The US LHCNet Project

ICFA Workshop, Sinaia
October 2006

Dan Nae
California Institute of Technology
Who We Are

- A transatlantic network designed to support the LHC and the U.S. HEP community
- Funded by the US DoE and CERN and managed by Caltech in collaboration with CERN
- Evolved from a network between US and CERN which dates back to 1985
- Our mission is to deliver a reliable network service to support the upcoming LHC experiments at CERN
- Designed to support the LHC three-tiered model and to deliver data directly to/from the US Tier 1’s (FNAL and BNL)
We offer Layer 3 services (IP) and we peer with all major US research networks

Layer 2 dedicated paths between CERN and the US Tier1’s

Layer 1 protected services coming soon

Redundant services using multiple paths across the Atlantic

Many layers of redundancy (equipment redundancy, path diversity, collaborations with other research networks)

Integrated monitoring with MonALISA *
LHCNet

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
US LHCNet

- 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
Co-operated by Caltech and CERN engineering teams

- Force10 platforms, 10GE WANPHY
- New PoP in NY since Sept. 2005
- 10 Gbps path to BNL since April 2006
- Connection to US Universities via UltraLight (NSF & university funded) backbone
RSTP Backup and Load Sharing

- Fast recovery time (enough to keep the CERN BGP peerings from resetting during a failure)
- But not so fast (sometimes they do)
- Cannot deal with flapping links
Future backbone topology

- **GVA-CHI-NY triangle**
- **New PoP in Amsterdam**
  - GEANT2 circuit between GVA and AMS
  - Access to other transatlantic circuits ➔ backup paths and additional capacity
  - Connection to Netherlight, GLIF (T1-T1 traffic and R&D)
New Topology deployment

Atlantic Ocean

NY-MANLAN
CHI-Starlight

AC-1 South
NY 111 8th

VSNL North
NY 60 Hudson

AMS-SARA

VSNL

GVA-CERN

AC-2

Global Crossing
Qwest
Colt
GEANT

In Production
November 1st
January 2007
January 2007
Multiple Fiber Paths: Reliability Through Diversity

- **Unprotected circuits (lower cost)**
- **Service availability from provider’s offers:**
  - Colt Target Service Availability is 99.5%
  - Global Crossing guarantees Wave Availability at 98%
  - Canarie and GEANT: No Service Level Agreement (SLA)

**LCG Availability requirement:** 99.95%
Next Generation LHCNet: Add Optical Circuit-Oriented Services

- Based on CIENA “Core Director” Optical Multiplexers
- Robust fallback, at the optical layer
- Circuit-oriented services: Guaranteed Bandwidth Ethernet Private Line (EPL)
- Sophisticated standards-based software: VCAT/LCAS.
USLHCNet NOC

- The CERN Network Operation Center (NOC)
  - Delivers the first level support 24 hours a day, 7 days a week.
  - Watch out for alarms
  - A problem not resolved immediately is escalated to the Caltech network engineering team.

- USLHCnet engineers “on call” 24x7
  - On site (at CERN) in less than 60 min
  - Remote hand service at MANLAN and StarLight is available on a 24x7 basis with a four hour response time.
Operations & management assisted by agent-based software (MonALISA)

500 TB of data sent from CERN to FNAL over the last two months
LHCNet Utilization during Service Challenge

- **Service challenge**
  - Achieving the goal of a production quality world-wide Grid that meets the requirements of LHC experiments
  - Prototype the data movement services
  - Acquire an understanding of how the entire system performs when exposed to the level of usage we expect during LHC running

- CERN-FNAL traffic during SC3 (April 2006)
- Disk-to-Disk
Circuit failure during SC2

1. LHCnet trans-Atlantic link cut

2. CMS SC2 traffic fails over to GEANT path & FNAL production OC12

3. CMS SC2 traffic rerouted by ESnet to FNAL Starlight 1 GE overflow link

4. LHCnet trans-Atlantic link returned to service
Additional Slides
LHCNet configuration (2007)
LHCNet connection to Proposed ESnet
Lambda Infrastructure Based on National
Lambda Rail: FY09/FY10

NLR wavegear sites
NLR regeneration / OADM sites

ESnet via NLR (10 Gbps waves)
LHCNet (10 Gbps waves)

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

- Prototype data movement services between CERN and the US
- High speed disk-to-disk throughput development
  - New end-systems (PCI-e; 64 bit cpu; New 10 GE NICs)
  - New data transport protocols (FAST and others)
  - Linux kernel patches; RPMs for deployment
- Monitoring, Command and Control Services (MonALISA)
- “Optical control plane” development
  - MonALISA services available for photonic switches
  - GMPLS (Generalized MPLS); G.ASON
  - Collaboration with Cheetah and Dragon projects
- Note: Equipment loans and donations; exceptional discounts
Milestones: 2006-2007

- May to September 2006: Service Challenge 4 - completed
- August 2006: Selection of telecom provider(s) from among those responding to the call for tender - completed
- 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
- Fall: Start of LHC operations
Primary Milestones for 2007-2010

- Provide a robust network service without service interruptions, through
  - Physical diversity of the primary links
  - Automated fallback at the optical layer
  - Mutual backup with other networks (ESnet, IRNC, CANARIE, SURFNet etc.)

- Ramp up the bandwidth, supporting an increasing number of 1-10 Terabyte-scale flows

- Scale up and increase the functionality of the network management services provided

- Gain experience on policy-based network resource management, together with FNAL, BNL, and the US Tier2 organizations

- Integrate with the security (AAA) infrastructures of ESnet and the LHC OPN
Additional Technical Milestones for 2008-2010

Targeted at large scale, resilient operation with a relatively small network engineering team

◆ **2008: Circuit-Oriented services**
  - Bandwidth provisioning automated (through the use of MonALISA services working with the CIENAs, for example)
  - Channels assigned to authenticated, authorized sites and/or user-groups
  - Based on a policy-driven network-management services infrastructure, currently under development

  - Extend advanced planning and optimization into the networking and data-access layers.
  - Provides interfaces and functionality allowing physics applications to interact with the networking resources
Conclusion

- **US LHCNet**: An extremely reliable, cost-effective High Capacity Network
- **A 20+ Year Track Record**
- **High speed inter-connections with** the major R&E networks and US T1 centers
- **Taking advantage of rapidly advancing network technologies to meet the needs of the** LHC physics program at moderate cost
- **Leading edge R&D projects as required**, to build the next generation US LHCNet