LHC Open Network Environment Proposal "of" the LHCT2S

10 February 2011

Eric Boyd, Internet2 Deputy CTO

1-2. Exec Summary - Background

- Bos-Fisk report on Tier 2 (T2) requirements for the Large Hadron Collider (LHC)
  - https://twiki.cern.ch/twiki/bin/view/LHCOPN/T2sConn
  - "Unstructured" T2 traffic will be the norm in the future and is already showing up in significant ways on the global R&E network infrastructures
  - T2s and T3s have widely varying capabilities
- 13 January 2011 LHCT2S Meeting at CERN
  - http://indico.cern.ch/conferenceDisplay.py?confId=116636
  - 4 proposals
  - All 4 served as a basis for the discussion

1. Exec Summary - Proposal

- LHC Open Network Environment (LHC ONE) is proposed
  - Builds on the idea of distributed exchange points
  - Intended to complement LHC OPN
  - Provide a collection of access locations that are effectively entry points into a network that is private to the LHC T1/2/3 sites
- Unstated, but implied …
  - Architectural description
    - Not an RFP from LHCT2S
    - Not a proposal to the LHCT2S
  - Intended to enable different approaches in different continents / regions
  - Not intended to be proscriptive; each continent / region will have to figure out how to implement access to LHC ONE
  - Not intended to be a complete description of LHC ONE and environment, just LHC ONE, at a high level

3. Data Intensive Science Environment

- It is recognized that the LHCONE concept may be of considerable value to other data-intensive science disciplines.
- Over time, it is possible that LHCONE may become a specific virtual instance on a general data-intensive science open network exchange.
- If this possibliity comes to pass, no policy would exclude use of the physical infrastructure by other science disciplines.
- However, the LHC community would be provided with a Service Level Agreement that guarantees at least the prescribed bandwidth available to the Tier1/2/3s.

4. LHC End User Environment

- Key observations from the Bos-Fisk report regarding the evolution of the end user environment:
  - Connectivity
  - Diversity
  - Flexibility
  - Monitoring
  - Trend of traffic becoming unstructured.

5. LHCONE Design Considerations

- 1. LHCONE complements the LHCOPN by addressing a different set of data flows
- 2. LHCONE enables high-volume data transport between T1/2/3s
- 3. LHCONE separates LHC-related large flows from the general purpose routed infrastructures of R&E networks
- 4. LHCONE incorporates all viable national, regional and intercontinental ways of interconnecting T1/2/3s
- 5. LHCONE uses an open and resilient architecture that works on a global scale
- 6. LHCONE provides a secure environment for T1/2/3 data transport
- 7. LHCONE provides connectivity directly to T1/2/3s and to various aggregation networks that provide connections to the T1/2/3s
- 8. LHCONE is designed for agility and expandability
- 9. LHCONE allows for coordinating and optimizing transoceanic data flows, ensuring the optimal use of transoceanic links using multiple providers by the LHC community

6. Definitions

- Tier 1s, Tier 2s and Tier 3s are collectively referred to as "**T1/2/3**."
- Any network that provides connections to T1/2/3s, and then in turn connects to LHCONE and any network that aggregates aforementioned networks is referred to as an "**aggregation networks**."
- The term "**connector**" refers to any entity that can connect to LHCONE: T1/2/3 and aggregation networks.
- The term "**exchange point**" refers to the hardware and physical facilities that provide the access points for LHCONE and the interconnect fabric of LHCONE. From the point of view of the organizations that provide the exchange points, those exchange points themselves may be a distributed exchange point.
- By "**distributed exchange point**" we understand a geographically distributed collection of network nodes under a single administrative authority, which to a connector appear as one single unit, administratively and operationally.

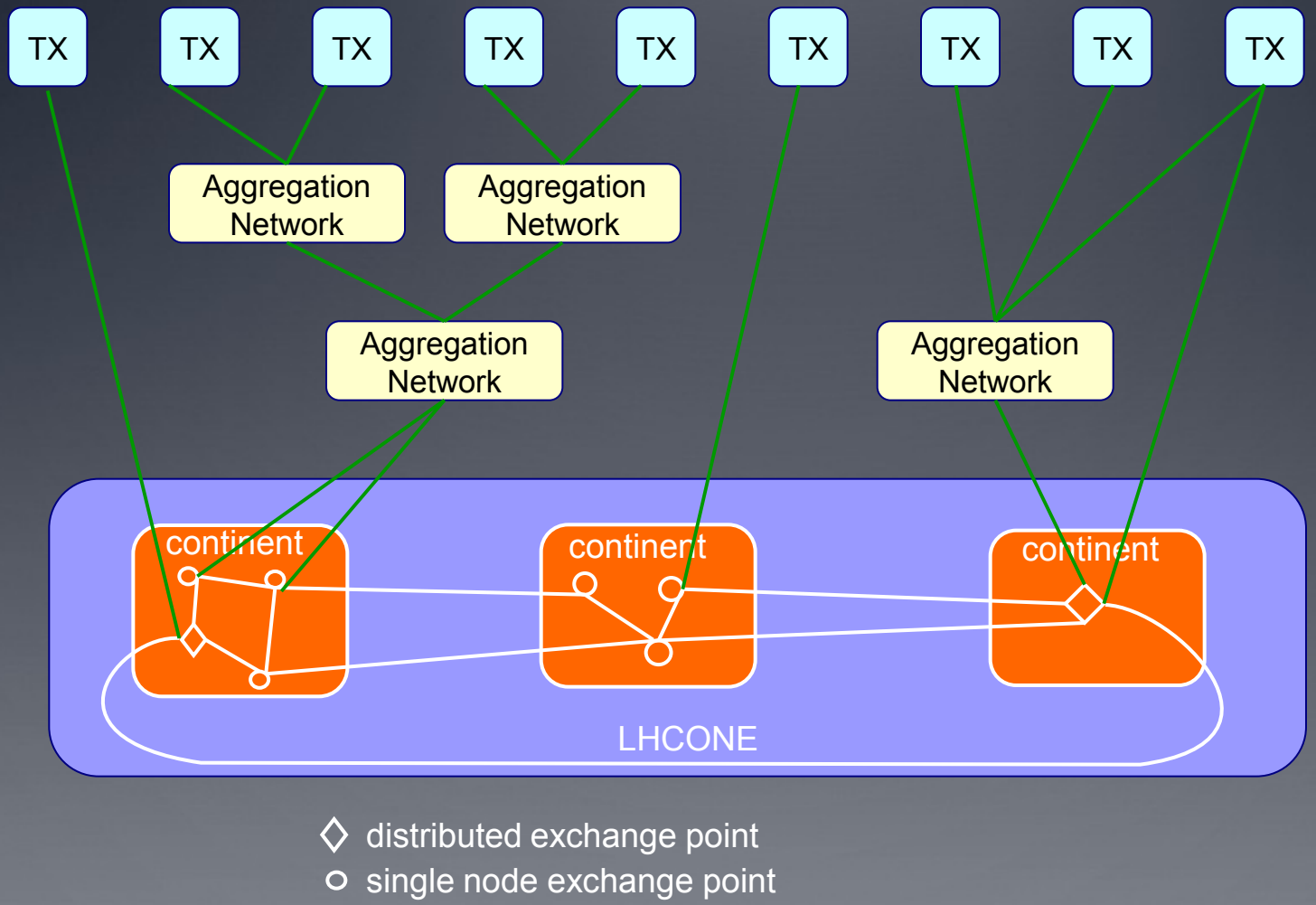
7. LHCONE Architecture

- LHCONE builds on the hybrid network infrastructures and open exchange points provided today by the major R&E networks on all continents 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 networks
- LHCONE architecture is based upon the following building blocks:
  - Single node exchange points
  - Continental / regional distributed exchange points
  - Interconnect circuits between exchange points
- Continental / regional exchange points are likely to be built as a distributed infrastructure with points of presence (access points) located around the region in ways that facilitate access by the LHC community.
- Continental exchange points are likely to be connected by allocated bandwidth on various (possibly shared) links to form LHCONE

5.2. LHCONE Access Methods

- Access method is up to the Tier1/2/3s, but may include:
  - dynamic circuits,
  - dynamic circuits with guaranteed bandwidth
  - fixed lightpath
  - connectivity at Layer 3.
- We envisage that many of the TierXs may connect to LHCONE through aggregation networks

7. LHCONE Example Implementation



8. LHCONE Services (Offered to T1/2/3s)

- Shared Layer 2 domains (private VLAN broadcast domains)
  - LHCONE provides IPv4 and IPv6 addresses on shared layer 2 domains that include all connectors.
  - LHCONE provides private shared layer 2 domains for groups of connectors (Aggregation networks or T1/2/3s) that only want to communicate among themselves.
  - Layer 3 routing is up to the connectors
    - Use of BGP, public IP addresses and public AS numbers is recommended, allowing every site to reach every site if all parties along the path so agree .
    - A Route Server per continent is planned to be available. In order to simplify configuration of the routers of the connector's members can decide to peer only with the Route Servers and get from them all the available prefixes.
- Point-to-point layer 2 connections
  - VLANS without bandwidth guarantees can be set up between pairs of connectors
- Lightpath / dynamic circuits with bandwidth guarantees
  - Lightpaths can be set up between pairs of connectors subject to a resource allocation policy agreed on by the community.
  - LHCONE provides a DICE IDC Version 1.1 protocol compatible circuit management system to facilitate lightpath / dynamic circuit setup, with the expectation this will eventually migrate to be compatible with the OGF NSI WG protocol when it emerges
- A perfSONAR archive provides LHCONE measurements and make them available, with the expectation this will eventually migrate to be compatible with the OGF NMC WG protocol when it emerges.
  - The presented statistics include current and historical bandwidth utilization values and link availability statistics for any past period of time.
  - LHCONE encourages each Tier1/2/3 and each aggregation network to install a perfSONAR node for both measurement and testing. LHCONE encourages publishing to the LHCONE perfSONAR archive.

This list of services is a starting point for LHCONE and not necessarily exclusive.

LHCONE does not preclude the continued use of the general R&E network infrastructure by the Tier1/2/3s, as is done today.

9. LHCONE Policy

- Important to have a consistent policy across the participants
- Expected that LHCONE policy will be defined and may evolve over time in accordance with the governance model
- Policy Recommended to LHCONE governance:
  - Any Tier1/2/3 can connect to LHCONE through one or more aggregation networks, and/or exchange points.
  - Between the regional exchange points that make up LHCONE, transit is provided to anyone in the Tier1/2/3 community that is part of the LHCONE environment such that they can freely interchange traffic among Tier1/2/3s connected to the LHCONE.
  - Exchange points must carry all LHC traffic offered to them (and only LHC traffic), and be built in carrier-neutral facilities so that any connector can connect with their own fiber or using circuits provided by any telecom provider.
  - Distributed exchange points must carry all LHC traffic offered to them (and only LHC traffic),, and be built in carrier-neutral facilities so that any connector can connect with their own fiber or using circuits provided by any telecom provider and the interconnecting circuits must carry all the traffic offered to them.
  - No additional restrictions can be imposed on LHCONE by the LHCONE component contributors.

9. LHCONE Policy Scope

- The scope of this policy framework is restricted to LHCONE.
- The policies for Tier1/2/3s to connect to aggregation networks are outside the scope of this document.
- The aggregator networks and/or the Tier1/2/3s might impose additional policy constraints on their own connections.
- Security is the responsibility of the aggregation networks and the Tier1/2/3s and is not the responsibility of LHCONE.

10. LHCONE Operations

- Existing modus operandi in the LHCOPN as well as work on federated operations happening at various locations around the world will be the initial guidance for organizing the operations for LHCONE

11. LHCONE Implementation Guidance

- Access Switches
  - Devices that provide the LHCONE Layer2 Ethernet connectivity with 1G and 10G Ethernet ports
  - 40G, 100G Ethernet ports are expected to be available in the future
  - Access switches are expected to be located at the Exchange Points
- Access Links
  - Ethernet-framed point-to-point links connecting a connector's device to one of the LHCONE Access Switches
  - Links are purchased and operated by the connectors and are not under the responsibility of LHCONE
  - Any connector may optionally connect to two (or even more) different Access Switches, for resiliency reasons

10. LHCONE Governance

- Similar to LHCOPN, LHCONE is a community effort; thus a similar governance model is proposed, where all the stakeholders meet regularly to review the operational status, propose new services and support models, tackle issues, design, agree and implement improvements.
- LHCONE governance defines the policies of LHCONE and requirements for participation. It does not govern the individual participants.
- LHCONE governance includes connectors, exchange point operators, CERN, and the experiments, in a form to be determined. In needs to be determined how T2s and T3s that do not connect directly to LHCONE have a voice in governance.
- LHCONE governance is responsible for defining how the costs are shared. Costs include, but are not limited to, port costs to connect to LHCONE, the operating and capital costs of LHCONE components, and the operating and capital costs of the links interconnecting the LHCONE components.
- LHCONE governance is also responsible for defining how the resources on LHCONE are allocated.

LHCONE Next Steps

- LHCONE V1.2: Small group -> LHCT2S
- LHCONE V2.0: LHCT2S -> LHC OPN
- Many details need to be fleshed out
- If this framework is agreed to, what is the right format to work out questions such as:
  - Who is refining the architecture?
  - Who is refining the services?
  - Who is creating the governance structure?
    - Proposed policy -> initial policy
  - Who is defining the operational processes?
  - Who is refining the implementation guidance?
  - Who is building/operating the various components (early / late stages)?