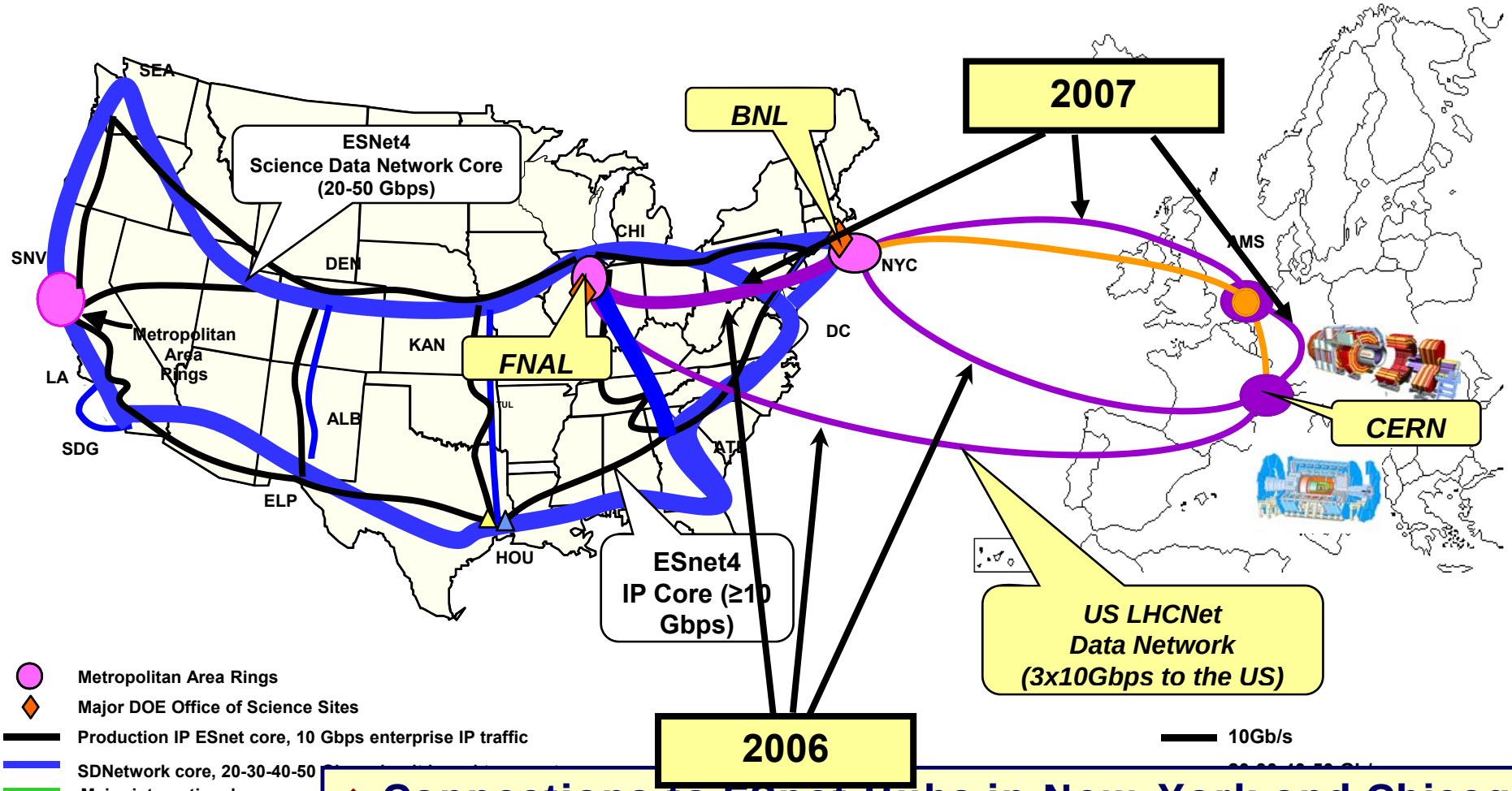
US LHCNet Update

Dan Nae

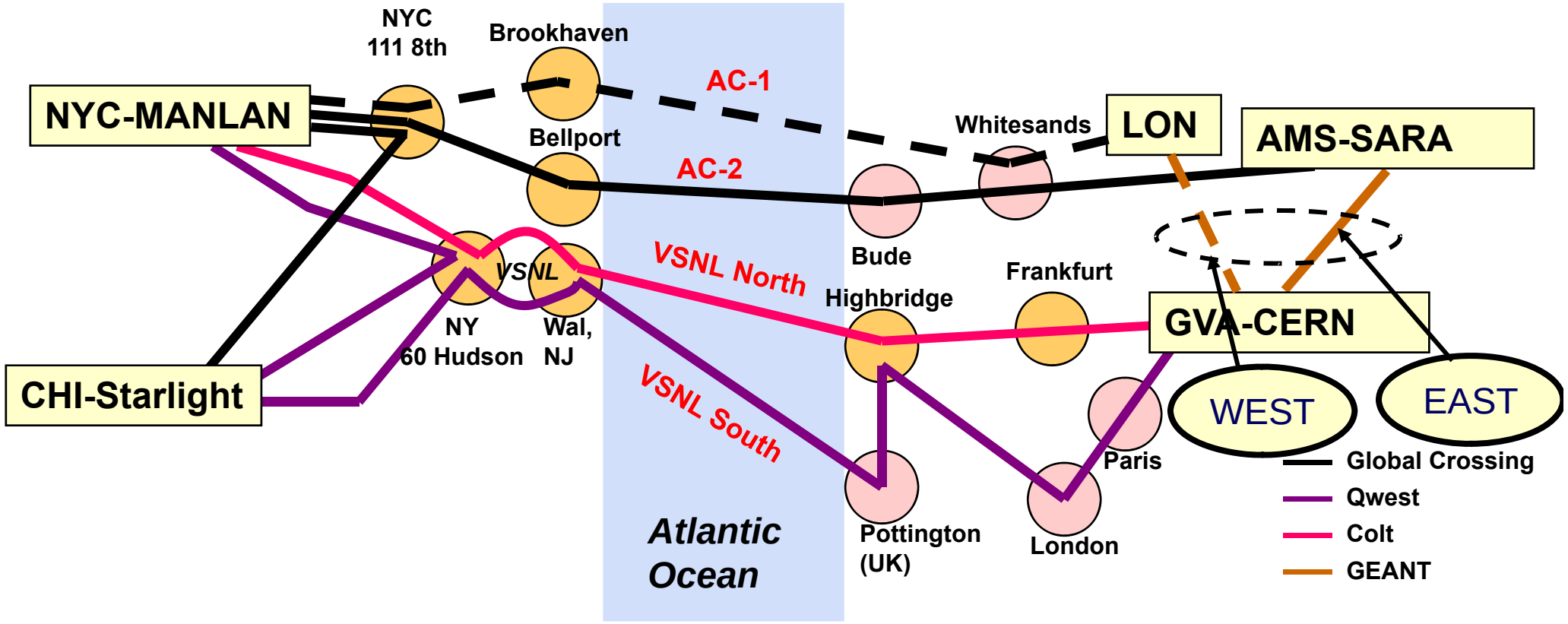
California Institute of Technology

US LHCNet + ESnet (2007)



- Connections to ESnet Hubs in New-York and Chicago
- Redundant “light-paths” to BNL and FNAL
- Redundant 10 Gbps peering with Abilene
- Access to USNet / HOPI for R&D

Multiple Fiber Paths: Reliability Through Diversity



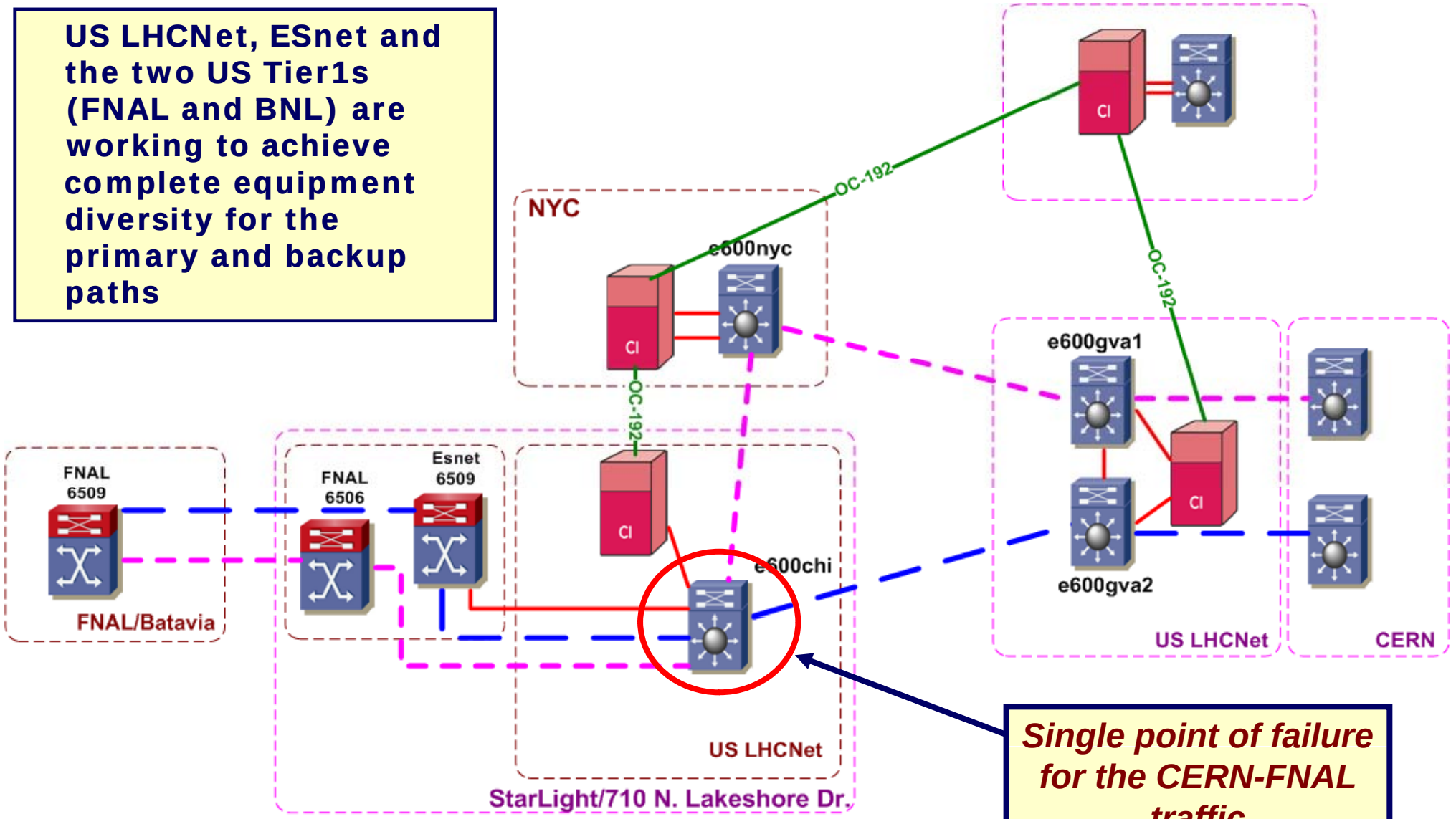
- **Four providers:**
  - **Colt**
  - **Qwest**
  - **Global Crossing**
  - **GEANT**

LCG Availability requirement: 99.95%

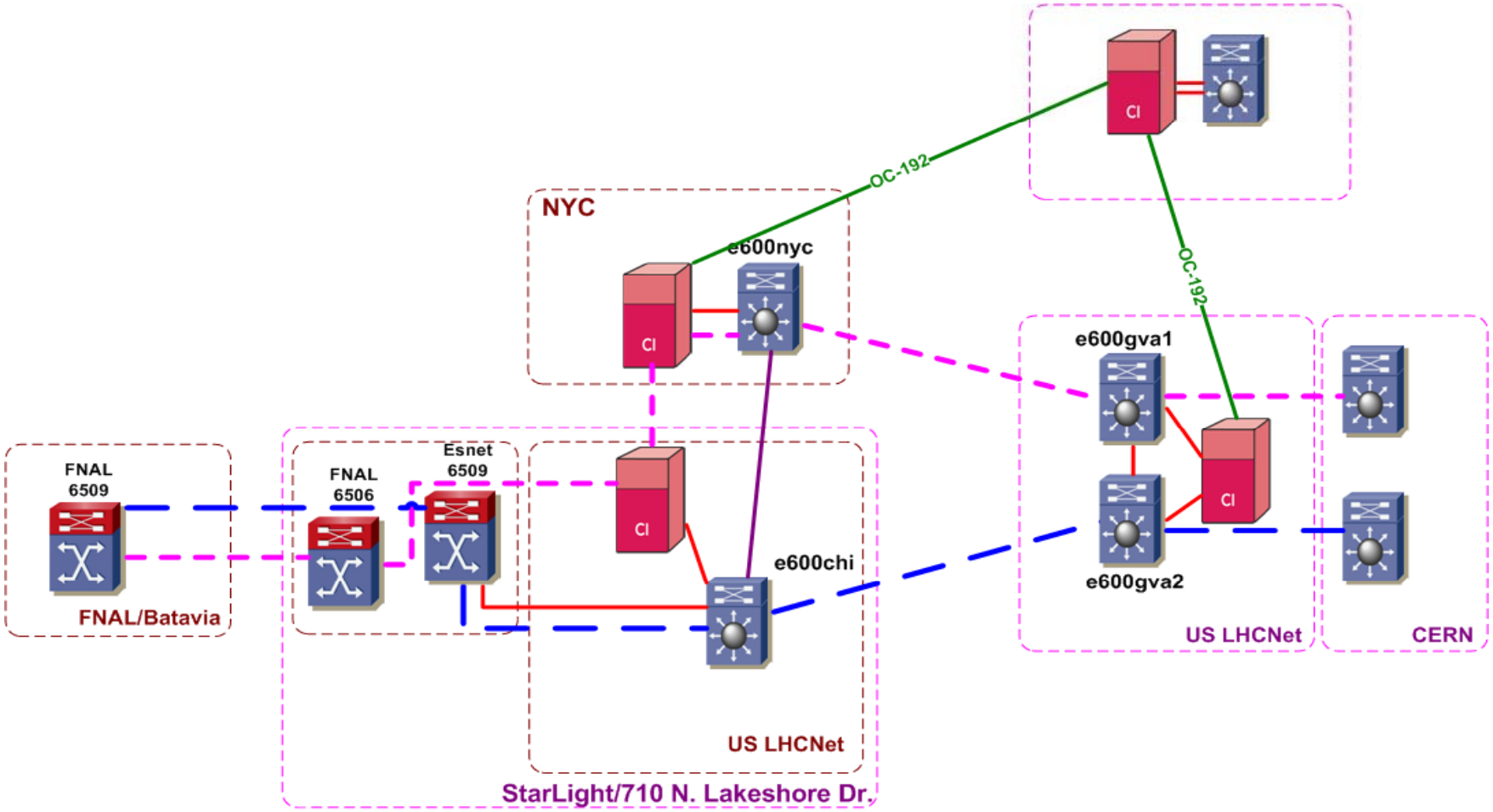
Additional Slides

Equipment Diversity

US LHCNet, ESnet and the two US Tier1s (FNAL and BNL) are working to achieve complete equipment diversity for the primary and backup paths



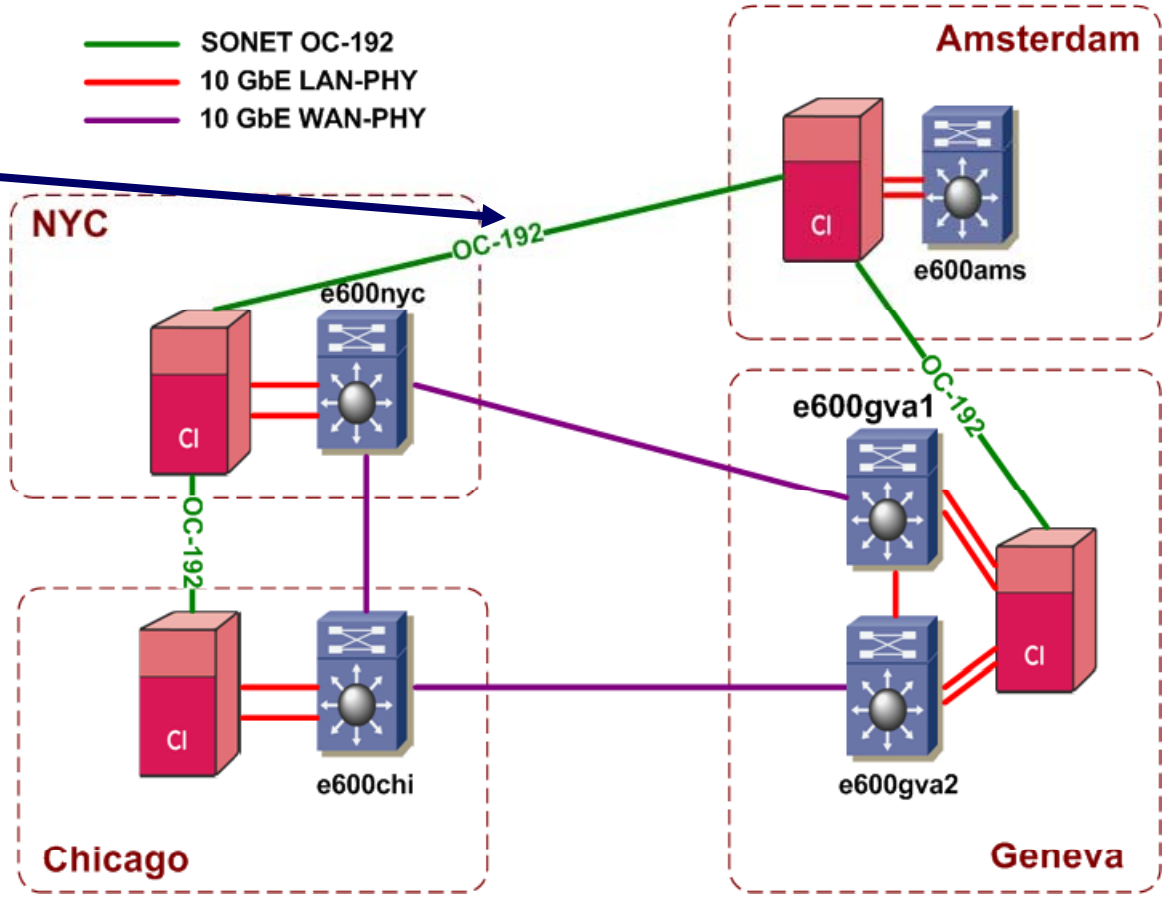
Equipment Diversity (cont.)



- The new setup allows for independent paths and can survive the failure of any single piece of equipment
- Great advantage in case of hardware of software maintenance
- Similar setup for the CERN-BNL connection

Ciena Transitions (Today)

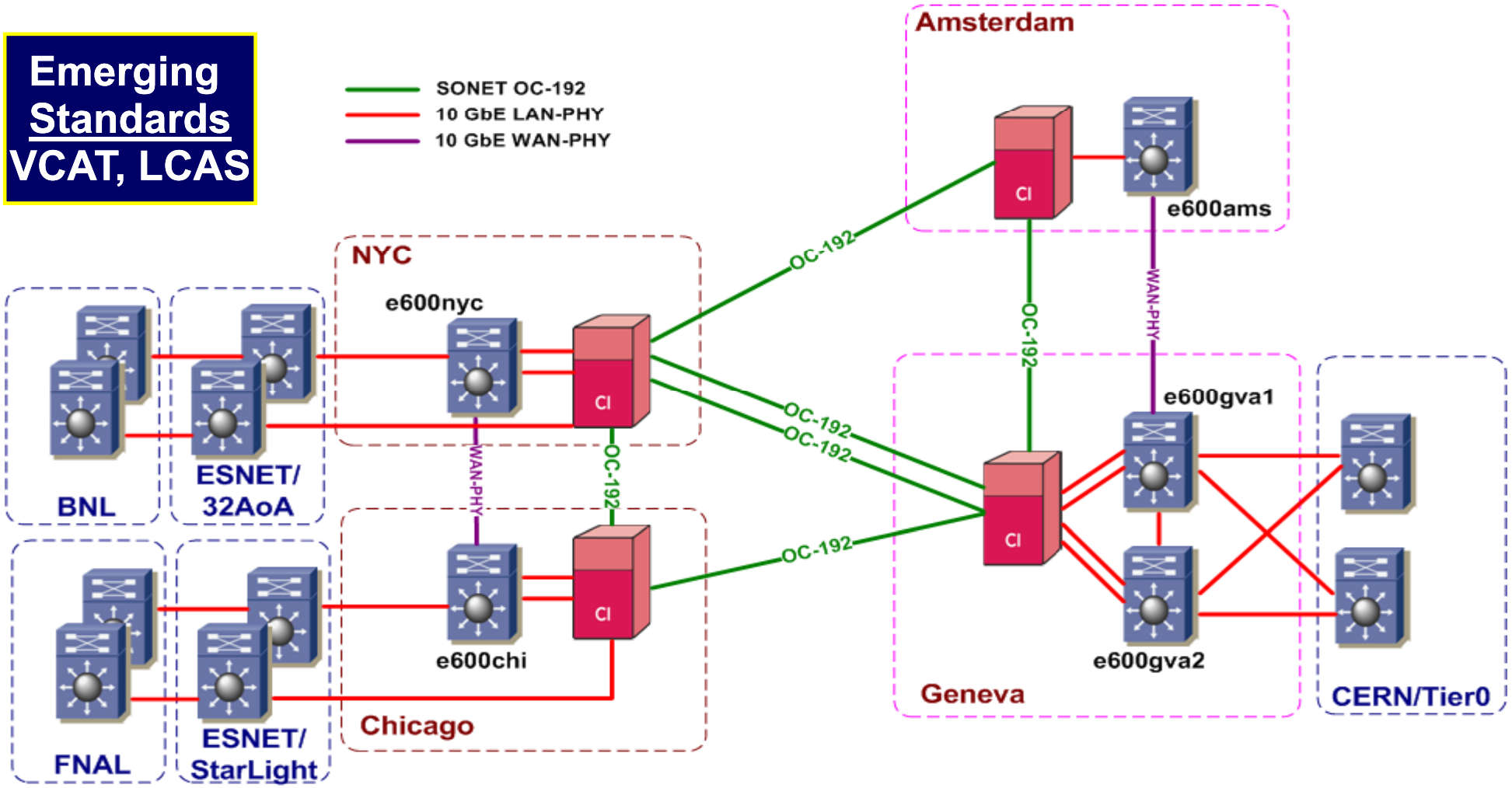
GEANT-ESnet Peering, FNAL-GridKa, BNL-GridKa



- Two parallel networks, one Force10 and one Ciena
- Today the main links (CERN-BNL, CERN-FNAL) go over the Force10s (proven reliability, stable configuration)
- Circuit oriented services development for the Cienas

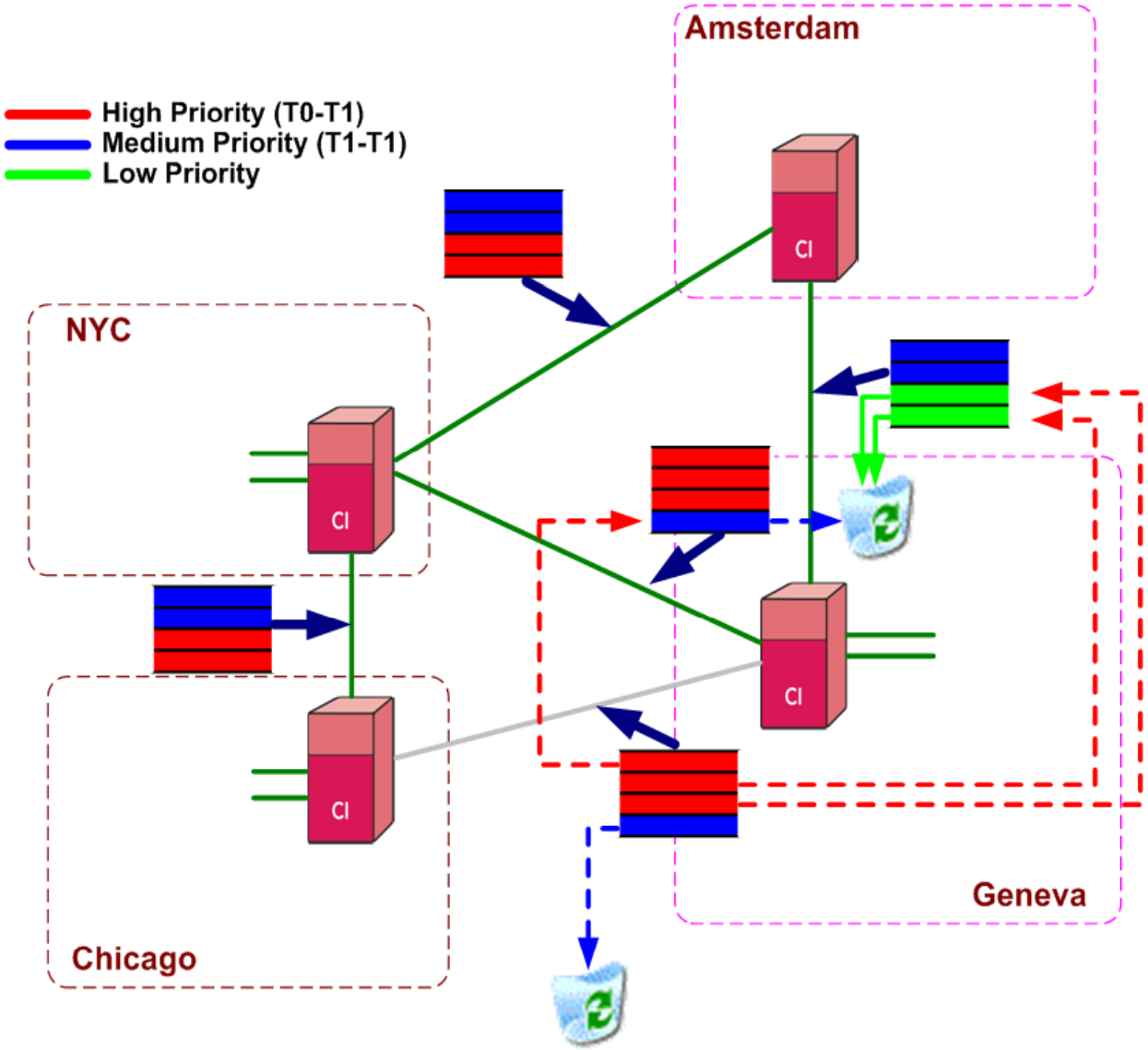
Planned Configuration (2008)

Emerging Standards
VCAT, LCAS



Robust fallback at layer 1 + next-generation hybrid optical network:
Dynamic **circuit-oriented network services with BW guarantees**

Ciena "Mesh Restoration" of a Circuit



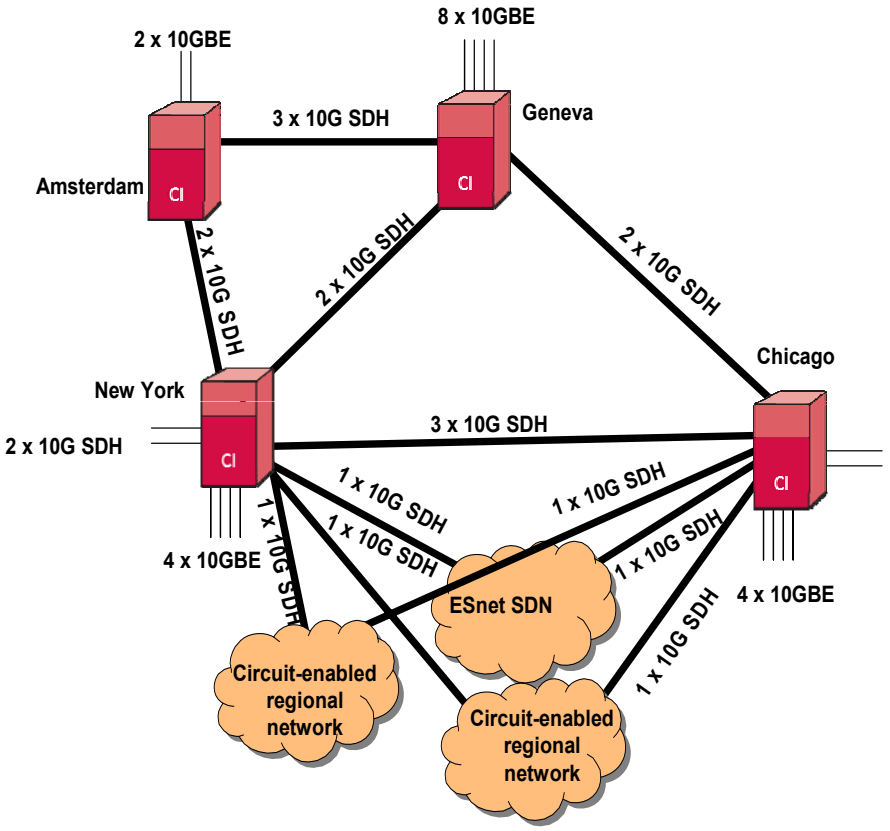
Provisioned circuits over a failed SONET link can be re-routed according to priorities and preempt lower priority circuits.

Fallback is automatic and very fast

(<50ms once failure is detected)

Network Forecast

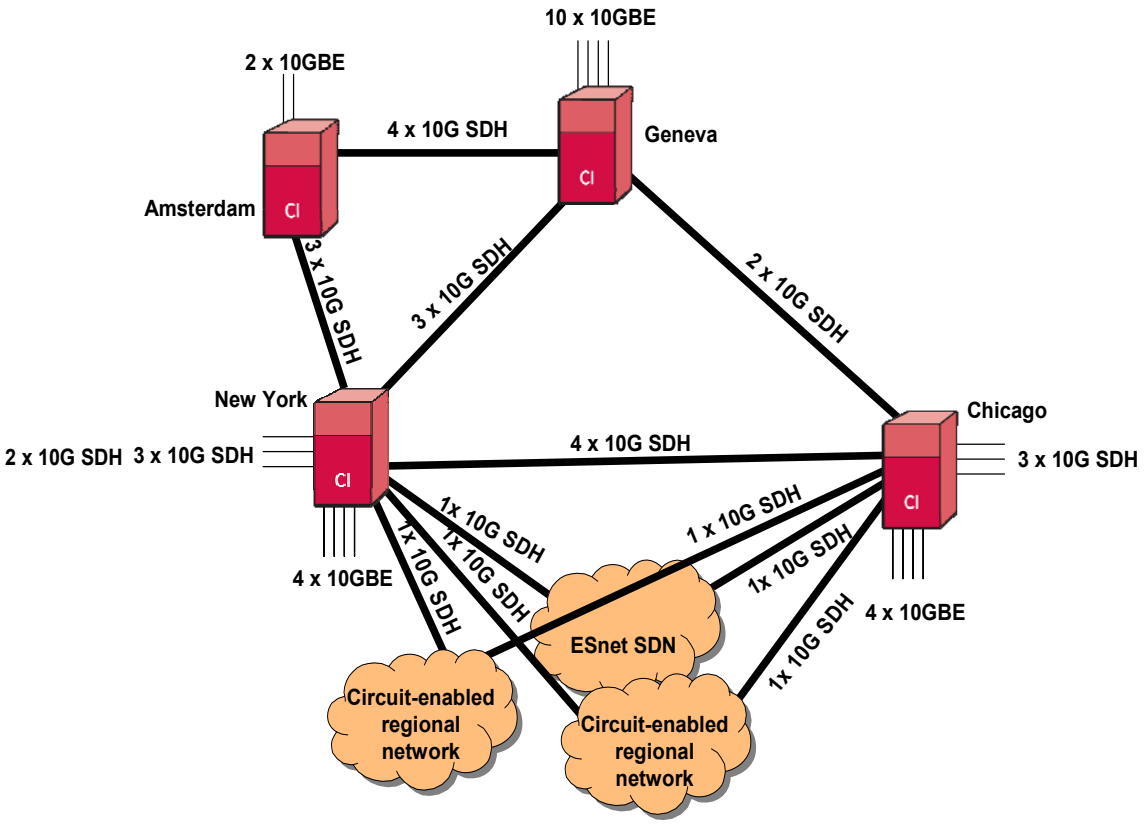
2009



Ports at Each PoP	10 GbE Ports	OC-192/STM-64 Ports
		7
		9
		16
CHI	4	12

Two port 10 GbE card

2010

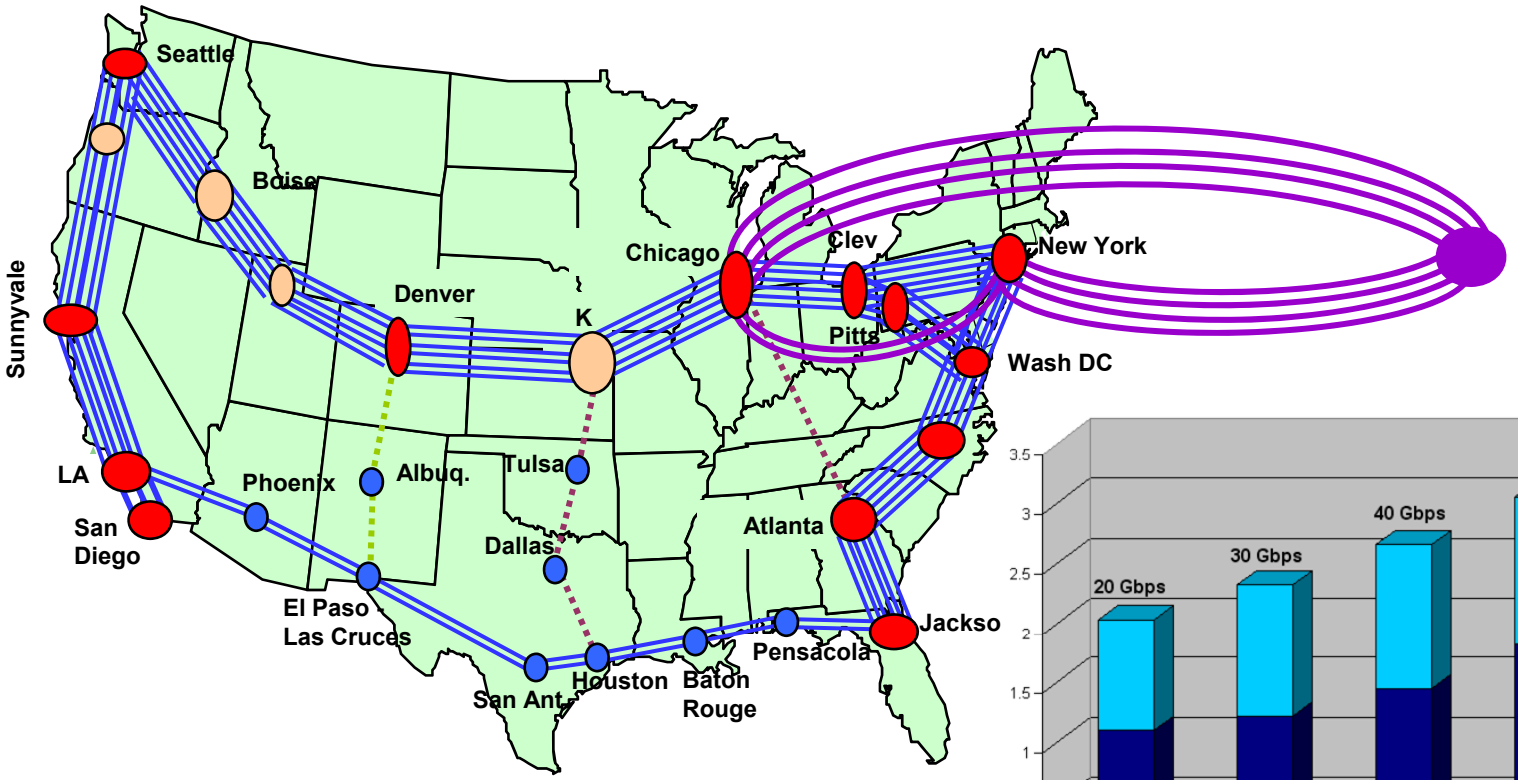


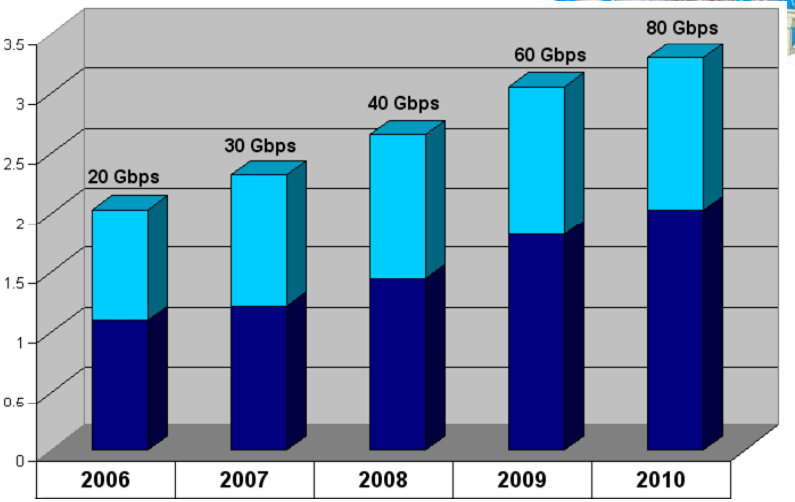
Ports at Each PoP	10 GbE Ports	OC-192/STM-64 Ports
	2	
	8	
	4	
CHI	4	10

Double density cards and matrix

40 Gbps wavelengths?

LHCNet connections to ESnet: FY09/FY10



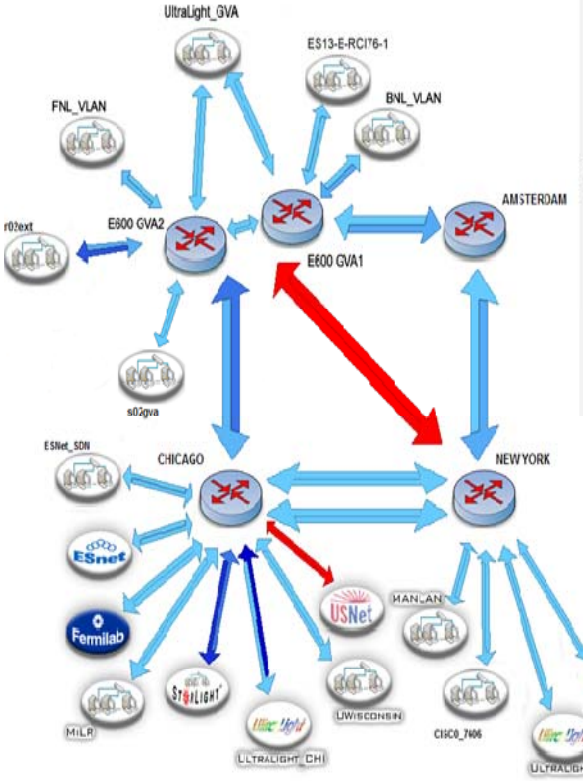


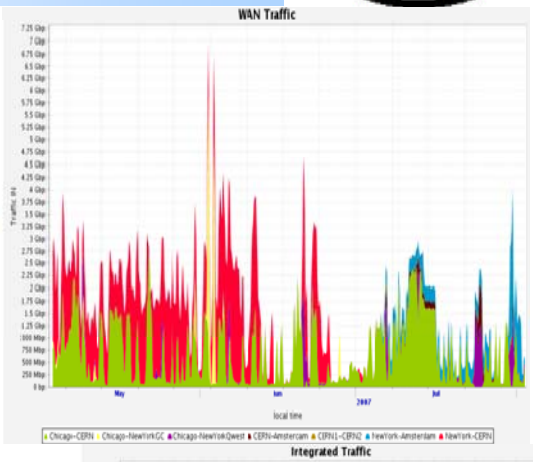
- NLR regeneration / OADM sites
- NLR wavegear sites
- ESnet via NLR (10 Gbps waves)
- LHCNet (10 Gbps waves)

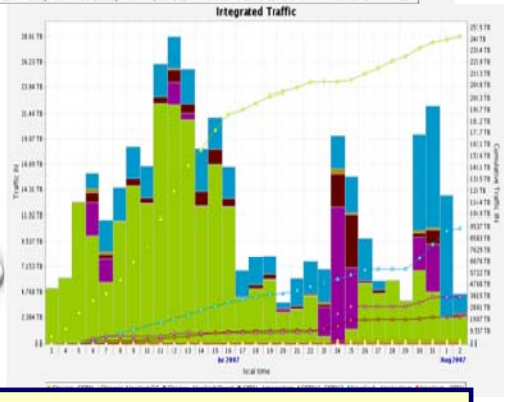
- LHCNet: To ~80 Gbps by 2009-10
- Routing + Dynamic managed circuit provisioning

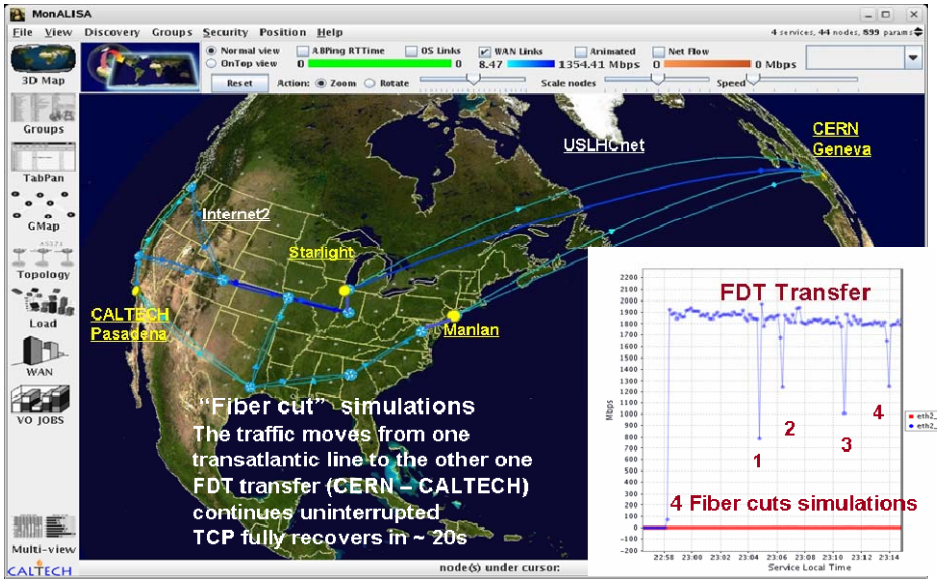
Network Monitoring

- MonALISA (TL1 Module)
- Spectrum (CERN first line)
- Various open source tools (cricket, nagios, rancid, syslog-ng, etc)
- perfSONAR (GEANT E2ECU)









- True end-to-end (host-to-host) monitoring using MonALISA
- "Network intelligence" or the ability to reconfigure the circuits based on performance, changing network conditions or high-priority scheduled transfers

US LHCNet Working Methods

Production Network

Develop and build next generation networks

High performance
High bandwidth
Reliable network

Pre-Production

N x 10 Gbps transatlantic testbed

New Data transport protocols

Interface and kernel setting

HOPI / UltraScience Net / Ultralight / CHEPREO / LambdaSation

Lightpath technologies

Vendor Partnerships

HEP & DoE Roadmaps

Testbed for Grid Development

Networks for Research

D0, CDF, BaBar, CMS, Atlas

GRID applications PPDG/iVDGL, OSG, WLCG, DISUN

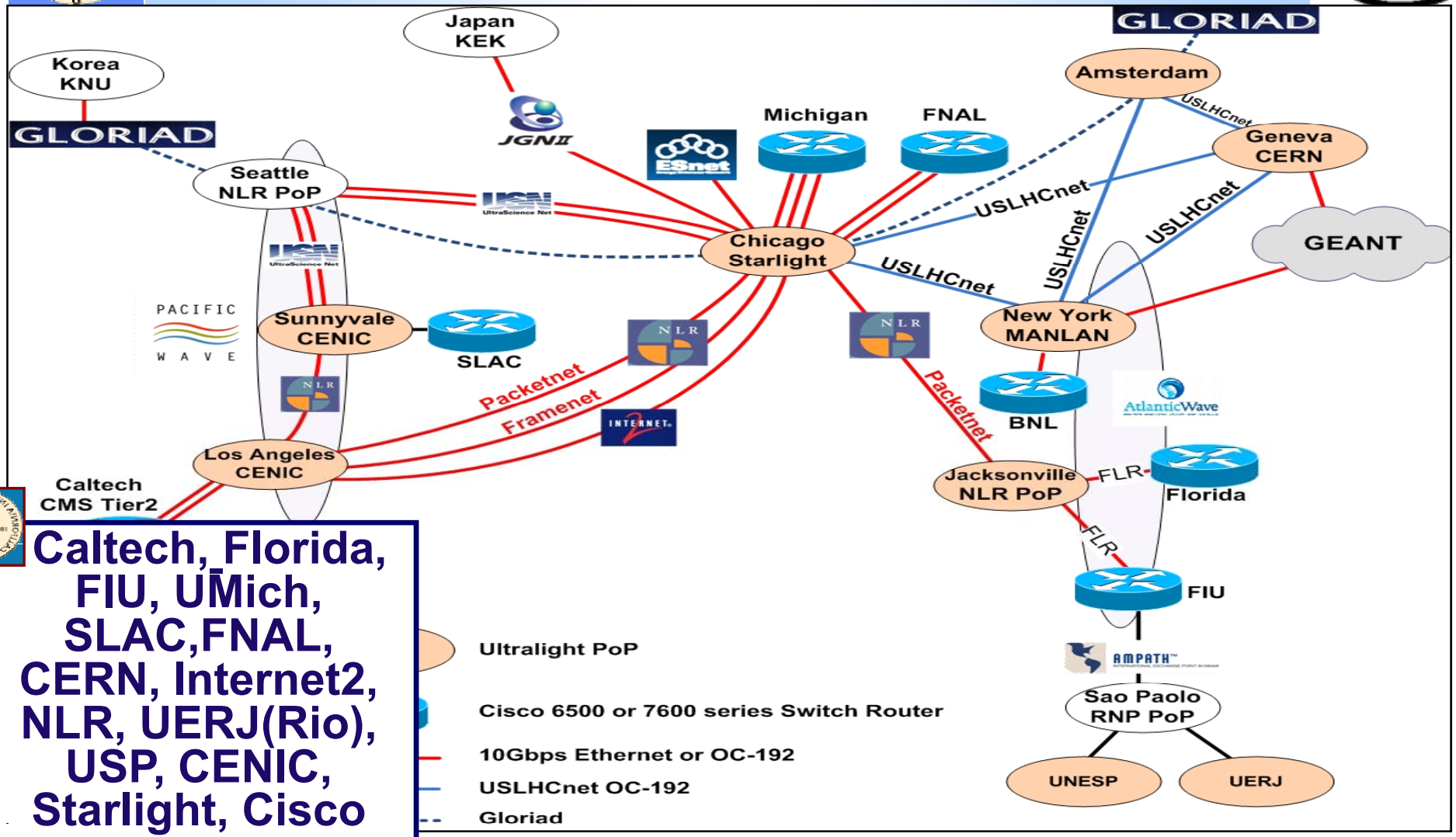
LHCOPN

Interconnection of US and EU Grid domains

VRVS/EVO

http://ultralight.caltech.edu

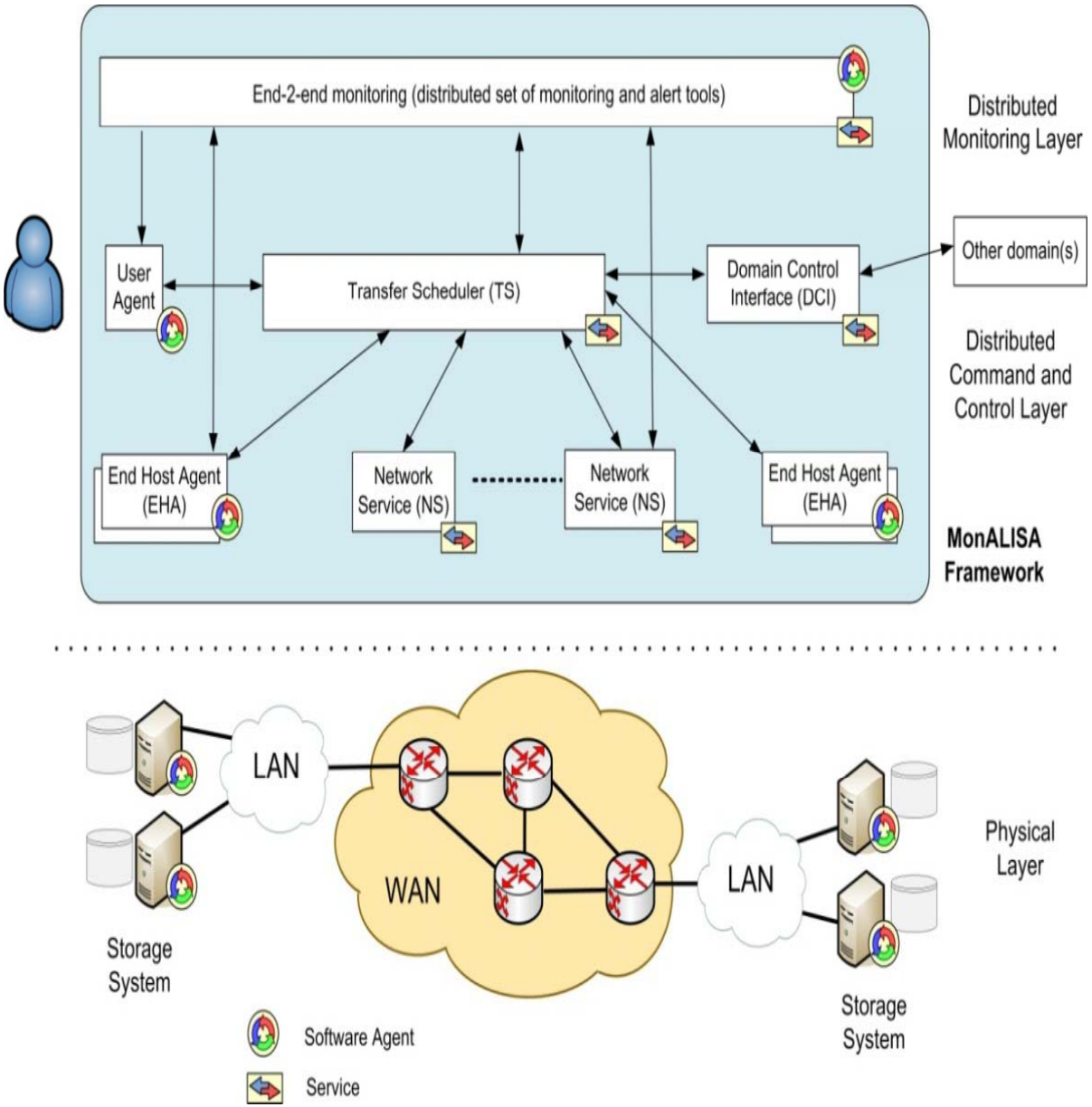
Four Continent Testbed and Facility



Caltech, Florida, FIU, UMich, SLAC,FNAL, CERN, Internet2, NLR, UERJ(Rio), USP, CENIC, Starlight, Cisco

Building a global, network-aware end-to-end managed real-time Grid

Network Services for Managed End-to-End Data Transfers

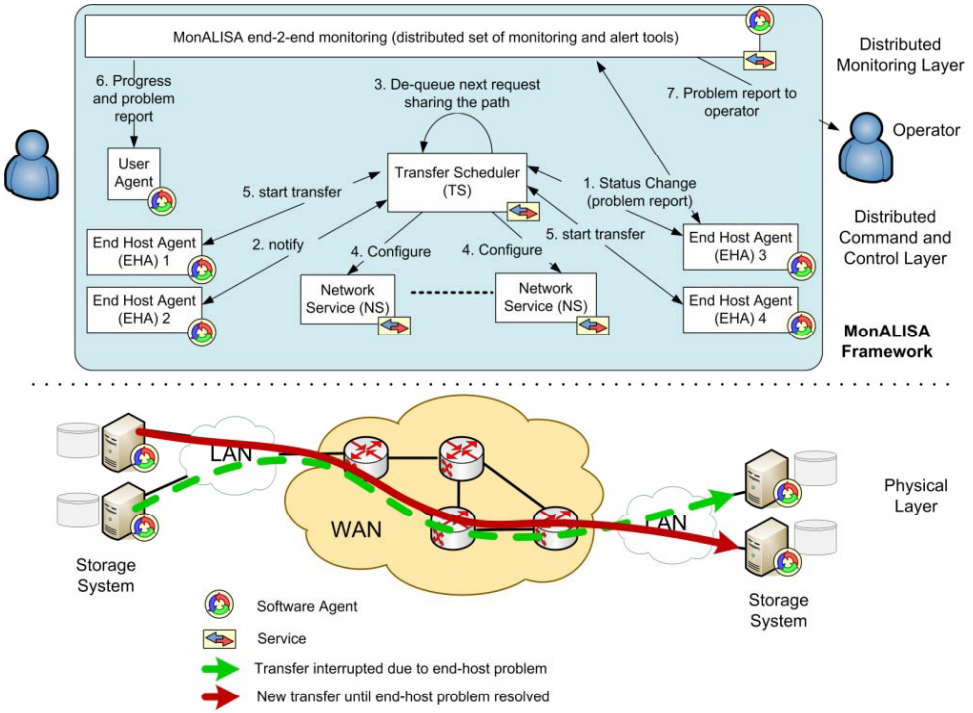


Robust Network Services based on

- Bandwidth guarantees
  - Virtual Circuits
- Scheduled Transfers
  - Transfer Classes
  - Priorities
- Monitoring of all components end-to-end
  - Network Elements
  - End-Hosts
- Interface to other circuit-oriented systems
  - Be part of heterogeneous end-to-end infrastructure

Problem Finding and Resolution

- **Problems encountered today are hard to track due to missing the *global view* of the system**



➡ **<u>Example situation:</u> the system recognizes an end-host problem during the transfer and takes mitigating actions, re-scheduling transfers and notifying operators**

End-to-end Monitored <u>Managed Transfers</u>

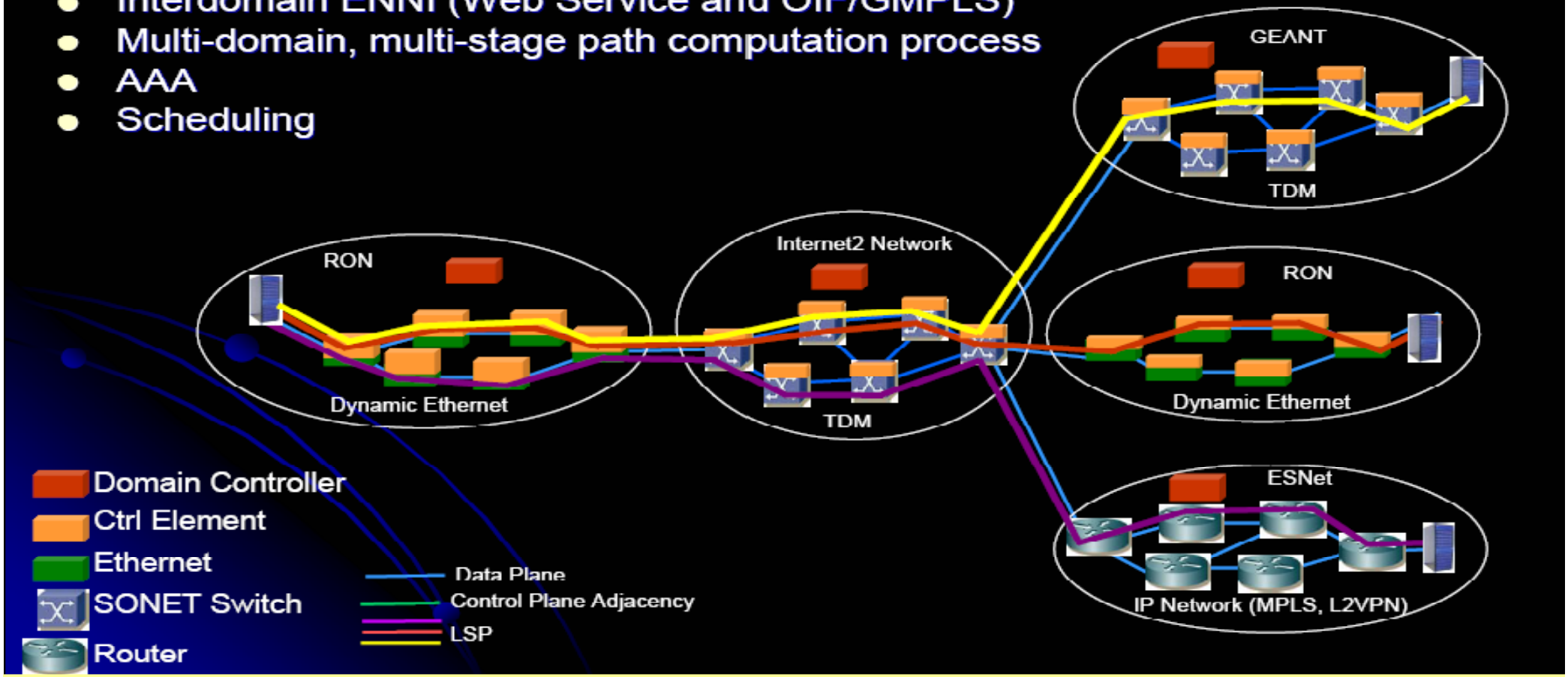
- **Track problem to the source**
  - ✸ **Network / End-host**
- ➡ **Take appropriate action**
  - ✸ **Change transfer path**
  - ✸ **Adjust end-host parameters**
  - ✸ **Re-schedule transfer**
- ➡ **Provide experts with relevant (real-time) information**
- **Keep the user/application up-to-date on transfer progress**
- **Progessive automation: Target optimal resource utilization**

- **Developed in the field-proven MonALISA Framework**

GLIF Concept

Multi-Domain Control Plane
The (near-term) big picture

- Multi-Domain Provisioning
- Interdomain ENNI (Web Service and OIF/GMPLS)
- Multi-domain, multi-stage path computation process
- AAA
- Scheduling



Tom Lehman, GLIF 2007 Winter Workshop
http://www.glif.is/meetings/2007/winter/controlplane/lehman-dynamic-services.pdf

US LHCNet Milestones

- **"Pre-production": The new infrastructure initially deployed 2007 will offer circuit-based services intended to provide redundant paths and on-demand, high bandwidth end-to-end dedicated circuits. Circuit-switched services will be used to directly interconnect the DOE laboratories to CERN and will be available on demand to policy-driven, data-intensive applications, managed by MonALISA services**
- **End of 2007: initial deployment of our circuit oriented network services on US LHCNet; simple scheduler with fixed bandwidth circuits for site to site on-demand data set transfers.**
- **Spring 2008: interaction with the data transfer application of the experiments, as well as with other intra-domain and inter-domain (LambdaStation, TeraPaths, DRAGON, Oscars) control plane services in order to provide an end-to-end path reservation.**
- ***LHC Startup:* July 2008: We will begin to exercise the network and services with real data, in close cooperation with the LHC experiments.**