100Gbps Networks across the Oceans

Artur Barczyk
California Institute of Technology

CHEP 2012 conference
New York, May 23rd, 2012

OUTLINE

The challenges of 100G (and beyond)

The technologies

Terrestrial Deployments

Application View

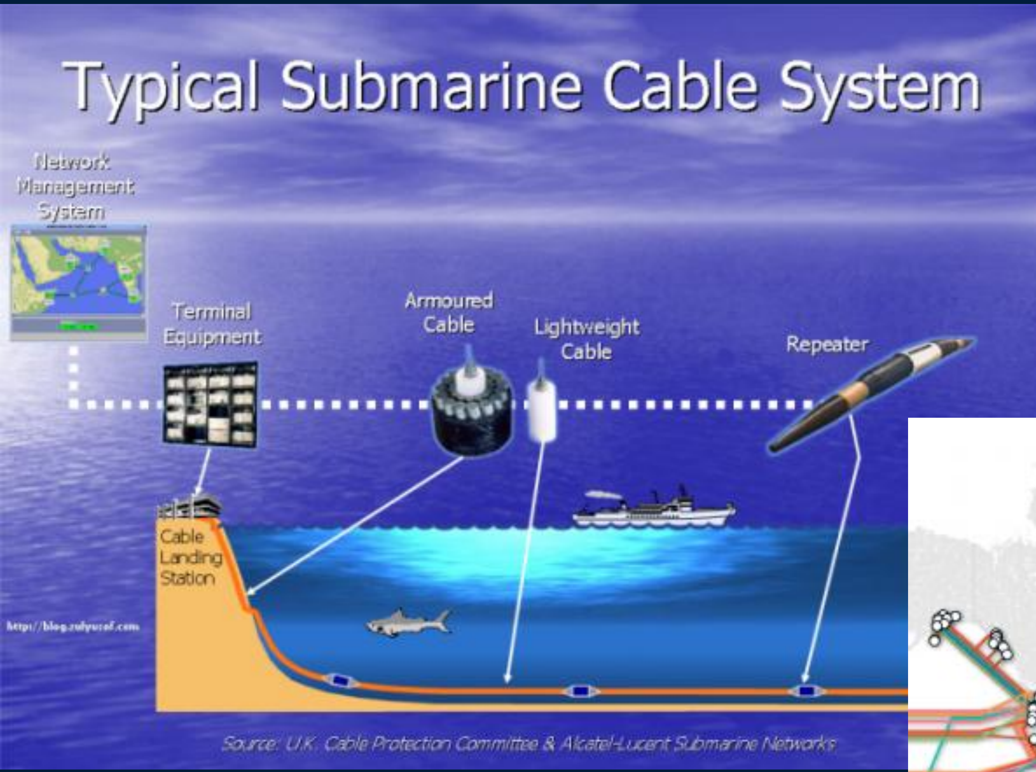
Submarine Systems

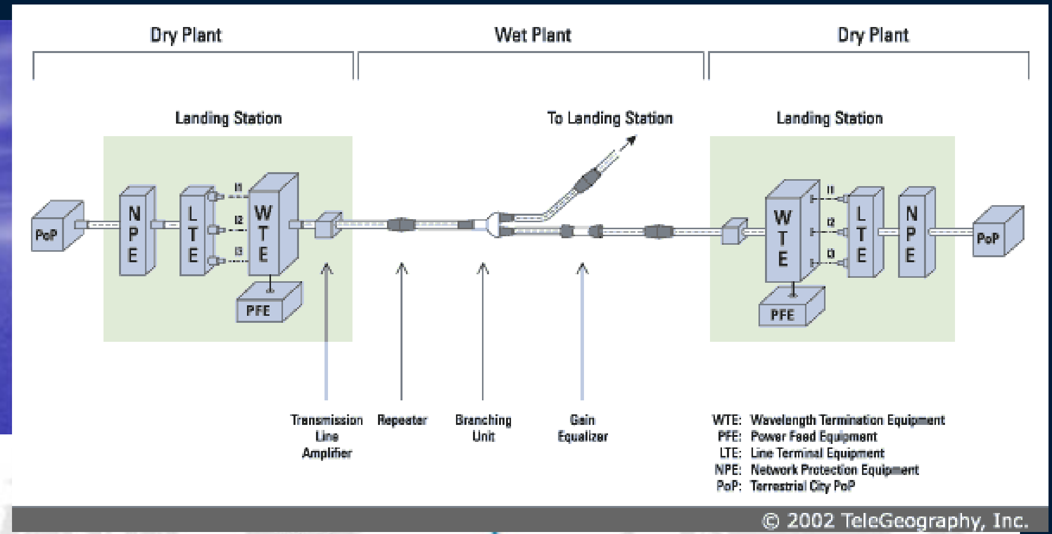
THE CHALLENGES OF ≥100 GBPS ULH TRANSMISSION

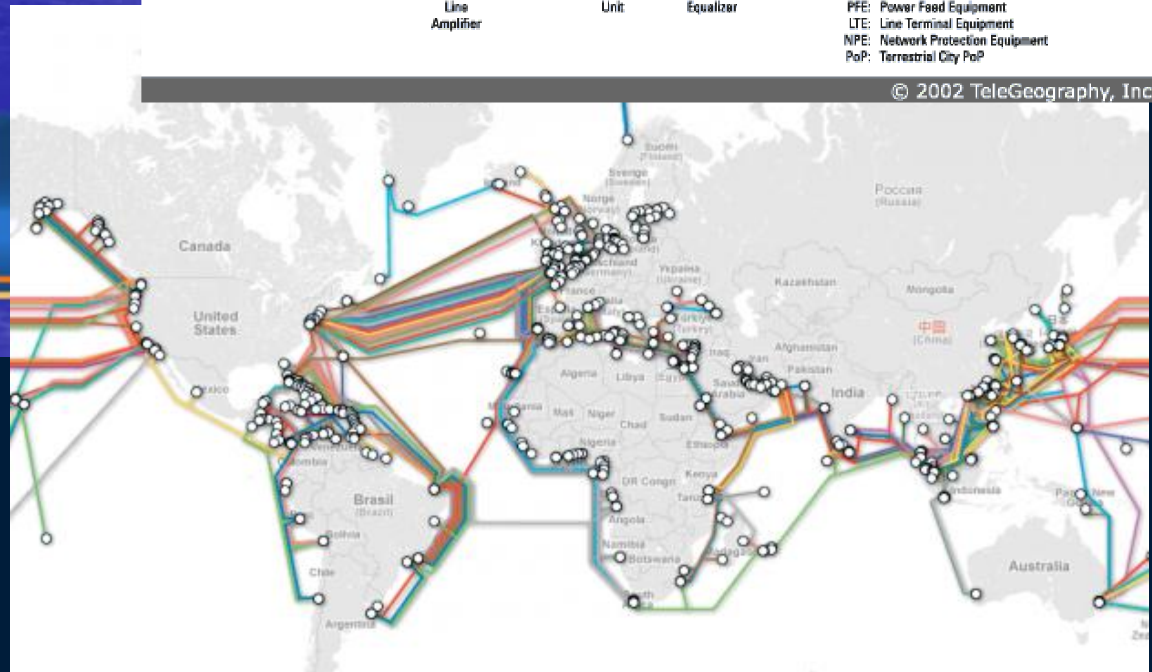
Increasing capacity

Ultra long-haul; crossing the oceans

Submarine Communication Systems





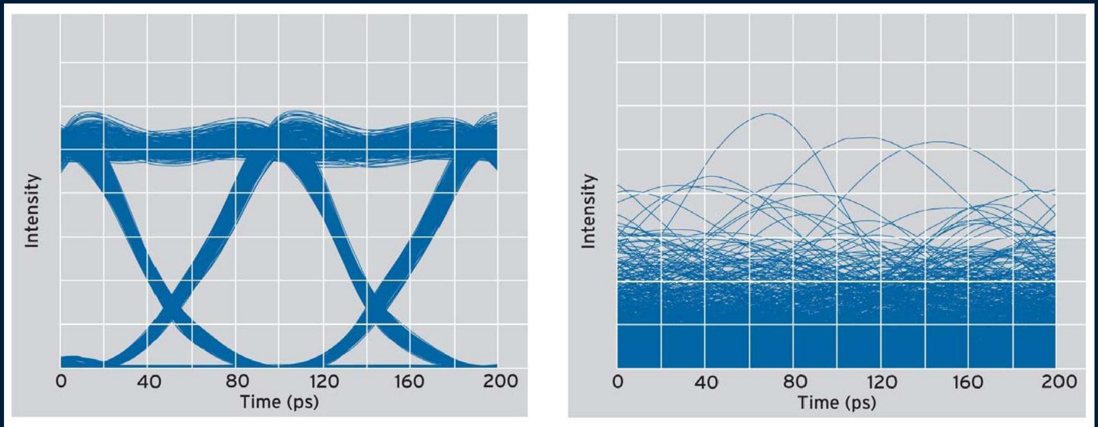


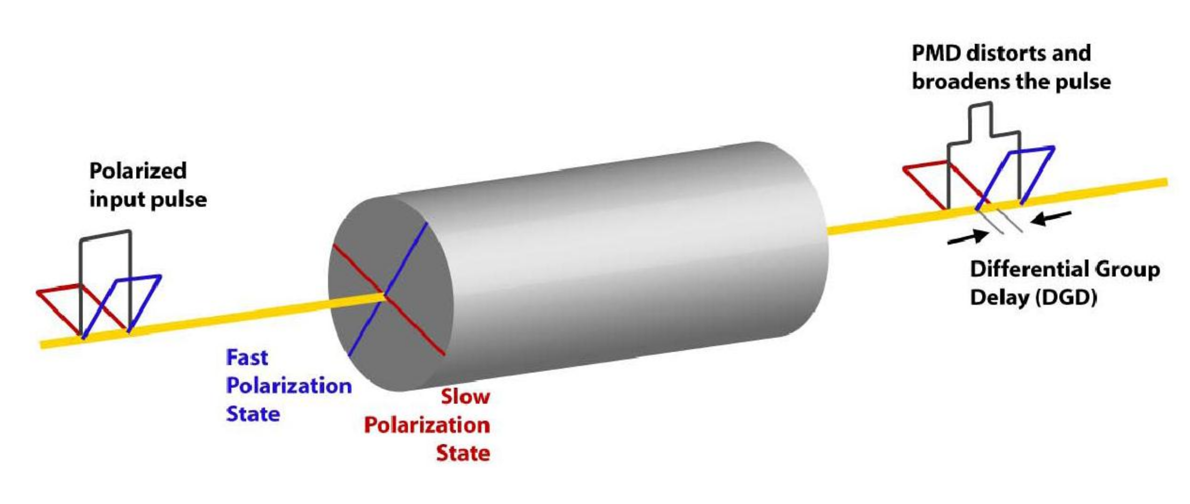
Transmission distances:
>6000 km (Atlantic)
>10000 km (Pacific)

http://www.submarinecablemap.com/

The challenge going above 10G

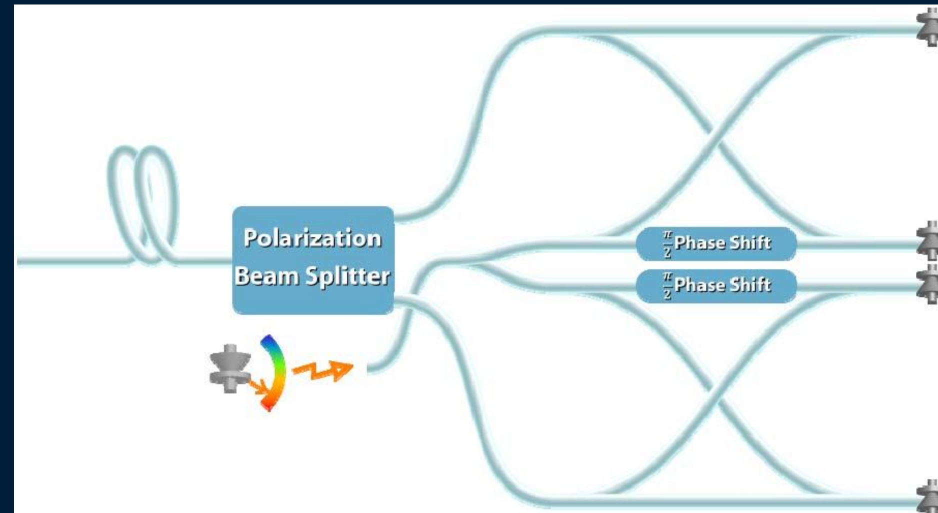
- **Challenges in (U)LH transmission – reach limiting factors:**
- **Signal attenuation; OSNR**
- **Dispersion**
  - Chromatic
  - Polarization
- **Nonlinear effects**
  - Stimulated Raman Scattering
  - Stimulated Brillouin Scattering
  - …

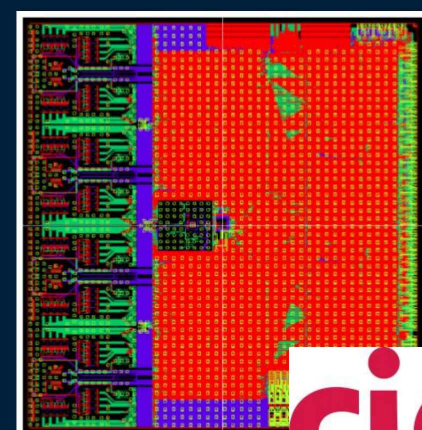




Increasing Bit Rate; Enabling Technologies

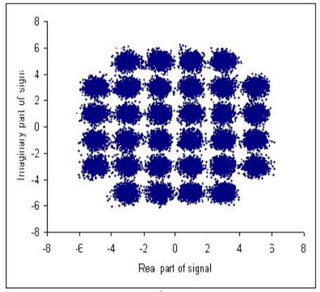
- **Coherent detectors**
  - Coupled local oscillator signal
  - enabling technology for phase keying in optical domain
- **High speed ADCs**
- **Digital Signal Processing**
  - Modulation, baud rate reduction
  - Dispersion compensation
- **Forward Error Correction**
  - Increase reach at lower launch power
- **Use two polarization states**
  - Baud rate reduction





Increasing Link Capacity

QAM, M-ARY

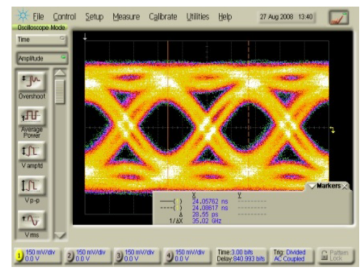


Three mechanisms to grow capacity

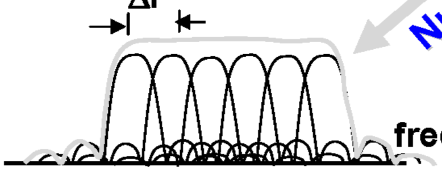
- → Challenge the baud rate
- → Challenge the bit/symbol
- → Challenge fixed λ spacing

Bits per symbol

Symbols per second



Number of carriers



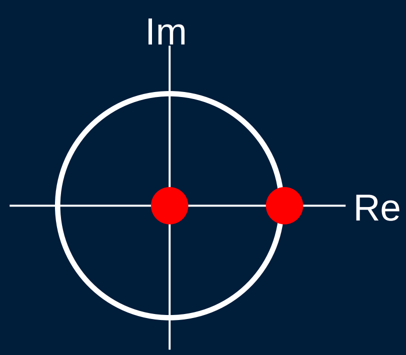
From Jan Willem Elion; CIENA

Flexibility in Spectrum Optimization Enabled by DSP

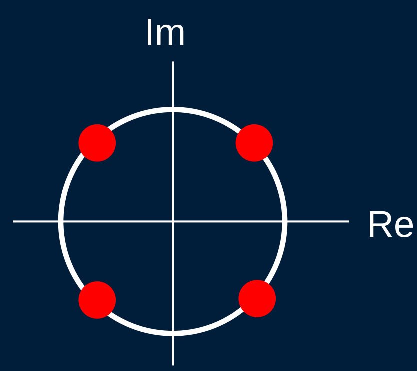
Modulation for >10G

- **On/Off keying is hard above 10Gbps**
  - 100Gbps:10ps bit interval, optical non-linearities
  - Increased sensitivity to dispersion (chromatic, PMD)

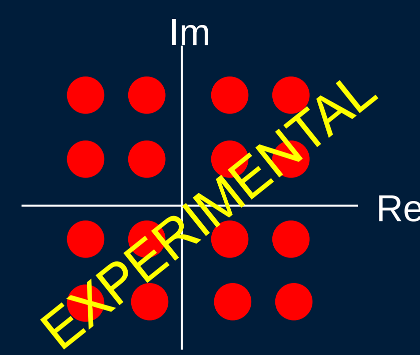
- **Complex modulation techniques known from wireless/radio transmission**
  - Use Amplitude and Phase



ON/OFF



Quadrature phase-shift keying
2 bits per symbol



DP 16-QAM
8 bits per symbol

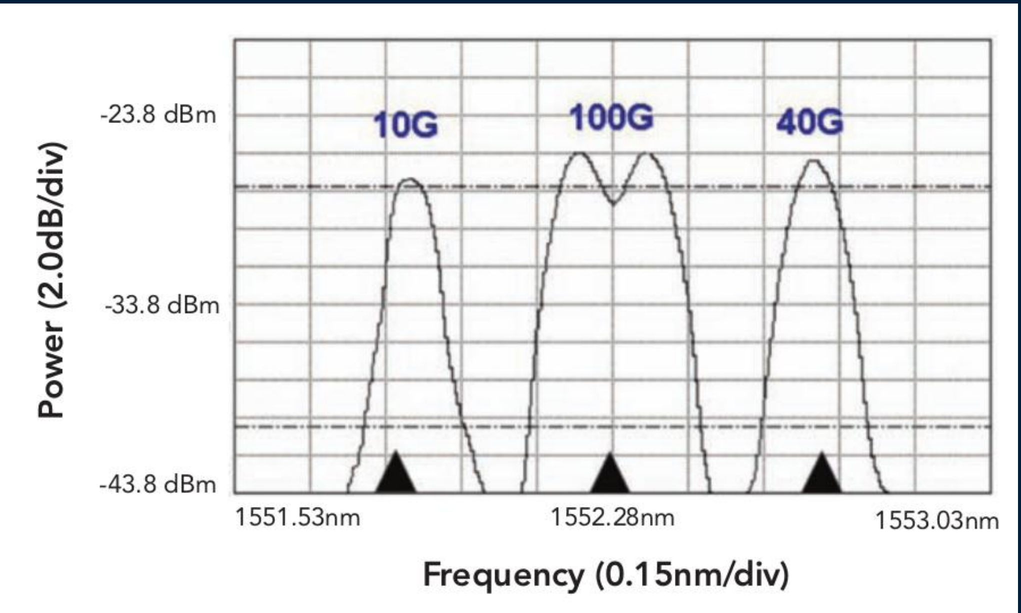
Dual-carrier and Flexible Grid

Use of two sub-carriers reduces symbol rate by half

Dual-polarized,
Dual sub-carrier

Single-polarized

Dual-polarized

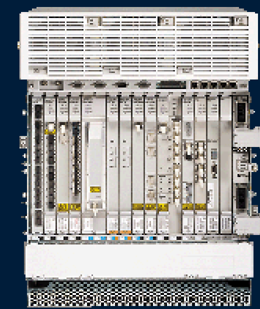


50GHz ITU Grid

1 Tbps transmission will require flexible frequency grid

Commercial 100G Optical Transport and Switching

- **Transport gear widely available**
  - Ciena (pioneering the field, as did Nortel, now also Ciena)
  - Infinera, Huawei, Fujitsu, Adva, Cisco, …



- **Routing and switching**
  - Brocade, Juniper, Cisco, Alcatel-Lucent, Ciena (switching), …

- **100G is being deployed today in commercial as well as R&E backbones**

100G TERRESTRIAL DEPLOYMENTS

100Gbps in NRENs

ESnet5 Routed Network November 2012
DRAFT

- SUNN — ESnet PoP/hub locations
- (100) ESnet managed 100G routers
- (10) ESnet managed 10G router
- (10) (100) Site managed routers
- LOSA — ESnet optical node locations (only some are shown)
- ESnet optical transport nodes (only some are shown)
- commercial peering points
- R&E network peering locations
- LBNL — Major Office of Science (SC) sites
- LLNL — Major non-SC DOE sites

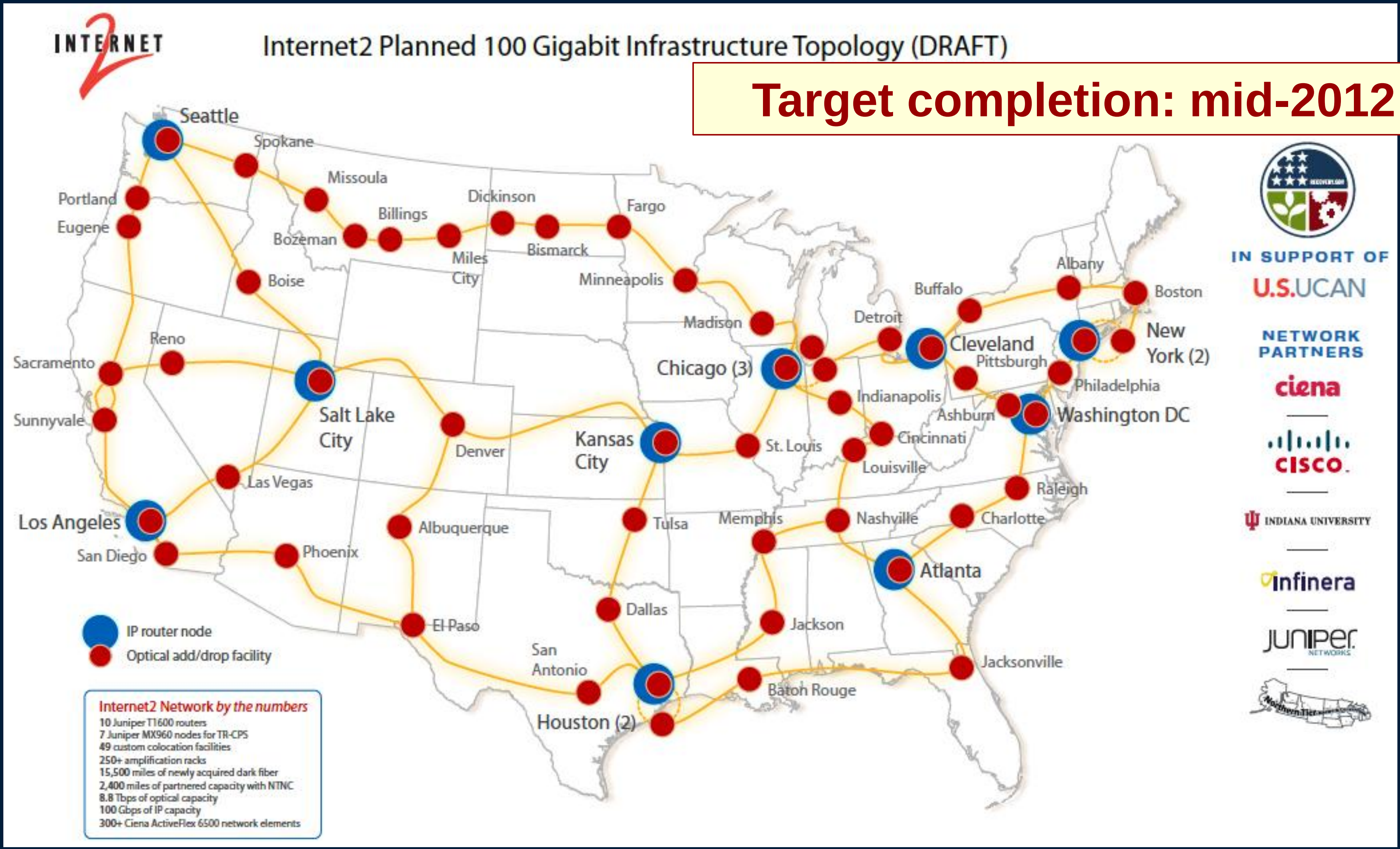
- Routed IP 100 Gb/s
- Routed IP 4 X 10 Gb/s
- 3rd party 10Gb/s
- Express / metro 100 Gb/s
- Express / metro 10G
- Express multi path 10G
- Lab supplied links
- Other links
- Tail circuits

Joe Metzger, ESnet

Geography is only representational

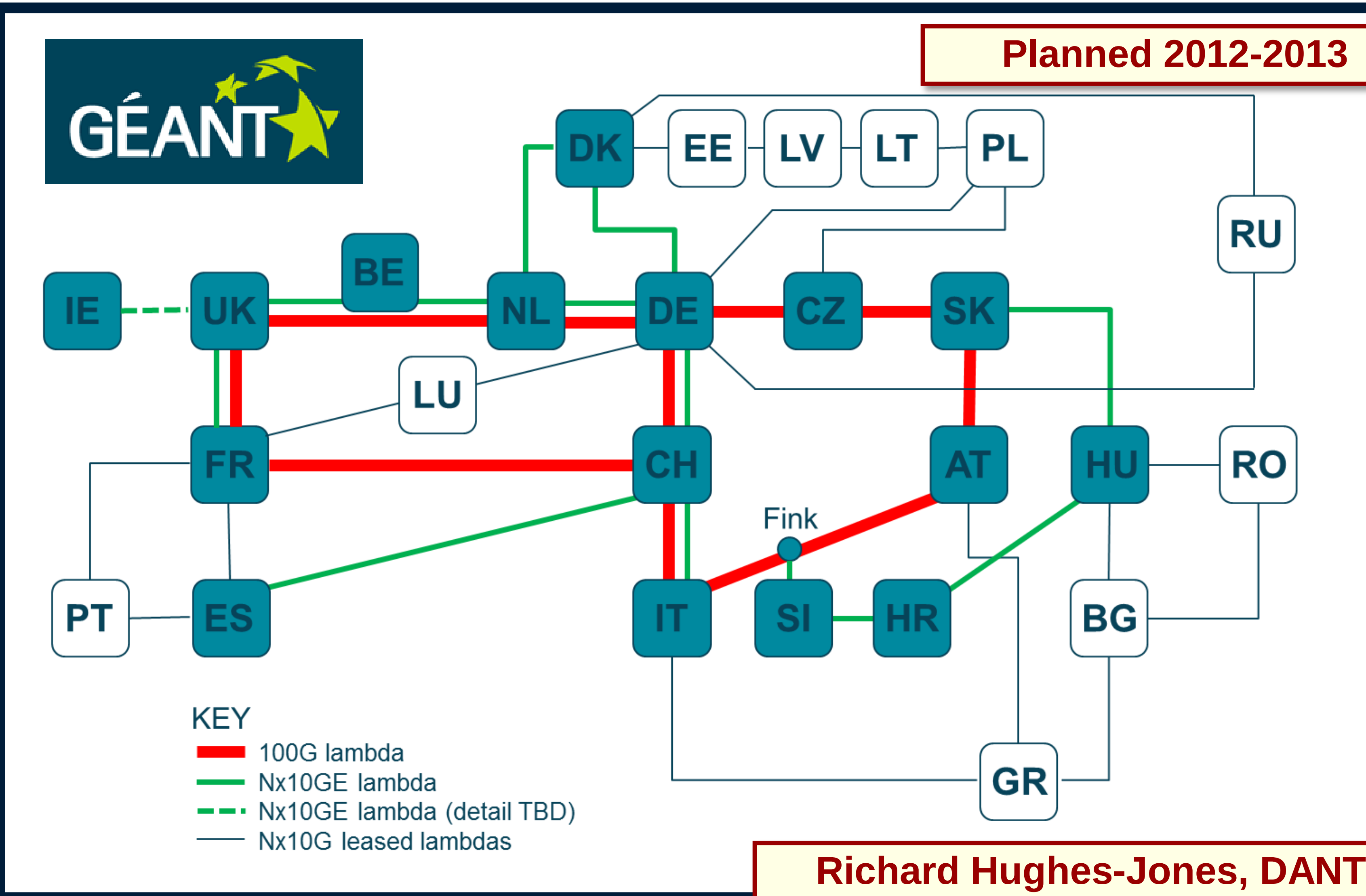
Internet2 100G Backbone Deployment, ongoing

Target completion: mid-2012



100G across Europe

Planned 2012-2013



Richard Hughes-Jones, DANTE

100G Deployement in R&E Networks (cont.)

- **DOE ESnet**
- **Internet2**
- **CANARIE**
- **SURFnet/CERN Dark Fibre link between Geneva and Amsterdam**
  - SURFnet's PoP at CERN
  - 100G wave between Open Lightpath Exchanges: Nertherlight-CERNLight
  - In addition to multiple 40G waves on same fibre
- **RENATER**
- **RoEduNet**
- **Future: CERN-Wigner Institute (Budapest)**
  - Planned 2 x 100G connection for remote Tier0 connection

Geneva
(not to scale)

APPLICATIONS OF 100G NETWORKS

The advantages of 100G

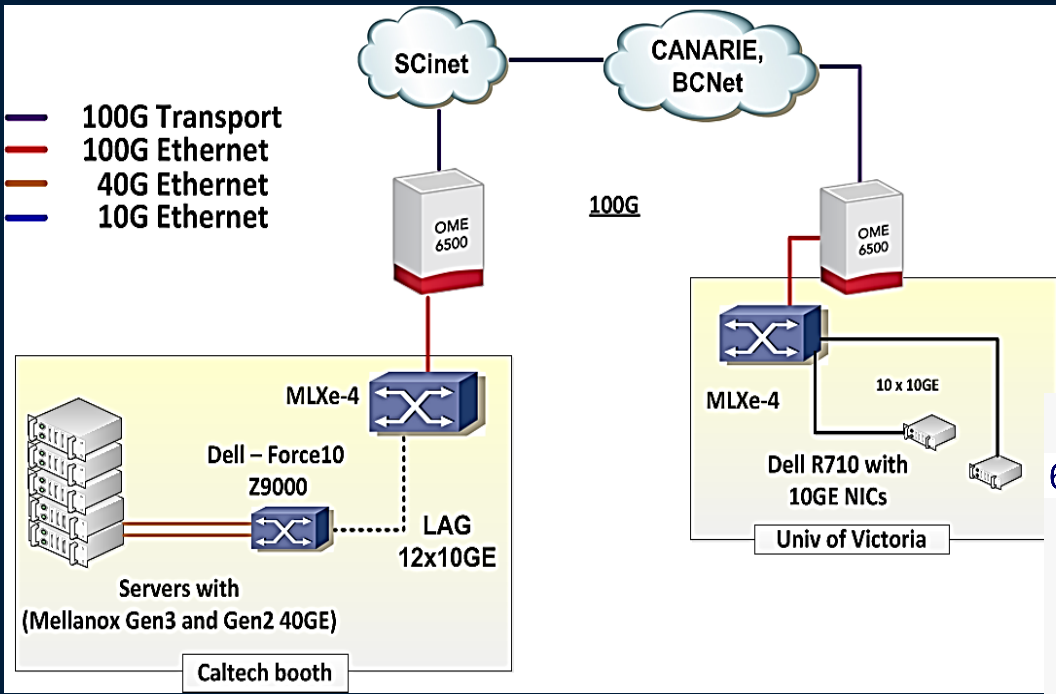
- **Bandwidth**
  - E.g. (HD) Video on demand: YouTube, Netflix, etc. creates unprecedented increase in bandwidth in the Internet
  - New scientific instruments and applications (genomics, climate modeling, SKA, …)
- **Latency**
  - E.g. automated trading
- **Management; OPEX**
  - Replacing N 10G links with n 100G links (N>n) reduces operational overhead and complexity
- **For operators: spectral efficiency**
- **CAPEX**
  - 100G interface/wave service cheaper than 10x 10G

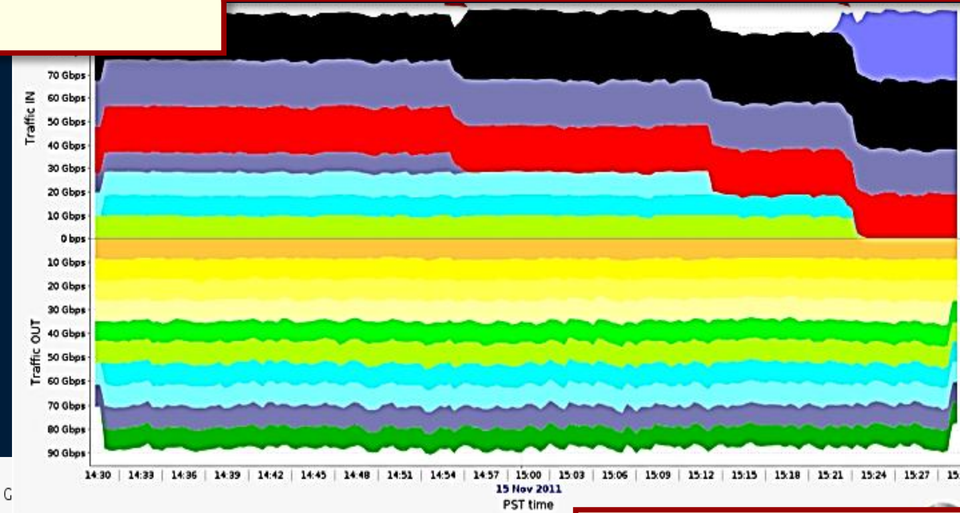
100G Demo at SuperComputing 2011

US LHCNet

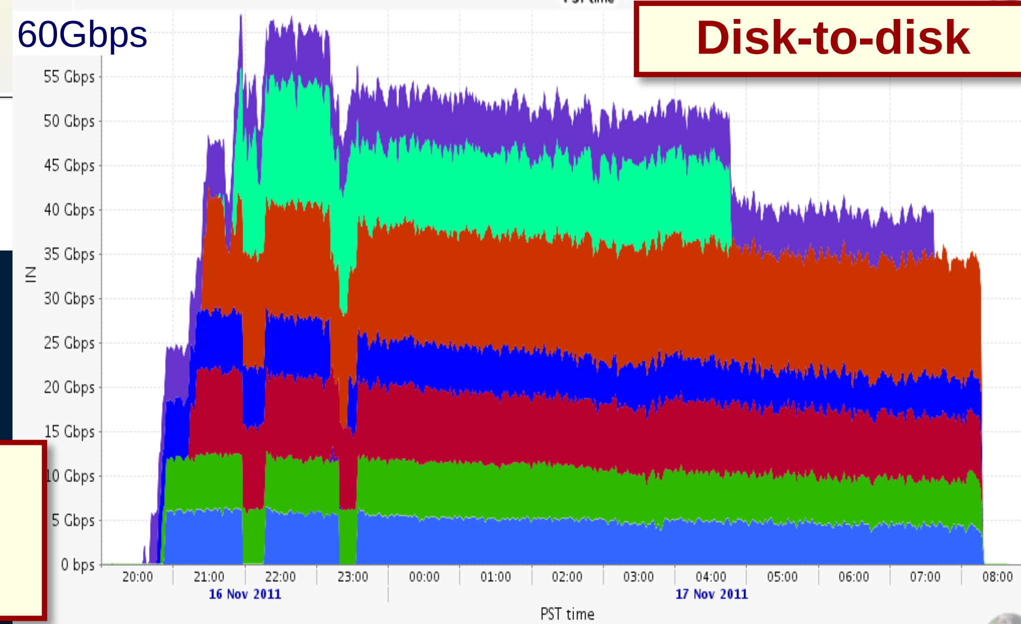
Caltech and University of Victoria (with partners):
100G demo at SuperComputing 2012

Memory-to-memory: 186 Gbps





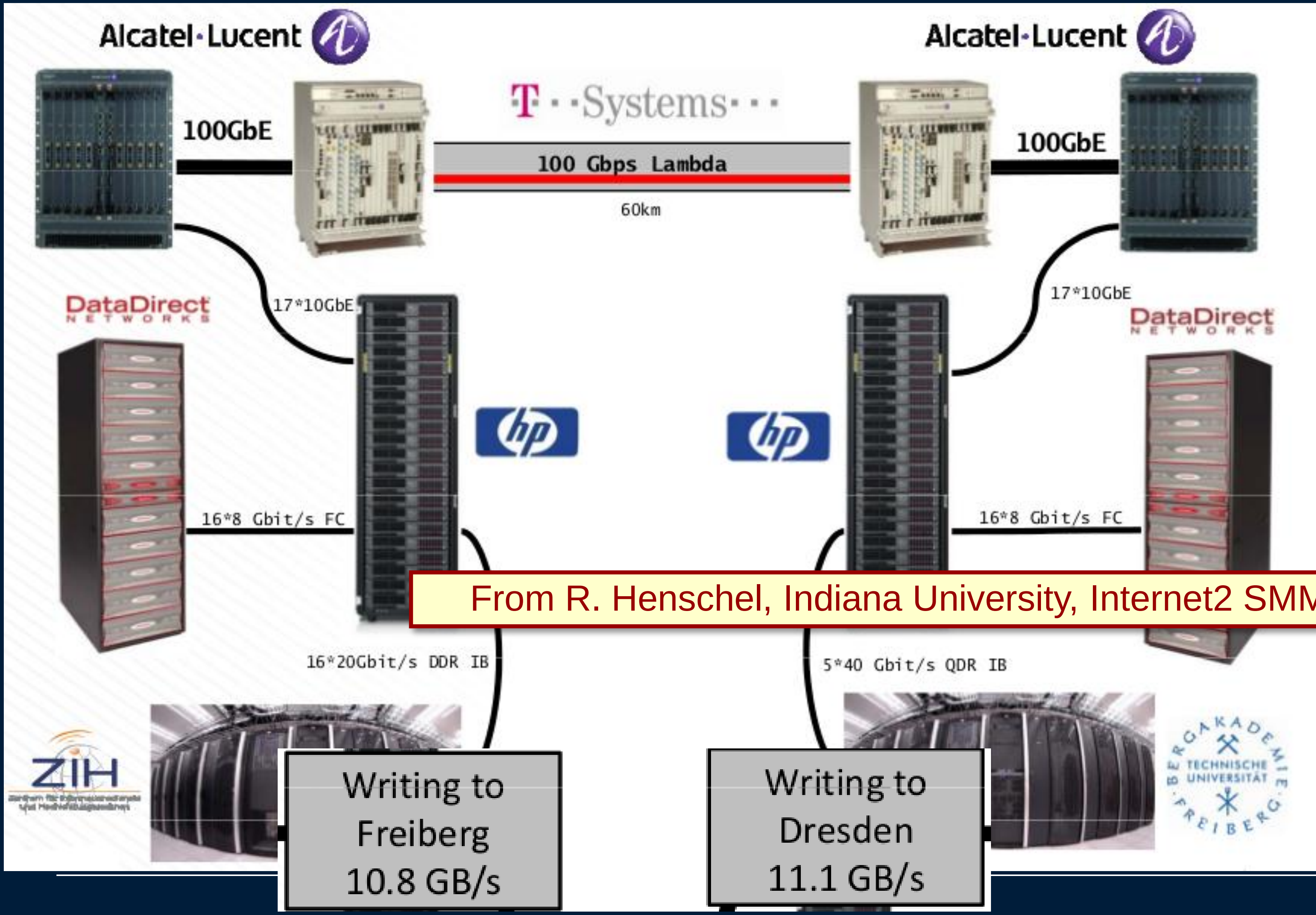
Disk-to-disk



For details: see Poster by Ian Gable et al. at this conference

Used servers with 40GE NICs with PCIe v3: 38Gbps per NIC

Lustre @ 100G Demo
between Dresden and Freiberg (60km)



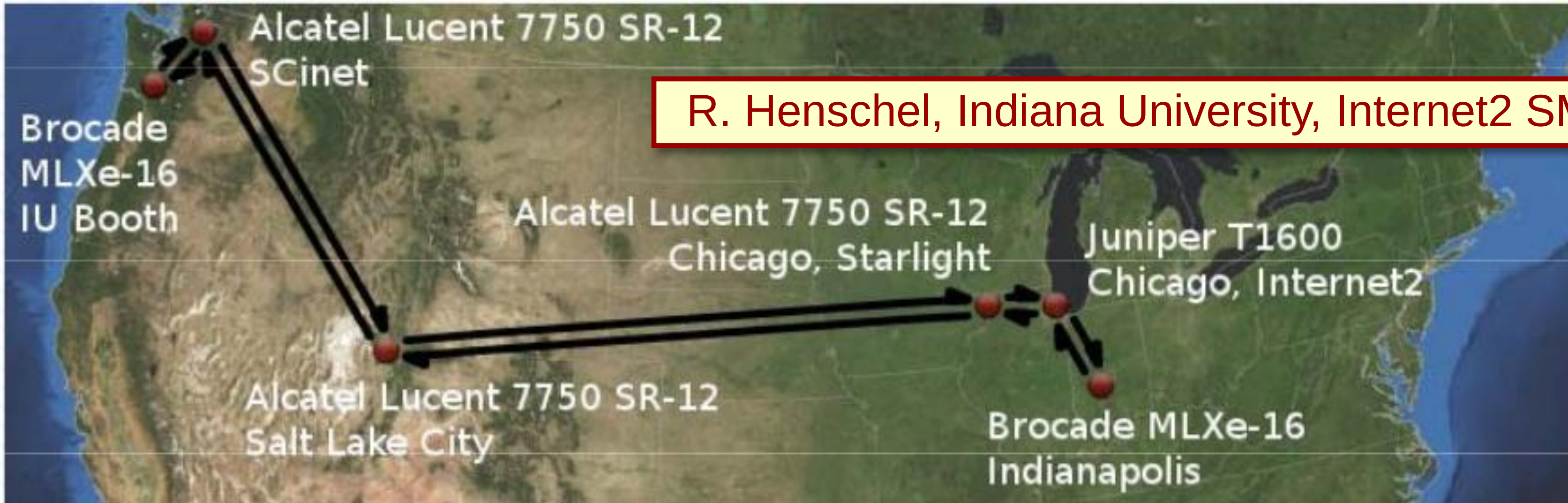
From R. Henschel, Indiana University, Internet2 SMM 2012

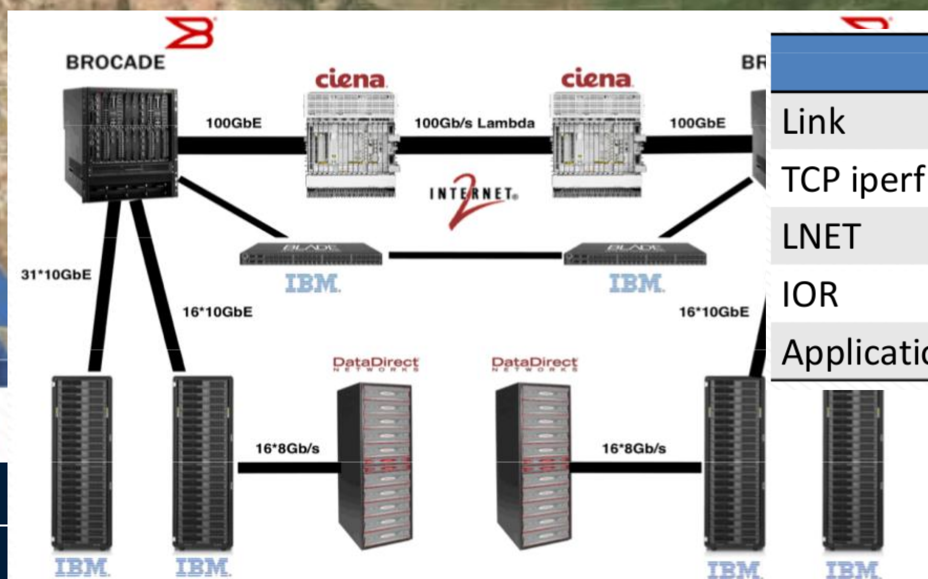
Lustre @ 100G Demo at SC'11 IU+ESnet+Internet2, 2100km

US LHCNet

- Internet2 and ESnet, 50.5 ms RTT

R. Henschel, Indiana University, Internet2 SMM 2012





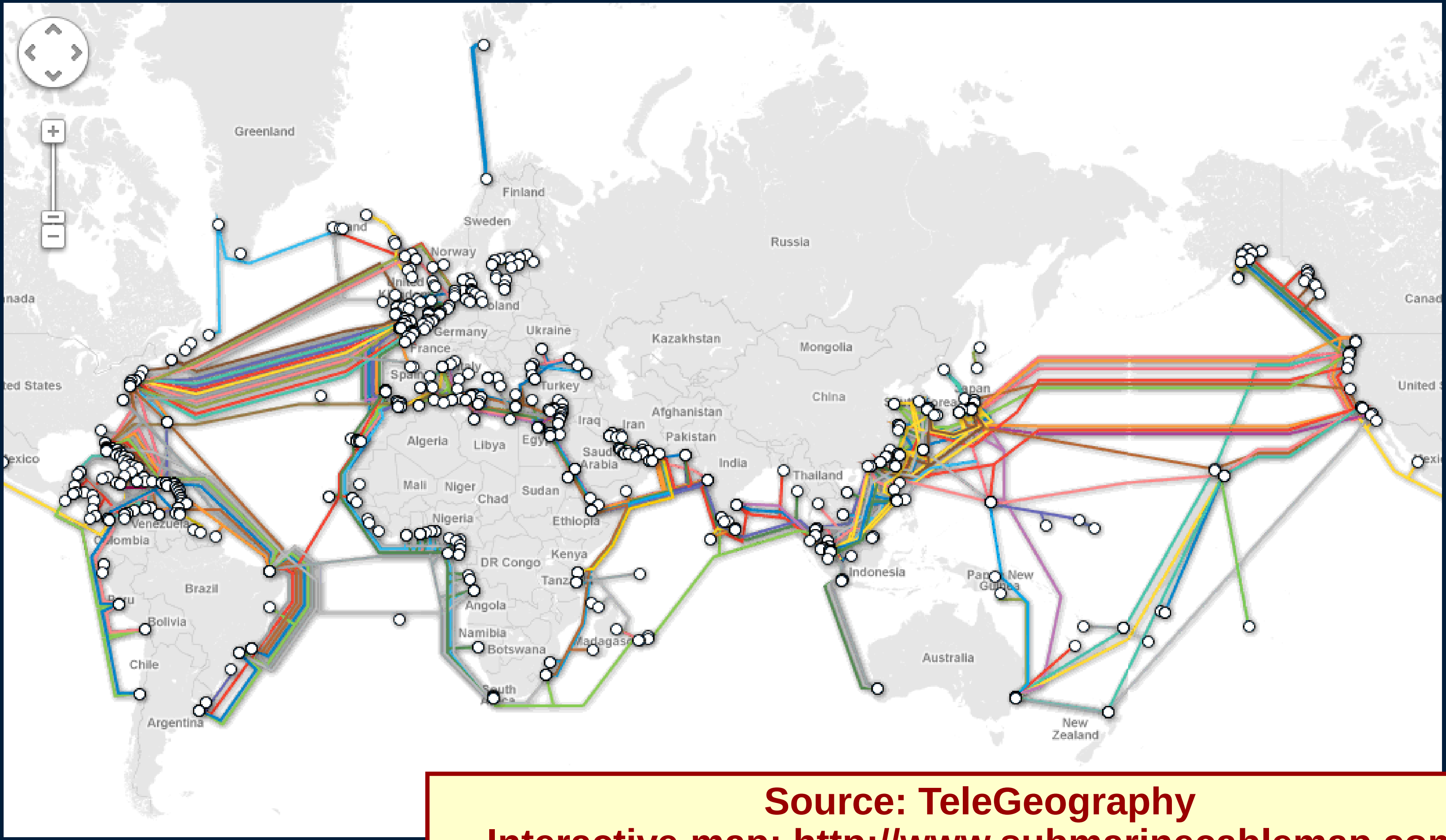
	Measurement	Efficiency	Test Time
Link	50.5 ms		4
TCP iperf	96 Gb/s	96%	6
LNET	9.4 GB/s	77%	2
IOR	6.5 GB/s	52%	2
Applications	6.2 GB/s	50%	1

RESEARCH TECHNOLOGIES

Robert Henschel

BACK TO SUBMARINE DEPLOYMENTS

Submarine Cable Systems



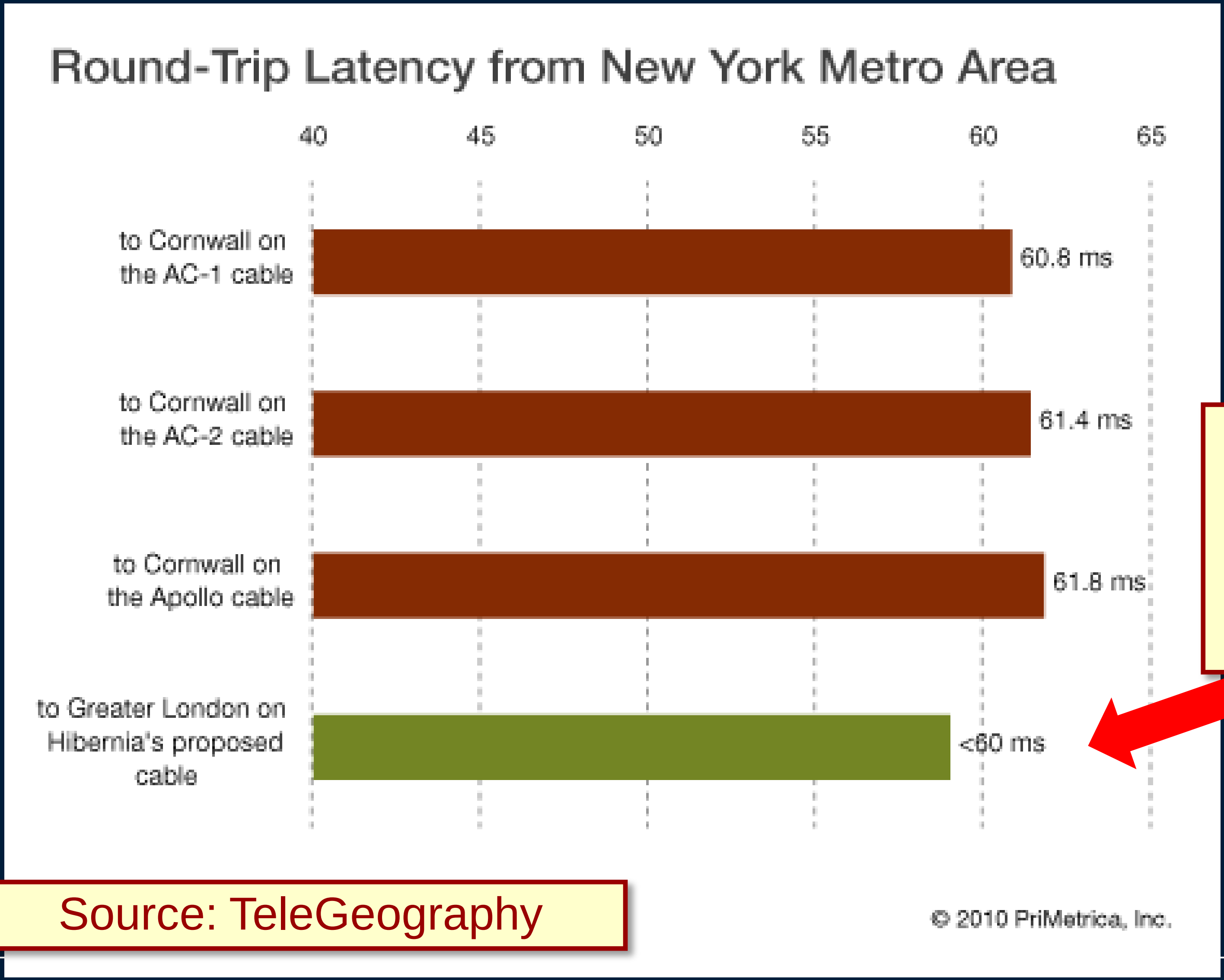
Source: TeleGeography
Interactive map: http://www.submarinecablemap.com/

40G Deployments

- **Question of reach, mainly, but also technology**
  - Some cable systems could upgrade to 40; not yet 100G
  - 40G: DP-QPSK → baud rate similar to 10G NRZ
- **Today 40G installed on at least 3 transatlantic systems**
  - Hibernia pioneered the deployment in 2009, targeting mainly low-latency trading business
  - Global Crossing and Tata followed suit
  - Mainly to increase total cable capacity

- **Hibernia successfully trialed 100G single wave in 50GHz spacing between New York and London**
- **Paving way for upgrade of existing Hibernia cable in 2013**

New cable systems: the quest for low latency



Hibernia Express

Target: financial sector, automated trading

Source: TeleGeography

New Transatlantic Systems under construction

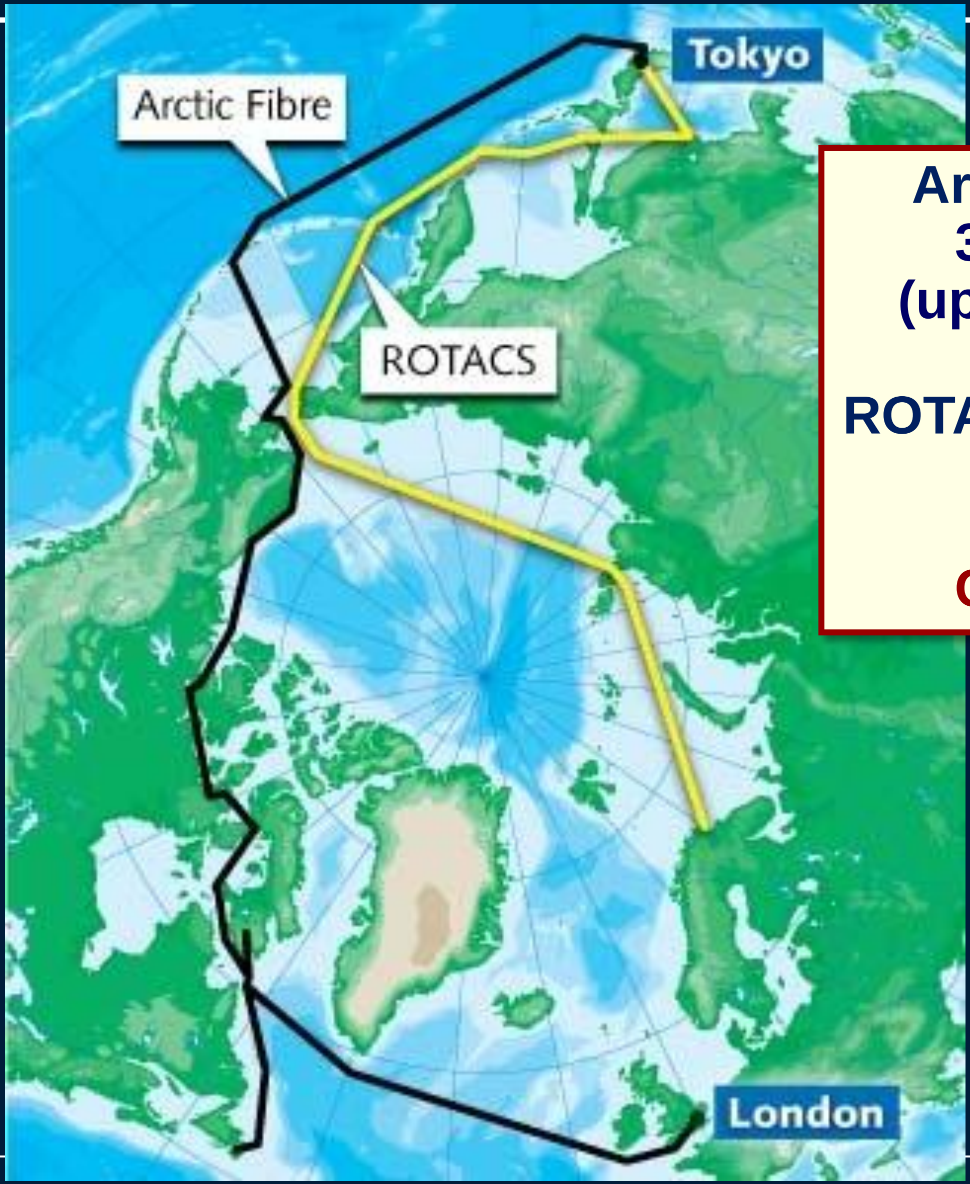
US LHCNet

Hibernia Atlantic: Hibernia Express
RTT < 60 ms
Expected Service Start date: Early 2013

Emerald Networks: Emerald Express
RTT < 62 ms
Capacity: 6x 100 waves x 100 Gbps
Expected Service Start date: Spring 2013

Still doubt Global Warming…?

US LHCNet



Arctic Fibre: Northwest Passage
3 pairs x 80 waves x 40 Gbps
(upgradeable to 100 Gbps waves)

ROTACS: (closer to) great circle route
6 x 100 waves x 100Gbps

Completion expected in 2014

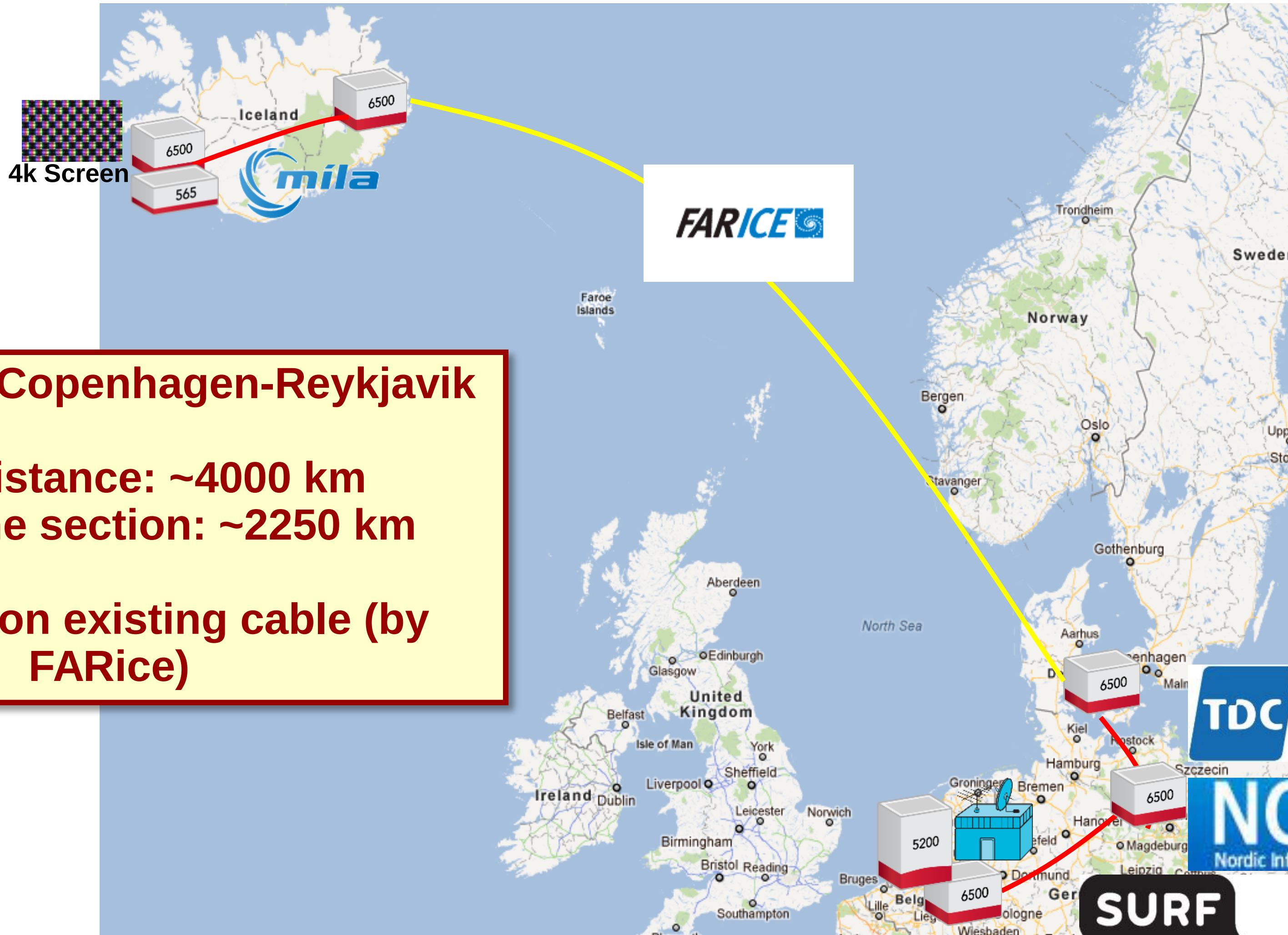
Tokyo-London latency:

Today: ~250 ms
ROTACS: ~155 ms
Arctic Fibre: ~168 ms

… but when will I get it?

- **Transition to use of 100G circuits driven by**
  - Capacity needs (obvious)
  - Economic factors
    - Compared to 10x 10G
    - Necessary hardware upgrade } → TCO
    - Lower operational costs
  - Diversity
    - One 100G circuit is big investment, what if it fails?
    - Protected service (expensive)
    - Multiple lines (capacity justification?)
    - Backup with (Nx) 10G? (service levels?)

This week: TERENA conference demo - 100G Denmark-Iceland



Amsterdam-Copenhagen-Reykjavik

Total distance: ~4000 km
Submarine section: ~2250 km

Deployed on existing cable (by FARice)

Summary

- **Terrestrial deployment of 100G in the WAN is growing fast**
  - In many R&E networks' backbones
- **Ultra-long haul 100Gbps transmission has been a challenge**
  - Coherent detection key enabling technology
- **New transatlantic 100G cable systems under construction**
  - Planned upgrades to 100G on other, existing systems
- **Transatlantic: will emerge in 2013**
- **End-systems demonstrated ability to use 100G efficiently**

100Gbps networks are reality. In LAN, MAN, WAN.
They are key component in the design of modern, distributed scientific instruments. 100G alone is not enough to guarantee adequate service levels in the age of exascale computing.
Watch out for other key developments; new emerging standards, Software Defined Networking, NSI, etc. – bringing networks closer to the application.

THANK YOU!

Artur.Barczyk@cern.ch