LHCONE
(SOME) THOUGHTS ON 100G

Artur Barczyk/Caltech

Denver, November 2013

LHC Open Network Environment

- In brief, LHCONE was born to address two main issues:
  - **ensure** that the services to the science community maintain their **quality and reliability**
  - **protect** existing R&E infrastructures against potential "threats" of very large data flows
- LHCONE goals (from ~2011)
  - Provide some guarantees of performance
    - Large data flows across managed bandwidth that would provide better determinism than shared IP networks
    - Segregation from competing traffic flows
    - Manage capacity as # sites x Max flow/site x # Flows increases
  - Provide ways for better utilization of resources
    - Use all available resources
    - Provide Traffic Engineering and flow management capability
  - Leverage investments being made in advanced networking

LHCONE Overview

Current activities split in several areas:

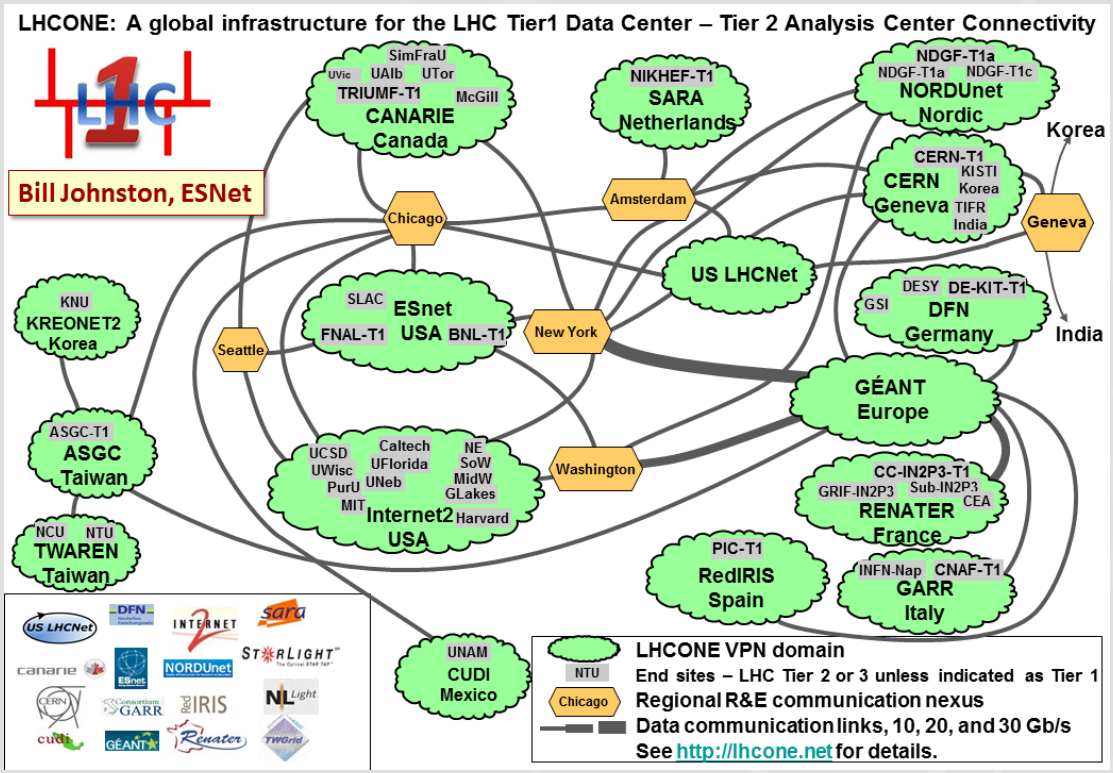
- Multipoint connectivity through **L3VPN**
  - Routed IP, virtualized service
- **Point-to-point dynamic circuits**
  - R&D, service prototype under development
- Common to both is logical separation of LHC traffic from the General Purpose Network (GPN)
  - Allows (in theory) traffic engineering
  - Allows trusted connections and firewall bypass
    - e.g. through Science DMZ-like site configurations
- For tasks which cannot be done with traditional methods
  - More **R&D in SDN/Openflow** for LHC traffic

LHCONE Today

CALTECH HEP NETWORKING | US LHCNet

Routed L3VPN Service, VRF

- Based on Virtual Routing and Forwarding (VRF)
- BGP peerings between the VRF domains
- Currently serving 44 LHC computing sites



- **In addition, the Point-to-point Multidomain Service Prototype**

Motivation for 100G

- Expect 10-fold data rates increase between now and 2021
- HEP experiments are moving towards more dynamic workflows and data management
  - reduce storage costs, consider less copies
  - dynamic caching, data movement on demand, remote data access
- This will increase network utilisation
- With deployment of 40GE server interfaces, capability to create single flow of more than 10GE
  - 10Gbps based infrastructure becomes a bottleneck
  - even with many 10G links
    - can alleviate to some extent by multi-path techniques
    - (not obvious in production, topic for R&D)

100G in production use

- For production use, 100+ Gbps based infrastructure must be just as resilient as we're used to from 10G
  - WLCG: "Network is our most reliable resource"
- This typically means multiple, redundant paths
  - cost factor (both circuits and equipment costs)
- Some challenges might remain, e.g. efficient use of multiple paths
- System(s) to manage the capacity are (still) needed
  - just because we have 100G, does not mean we can rely on overprovisioning
- …

Some points for investigation

- Not just capacity - how can we leverage the increase in capacity in a smart way?
  - Think SDN@100G here
  - Network-Application interface @100G
  - Recent techniques like e.g. RDMA-based file transfers
- From data processing point of view:
  - What impact will reduced transfer latency have on a CDN-style system like one for LHC?
  - on workflows and their management?
  - how dynamic can we make the data movement?
    - e.g. if we can rely on data replication on demand, can we think of having only one (global?) master copy?
- There's many other questions, undoubtedly, once we start thinking of integrating 100+ G networks into the system

100G for LHCONE

- 100G networking between the continents is important
  - LHC computing is massively distributed on a global scale

- Interest in combining 100G with SDN for
  - optimal resource utilization in the network
  - optimal resource utilization in the end-sites/systems

- Not to forget: Need a reasonable level of resiliency for production use

THANK YOU!

Artur.Barczyk@cern.ch