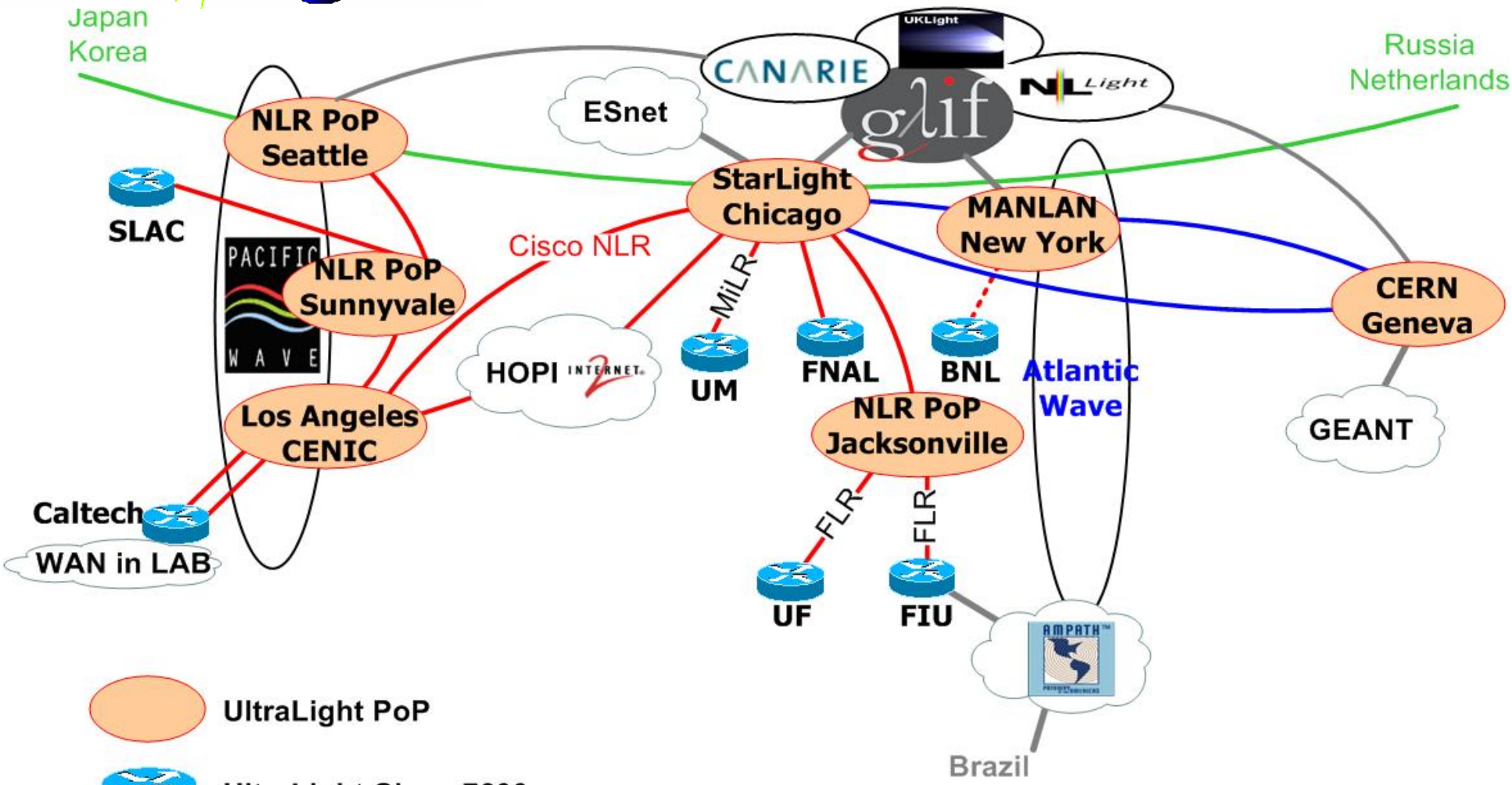
UltraLight Status Report



UltraLight PoP

UltraLight Cisco 7600

UltraLight 10 Gbps

US LHCNet 10 Gbps

Gloriad

Shawn McKee / University of Michigan

USATLAS Tier2 Meeting

August 17, 2006 - Harvard

Reminder:The UltraLight Project

UltraLight is

- A four year $2M NSF ITR funded by MPS.
- Application driven Network R&D.
- A collaboration of BNL, Buffalo, Caltech, CERN, Florida, FIU, FNAL, Internet2, Michigan, MIT, SLAC, Vanderbilt.
- Significant international participation: Brazil, Japan, Korea amongst many others.

Goal: Enable the network as a managed resource.

Meta-Goal: *Enable physics analysis and discoveries which could not otherwise be achieved.*

Status Update

There are three areas which I want to make note of for the Tier-2s

1. Work on new UltraLight kernel
2. Development of VINCI/LISA/Endhost agents (US ATLAS test of this in Fall…)
3. Work on FTS (either with FTS developers or as an equivalent project)
4. …and one addendum on US LHCNet…

UltraLight Kernel Development

Having a standard tuned kernel is very important for a

An Ultrascale Information System for Data Intensive Research

Main page | Monitor Ultralight | News | GAE | iGrid | SC05

Network Workgroup KERNEL Actitivies

We have an "UltraLight" kernel set of RPMS available which incorporate Web100 and FAST TCP patches courtesy of David Wei (Caltech FAST Team). There are two sets of RPMS created for Kernel **2.6.12.3**:

- i686 RPMS for 2.6.12.3
- x86_64A RPMS for Opteron_64 2.6.12.3
- x86_64I RPMS for EM64T 2.6.12.3

The src rpm is available kernel-2.6.12-3FastNFSv4.src.rpm.
NOTE: to use FAST you must have the `fast.ko` module installed. You can download that from the link near the bottom of the page.

Details

We have been working on developing a new linux kernel for use within UltraLight (October 10, 2005). We chose to start from a 2.6.12.3 kernel initially and add in the following steps below. As of November 12, 2005 we have a 2.6.14.1 kernel (no kernel debug) which can be

Optical Path Plans

Emerging "light path" technologies are becoming popular in the Grid community:

- They can extend and augment existing grid computing infrastructures, currently focused on CPU/storage, to include the network as an integral Grid component.
- Those technologies seem to be the most effective way to offer network resource provisioning on-demand between end-systems.

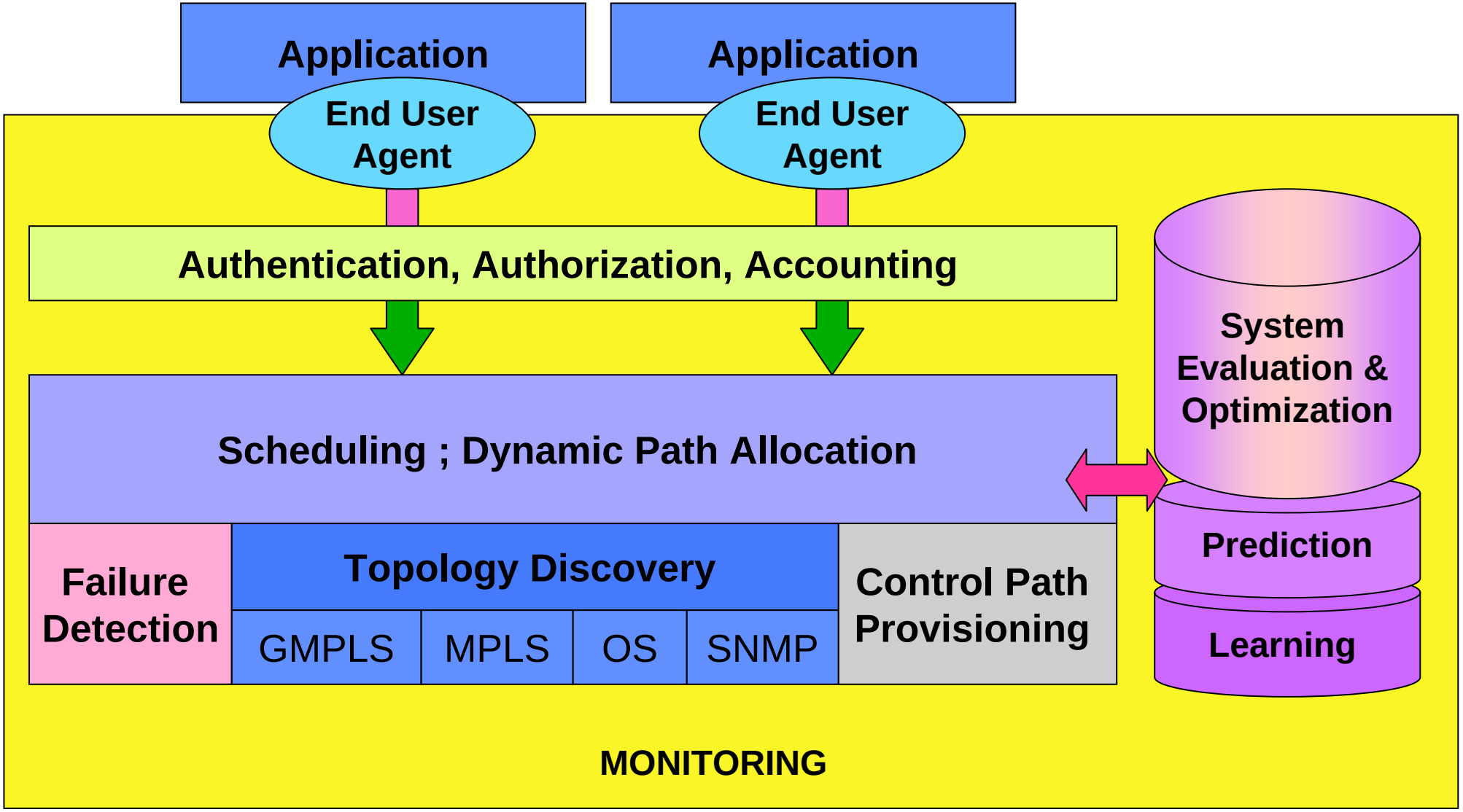
A major capability we are developing in Ultralight is the ability to dynamically switch optical paths across the node, bypassing electronic equipment via a fiber cross connect.

The ability to switch dynamically provides additional functionality and also models the more abstract case where switching is done between colors (ITU grid lambdas).

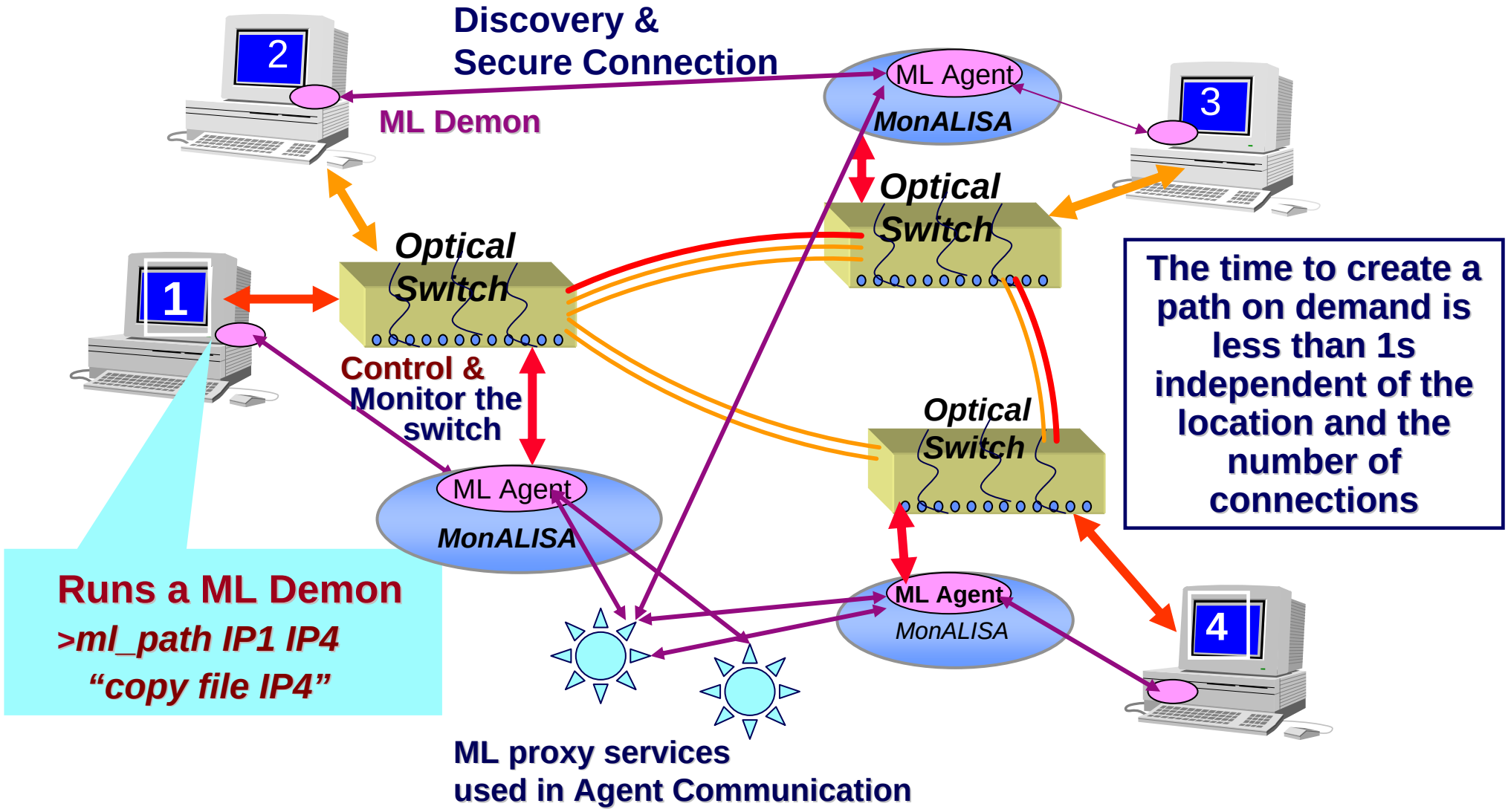
VINCI: A Multi-Agent System

- VINCI and the underlying MonALISA framework use a system of autonomous agents to support a wide range of dynamic services
- Agents in the MonALISA servers self-organize and collaborate with each other to manage access to distributed resources, to make effective decisions in planning workflow, to respond to problems that affect multiple sites, or to carry out other globally-distributed tasks
- Agents running on end-users' desktops or clusters detect and adapt to their local environment so they can function properly. They locate and receive real-time information from a variety of MonALISA services, aggregate and present results to users, or feed information to higher level services
- Agents with built-in "intelligence" are required to engage in negotiations (for network resources, for example), and to make pro-active run-time decisions, while responding to changes in the environment

The Main VINCI Services



Agents to Create on Demand an Optical Path or Tree



DEMO: MonALISA and path-building

An example of optical path building using MonALISA is shown at: http://ultralight.caltech.edu/web-site/gae/movies/ml_optical_path/ml_os.htm

One of the focus areas for UltraLight is being able to dynamically construct point-to-point light-paths where supported.

We still have a pending proposal (PLaNetS) focused on creating a managed dynamic network infrastructure…

LISA, EVO and Endhosts

Many of you are familiar with VRVS. Its successor is called EVO (Enabling Virtual Organizations). It improves on VRVS in a number of ways:

- Support for H.263 (capture and send your desktop as another video source for a conference)
- IM like capability (presence/chat)
- Better device / OS / Language support
- Significantly improved reliability and scalability

- Related to this last point is a the "merger" of MonALISA and VRVS in EVO.
- Endhost agents (LISA) are now an integral part of EVO. Endhost agents monitor the user's hosts and react to changing conditions
- Something like this is envisioned as a component of deploying a 'managed network'
- Prototype testing of network agent this fall?

FTS and UltraLight…

To date there has been little interaction between people working on the network and those working on data transport for ATLAS (or LHC in general)

There is a significant amount of work architecting, developing and hardening the data management (and transport) for ATLAS…little time (or understanding of possibilities) for the network.

A dynamic managed network introduces new possibilities. Research efforts in networking need to be fed into the data transport architecting.

UltraLight is planning to engage the FTS developers and try to determine their understanding of (and plans for) the network.

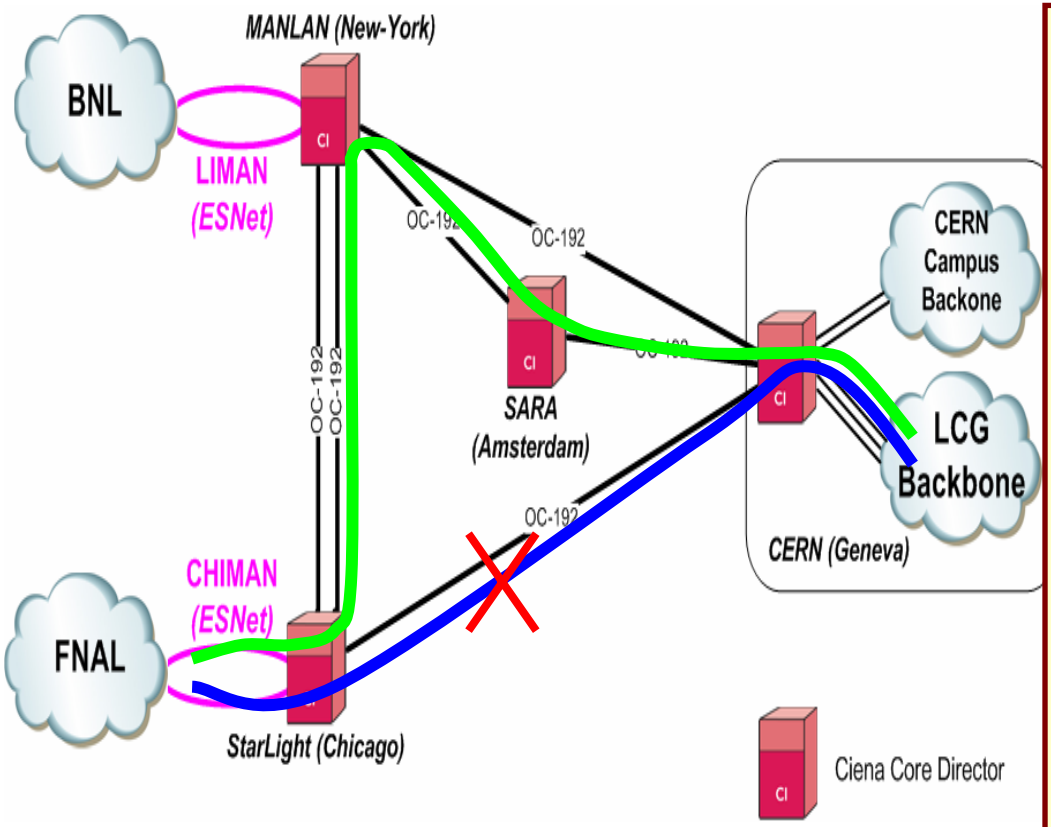
GOAL: Account for the network and improve robustness and performance of data transport and the overall infrastructure.

Aside: US LHCNet Status and Plans

The following 7 slides (from Harvey Newman) provide some details about US LHCNet and its plans to support LHC scale physics requirements.

Details are provided for reference but I won't cover them in my limited time.

Next Generation LHCNet: Add Optical Circuit-Oriented Services



CERN-FNAL Primary EPL

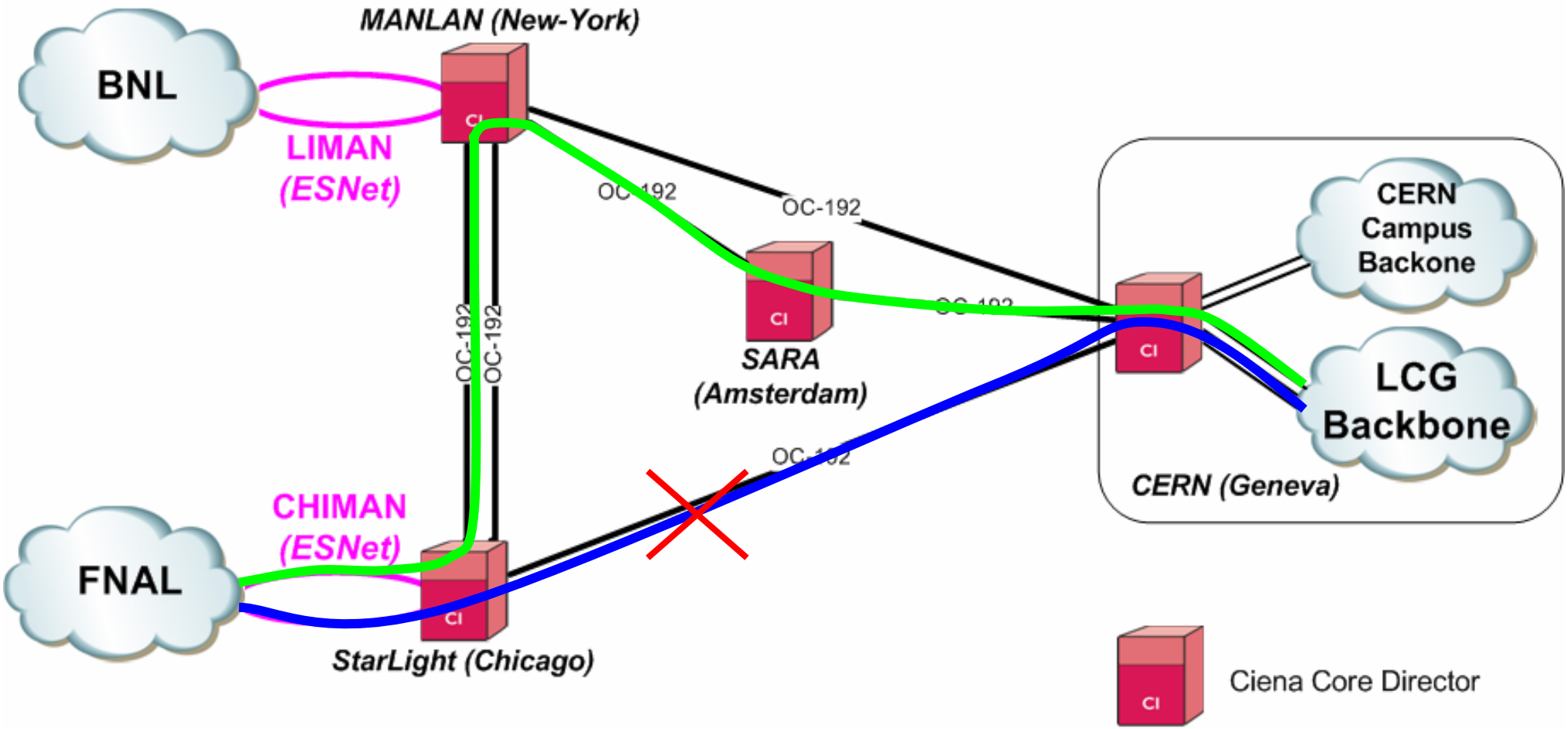
CERN-FNAL Secondary EPL

Based on CIENA "Core Director" Optical Multiplexers

- Highly reliable in production environments
- Robust fallback, at the optical layer
- Circuit-oriented services: Guaranteed Bandwidth Ethernet Private Line (EPL)
- Sophisticated standards-based software: VCAT/LCAS.
  - VCAT logical channels: highly granular bandwidth management
  - LCAS: dynamically adjust channels
- Highly scalable and cost effective, especially for many OC-192 Links

- This is consistent with the directions of the other major R&E networks such as Internet2/Abilene, GEANT (pan-European), ESnet SDN

Next Generation LHCNet: Add Optical Circuit-Oriented Services

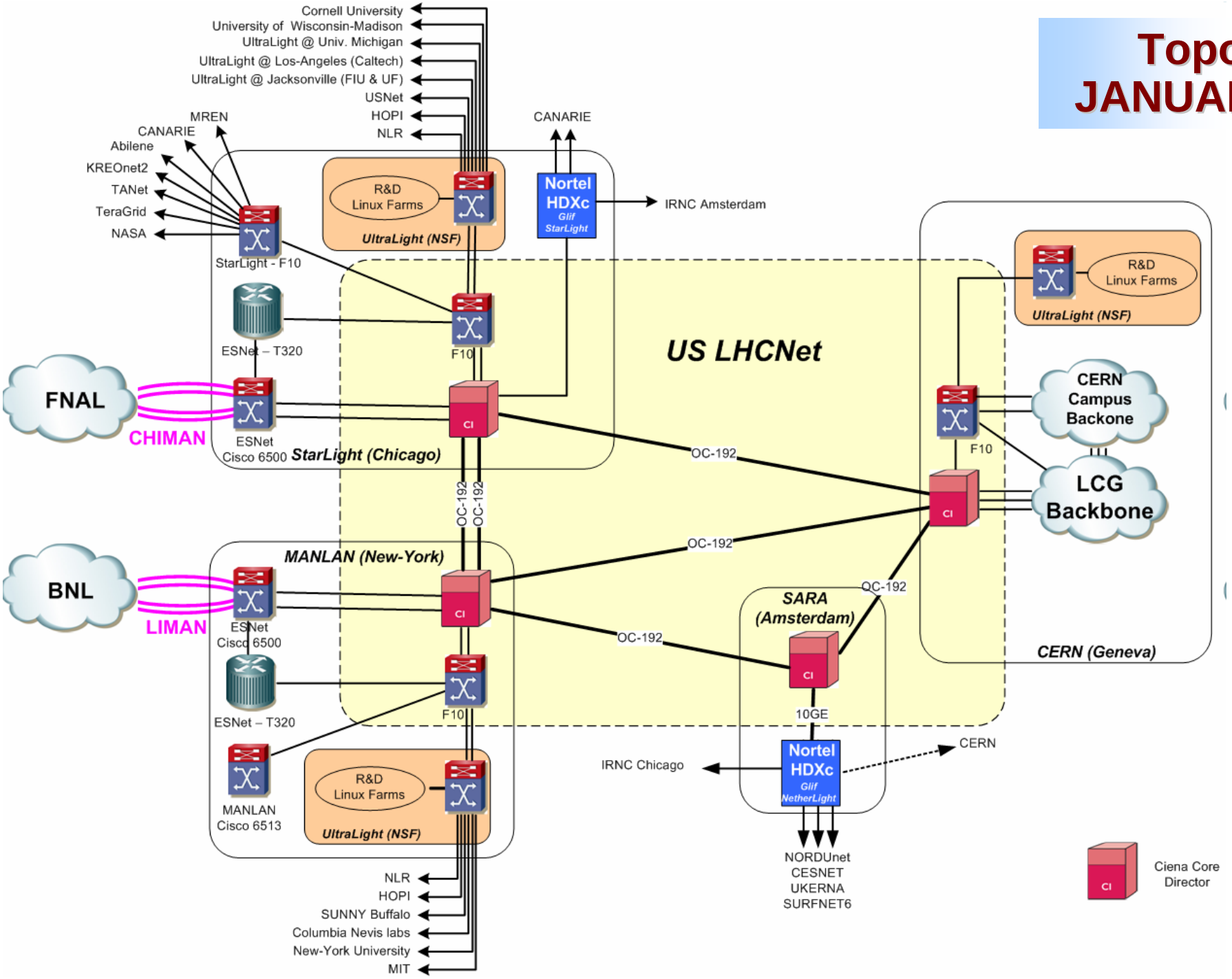


CERN-FNAL Primary EPL

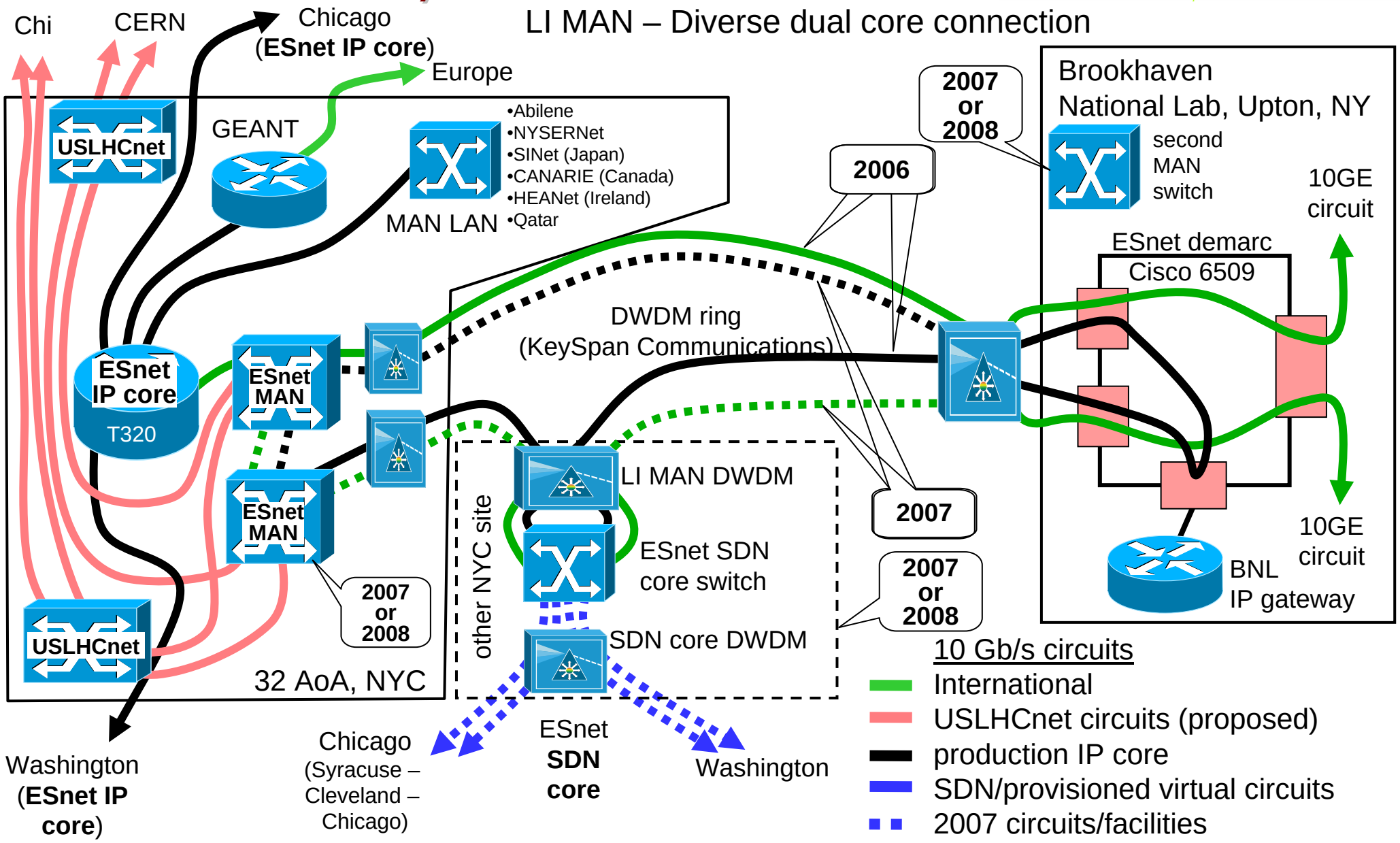
CERN-FNAL Secondary EPL

Force10 switches for Layer 3 services

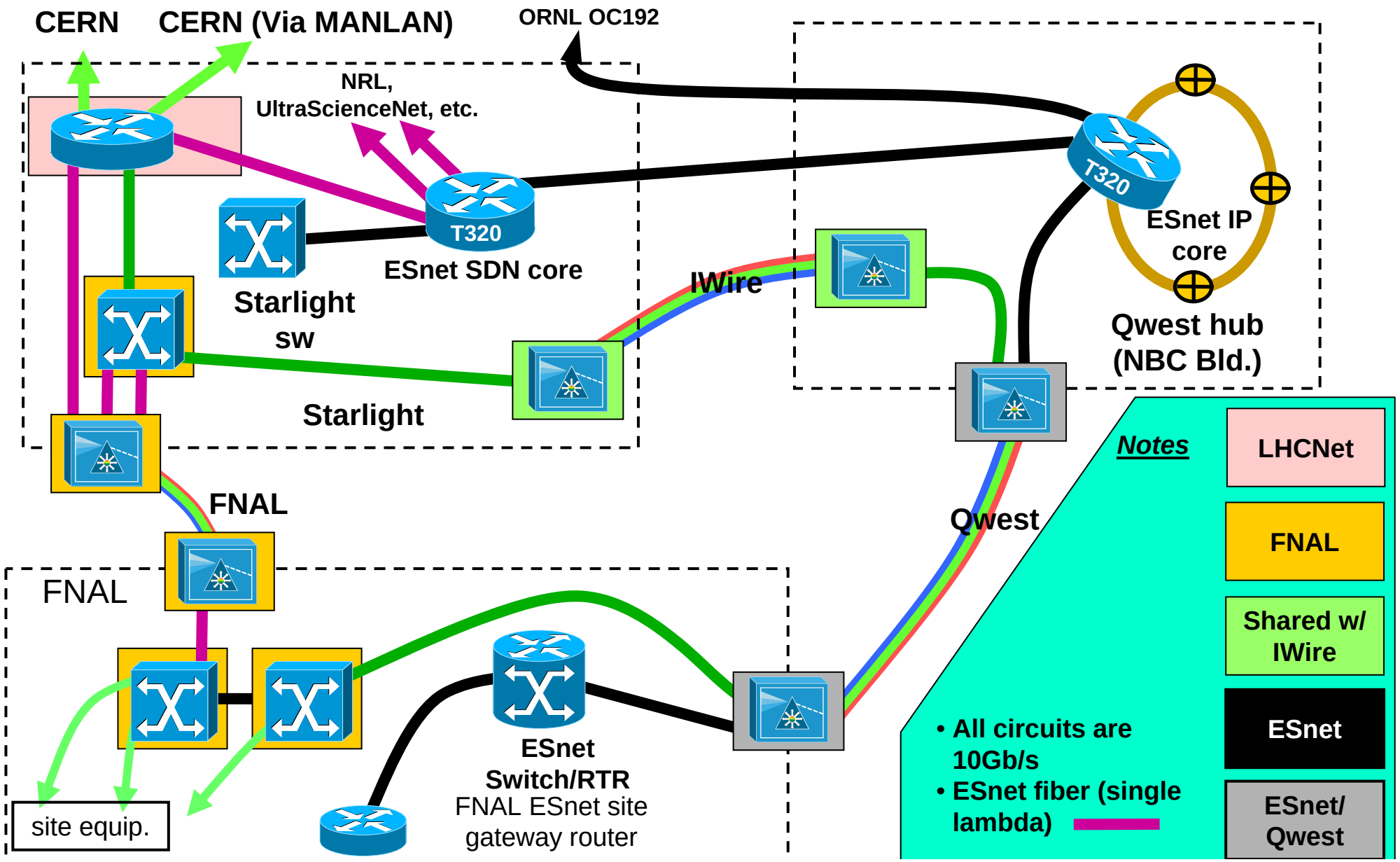
Topology JANUARY 2007



BNL and Long Island MAN Ring - Feb., 2006 to 2008



FNAL and Chicago MAN Ring



Next Generation LHCNet

- Circuit oriented services
  - Bandwidth guarantees at flexible rate
  - Provide for data transfer deadlines for remote data analysis
  - Traffic isolation for unfriendly data transport protocols
  - Security
- CIENA Platforms
  - OSRP *Optical Signaling and Routing Protocol*
    - Distributed signal and routing protocol which abstracts physical network resources
    - Based on G.ASON
    - Advertise topology information and capacity availability
    - Connection management (Provisioning/Restoration)
    - Resource discovery and maintenance
  - Ethernet Private Line (EPL) – Point-to-Point
    - Dedicated bandwidth tunnels: guaranteed end-to-end performance
    - VCAT/LCAS/GFP-F allows for resilient, right-sized tunnels
    - Automated end-to-end provisioning
- Technology and bandwidth roadmap in line with ESNet (SDN), Internet2 (HOPI/NEWNET) and GEANT plans

Milestones: 2006-2007

- May to September 2006: Service Challenge 4
- August 2006: Selection of telecom provider(s) from among those responding to the call for tender
- October 2006: Provisioning of new transatlantic circuits
- Fall 2006: Evaluation of CIENA platforms
  - Try and buy agreement
- End 2006: 1st Deployment of Next-generation US LHCNet
  - Transition to new circuit-oriented backbone, based on optical multiplexers.
  - Maintain full switched and routed IP service for a controlled portion of the bandwidth
- Summer 2007: Start of LHC operations

Conclusion

A number of developments are in progress.

Tier-2's should be able to benefit and can hopefully help drive these developments via testing feedback

Networking developments need to be feed into existing and planned software for LHC

US LHCNet is planning to support LHC scale requirements for connectivity and manageability