# DYNES (DYnamic NEtwork System) & LHCONE (LHC Open Network Env.)

Shawn McKee
University of Michigan
Jason Zurawski
Internet2

**USATLAS Facilities Meeting**
October 11, 2011
Southern Methodist University
Dallas, Texas

# Overview/Outline

- I want to quickly review DYNES and LHCONE so everyone understands what they are, how they are related and be able to consider how we (USATLAS) might take advantage of them.
- Then I want to raise some discussion points
  - DYNES integration within the facility…how?
  - LHCONE participation and planning…when?
  - DYNES-LHCONE interactions…why?

# DYNES Summary

- NSF MRI-R2: DYnamic NEtwork System (DYNES, NSF #0958998)
- **What** is it?:
  - A nationwide cyber-instrument spanning up to ~40 US universities and ~14 Internet2 connectors
    - Extends Internet2s ION service into regional networks and campuses, based on OSCARS implementation of IDC protocol (developed in partnership with ESnet)
- **Who** is it?
  - A collaborative team including **Internet2**, **Caltech**, **University of Michigan**, and **Vanderbilt University**
  - Community of regional networks and campuses
  - LHC, astrophysics community, OSG, WLCG, other virtual organizations
- What are the **goals**?
  - Support large, long-distance scientific data flows in the LHC, other leading programs in data intensive science (such as LIGO, Virtual Observatory, and other large scale sky surveys), and the broader scientific community
  - Build a distributed virtual instrument at sites of interest to the LHC but available to R&E community generally

# LHCONE Summary

- LHCONE - **LHC Open Network Environment**
- Results of LHC Tier-2 network working group convened summer 2010. A merger of 4 "whitepapers" from the CERN LHCT2 meeting in January 2011 (See http://lhcone.net )
- LHCONE builds on the hybrid network infrastructures and open exchange points provided today by the major R&E networks on all continents
- **Goal: To build a global unified service platform for the LHC community**
- By design, LHCONE makes best use of the technologies and best current practices and facilities provided today in national, regional and international R&E network

# DYNES Concepts

- Solutions:
  - **Dedicated bandwidth** (over the entire end-to-end path) to move scientific data
  - Invoke this "**on demand**" instead of relying on permanent capacity (cost, complexity)
  - **(Co-)Exists** in harmony with traditional IP networking
  - Connect between facilities that scientists need to access
  - **Integration with data movement applications**
    - Invoke the connectivity when the need it, based on network conditions
- Prior Work
  - "Dynamic Circuit" Networking – creation of Layer 2 point to point VLANs
  - Transit the Campus, Regional, and Backbone R&E networks
  - Software to manage the scheduling and negotiation of resources

# DYNES: Why Not Static Circuits or Traditional, General Purpose Networks ?

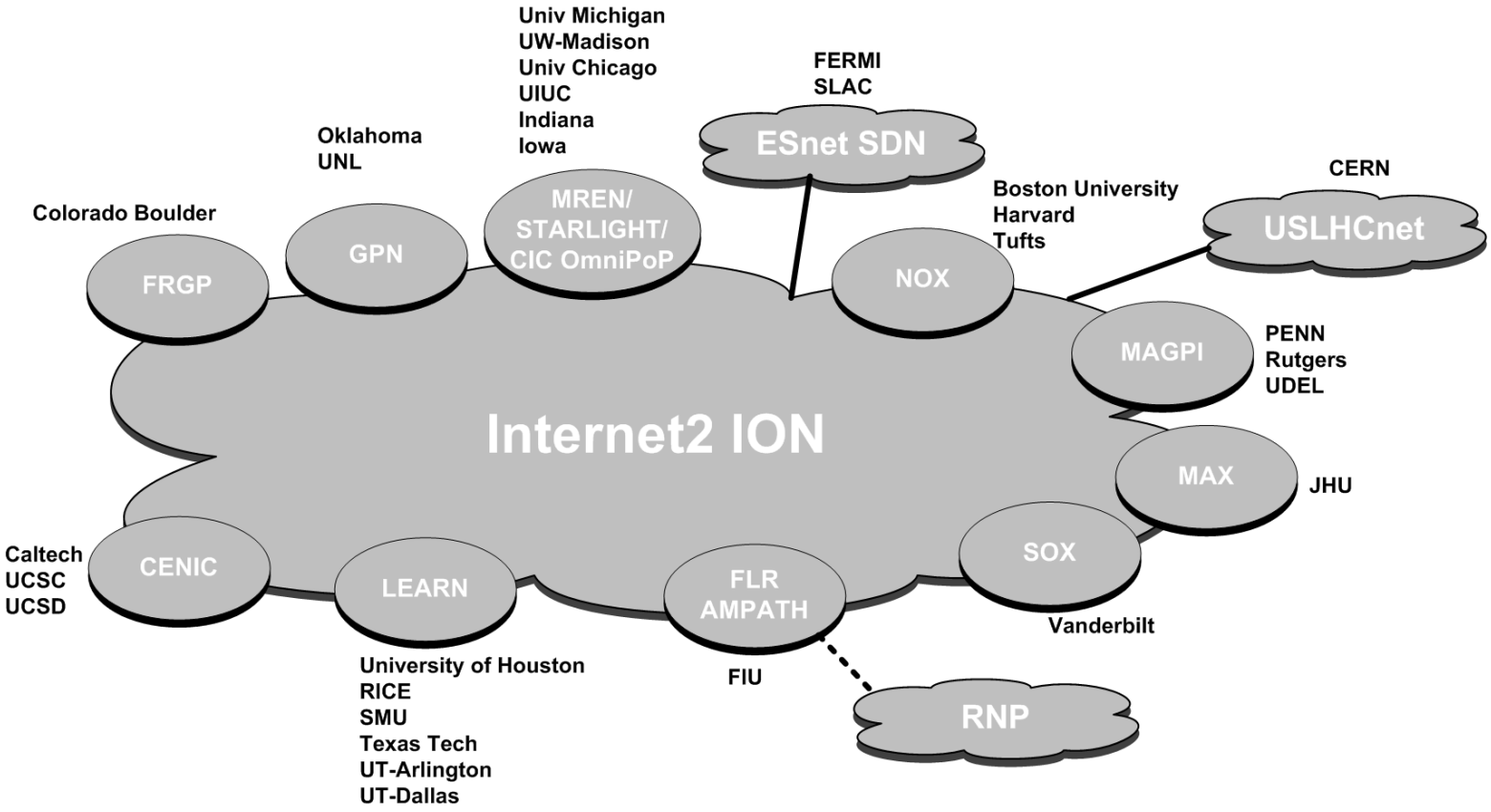
- Separation (physical or logical) of the dynamic circuit-oriented network from the IP backbone is **driven** by the need to meet different functional, security, and architectural needs:
  - Static "nailed-up" circuits **will not scale.**
  - GP network firewalls incompatible with enabling large-scale science network dataflows
  - Implementing many high capacity ports on traditional routers would be **very expensive**
    - Price balance: Worse in the next generation: 40G and 100G general purpose router ports are several hundred k$ each.

# DYNES Scope

- Initial Deployment Locations:
  - 30 End Sites
  - 8 Regional Networks
  - Collaboration with like minded efforts (DoE ESCPS and StorNet)
- Accepting additional applications
  - dynes-questions@internet2.edu
- Supporting all data-intensive, distributed science - early focus on Physics (LHC) sites

# DYNES Infrastructure Overview

- DYNES Topology
  - Based on Applications received
  - Plus existing peering wide area Dynamic Circuit Connections (DCN)

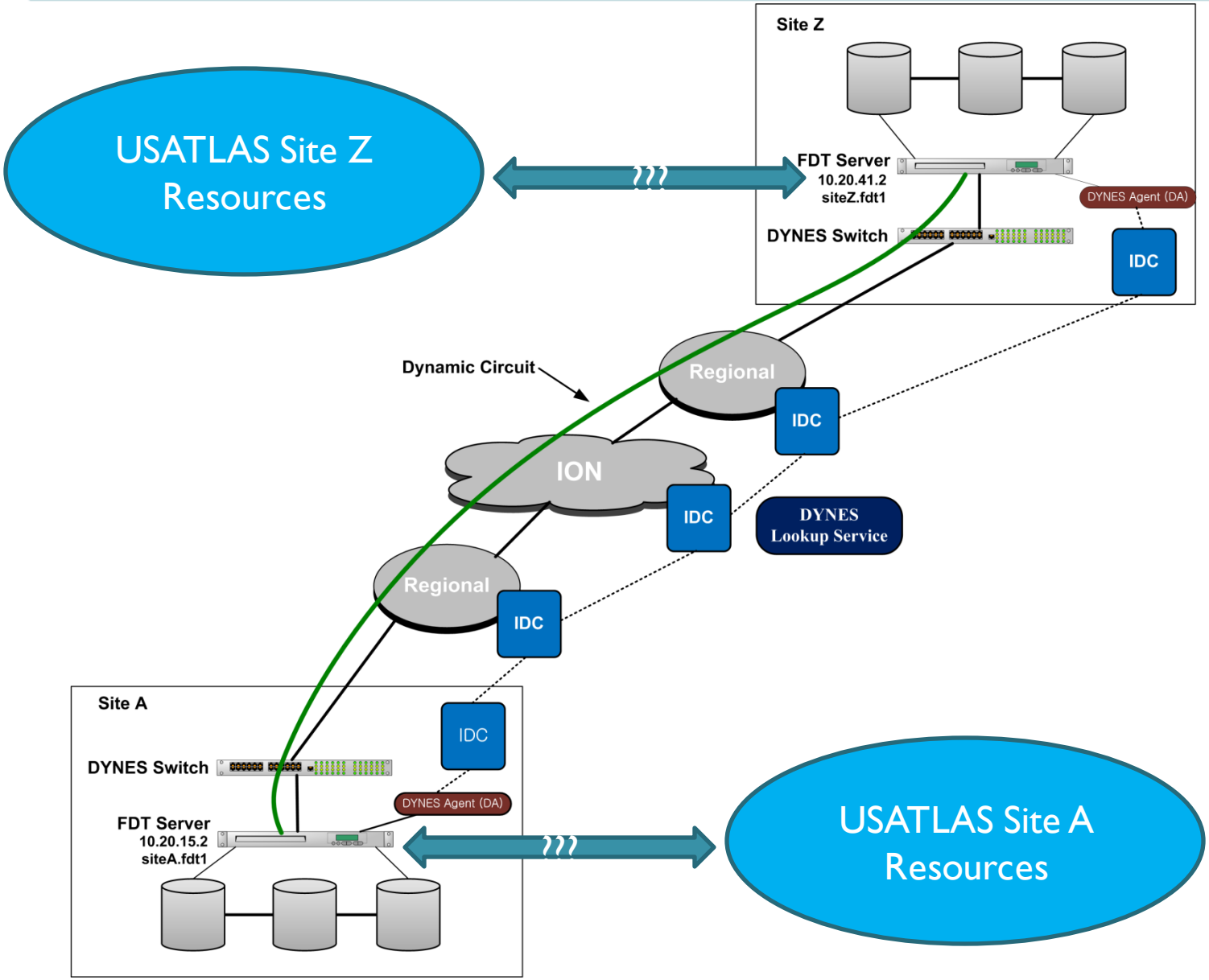


# DYNES Standard Equipment and Software

- Inter-domain Controller (IDC) Server and Software
  - IDC creates virtual LANs (VLANs) dynamically between the FDT server, local campus, and wide area network
  - Dell R410 (1U) Server
  - OSCARSv0.6 and DRAGON Software
- Fast Data Transfer (FDT) server
  - Fast Data Transfer (FDT) server connects to the disk array via the SAS controller and runs the FDT software
  - Dell R510 (2U) Server
- DYNES Ethernet switch options:
  - Dell PC6248 (48 1GE ports, 4 10GE capable ports (SFP+, CX4 or optical)
  - Dell PC8024F (24 10GE SFP+ ports, 4 "combo" ports supporting RJ45; SFP+ supporting optical)
- perfSONAR Monitoring

***Many thanks to Dell for supporting DYNES with LHC pricing!***

# DYNES Data Flow Overview



# DYNES Current Status

- 4 Project Phases
  - Phase 1: Site Selection and Planning (***Completed in Feb 2011***)
  - Phase 2: Initial Devel./Deployment (***Feb 2011 through July 2011***)
  - Phase 3: Scale up to Full Deployment (***July 2011 through Aug 2012***)
    - **DYNES Participant Deployment (*July 2011-November 2011*)**
    - Full-scale System Development, Testing & Evaluation (*November 2011 – August 2012*)
  - Phase 4: Integration at Scale; Transition to Routine O&M (***Aug 2012 through August 2013***)
  - ***Details in supplemental slides at end of this talk***

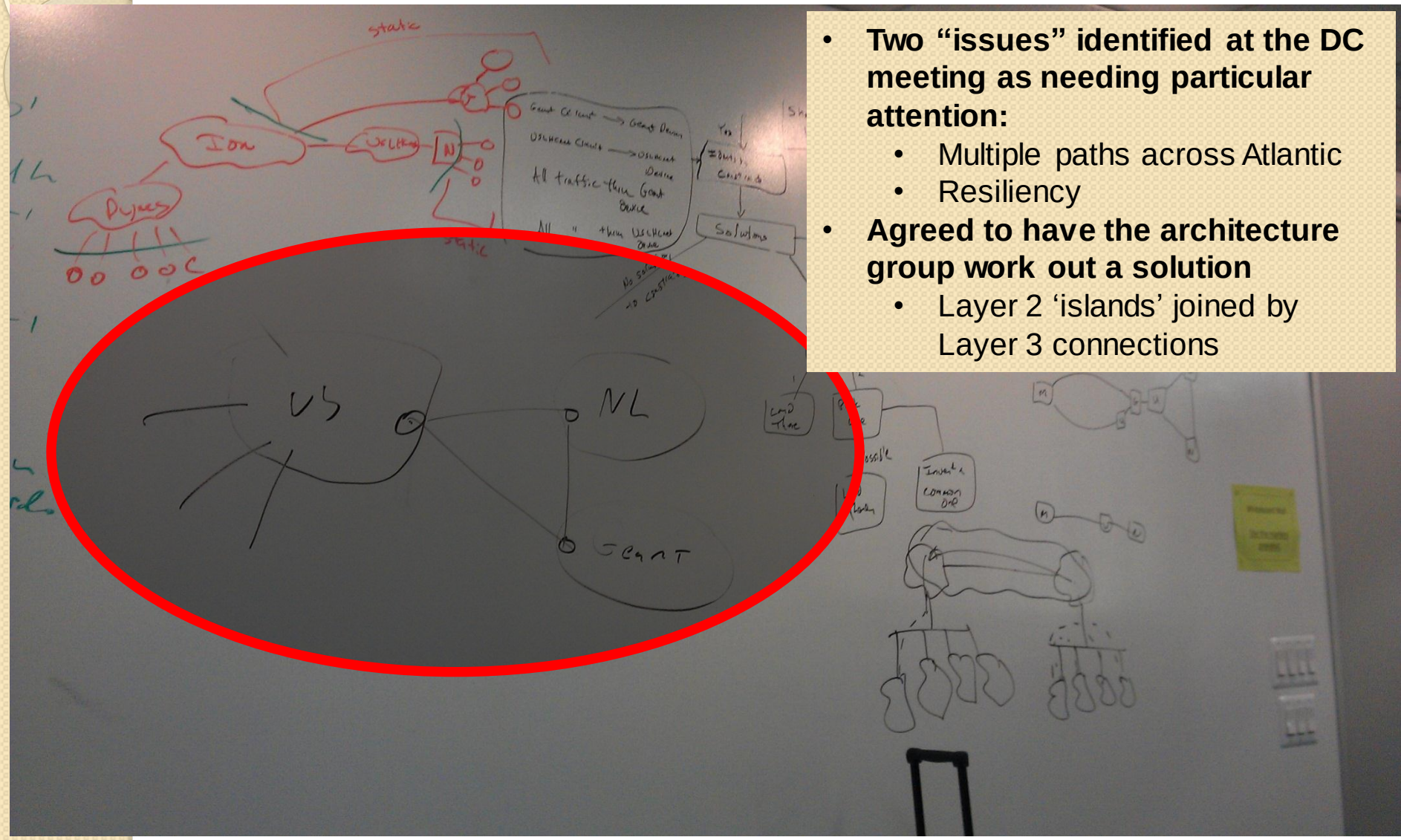
- A DYNES Program Plan document along with many other documents are available at:
  - http://www.internet2.edu/dynes

- Questions can be sent to the mailing list:
  - dynes-questions@internet2.edu

# LHCONE Goals

- The LHCONE effort has been moving ahead via meetings during 2011
- **GOALS**: Identify, organize and manage LHC related flows from the Tier-2 and Tier-3 sites
- Having a way to identify LHC related flows helps by:
  - Allowing this traffic to be "engineered" by whatever means exist within the infrastructure
  - Makes monitoring much more straightforward
  - Enables quicker problem isolation and resolution
  - Motivates additional resources for LHC needs
- How to get LHCONE moving?

# “Joe’s Solution” – Result of June 2011 Meeting

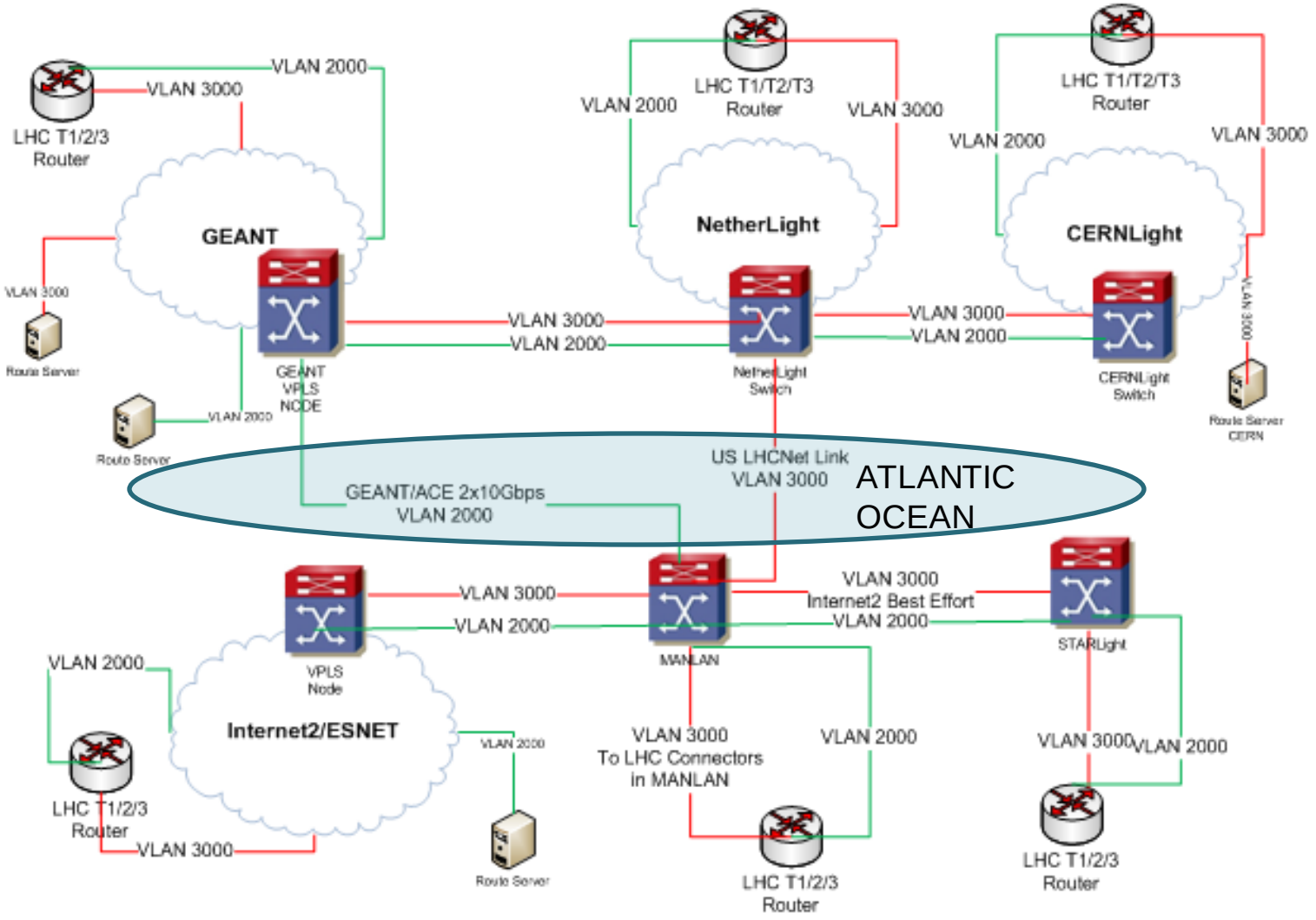



- **Two “issues” identified at the DC meeting as needing particular attention:**
  - Multiple paths across Atlantic
  - Resiliency
- **Agreed to have the architecture group work out a solution**
  - Layer 2 ‘islands’ joined by Layer 3 connections

# LHCONE Pilot

- Multipoint:
  - Domains interconnected through Layer 2 switches
  - Two VLANs (nominal IDs: 3000, 2000)
    - VLAN 2000 configured on GEANT/ACE transatlantic segment
    - VLAN 3000 configured on US LHCNet transatlantic segment
  - Allows to use both TA segments, provides TA resiliency
  - 2 route servers per VLAN
    - Each connecting site peers will all 4 route servers
  - Enables up to 25G on the Trans-Atlantic routes for LHC traffic.
- Point to Point:
  - Suggestion: Build on efforts of DYNES and DICE-Dynamic service
  - DICE-Dynamic service being rolled out by ESnet, GÉANT, Internet2, and USLHCnet
    - Remaining issues being worked out
    - Planned commencement of service: October, 2011
    - Built on OSCARS (ESnet, Internet2, USLHCnet, RNP) and AutoBAHN (GÉANT), using IDC protocol

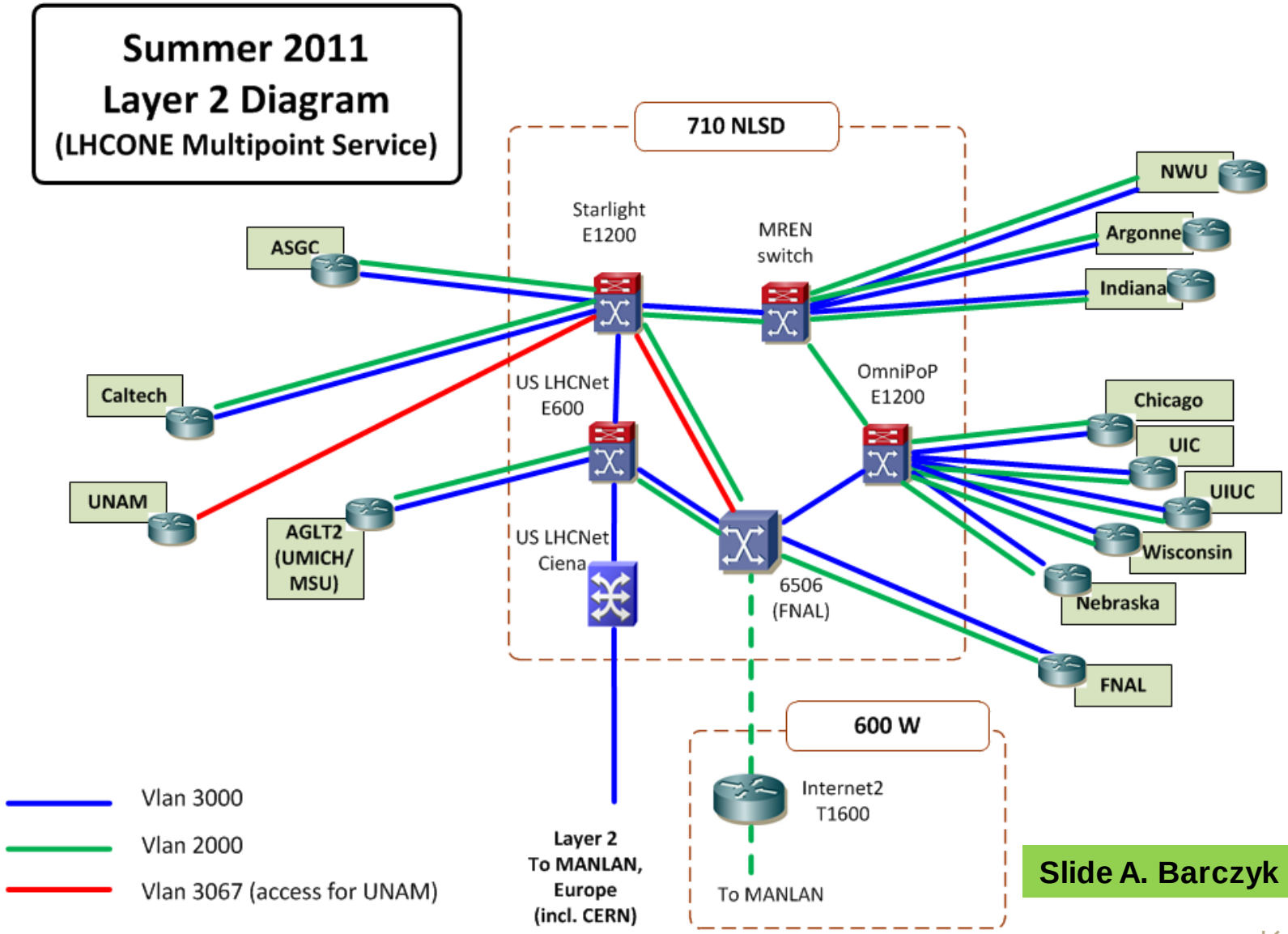
# LHCONE Interim Solution




Loop free L2 topology for P2MP service

Slide A. Barczyk

# LHCONE in Starlight (Interim)




Slide A. Barczyk

# LHCONE Pilot … But Where Are We Really?

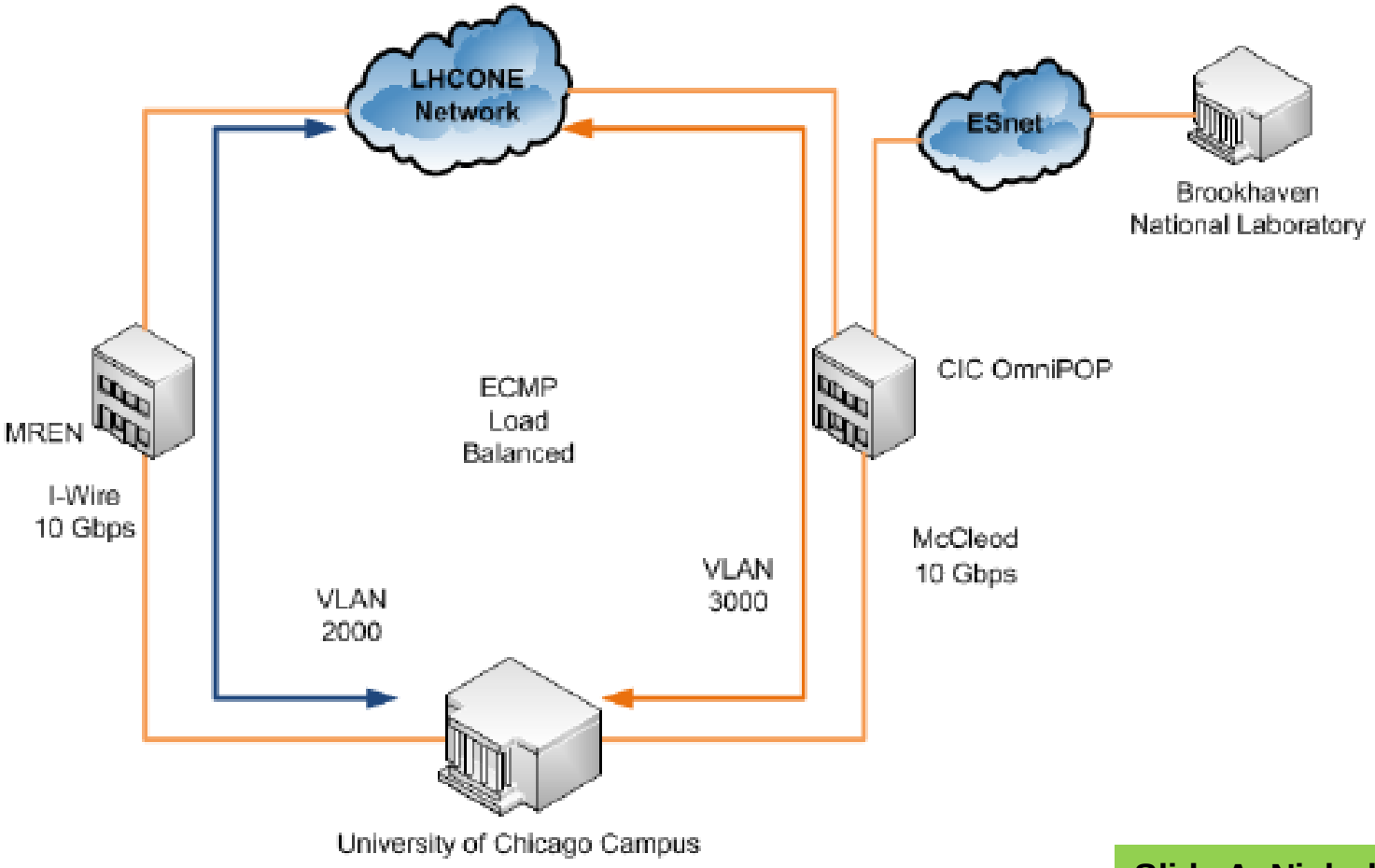
- Continents Connected?
  - Access Switches in the US and Europe, work to get Asia connected
  - VLANs stretched about as far as they can stretch
  - Capacity identified (e.g. NA, EU and TA) – but note that in many cases this is not as much as would be traditionally available (e.g. it's a pilot built from 'spare' parts)
  - Basic functionality in place, end sites are joining – can success be measured?
- Technology "Sound"?
  - Lots of questions about this, Layer2 Islands w/ Layer3 connections is hard to manage (and debug)
  - Already routing loop situations (e.g. the '2 VLAN Solution' is not really a sound solution)
  - Serious doubts about what happens as more TA links get added – more sites in the US/EU come online with multiple access locations (e.g. the situatuon in the Chicago region is hard to manage…)

# LHCONE Pilot … But Where Are We Really?

- Customers Happy?
  - A very relevant and timely question asked by Michael at the SARA Meeting
  - Oversight is lacking
  - Not a lot of information being shared with the 'stakeholders'. How to fix?
- Architectural Changes?
  - Some proposals to scrap the whole thing and start over
  - Some questions about if the scalability issues we see today will cloud the overall concept

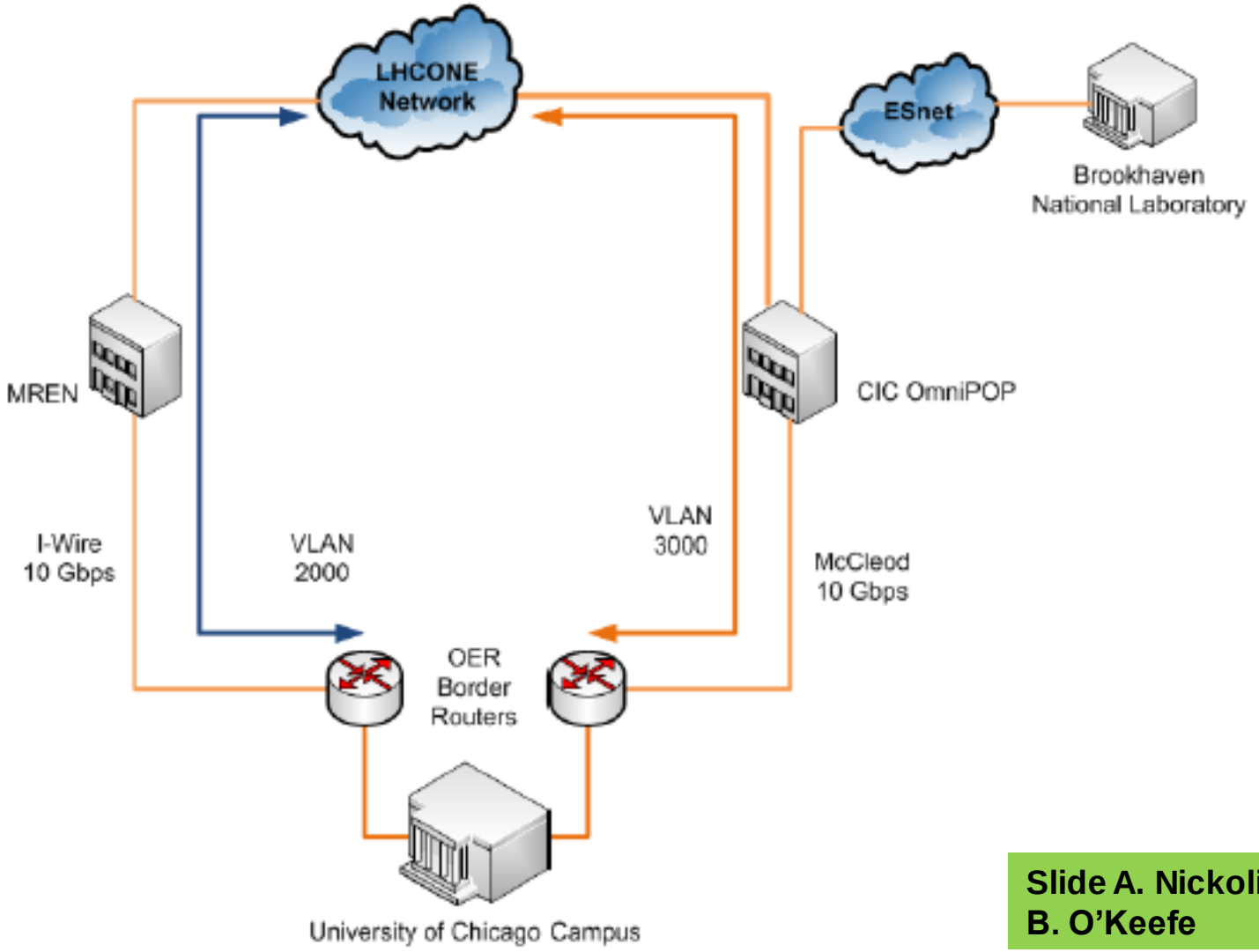
- **It is critical that we stay engaged in LHCONE while realizing the current limitations implied by building something from donated/reused components and effort.**

# LHCONE MWT2_UC L2 Example




Slide A. Nickolich
B. O'Keefe

# LHCONE MWT2_UC L3 Example




**Slide A. Nickolich**
**B. O'Keefe**

# LHCONE Planned "Features"

- LHCONE should be a well defined architecture once it is put into production
- It has always included the following concepts:
  - Multipoint VLAN (extent is under debate)
  - Point-to-Point connections
  - A routed IP component (extent is under debate)
  - Traffic-engineering (L1, L2 and L3 options)

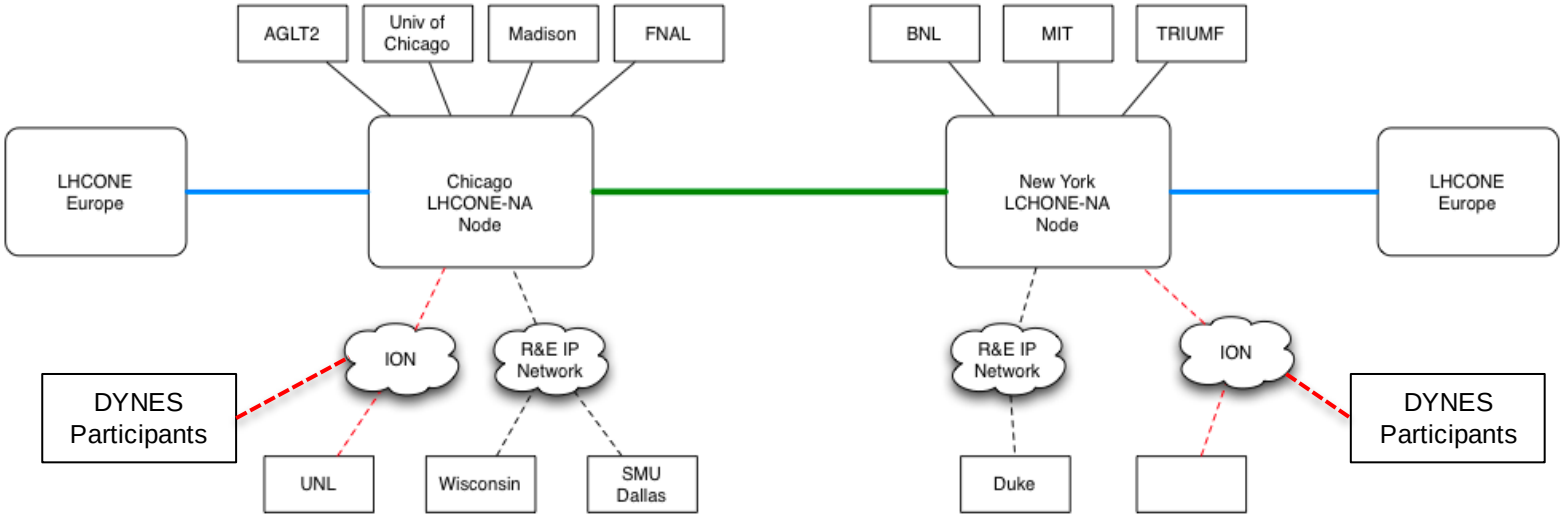
# LHCONE Planning for USATLAS

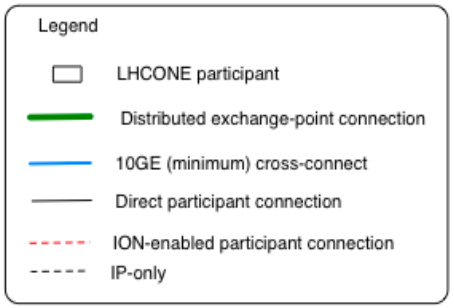
- For USATLAS, we need to determine how best to participate within the LHCONE effort.
- Practically, this means how and when our sites might enable LHCONE.
- We need to have well-defined metrics in place which we gather before and after adding a site to LHCONE
- We can discuss this in a bit…

# DYNES Integration in USATLAS

- One goal of DYNES is getting the needed network capabilities pushed out to user-sites (last-mile). This helps everyone to be able to utilize the new DCN capabilities.
- A natural primary use of DYNES for USATLAS is to allow Tier-3 sites to prioritize data flows as needed, primarily from Tier-2s or other Tier-3s.
  - We need to start testing **how** we can integrate DYNEs capabilities *semi-transparently* for USATLAS users. (First steps…)
  - Goal is to improve the end-user experience in getting larger amounts of data in a timely way compared to the current situation.
  - Eventually should be completely transparent and integrated with our tools such that DYNES sites automatically utilize DYNES.

# DYNES and LHCONE:
# DYNES as an "On-ramp" to LHCONE

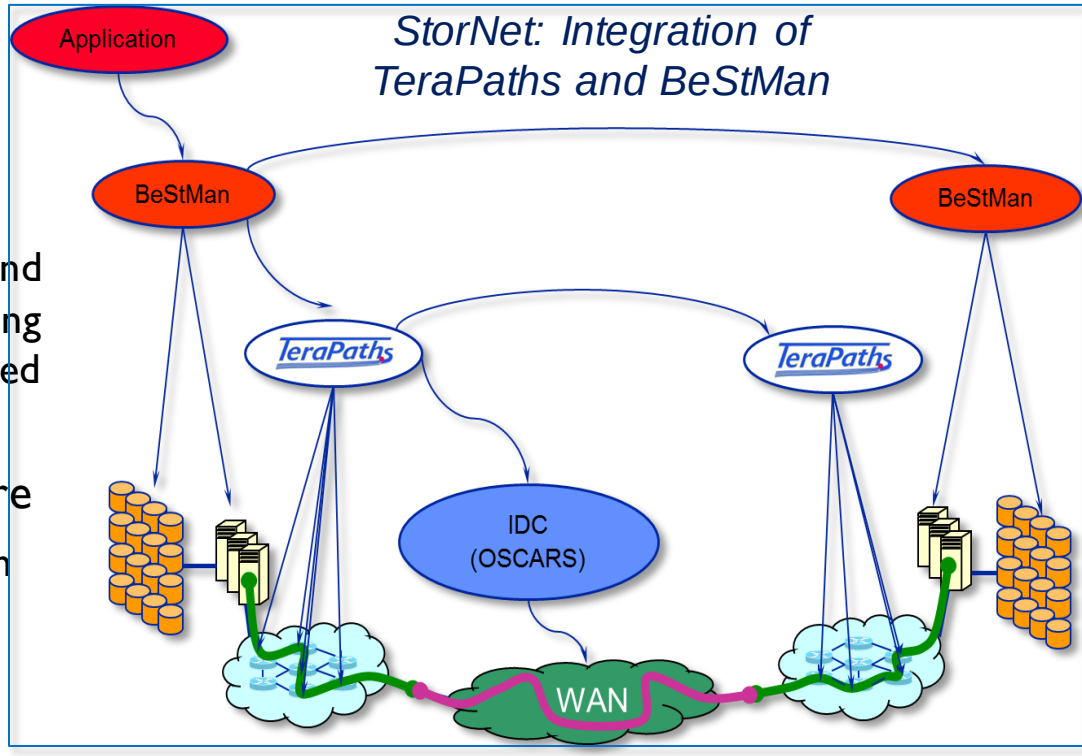


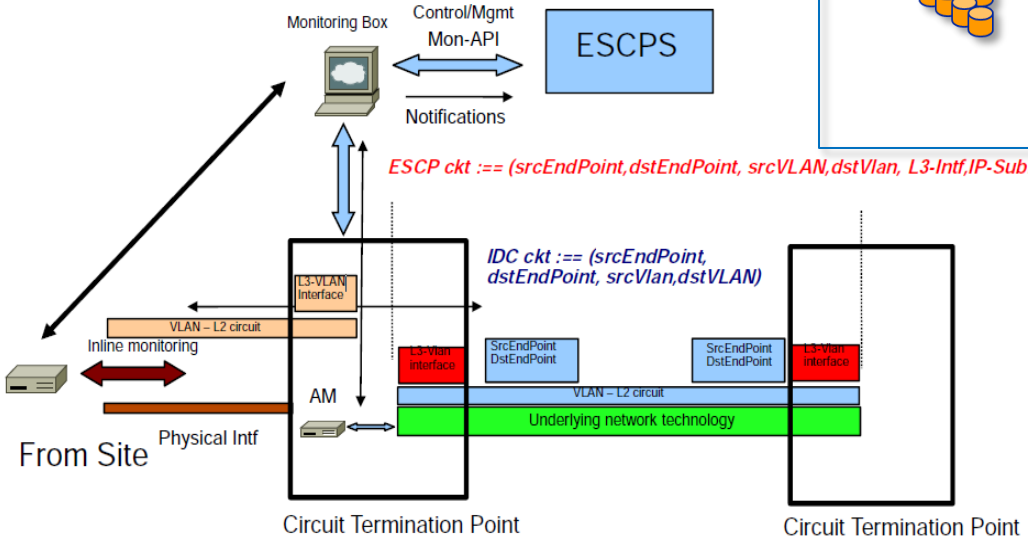




- DYNES Participants can dynamically connect to Exchange Points via ION Service
- DYNES operation is currently based on and "end-to-end" model, so will need to make some adjustment with respect to how to extend services/connections thru Exchange Point to service endpoints
  - Dynamic Circuits thru and beyond the exchange point?
  - Hybrid dynamic circuit and IP routed segment model?

# Other Projects Leveraging Dynamic Circuits at the HEP sites: StorNet, ESCPS

- StorNet – BNL, LBNL, UMICH
  - Integrated Dynamic Storage and Network Resource Provisioning and Management for Automated Data Transfers
- ESCPS – FNAL, BNL, Delaware
  - End Site Control Plane System




- **Building on previous developments in and experience from the TeraPaths and LambdaStation projects**

# Open Questions

- Integration into ATLAS/CMS software stack
  - Started a thread with PhEDEx developers, Shawn/Michael/Jason had a minor thread about DQ2
  - Should the 'intelligent' data movers know there is a network choice? Should the low level tool only be aware?
- Enabling at more T3s
  - More than in our current set. We will most likely have more funding available in the end – candidates?
- Reaching Europe, Asia, South America (Australia? Africa?)
  - Can plug into RNP/GEANT pretty easily.
  - APAN has OSCARS working to a certain extent in Asia
  - IRNC (NSF) international links have to support OSCARS, this gets the tech into South America/India/Australia/Asia/Europe, etc.
  - What is most important, and should be pursued?
- LHCONE
  - Lots more questions in this area than answers…
    - One positive thing is that DYNES will enable dynamic access into VLANs … but will LHCONE still be pushing the concepts discussed this year?
  - Is all the work beneficial to experiments?

# DYNES/LHCONE Discussion(1)

- How to start the DYNES integration into USATLAS?
  - Requires a deployed "instrument" (Feb 2012?)
  - Need to think about ATLAS DDM, Xrootd Federation, standard tools and how best to connect with DYNES…can influence/modify DYNES "design"
  - Familiarize users with DYNES first, then begin "integration"? Or start out now with a subset of the facility?

# DYNES/LHCONE Discussion(2)

- How best to test and influence LHCONE?
  - Current solution is temporary and perhaps not "production" level.
  - Testing for Tier-2's will be interesting to verify:
    - Performance doesn't decrease
    - Problems (esp. trans-oceans transfers) are less visible and/or easier to debug/fix
  - Suggest we define a specific timeline to have AGLT2 and/or MWT2 test LHCONE as is.
    - Benchmarks for current setup?
    - Comparison after transition to LHCONE participation

# Summary and Other Discussion

- LHCONE and DYNES are getting close to being useable for USATLAS.
- There are other resource projects that may prove useful once we have circuits.
- We need to be planning for how to utilize these capabilities for the facility
  - Our input can help ensure they will meet our needs
- Other issues we should cover? Questions?

# DYNES/LHCONE References

- DYNES
  - http://www.internet2.edu/dynes
- LHCONE
  - https://twiki.cern.ch/twiki/bin/view/LHCONE/WebHome
- OSCARS
  - http://www.es.net/oscars
- DRAGON
  - http://dragon.east.isi.edu
- DCN Software Suite (DCNSS)
  - http://wiki.internet2.edu/confluence/display/DCNSS/
- FDT
  - http://monalisa.cern.ch/FDT/

DYNES and LHCONE supplemental information

# ADDITIONAL SLIDES

# DYNES Demo

- Before Jason's talk let's see a quick example of how DYNES can work…

# DYNES IDC

- Inter-domain Controller (IDC) Server and Software
  - IDC creates virtual LANs (VLANs) dynamically between the FDT server, local campus, and wide area network
  - IDC software is based on the OSCARS and DRAGON software which is packaged together as the DCN Software Suite (DCNSS)
  - DCNSS version correlates to stable tested versions of OSCARS. The current version of DCNSS is v0.5.4.
  - Initial DYNES deployments will include both DCNSSv0.6 and DCNSSv0.5.4 virtual machines
    - Currently XEN based
    - Looking into KVM for future releases

# DYNES FDT

- The DYNES Agent (DA) will provide the functionality to request the circuit instantiation, initiate and manage the data transfer, and terminate the dynamically provisioned resources. Specifically the DA will do the following:
  - Accept user request in the form of a DYNES Transfer URLs indicating the data location and ID
  - Locates the remote side DYNES EndPoint Name embedded in the Transfer URL
  - Submits a dynamic circuit request to its home InterDomain Controller (IDC) utilizing its local DYNES EndPoint Name as source and DYNES EndPoint Name from Transfer URL as the destination
  - Wait for confirmation that dynamic circuit has been established
  - Starts and manages Data Transfer using the appropriate DYNES Project IP addresses
  - Initiate release of dynamic circuit upon completion