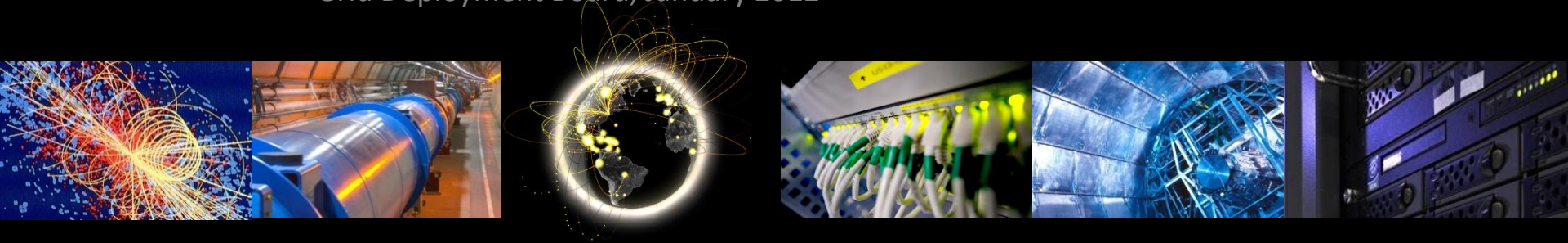


# LHCOPN / LHCONe Status Update

John Shade /CERN IT-CS

Summary of the LHCOPN/LHCONe meeting in Berkeley  
Grid Deployment Board, January 2012



# Introduction

- Agenda and slides from Berkeley meeting are at:  
<https://indico.cern.ch/conferenceDisplay.py?confId=160533>
- Other useful links:  
<http://lhcone.net>
- Mailing List (CERN e-group):  
LHCONE-interest
- Meeting was two days, first half-day devoted to LHCOPN.

# Berkeley and ESnet welcome

- Introduction to *Lawrence Berkeley National Laboratory* by Kathy Yelick
- Home of 12 nobel laureates!
- 800 students; 4200 employees



# LHCOPN News

- New CERN-ASGC (Amsterdam) backup link (2.5Gbps)
- CERN LHCOPN routers now Brocade MLXe16 with 100G Ethernet connections to the CERN LCG
- IPv6 peering between CERN and KIT  
Gridftp test traffic exchanged
- IPv6 can be configured between CERN and any T1

# LHCOPN Operations Report

- GGUS & LHCOPN GGUS now cross-reference each other & are searchable
- Decrease in number of service impacting events in 2011
- Backup link tests need to be scheduled and documented



# LHCOPN Monitoring Status

- Most Tier1s now have two perfSONAR-PS servers. Dashboard could be greener!
- Dashboard has been augmented with features for defining new clouds (matrices) & setting alarms on primitive services  
<https://perfsonar.usatlas.bnl.gov:8443/exda/?page=25&cloudName=LHCOPN>
- Jason offered a workshop on using the toolkit for troubleshooting - seeking interested parties
- Suggestion that WLCG should have a body (of people) for troubleshooting WAN issues

# LHCONE Agenda Bashing(!)

- During 2011, LHCONE consisted of two implementations, each successful in its own scope:
  - Transatlantic Layer 2 domain
    - Aka VLAN 3000, implemented by USLHCNet, SURFnet, Netherlight, Starlight
  - European VPLS domain
    - Mostly VLAN 2000, implemented in RENATER, DFN, GARR, interconnected through GEANT backbone (DANTE)
- In addition, Internet2 deployed a VPLS-based pilot in the US
- Problem: Connecting the VPLS domains at Layer 2 with other components of the LHCONE
- The new multipoint architecture therefore foresees inter-domain connections at Layer 3

# LHCONE Agenda (contd.)

- **Fork in the path forward:**

A solution for “now”

- To make sure the immediate needs are satisfied

A long-term view at the LHC shutdown time scale

- Leveraging next generation technologies
- Requires some R&D investment to assure global scalability

- **Time scale: coordinate with LHC schedule:**

2012: LHC run will continue until November

2013-2014: LHC shutdown, restart late 2014

2015: LHC data taking at full nominal energy (14 TeV)



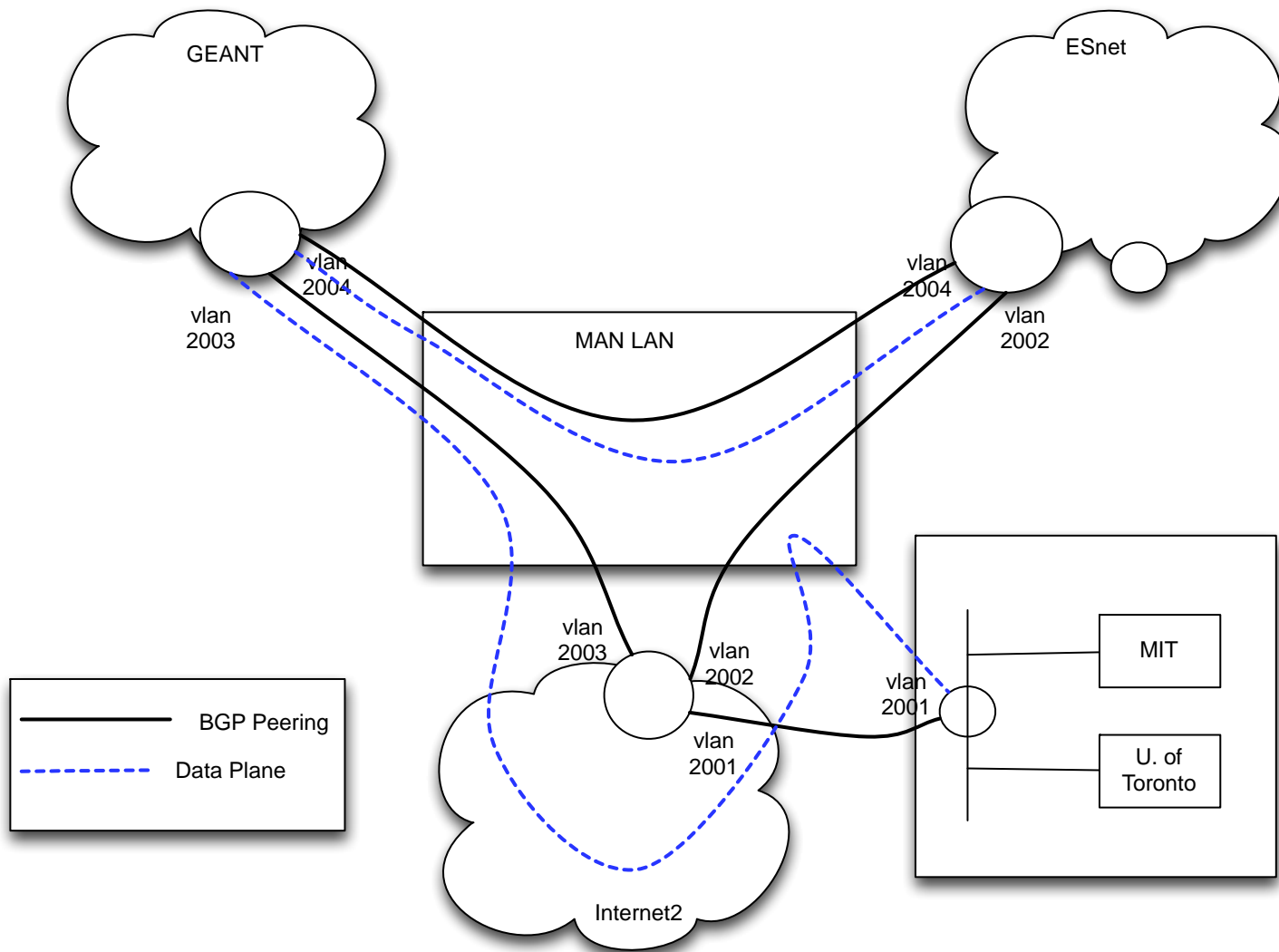
# LHCONE Agenda (contd.)

- The Amsterdam Architecture workshop (Dec. 2011) defined 5 activities:
  1. VRF-based multipoint service: a “quick-fix” to provide the multipoint LHCONE connectivity as needed in places today
  2. Layer 2 multipath: evaluate use of emerging standards like TRILL (IETF) or Shortest Path Bridging (SPB, IEEE 802.1aq) in WAN environment
  3. Openflow: There was wide agreement at the workshop that SDN is the probable candidate technology for the LHCONE in the long-term, however needs more investigations
  4. Point-to-point dynamic circuits pilot
  5. Diagnostic Infrastructure: each site to have the ability to perform end-to-end performance tests with all other LHCONE sites

# VRF Overview

- VRFs for Internet2, GEANT and Esnet are in place
- VLANs to establish peerings across MAN LAN have been established
  - 2002 Internet2-GEANT
  - 2003 Internet2-ESnet
  - 2004 GEANT-ESnet
- Proposal to separate control plane (BGP) from data plane for more optimal data transfers
- By setting up a shared VLAN within the exchange point, 3<sup>rd</sup> party next hop routing is enabled

# Sub-optimal Data Plane



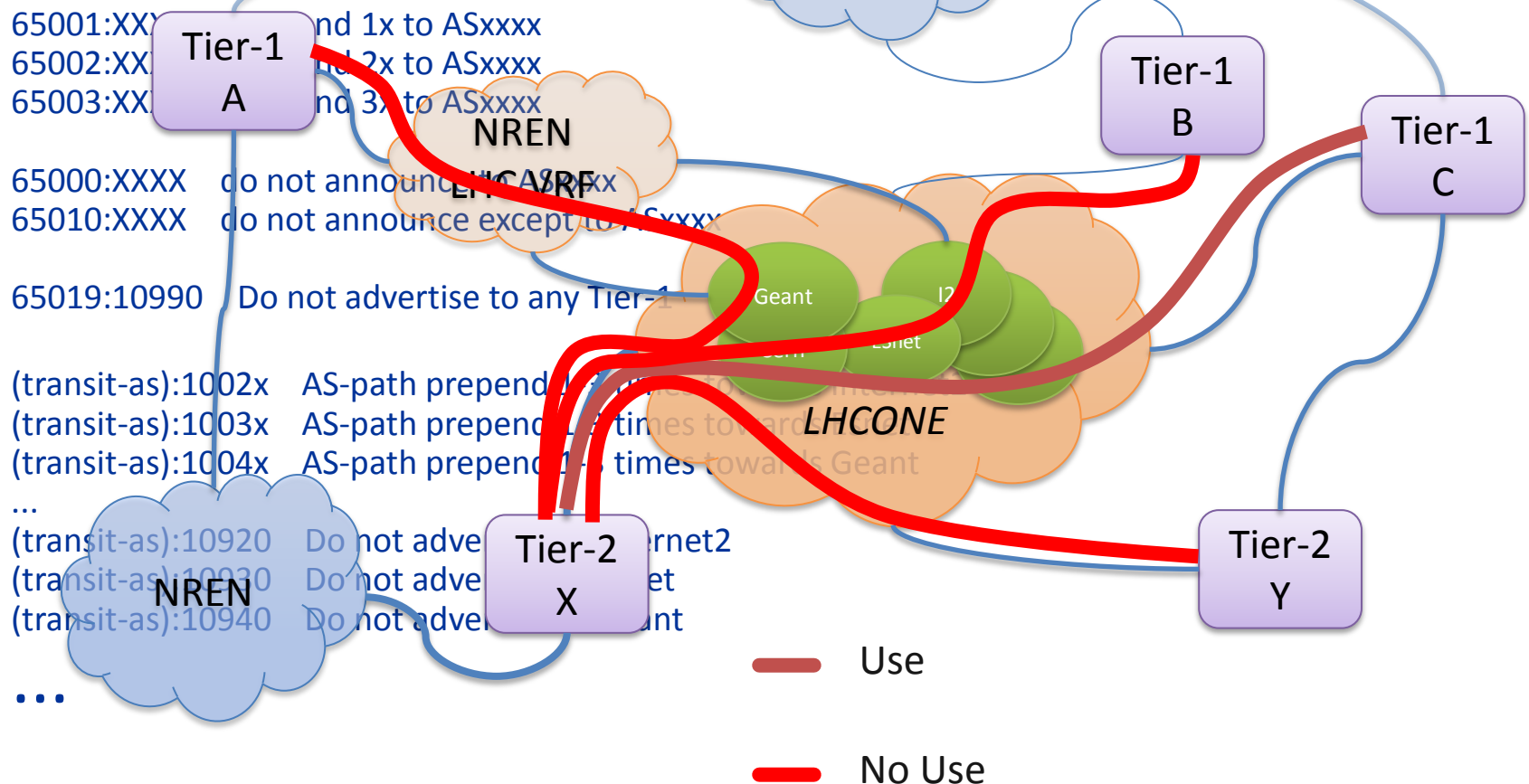
# Lunch Break & ALS Visit

- <http://www-als.lbl.gov/>



# LHCONE BGP communities

- Magnus Bergroth from Nordunet suggested some BGP communities (informational, operational communities)





# ACE/TP3 presentation

- ACE + TP3 are infrastructure projects funded by the US National Science Foundation
- America Connects to Europe (ACE) focuses on Europe.  
60Gbps deployed between EU-US
- TransPAC3 (TP3) focuses on Asia.  
30Gbps deployed between US and Asia
- Infrastructure is deployed and operational and is available for shared use by LHCONE



# perfSONAR MDM update

- Now supported by GEANT service desk
- New release, entirely package based, running on any standard RHEL, 32 or 64 bits server
- Nice looking web-interface (password protected ☹)
- DANTE-suggested choice for EU Tier2s

N.b. The two versions of perfSONAR are compatible at the protocol level, and work is underway to ensure compatibility of the Measurement Archives

# Diagnostic infrastructure status

- Minutes have links to latency & bandwidth graphs
- Richard explained what is required to monitor VRF access links & suggests building on perfSONAR (both flavours)
- Cacti traffic utilisation plots on GEANT/NREN VRF access links
- Eric showed a strawman proposal for LHCONE diagnostic service with the goals of:
  1. monitoring the health of LHCONE
  2. diagnosing and fixing problems detected by LHCONE
- On-demand tests, sparse mesh of regularly scheduled tests between sites; regular tests between VRFs (networks)
- one-way delay and latency measurements w/ history
- Start with the DICE Diagnostic Service definition in use and deployed by ESnet, GEANT, and Internet2

# Overview of TRILL/802.1aq

- TRILL=TRansparent Interconnection of Lots of Links (IETF)

Tries to solve spanning tree limitations (idea of Radia Perlman, inventor of the Spanning Tree Protocol)

May may be useful in L2 Ethernet domains within LHCONE

- 802.1aq - Shortest Path Bridging (SPB)
- Need to evaluate merits of both - research project

# Dinner at Caffè Venezia

- <http://www.caffevenezia.com/>





# Point-to-point virtual circuits pilot

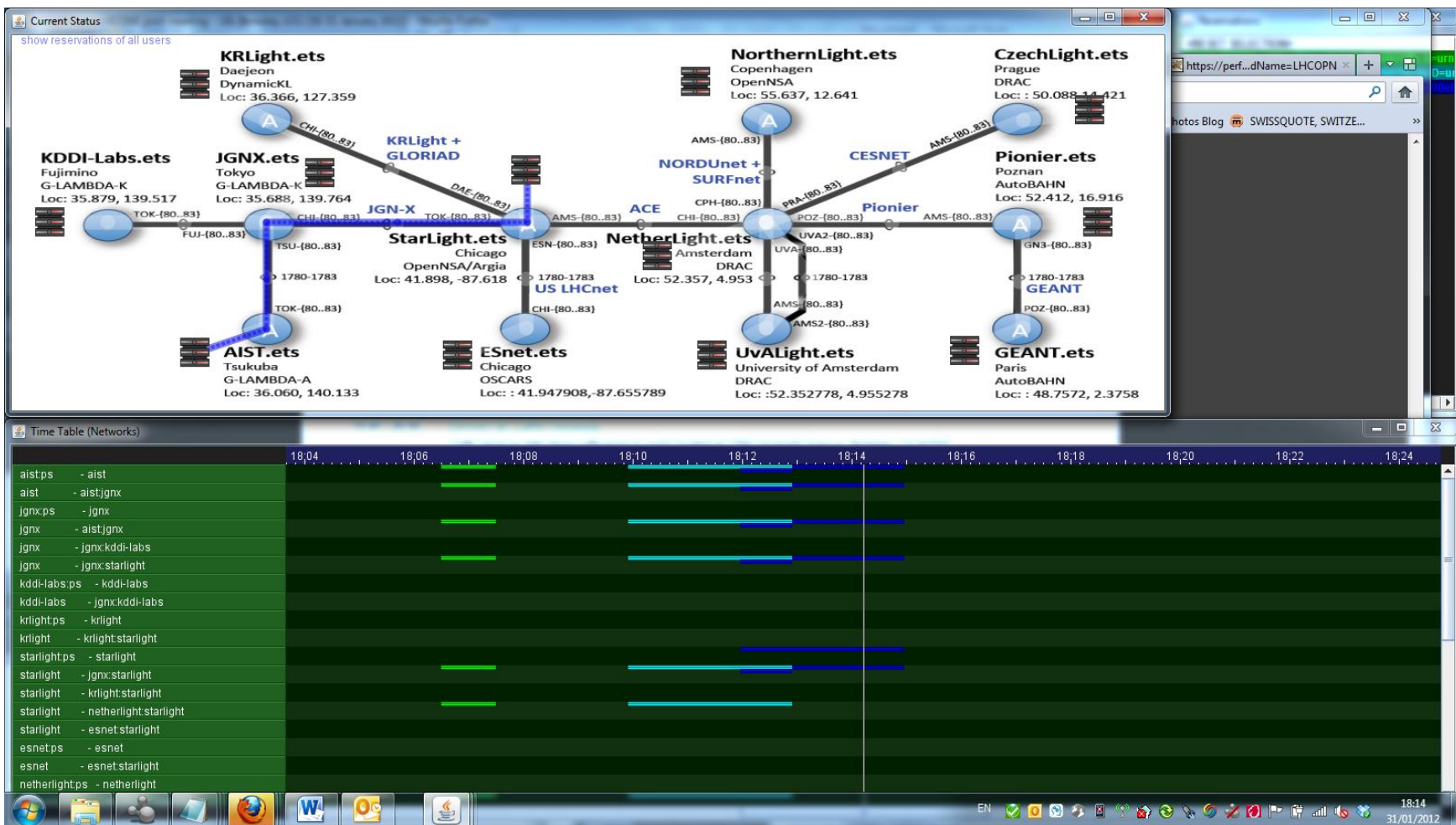
- Jerry Sobieski (NORDUnet) gave presentation on GLIF Automated GOLE Pilot Project and NSI

The GLIF Automated GOLE Pilot Project was initiated in 2010 to provide a global fabric of Open Lightpath Exchanges for the express purpose of maturing dynamic provisioning software

- NSI - Network Services Interface (requesting agents & provider agents)
- Example of NEXPreS given (astronomy)
- Need to understand how to run LHC applications on a BoD infrastructure – application and middleware impact.

# Point-to-point virtual circuits demo

- <http://163.220.30.174:8070/monitor.jnlp>



GDB update on Berkeley LHCOPN/LHCONE  
meeting J.Shade

08-JAN-2012

20



# Short-term LHCONE Point-to-point service

- OSCARS and AutoBAHN have interoperable IDC (Inter-Domain Controller) implementations
- ESnet, GEANT, Internet2, USLHCnet have interoperable point-to-point service definitions
- GLIF has implemented a low-bandwidth test bed implementing NSI protocol
- Early stages of investigating Software Defined Networking “Peering” underway

# Eric's Proposal

- LHCONE should start using interoperable production service deployments (e.g. ESnet Science Data Network, GEANT AutoBAHN, Internet2 Advanced Layer 2 Services, Open DRAC, etc.) now
- LHCONE should experiment with tested service deployments (e.g. GLIF deployment of NSI) now
- LHCONE should migrate along with interoperable production service deployments over time

# Dynamic Network Services

- Not presented, but available under “more information”
- Vision is the ability to partition network to:
  - Enable isolation of large flows and small flows
  - Enable predictable performance
  - Enable virtual privacy
  - Enable incremental scalability of underlying resources
- Cloud model for computing and services likely to extend to network resources

# LHCONE monitoring update

- Extensive Twiki page available at:  
<https://twiki.cern.ch/twiki/bin/view/LHCONE/SiteList>
- 15-member Dashboard Matrix is at:  
<https://130.199.185.78:8443/exda/?page=25&cloudName=LHCONE>
- Need a "before" baseline that we can compare with after LHCONE is in place.
- Note that this is not a final or permanent setup for LHCONE monitoring!
- No plans by ATLAS for full matrix network monitoring of all their sites!!

# VRF discussion

- Emulate Internet Tier1 ISPs (defined by connections to all other Tier1s) & have a VRF default free-zone.
- LHCONe participants connect to the default free zone:
  - Directly by peering with one or more core VRFs (e.g. DFN)
  - Indirectly by peering with one or more core VRFs through a Layer 2 connection (e.g. LEARN via OS3E to Internet2)
  - Directly by peering with one or more core VRFs via an IXP (e.g. MIT at MAN LAN)
- Challenge is to avoid unnecessary loops in the data plane
- A patron acts as the control plane agent for a LHCONe connector to an IXP
- Need to sort out who supplies addresses & who works out what BGP policies should be.



