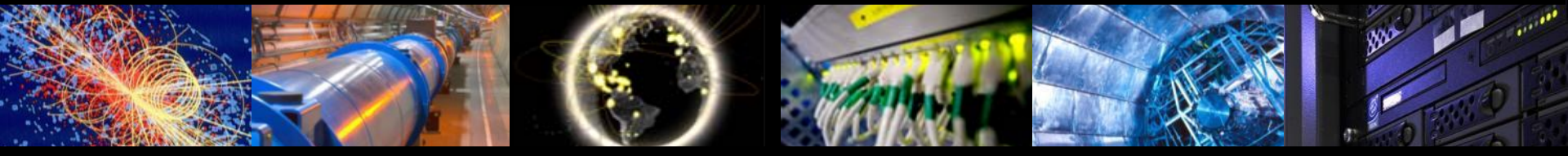


# WLCG Network WG (WNWG) – a status report

Michael Ernst  
Brookhaven National Laboratory



# From the April C-RRB Report

“The CRSG encourages close collaboration of the different centres with the experiments to continue the implementation of intelligent storage management policies to allow efficient and cost-effective access to data. **In particular the implications on network bandwidth for best-use of resources should be considered.** We consider this issue very relevant for the operation of the LHC experiments after 2014.”

<http://indico.cern.ch/getFile.py/access?contribId=16&sessionId=2&resId=0&materialId=slides&confId=174803>

- This can only succeed if the characteristics of the underlying network infrastructure are well understood by application developers and those operating central services of the experiments, and if the evolving requirements regarding data flows and their respective quantities are understood by the network providers
- Until recently there was no forum that supported the dialogue between application developers and network providers
  - LHCOPN working group not serving as a communication forum between application developers and network experts
  - With LHCONE and evolving network technology (e.g. “Intelligent Networks”) many more organizations (e.g. DANTE, EU NRENs, I2 with their operations and research branches) in addition to Tier-1 sites and CERN interact directly with LHC applications and sites

# WLCG Facilities

- Today in WLCG our primary concern regarding resources is CPU and storage capacities
  - All LHC experiments are increasingly relying on networks to provide connectivity between distributed facility elements across all levels
    - E.g. ATLAS is ramping up the usage of T2 storage for primary datasets
  - Wide Area Network infrastructure is complex
    - Besides the (mostly) predictable service of the OPN the vast majority of traffic (i.e. for ATLAS and CMS) has been using the GPN
    - With LHCONe there is a 3<sup>rd</sup> piece of network infrastructure
      - Meant to serve T1  $\leftrightarrow$  T2/T3, T2  $\leftrightarrow$  T2, T2  $\leftrightarrow$  T3 needs
      - Evolution & Deployment largely driven by network service providers
        - » LHC experiments not properly represented at planning meetings
  - With evolving Computing Models traffic flows change significantly (e.g. vanishing T1/T2 differences, i.e. in ATLAS)
    - Propose the WG conduct a study with the goal to determine what kind of network and how much networking is necessary, today and in the next ~5 years

# Issues

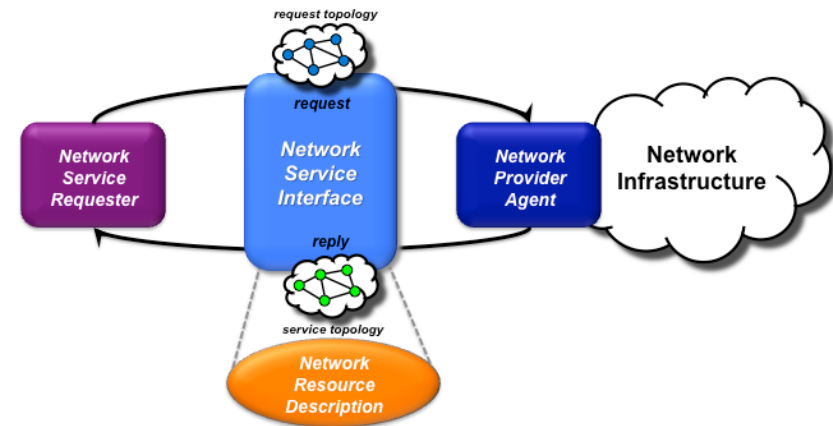
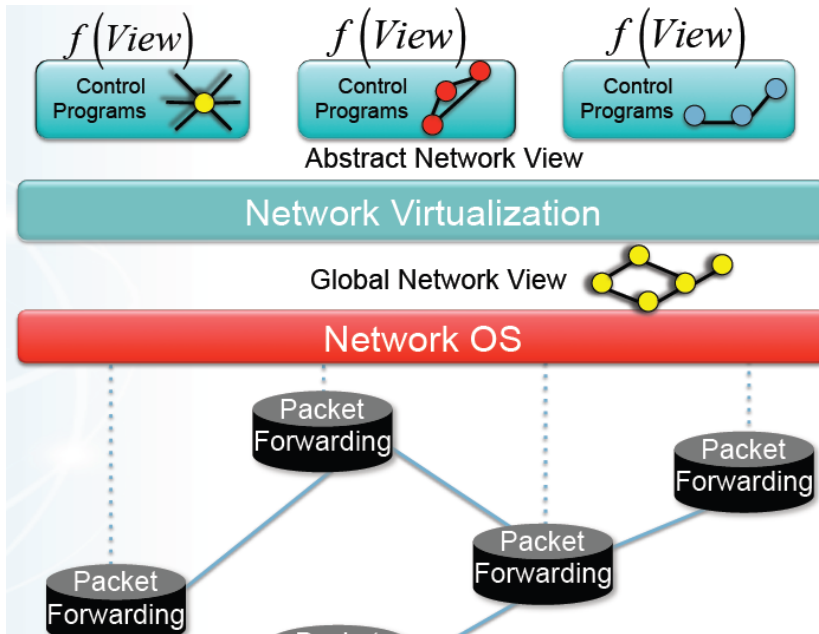
- Global Networking between sites has evolved into a service that is vital to the success of LHC Computing
  - Network providers are guessing how experiments are using and making assumptions as to how they would be using the service in the future
    - At LHCONe planning meetings I am hearing, frequently, requests for input but there is no authoritative/coherent information provided by the WLCG community
      - They are still referring to the “Bos/Fisk note”
        - » While an excellent foundation for a collaboration to start LHCONe, makes me wonder (if not nervous) how this can help them to implement our requirements in detail
      - Requests for resources from providers to their funding agencies on our behalf requires them to know what our needs are, and will be in the future
    - LHC Experiments realize that recent developments in networking go well beyond packet exchange through a “passive” pipe
      - Network service providers looking to engage us into traffic engineering
      - (Dynamic) Circuit technology (p2p service) w/ additional b/w for high priority data replication requires stakeholders’ active participation
      - Evolving service technologies are engaging applications in actively managing data flows (“Intelligent Network Services”, see next slides)
        - » While interfaces on the network side “ready”, potential benefit to LHC applications not understood and therefore not exploited

# WLCG Network Working Group (WNWG): Scope

- Focusing on network infrastructure and network services in support of distributed applications of the LHC Experiments
  - Aims at finding common solutions across the experiments
  - Includes definition & implementation of the interface between LHC community and network service providers
    - Meant to address high(er) level issues (vs. daily operational ones)
    - Not to interfere with procedures used by site/local network providers
    - Purpose is to address functionality & performance issues all along the vertical stack (from layer 2 to application)
    - Voice LHC requirements in consolidated & coherent fashion
  - Technical representatives for planning process and liaison on behalf of stakeholders for communication with providers and developers
    - Includes definition, setup and tracking of R&D activities
    - Foundation for informed choices
    - Agree on strategic directions and, if possible, develop a plan that will be presented and discussed at GDB and MB meetings

# Software Defined Networking (SDN) – This looks like a promising Technology

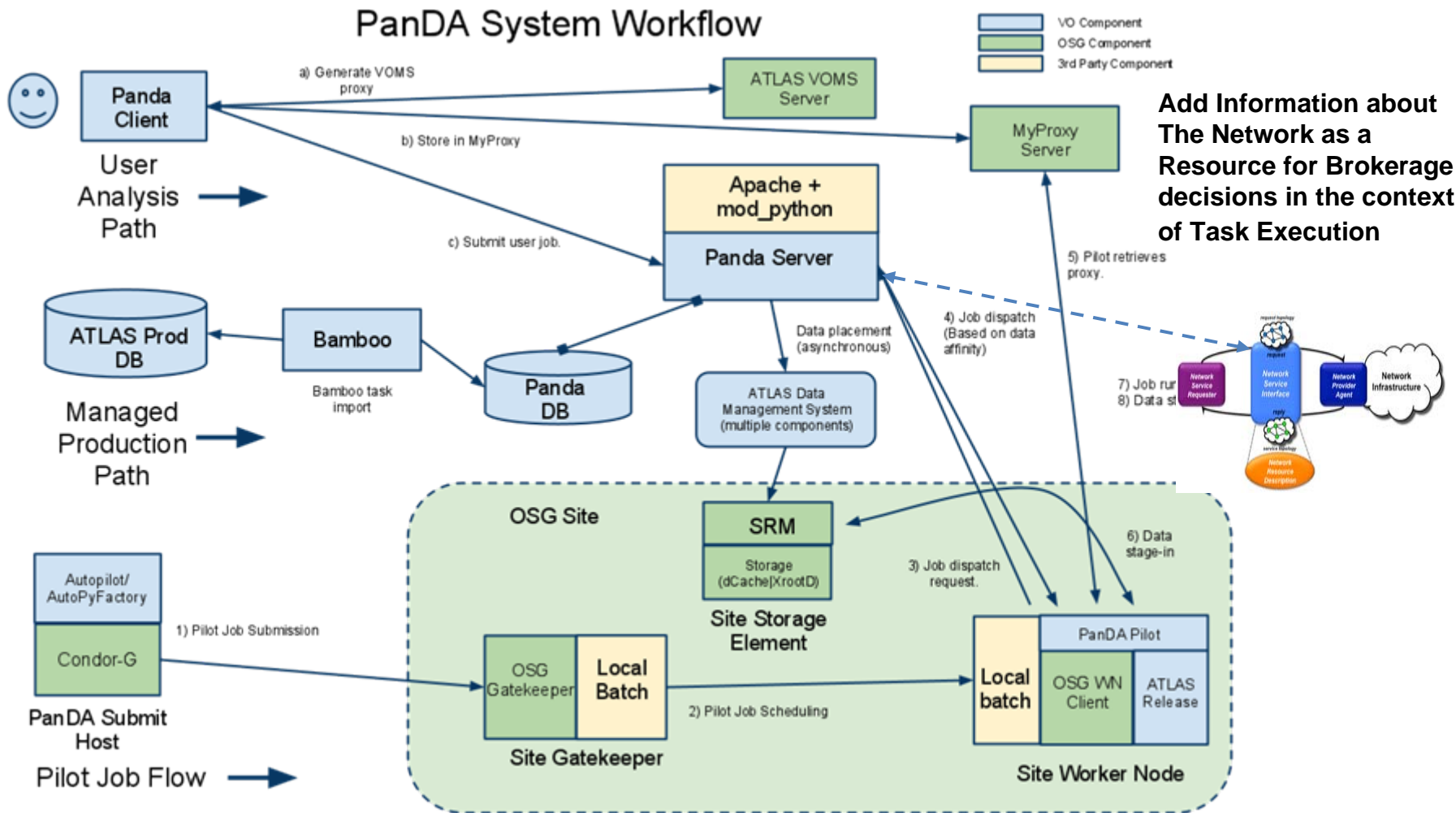
- SDN Paradigm - Network control by applications; provides an API to externally define network functionality
  - Enabler for applications to fully exploit available network resources
  - OpenFlow, a SDN implementation, widely adopted by Industry (Network Equipment Manufacturers, Google, Cloud Computing)



Network Services Interface Framework

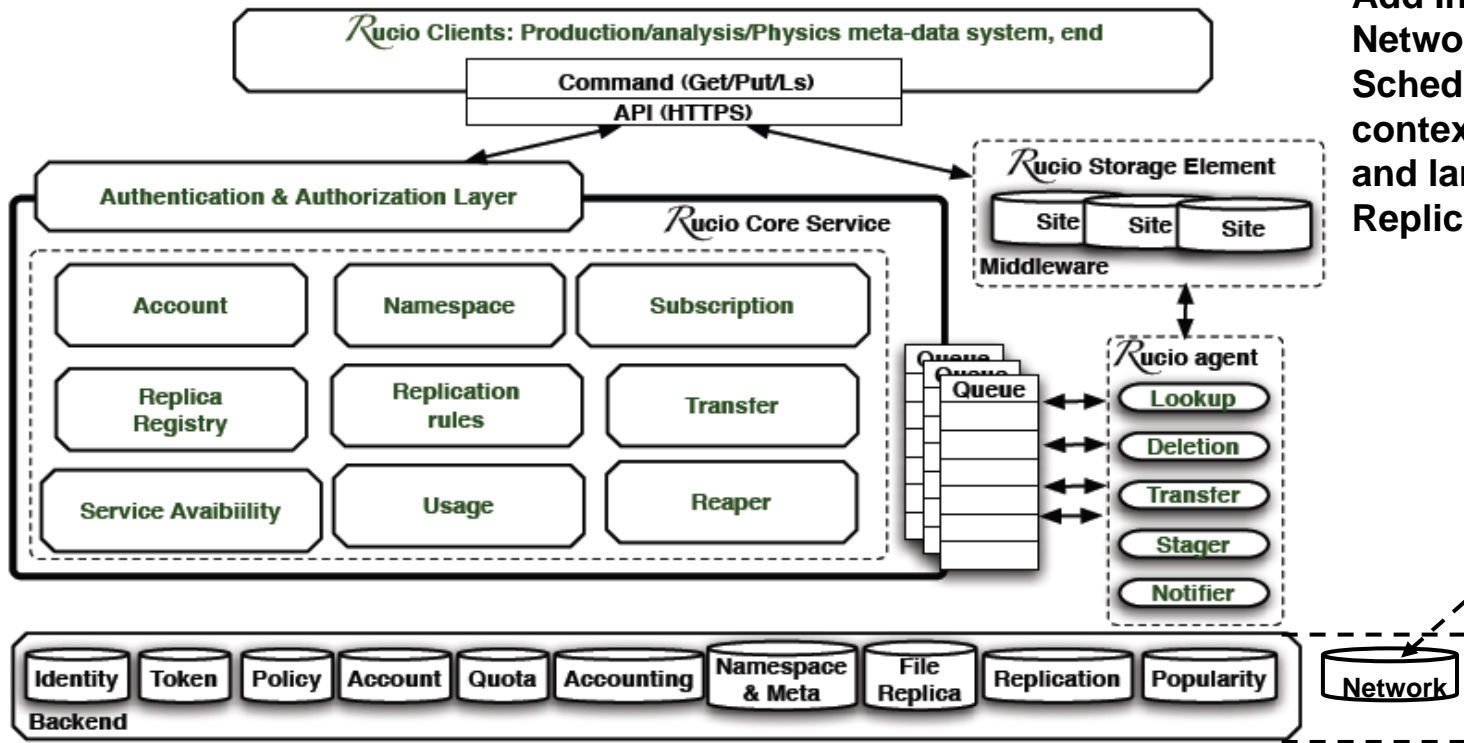
Open Networking Foundation: <https://www.opennetworking.org/index.php>

# Example: ATLAS Workload Management (works likewise w/ other WLM Systems)

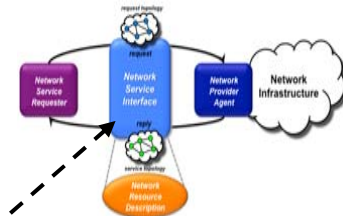


# Example: ATLAS Distributed Data Management

(works likewise w/ other DM Systems)



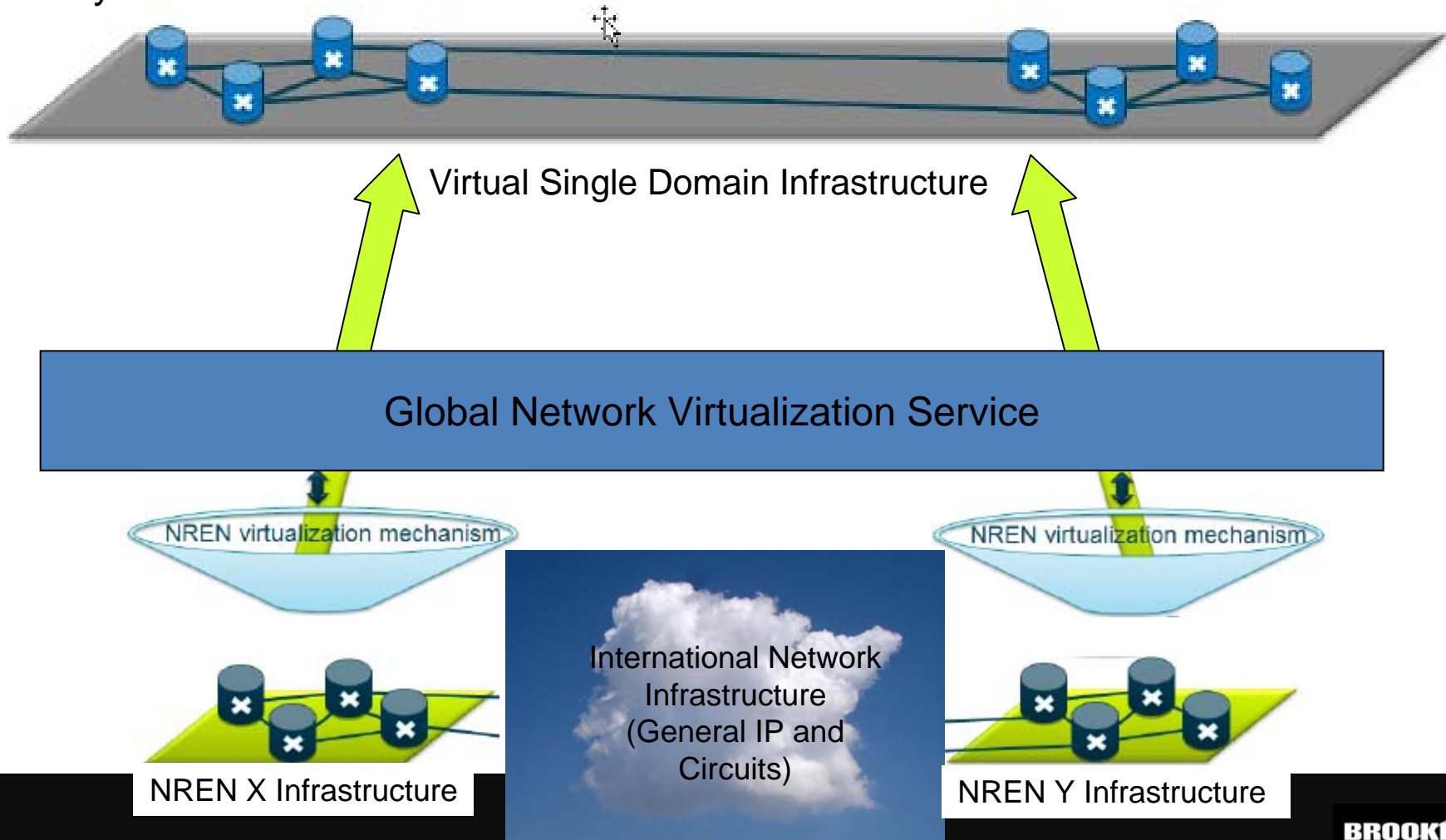
Add Information about the **Network** as a Resource for **Scheduling Decisions** in the context of **Data Placement** and large scale **Data Replication** activities





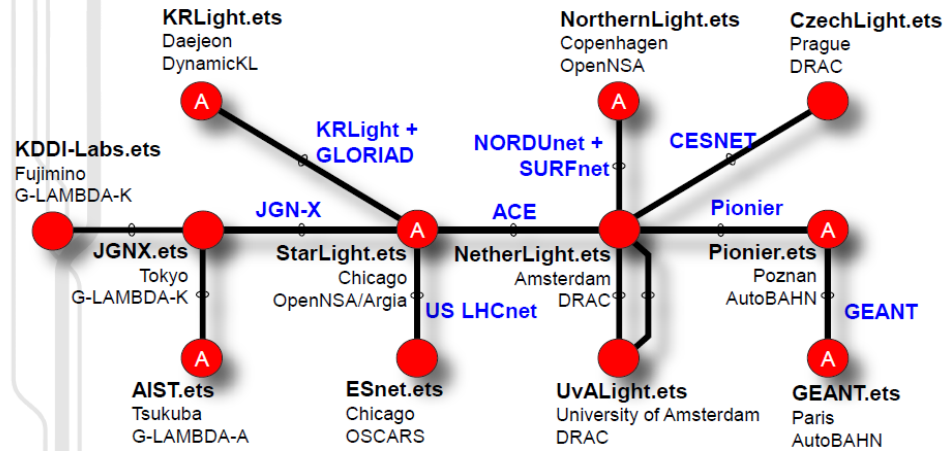
# Virtualized Network Infrastructure for Domain-Specific Applications (i.e. LHC Computing)

- Domain-specific networks as an overlay on top of shared infrastructure
  - Predictable Performance characteristics
  - Dynamic/on-demand creation and release



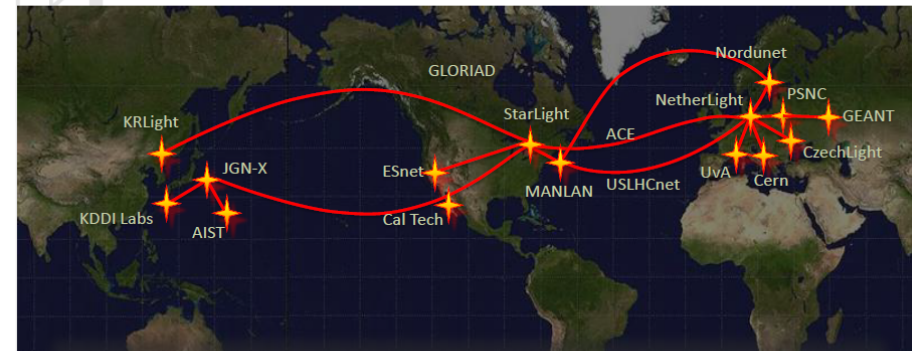
- Purpose of this presentation is to present the GLIF Automated GOLE facility to the LHCONE community
- And to propose the facility for testing and evaluation of the LHCONE “open exchange” architecture
- Outline:
  - GLIF Automated GOLE Pilot Project:
    - What is it? how does it work?
    - 2012 Goals for GOLEs
  - Network Service Interface:
    - What is it? how does it work? Why should LHCONE care?
  - AutomatedGOLE as a proving ground for LHCONE

**Demo NetworkSupercomputing 2011**



**NORDUnet The Automated GOLE Fabric**  
Nordic Infrastructure for Research & Education

**NORDUnet AutoGOLE Objectives for 2012...**  
Nordic Infrastructure for Research & Education



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, demonstrating the value of GOLEs to emerging network service models, and to develop a set of BCP for these services.

- Expand participation (pre-production service evaluation):
  - GLORIAD-US (SEA: eta CY12Q1)
  - WIX (WDC eta CY12Q2)
  - P-Wave/PNWG and RNP beginning testing now, in AG by ~H2 (?)
  - Would like to see:
    - ION, Esnet, USLHCnet, AMPATH,
- Build toward production BCPs:
  - NORDUnet plans a production NSI capability in CY2012...
  - SURFnet plans a production NSI capability in CY2012...
  - Anticipate others to do likewise as service models mature
  - Need to address \*service\* design and configuration
- Applications:
  - NEXPRES – EVLBI (testing from OSO to JIVE in CY12Q1)
  - CO-Universe – HD video
  - **LHCONE – HEP**
  - Others in works (under the radar) and still looking for others ...

# SDN – How far along is it ?

- Research in using OpenFlow for SDN has a great potential
  - Different directions for solutions pursued
  - How can the LHC Community participate in the development or should we better wait?
- Network abstractions is a key research topic
  - LHC applications need stable interface definitions
  - If interface doesn't exist today, how long will it take to create it ?
- A programming language for networking primitives is needed
  - Is there a place for LHC application developers to participate prior to the arrival of a (quasi) stable version?
    - Network Research community eager to have them review what exists and participate in the development
- I have raised above questions at the Oslo meeting and was told that researchers want the dialogue with the LHC application developers now

# A possible path forward

- We are close to a 2-year shutdown of the LHC machine
  - As we have developed a better understanding of our still evolving computing needs, and how we can best benefit from the middleware that supports flexible & nimble operations, we want to use the time to review the situation and participate in shaping the environment, including the network as a fully integrated facility component
  - We understand we have the providers' full attention
    - Define & implement Operational Interface between Users and Providers
      - E.g. the Working Group on LHCONE Operations starts work soon
      - Will be initially chaired by individuals from DANTE and Internet2
      - But WLCG needs something that covers OPN, LHCONE and the GPN
    - Besides providing production services most of the providers participate in pilot and research activities
      - with the motivation of providing new types of services to the user communities, and, if suitable, it's our role to contribute applications and the associated experts/expertise

# Status and Next Steps

- WLCG MB approved formation of the WG in July
  - Formation and formulation of agenda in progress since
- Find and appoint WG Members
  - Needs more time than anticipated because of the complex ecosystem
  - Experiments have appointed WG Members
    - ALICE/Iosif Legrand, ATLAS/Simone, CMS/Daniele & Lothar, LHCb/Peter Clarke
  - Solicit nomination for Site Representative(s)
    - In preparation, will ask the WLCG CB to select Tier-1 and Tier-2 site reps
    - Suggest rotation to foster diversity
  - Solicit nomination for Representatives from Network Service and Technology Provider community
    - Had discussions with some providers at the Oslo meeting
    - Difficult to understand their ecosystem – ask NREN coordination body for representatives?
    - Amenable to the concept of membership terms and rotation
- Items on the agenda
  - Collection and documentation of LHC experiment network requirements
  - Development and formulation of a network strategy
    - Including but not limited to future role and relation of LHCOPN, LHCONE, GPN
- Some activities in progress
  - Meeting on "Intelligent Network Services" at CERN on December 13 at CERN
    - Idea: bring together R&D activities carried out by providers with LHC application developers
  - Snowmass2013 has specific study group on Computing as one of the Frontiers
    - Chart the next ten years of U.S. High Energy Physics, viewed in a global context
    - I was invited to co-chair the network subgroup (prep meeting at FNAL Thu/Fri this week)

# Backup Material

# Proposed Composition of the WNWG

- Experts from the following areas
  - Applications (includes analysis models, workflows) and Distributed Data Management
    - 1-2 people per experiment
  - Facilities and Operations Coordination
    - Operations interface between WLCG entities and experiment services and network service providers
    - Provider representation could be delegate(s) from LHCOPN/LHCONE WG
    - 3-4 people representing T1/2/3 and the WLCG Ops Coordination Team
  - Network Services and Technologies
    - Up to 3 people representing Network Providers and emerging technologies
      - Provider representation could be delegate(s) from LHCOPN/LHCONE WG
      - Could be the same individuals as above
    - Representation from R&D community/organizations

# WNWG Charge (1/2)

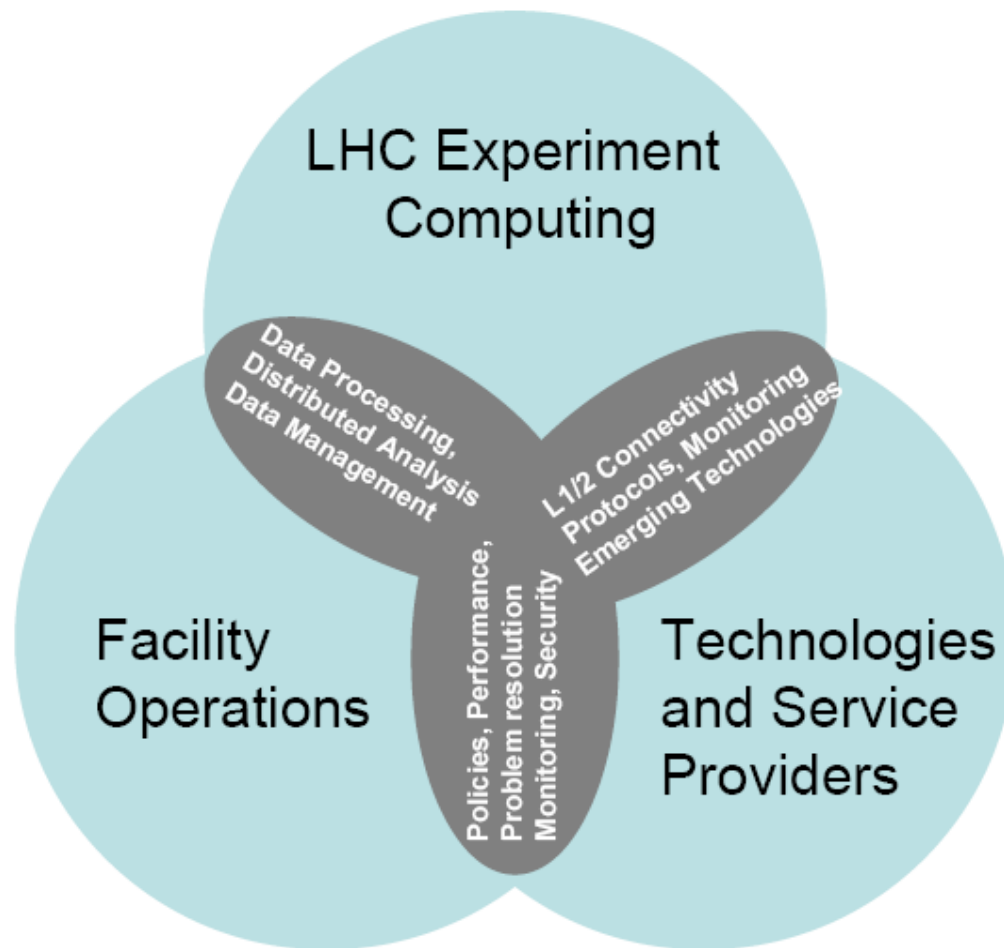
- Collect, document and update requirements.
- Propose and develop technical approaches and technical specifications to meet the functional, performance and operational requirements of the LHC Experiments
  - Common solutions as guiding principle
- Provide recommendations to WLCG management as to how the requirements w.r.t. functionality, commissioning and operations of network services could be met by network providers and sites
- Estimate amount of effort to implement them and ask WLCG Management to provide resources if they are adopted
- Establish a timeline of critical milestones and deliverables and track their implementation



# WNWG Charge (2/2)

- Communicate and collaborate with network service and technology providers and sites to provide:
  - Resolution to problems (daily operations issues are handled by the WLCG Operations Coordination Team (OCT))
    - » E.g. the deployment of perfSONAR monitoring at sites will be handled by the OCT
  - Quarterly reviews of the performance of network services relative to the original requirements and the actual evolving needs
    - » Actual Performance as measured to be compared against Metrics established by Experiments
  - Quarterly evaluations of the adequacy of the system and the performance of the providers
  - Recommendations for adjustments, improvements and modifications to be undertaken in the future

# WNWG Relations



# A possible path forward (2)

- We aim at making the network with its operational and development components part of the global resource planning and operations process that integrates worldwide distributed LHC computing facilities and the associated network requirements
  - The LHCOPN, LHCONe (as well as the GPN infrastructure we depend on) is part of the scope. In the larger context the network's role and implementation will be reviewed and integrated into the planning process.
- Instead of standing aside WLCG, with the LHC applications, will participate in research activities conducted by network providers, which should be ambitious in its approaches on the timeframe of the next n years. It should aim to bring an alternative to “traditional” large-scale IP infrastructures or demonstrate that it is not feasible, economic or advantageous to do so.
- WLCG will engage as an early adopter user community in a managed way with network service providers to demonstrate that new infrastructures can provide production services. This should bring new capabilities to us and not impact the current production services. This is a coordination task and should be run as a joint effort

# Motivation

- Rather than the Communities, which are viewing the currently evolving deployment of “LHCONE” production-grade, the Network Providers understand the current deployment of LHCONE as a first phase of a new network infrastructure to improve global T1  $\Leftrightarrow$  T2, T2  $\Leftrightarrow$  T2, T3  $\Leftrightarrow$  T2, ... communication
  - Current VRF-based approach is likely to evolve into something more ambitious
- Having access to network performance data (bandwidth, latency as obtained from perfSONAR measurements) allows the application to optimize usage of CPU & Storage provided by distributed facilities.
- Huge potential by using enhanced network features to optimize managed transfers (e.g. DDM/FTS for ATLAS) and task brokerage, and in storage federations with random/chaotic access across countries & continents
  - Application-managed data flows

# WNWG and R&D

- The WNWG shall actively be looking at results of R&D to identify, prioritize, and to promote initiatives that advance the usage and usability of network services as part of the WLCG computing infrastructure. Specific activities include
  - Opportunities & Risks
  - Adoption of new network technologies
  - Establish WGs for R&D
    - WGs will consult broadly and will work closely with experiments, network service & technology providers and sites

# Communication & Documentation

- Email/virtual meetings and Twiki
- Active participation in meetings like the joint LHCOPN/LHCONE workshops (every 6-9 months) and network-related workshops relevant to LHC applications
- Quarterly (or whenever necessary) reports to WLCG MB, GDB, and others if appropriate