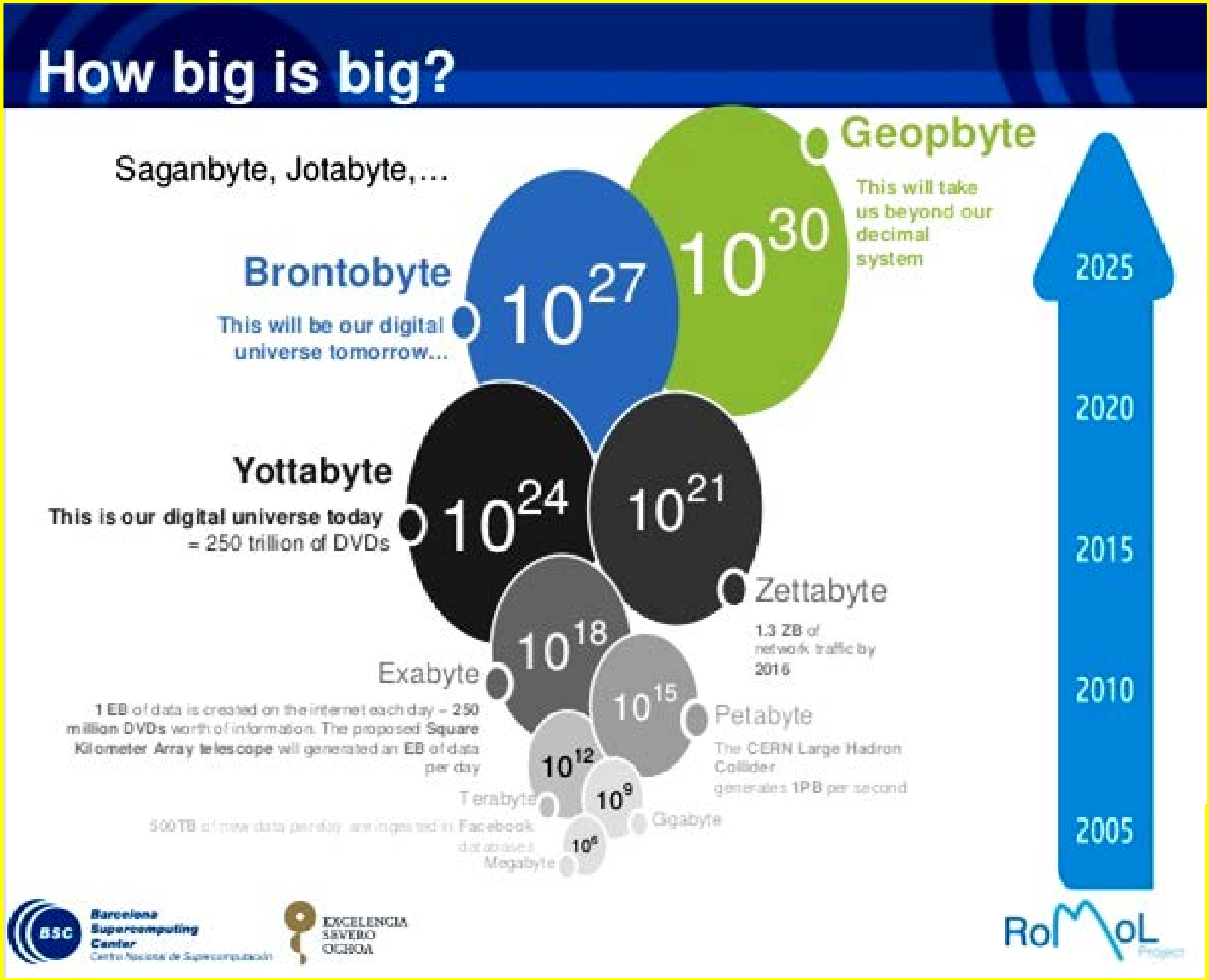


*Other includes Agriculture, Construction & Emergency Services

By 2019, Connected Homes will be largest
Connected Health will have fastest growth

Cisco VNI Global IP Traffic Forecast 2014-19

Internet of Everything (IoE)
Dominance of our Digital Future



How big is big?

Saganbyte, Jotabyte,...

Geopbyte

This will take us beyond our decimal system

Brontobyte

This will be our digital universe tomorrow...

Yottabyte

This is our digital universe today = 250 trillion of DVDs

Zettabyte

1.3 ZB of network traffic by 2016

Exabyte

1 EB of data is created on the internet each day – 250 million DVDs worth of information. The proposed Square Kilometer Array telescope will generated an EB of data per day

Petabyte

The CERN Large Hadron Collider generates 1PB per second

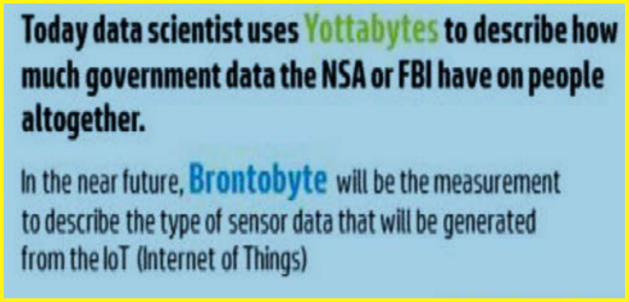
Terabyte

500TB of new data per day are ingested in Facebook databases

Gigabyte

Megabyte

- **1 Zettabyte:**
  2016 Network Traffic
- **1 Yottabyte:**
  Data in our digital universe today
- **1 Brontobyte**
  the IoT digital universe ~2023+
- **1 Geopbyte**
  the IoE digital universe in the HL LHC era ~2030+



Today data scientist uses Yottabytes to describe how much government data the NSA or FBI have on people altogether.

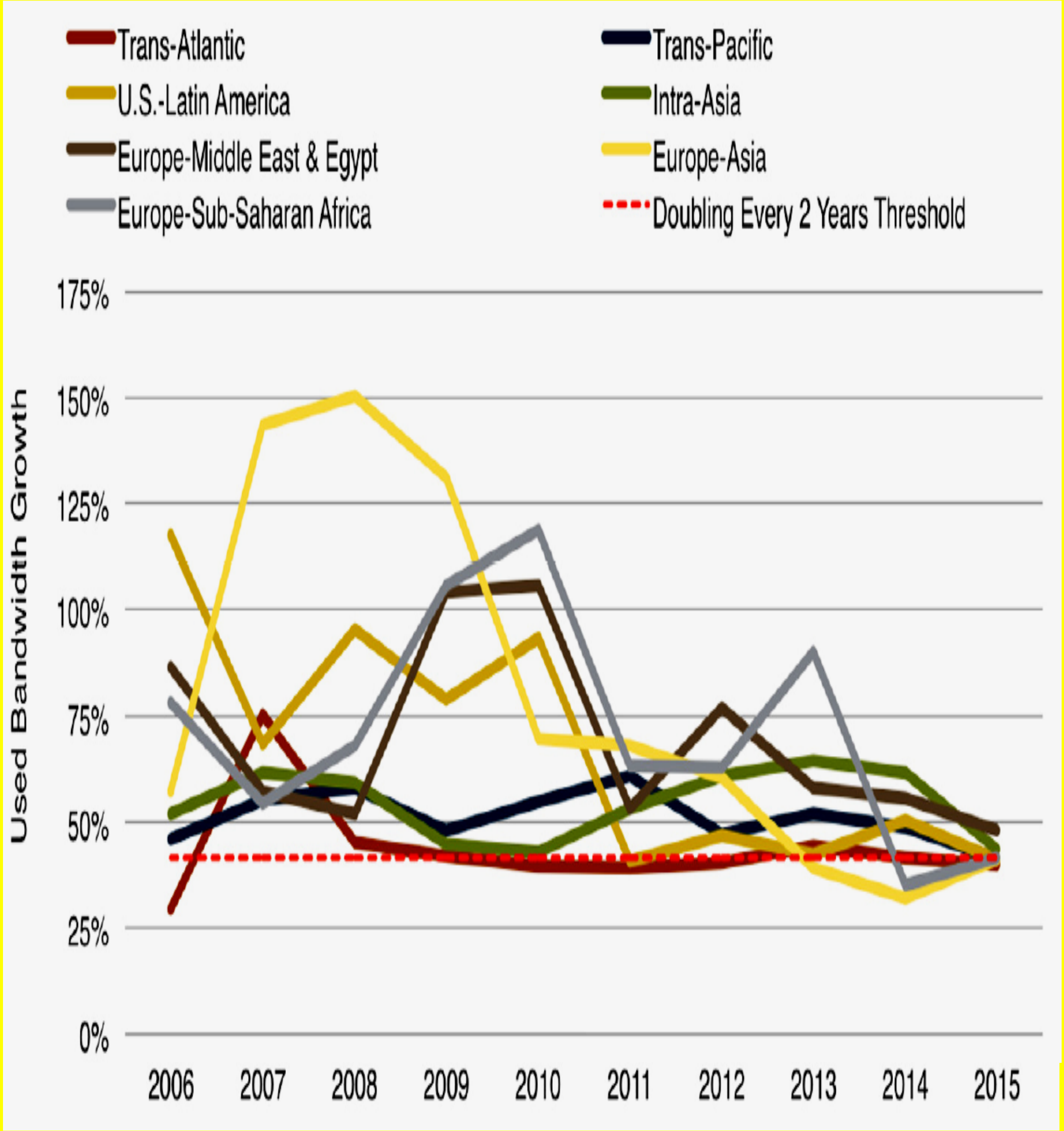
In the near future, **Brontobyte** will be the measurement to describe the type of sensor data that will be generated from the IoT (Internet of Things)

http://www.slideshare.net/FundacionAreces/mateo-valero-big-data-de-la-investigacin-cientfica-a-la-gestin-empresarial

Global International Bandwidth and Pricing Trends

TeleGeography at PTC 16 and 17

Bandwidth Growth by Region and Moore's Law

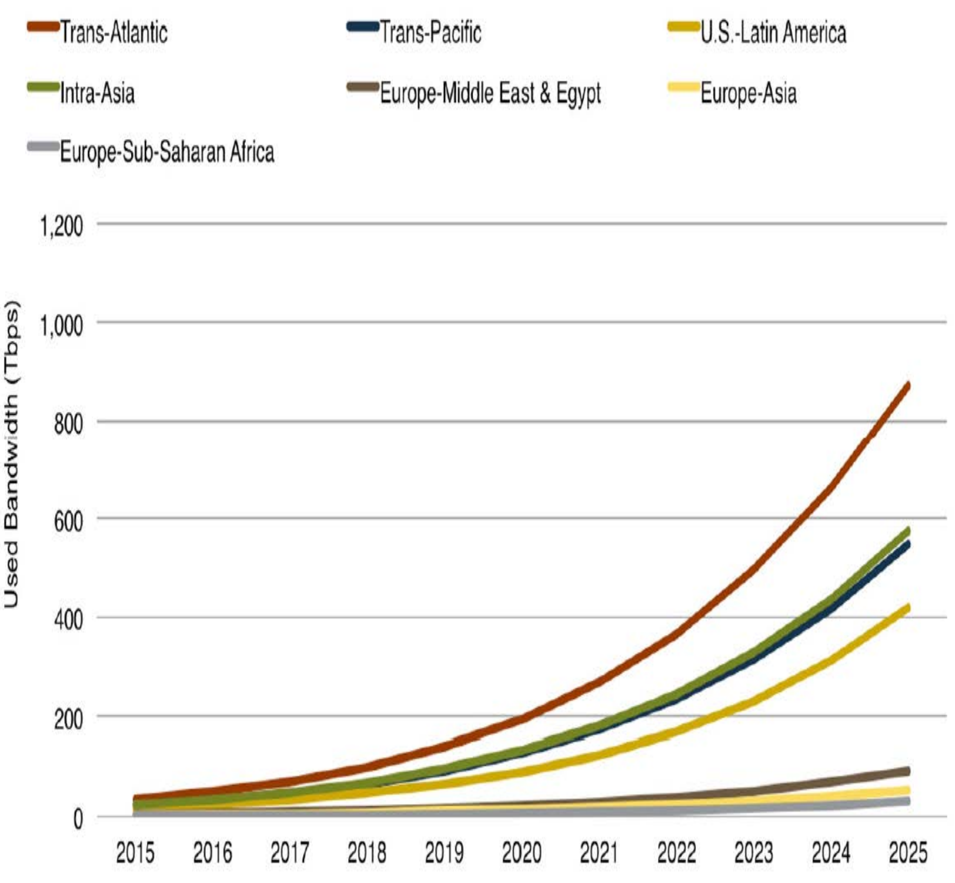


- Interesting convergence of the bandwidth growth in all regions towards the Moore's Law growth number
  - Moore's Law: "number of transistors on a circuit doubles approximately every 2 years"
  - Implied annual growth: 41%
- Also shows an interesting historical record of explosive growth of capacity on some routes in the last decade

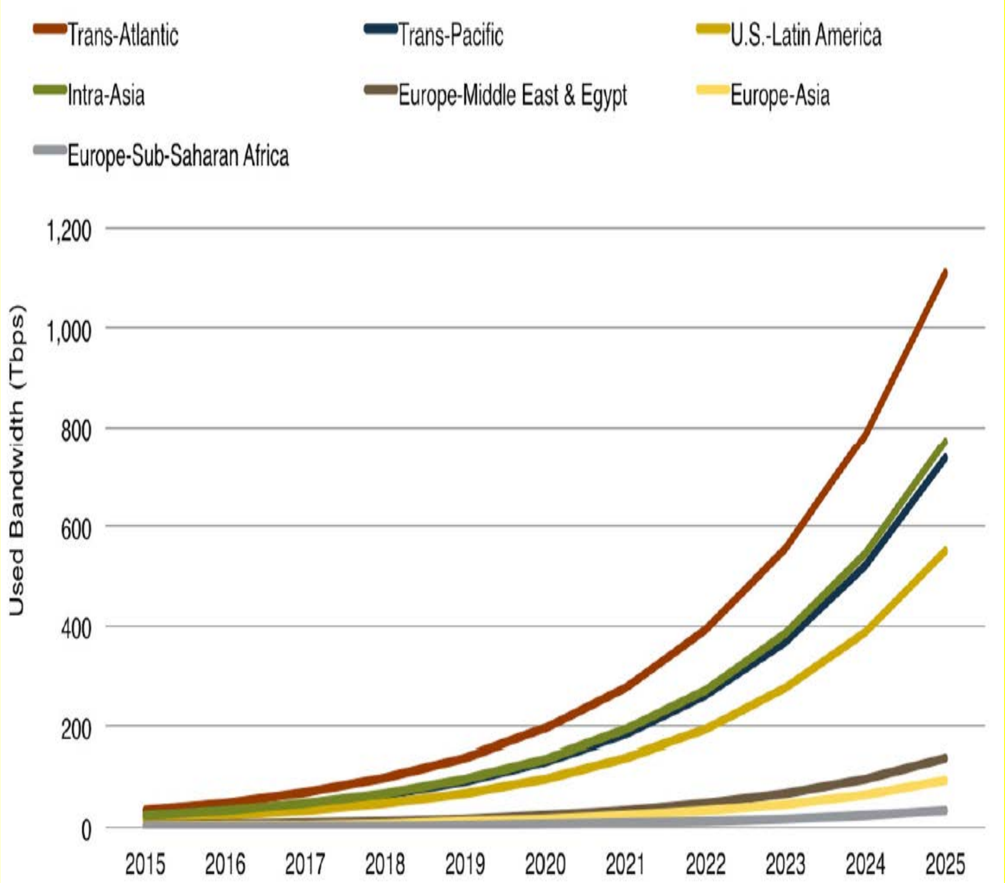
Telegeography; A. Maulding, PTC 16

Forecasts of Used Int’l Bandwidth Through 2025: Telegeography vs Moore’s Law

Telegeography Projection



Moore’s Law Projection



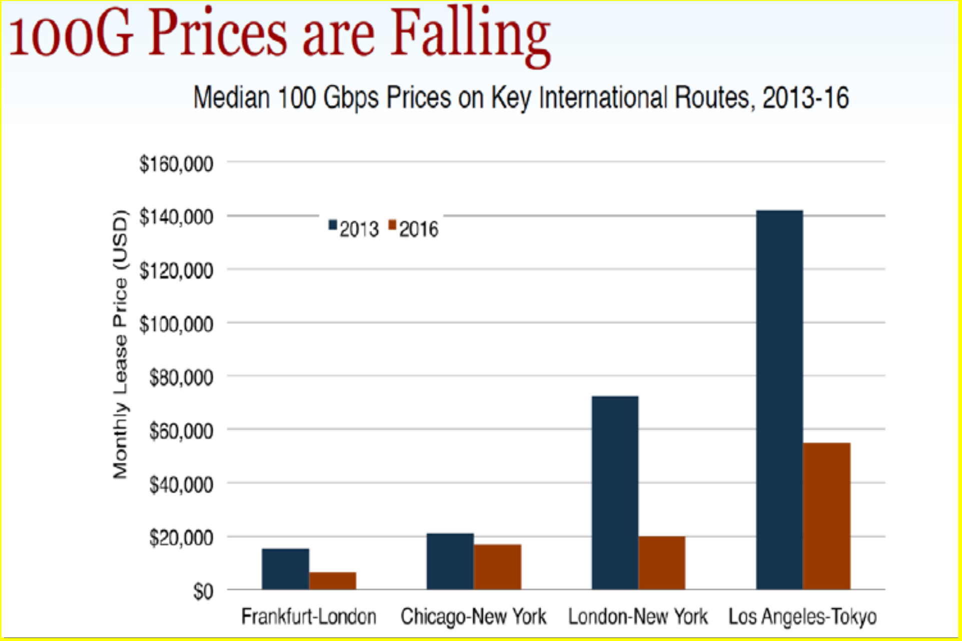
- Telegraphy forecast is close to Moore’s Law: > ~1 Pbps by HL LHC
- Note: *R&E Network Traffic Growth Rate is Larger*

Telegeography; A. Maulding, PTC 16

Wavelength Price Evolution

Telegeography
B. Boudreau PTC 17

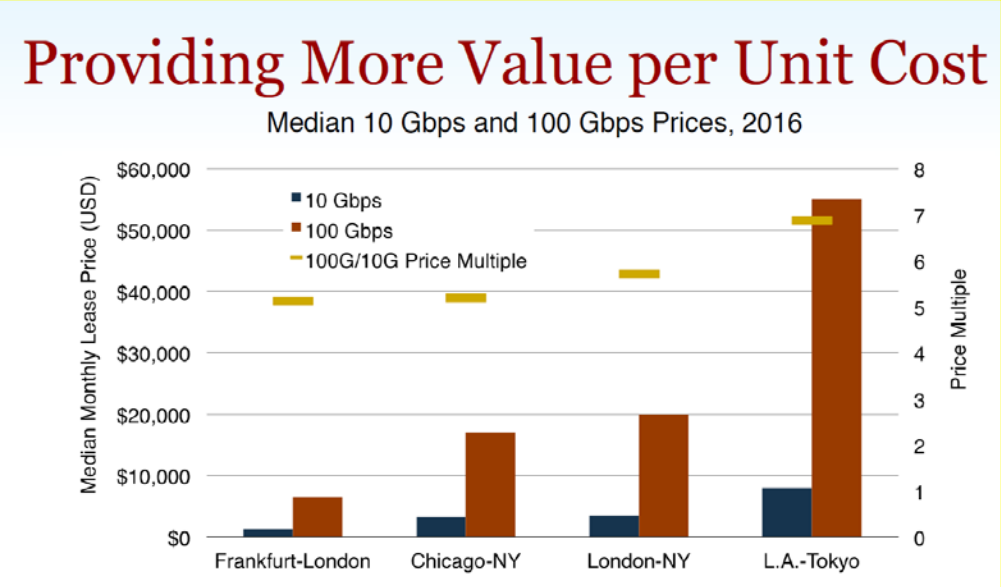
“100G is the New 10G”



100G Prices are Falling

Median 100 Gbps Prices on Key International Routes, 2013-16

Median 10G and 100G Prices, Multiple

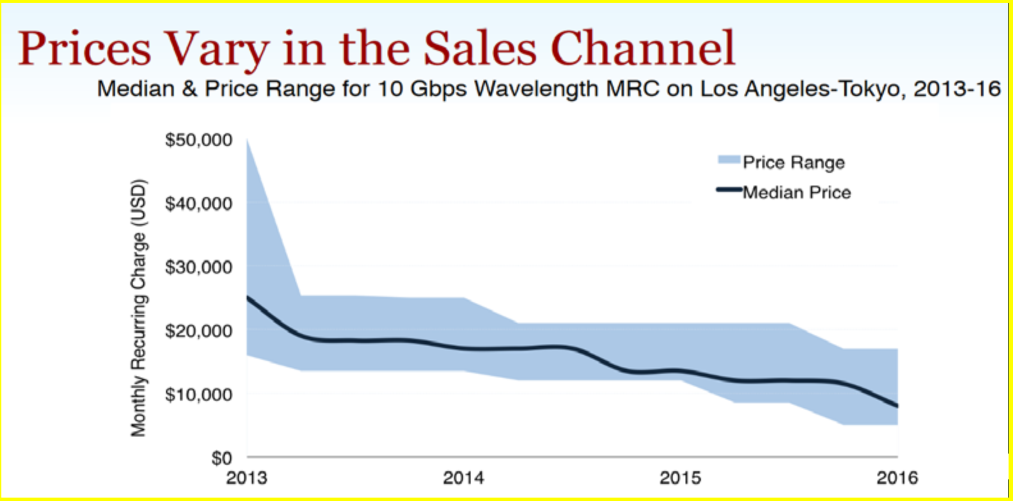


Providing More Value per Unit Cost

Median 10 Gbps and 100 Gbps Prices, 2016

1. Prices continue to decline
2. Different regions becoming “more similar” in price
3. But price differences are still striking

10G Price Evolution: Median and Range

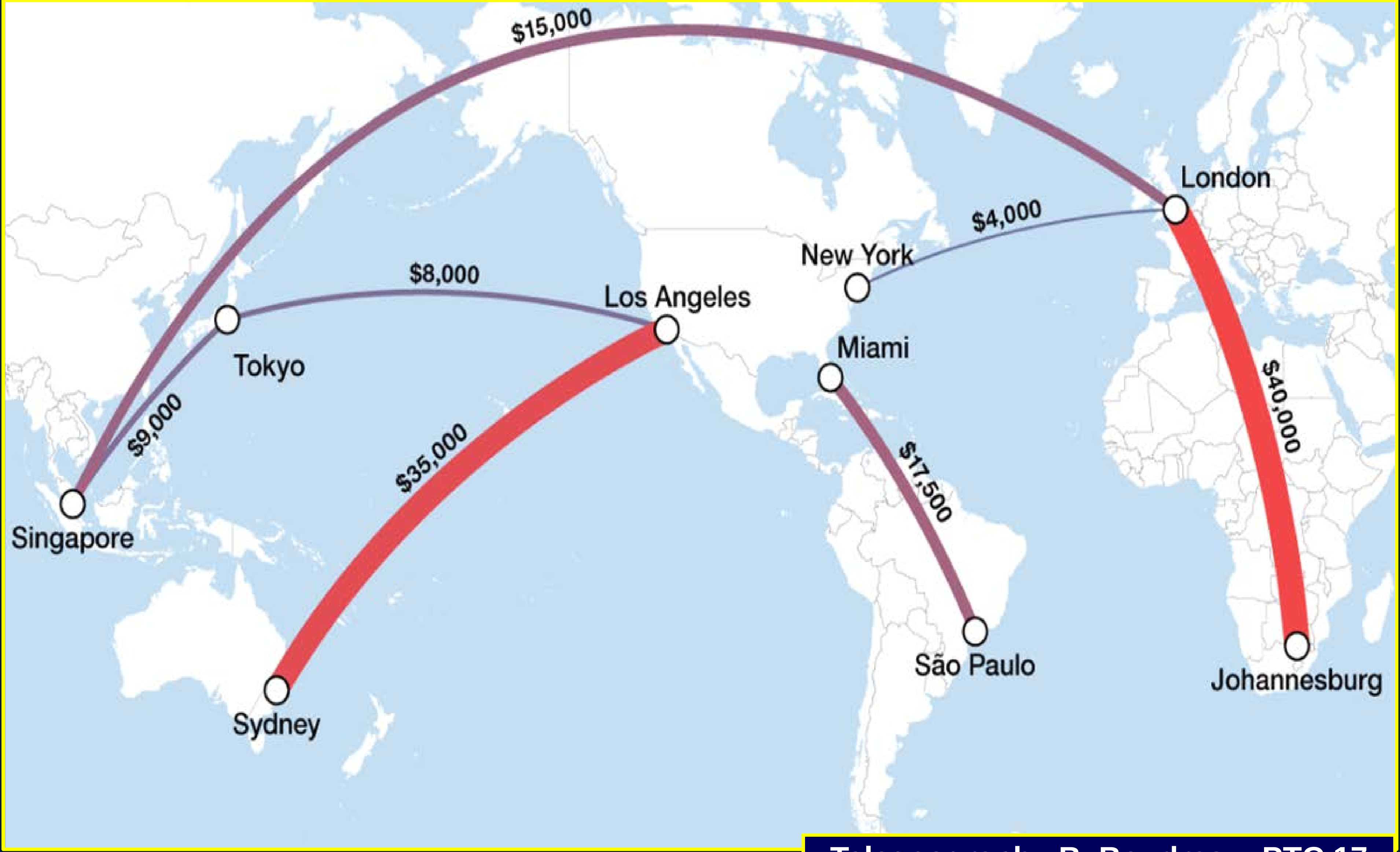


Prices Vary in the Sales Channel

Median & Price Range for 10 Gbps Wavelength MRC on Los Angeles-Tokyo, 2013-16

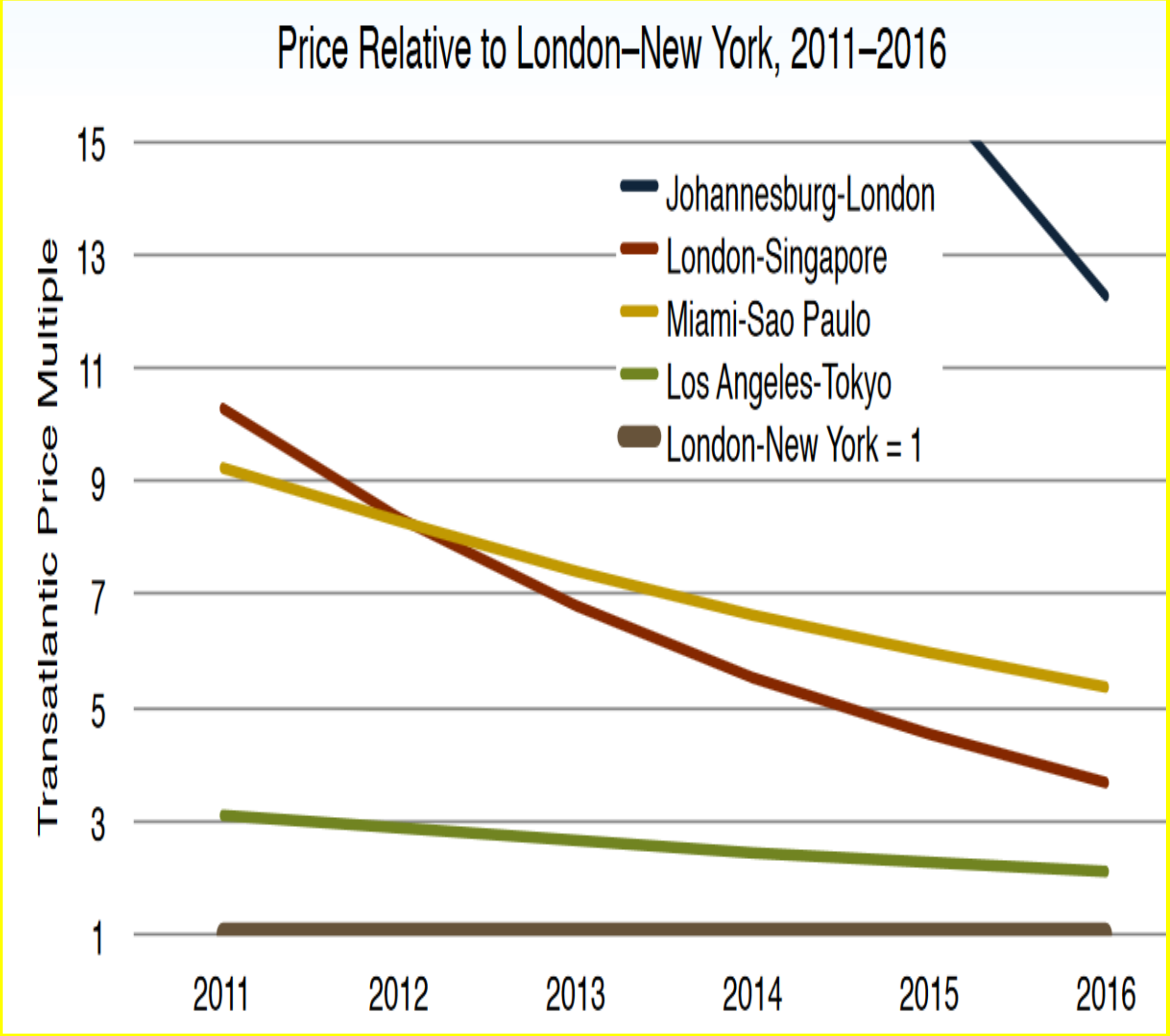
Bandwidth Prices Still Vary by Region

10G Leased Line Monthly Lease Price



Telegeography B. Boudreau, PTC 17

Global Prices Among Regions Are Converging



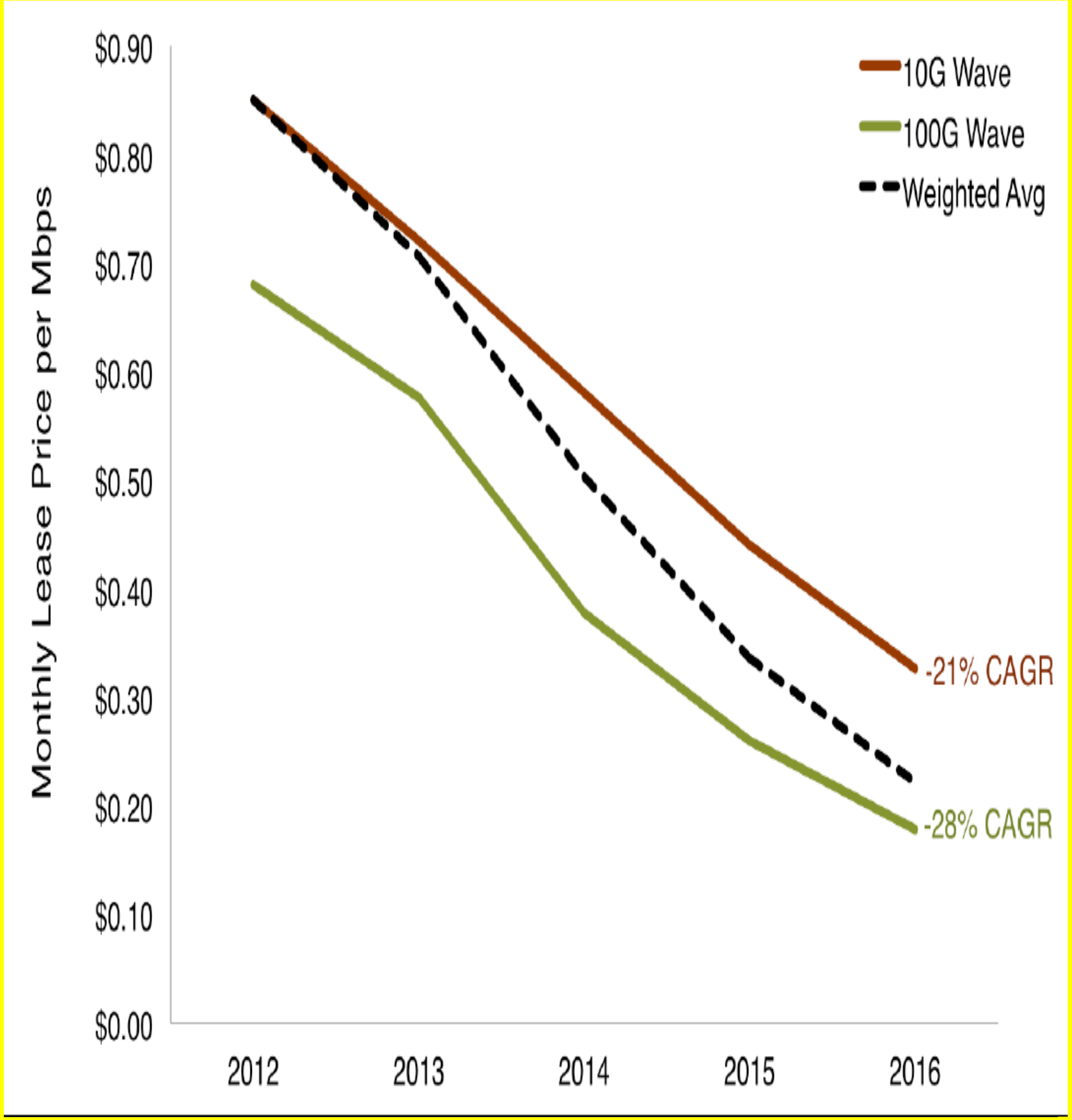
Telegeography B. Boudreau, PTC 17

Why they have converged

- Prices on high growth routes have declined more than on established routes
- More cables coming into service on under-developed routes fuel price erosion
- Technology advances lower unit costs

Evolution of 100G vs 10G Pricing

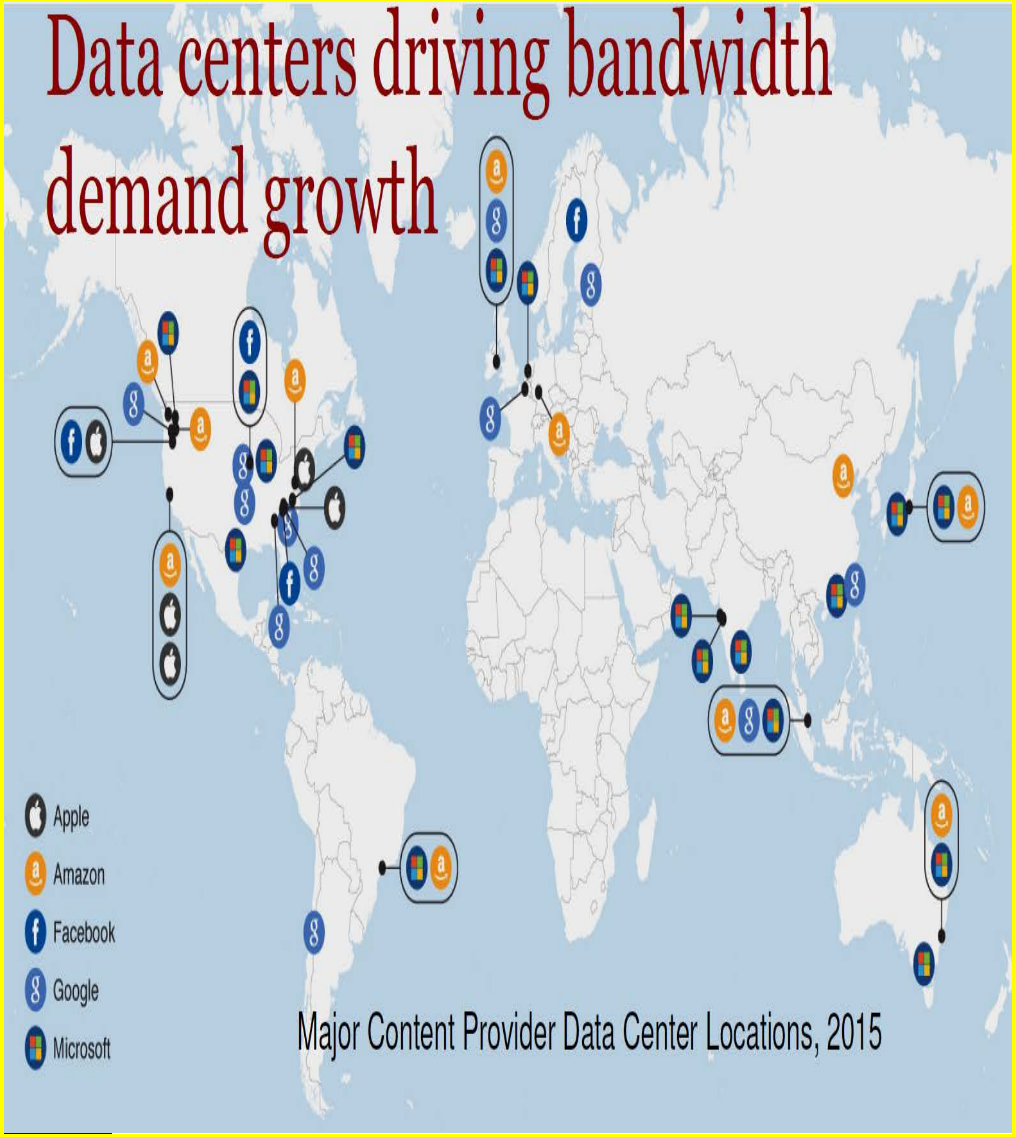
Effect of Shifting Buying Preferences



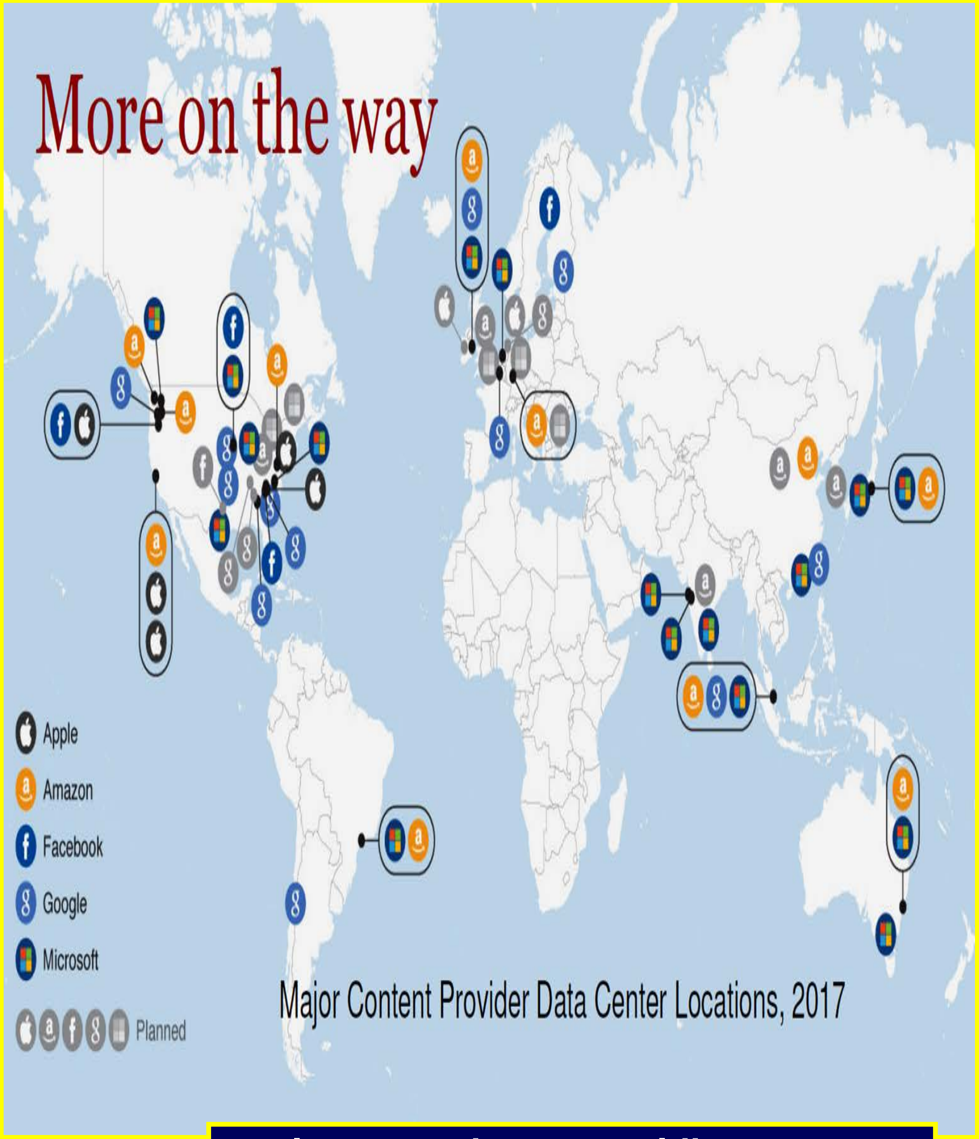
- 100 G prices are now declining faster than 10G prices
(-28% vs -21% CAGR)
- Shifting from 10G to 100G when affordable brings a faster effective drop in unit price per year
- As above 100G/10G price multiple is finally declining, to 5-6 X on main European and Transatlantic routes

Telegeography; Stronge PTC 17

Data Centers Driving the Bandwidth Growth



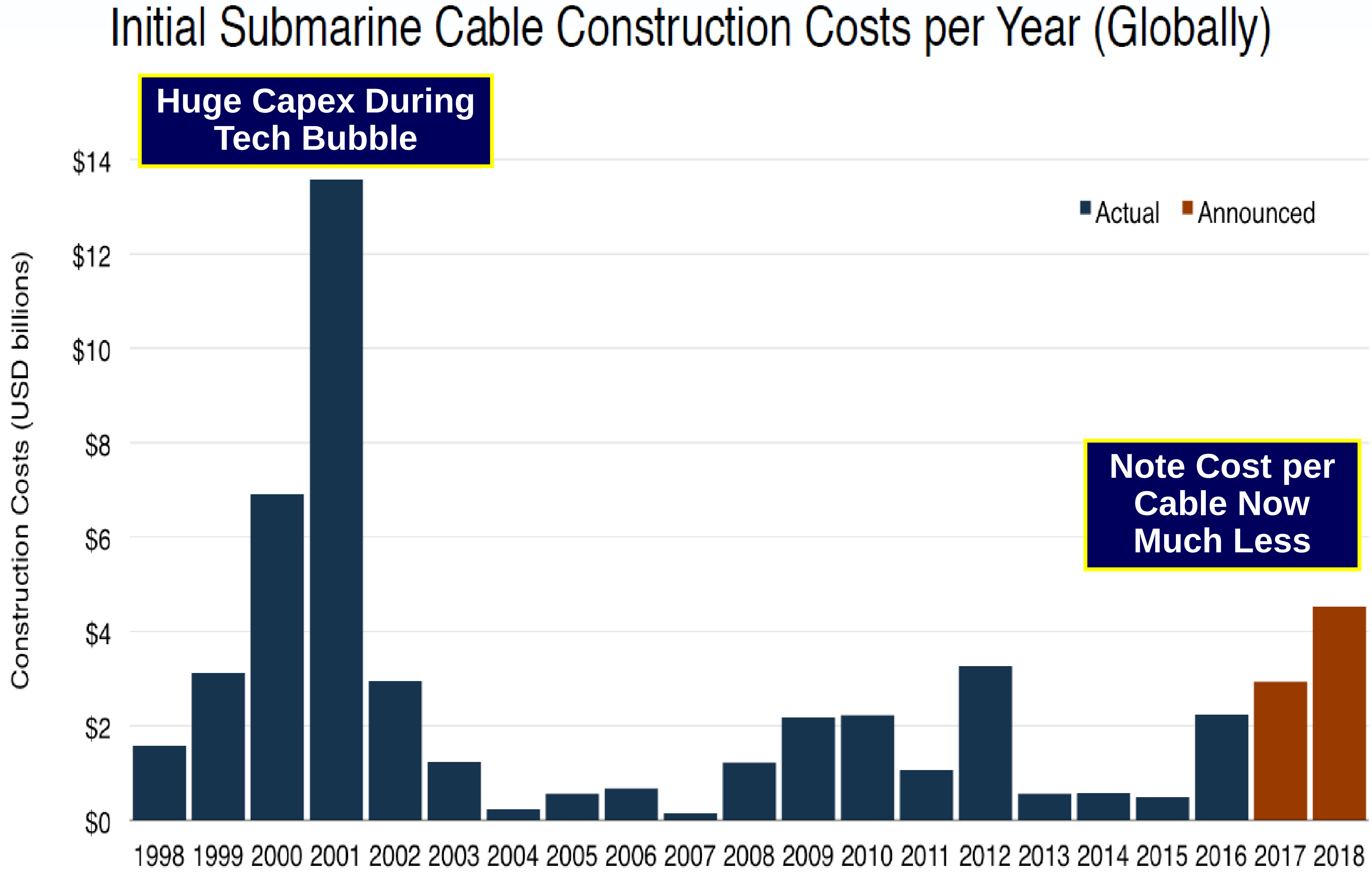
Major Content Provider Data Center Locations, 2015



Major Content Provider Data Center Locations, 2017

Telegeography; A. Maulding, PTC 16

Global Cable Construction On the Rise Again



Huge Capex During Tech Bubble

Note Cost per Cable Now Much Less

Telegeography Stronge, PTC 17

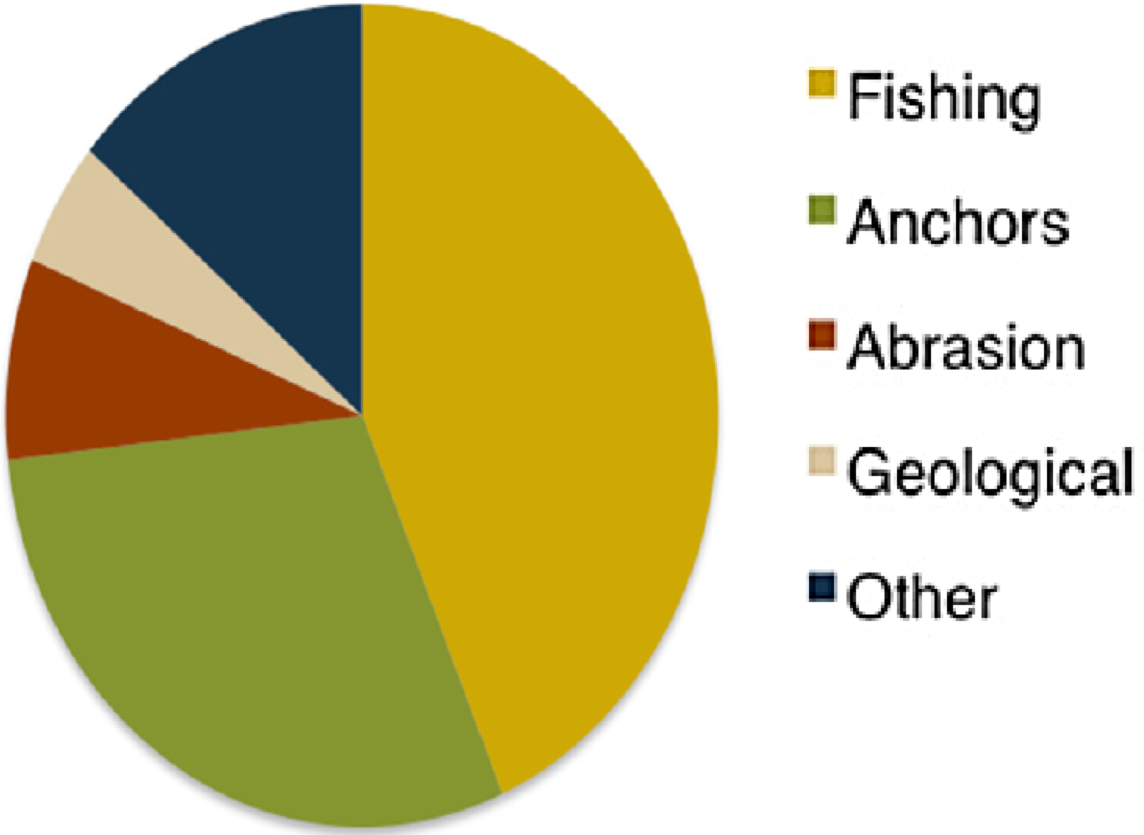
Submarine Cables

Targets of Violence

- Most frequent causes of damage are:
  - ✷ Fishing
  - ✷ Anchors

Not Sharks or Russian Submarines

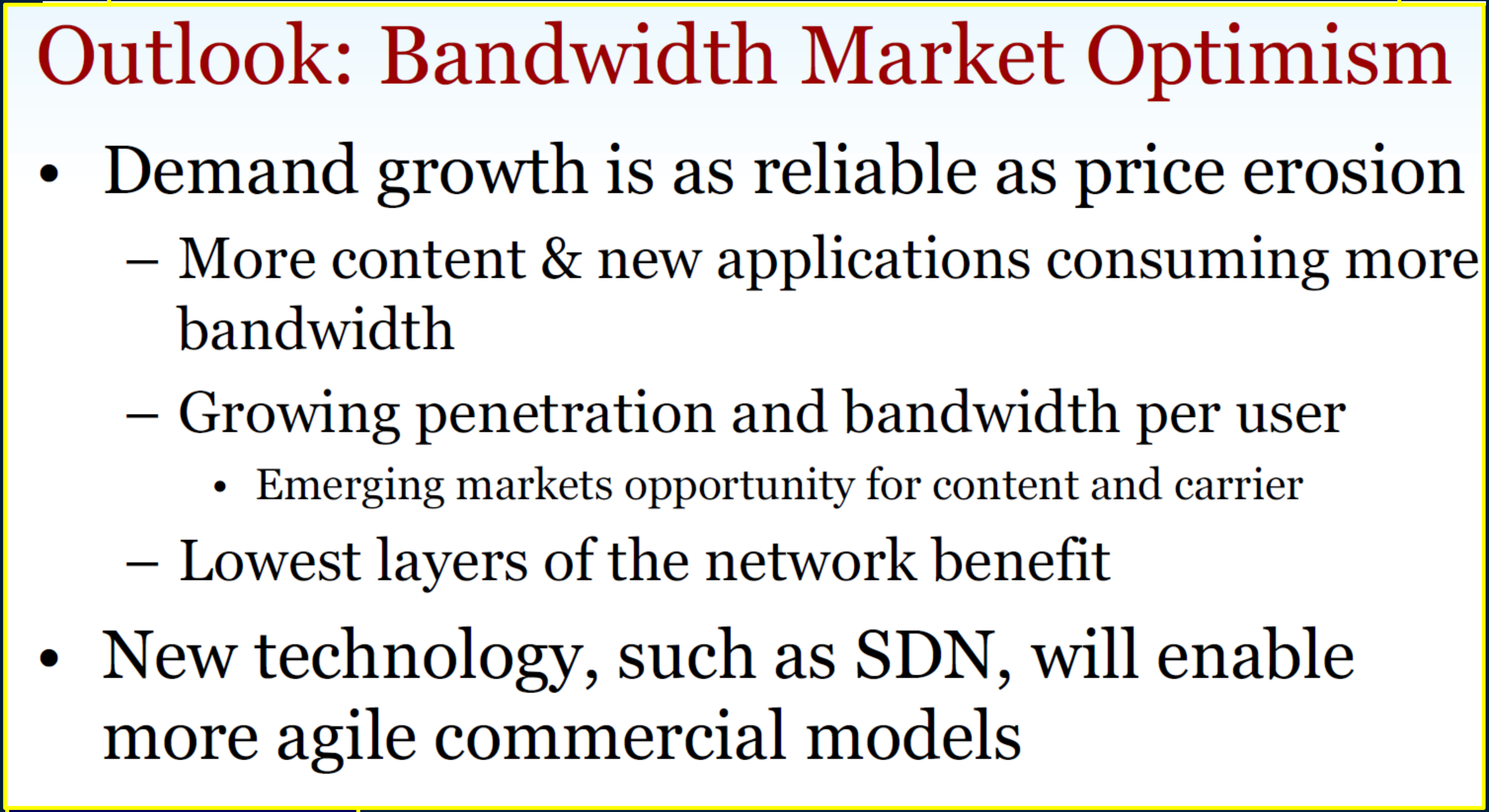
Cause of External Aggression Cable Faults, 2007-2009



Source: *Trends in Submarine Cable System Fault*, M. Kordahi, S. Shapiro, & G. Lucas

Telegeography: A. Maulding at PTC 16

Outlook: Bandwidth Market Optimism



Outlook: Bandwidth Market Optimism

- Demand growth is as reliable as price erosion
  - More content & new applications consuming more bandwidth
  - Growing penetration and bandwidth per user
    - Emerging markets opportunity for content and carrier
  - Lowest layers of the network benefit
- New technology, such as SDN, will enable more agile commercial models

Telegeography:
B. Boudreau PTC 17

How Networks Most Affect Daily Life in the US

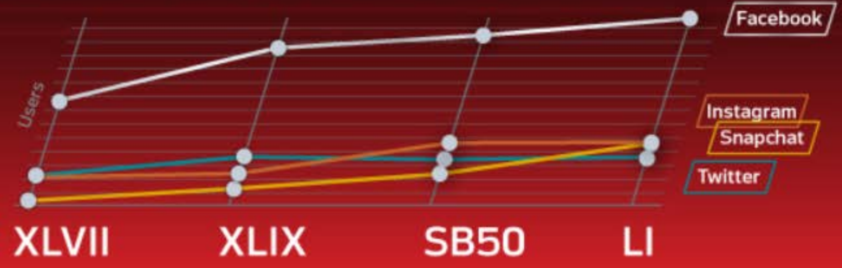
THE MOST SOCIAL SUPER BOWL

1.7 TB

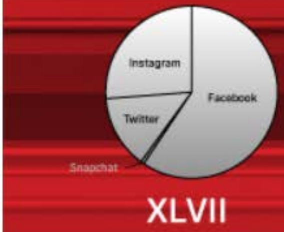
SOCIAL NETWORKING DATA TRANSFERRED

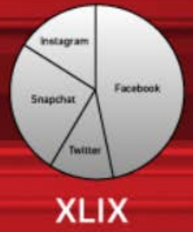
55% INCREASE OVER SUPER BOWL 50!

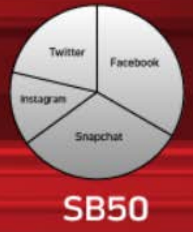
USERS PER SOCIAL NETWORK



SOCIAL NETWORK AGGREGATE BANDWIDTH



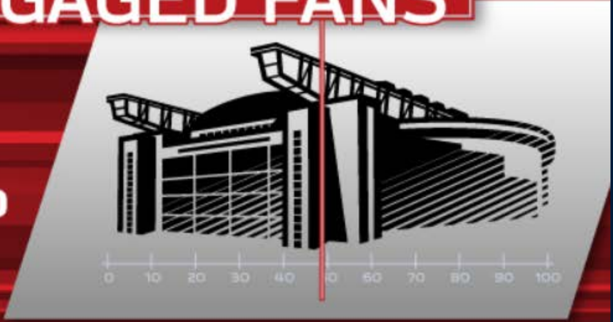




FACEBOOK AND SNAPCHAT COMBINED TO CONSUME ALMOST 10% OF THE TOTAL BANDWIDTH USED

TOTAL ENGAGED FANS

49%



35,430 FANS
WERE ON THE WI-FI NETWORK THROUGHOUT THE GAME

AT PEAK, THERE WERE
27,191
CONCURRENT USERS ON THE NETWORK.

41% MORE THAN SB50

53% MORE THAN XLIX

101% MORE THAN XLVIII

To learn how Extreme can boost your Wi-Fi infrastructure and network analytics experience visit
http://http://extr.co/SB51data

Extreme networks

OFFICIAL WI-FI SOLUTIONS PROVIDER OF THE NFL

DATA SETS COVER THE 8AM TO 1AM REPORTING INTERVAL AND ARE BASED ON PRELIMINARY ANALYSIS AFTER SUPER BOWL LI