LHCONE

January 30, 2012

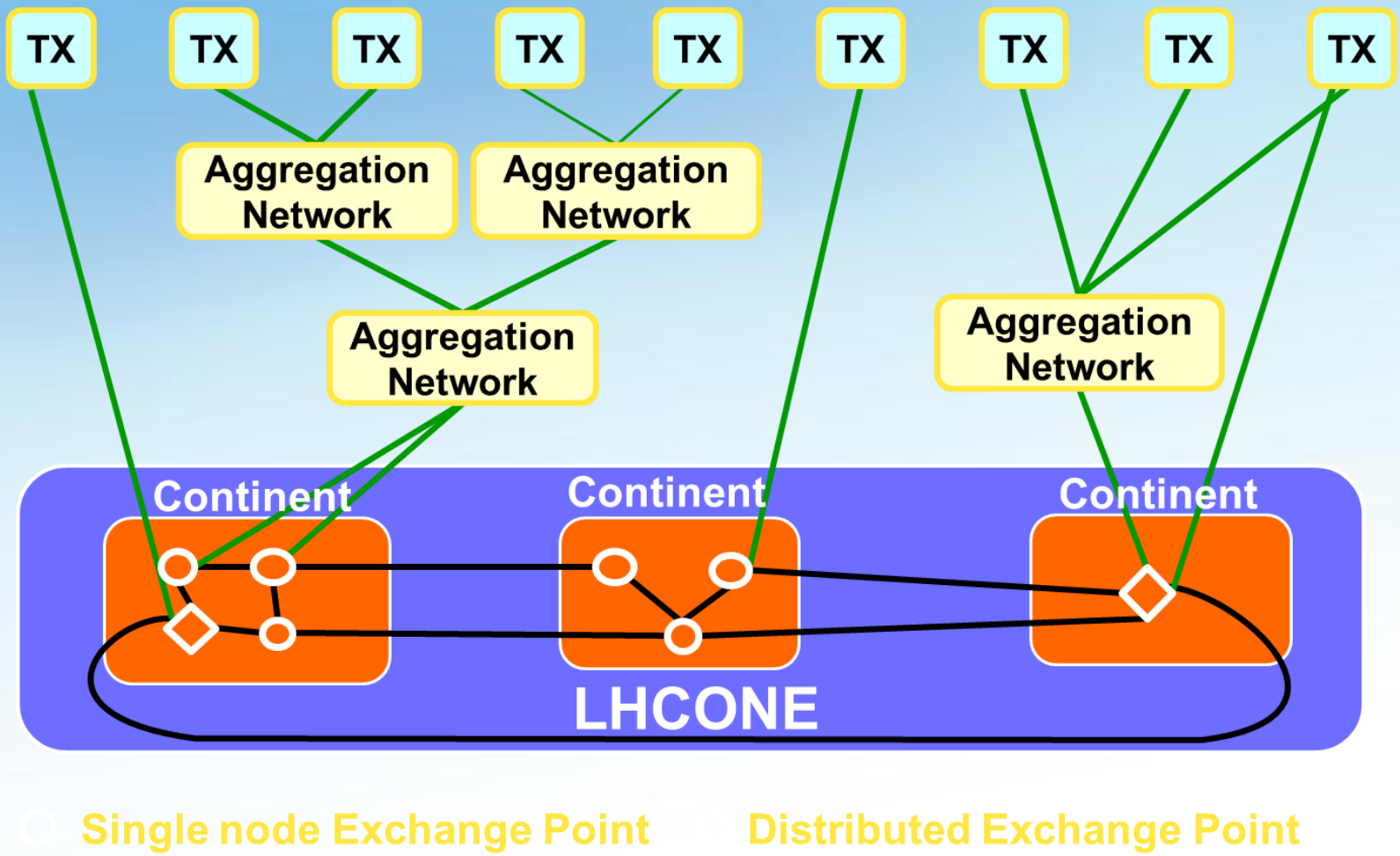
Eric Boyd, Internet2

CURRENT STATUS: VRF SOLUTION FOR LHCONE

LHCONE Goals

- Ease of Connection
  - Provide a collection of access locations that are effectively entry points into a network that is private to the LHC T1/2/3 sites.
  - As soon as a T1/2/3 site is connected to LHCONE, it ought to be able to easily exchange data with any other T1/2/3 site over an infrastructure that is sized to accommodate that traffic.
  - LHCONE must accommodate both IP connections and several variations of circuit-based connections.
  - T1/2/3 sites may connect directly or via their network provider (e.g. National Research and Education Network (NREN), US Regional Optical Network (RON), ESnet, etc.).
- Exploitation of Infrastructure
  - Build on the familiar idea of exchange points
  - Provide a mechanism to better utilize available transoceanic capacity.
- Ease of Operation
  - Provide an infrastructure with appropriate operations and monitoring systems to provide the high reliability (in the sense of low error rates) that is essential for the high bandwidth, high volume data transfers of the LHC community
  - Provide a test and monitor infrastructure that can assist in ensuring that the paths from the T1/2/3 sites to LHCONE are also debugged and maintained in the low error rate state needed for LHC traffic

LHCONE High-level Architecture

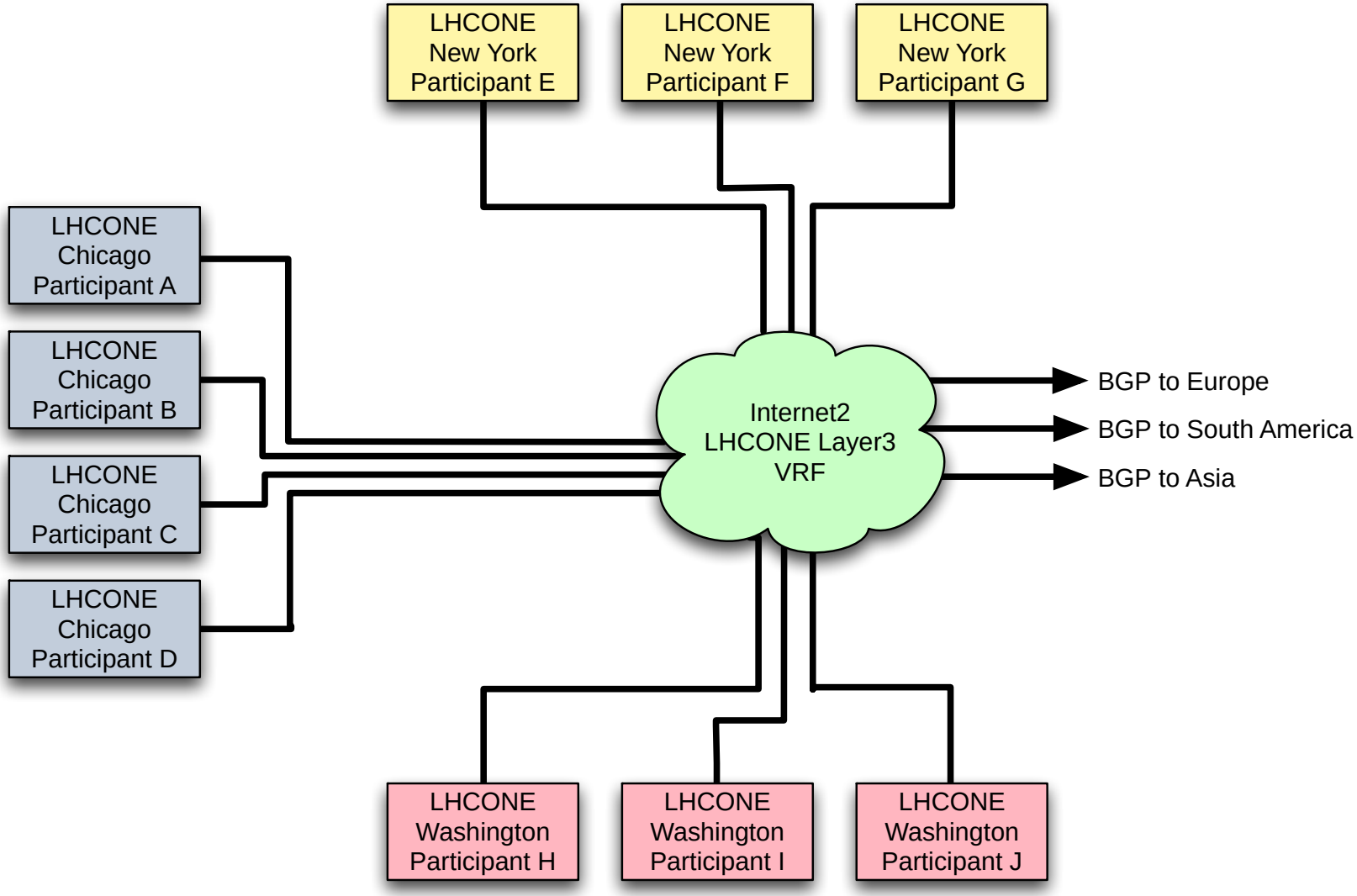


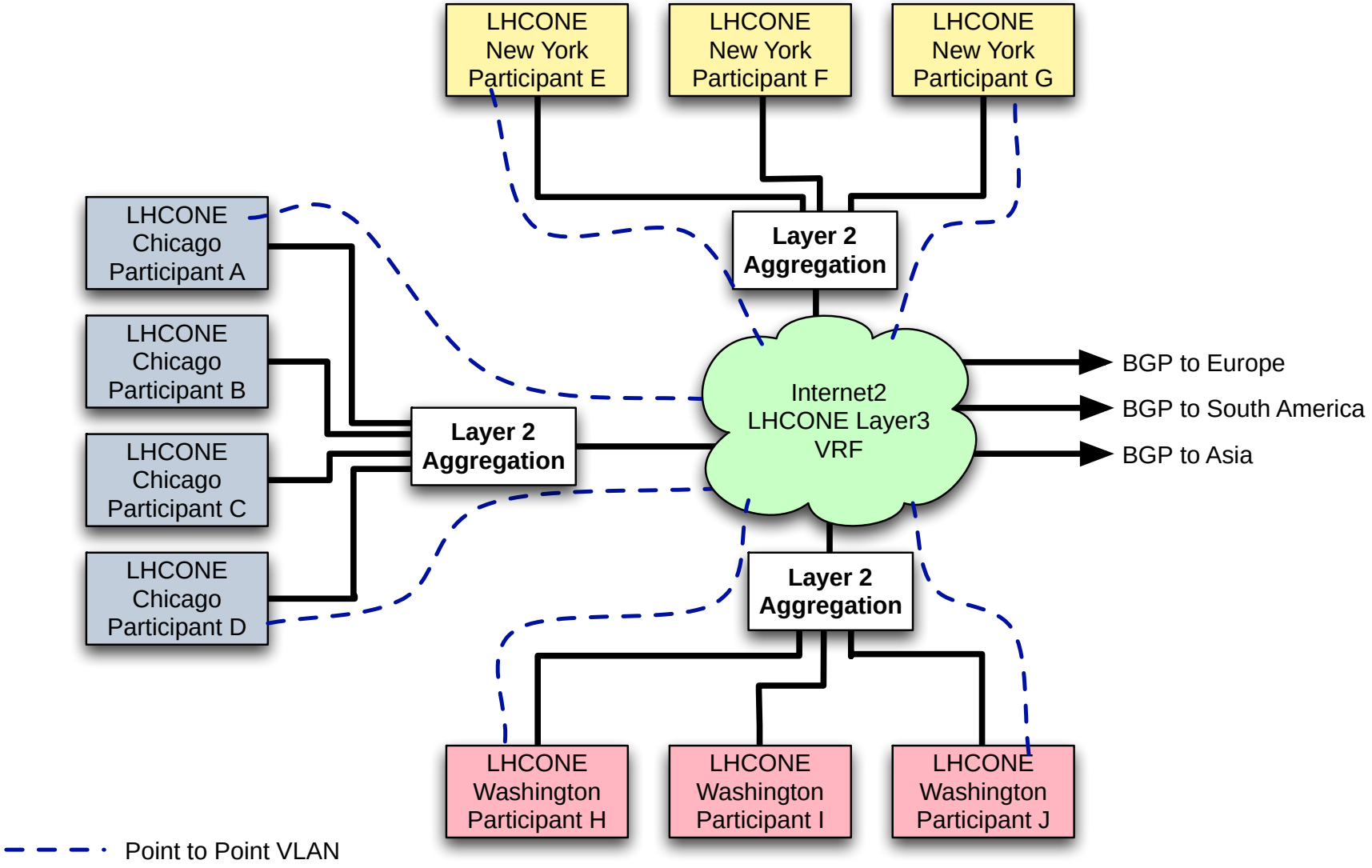
LHCONE Design Considerations

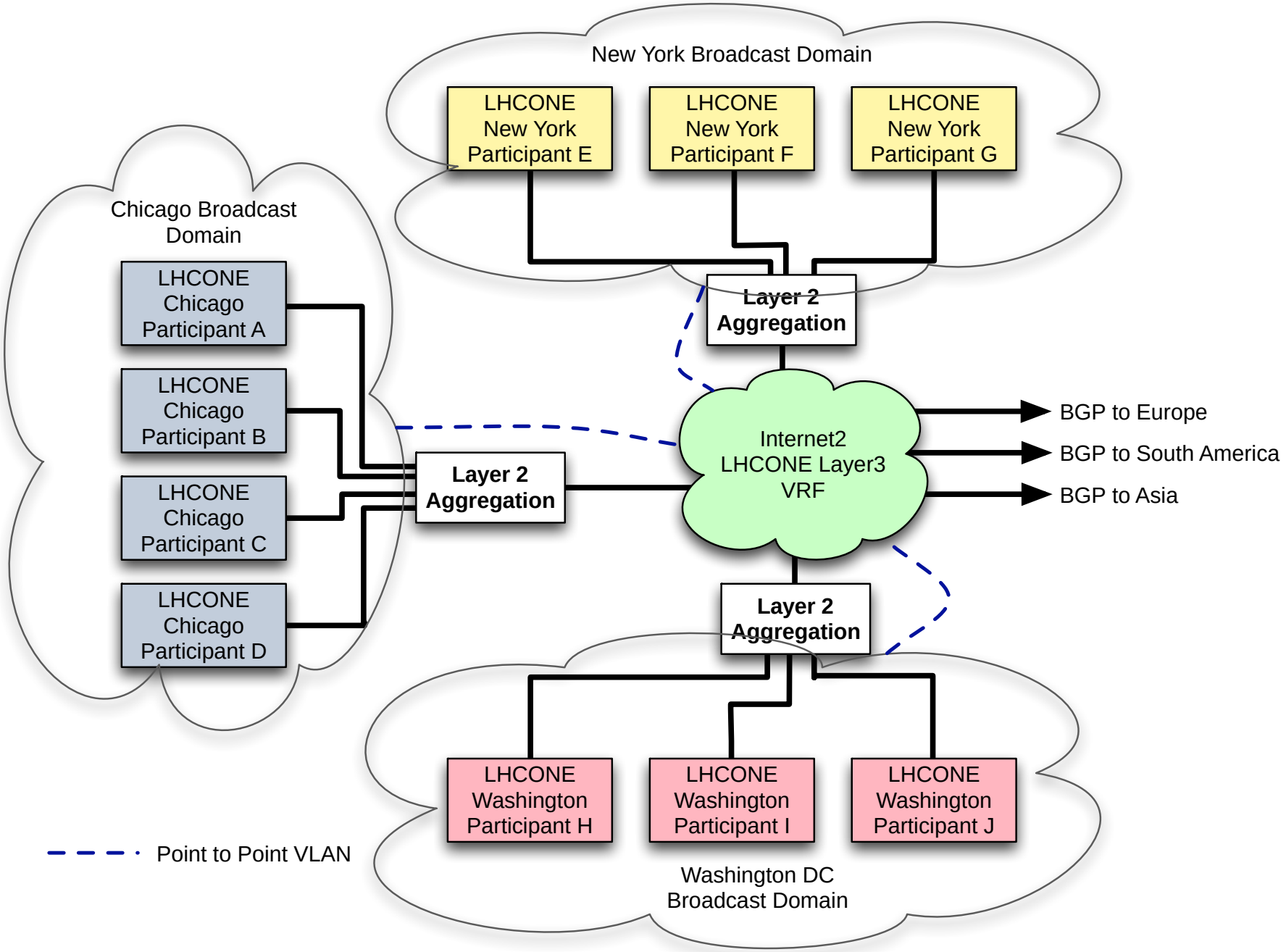
- LHCONE complements the LHCOPN by addressing a different set of data flows.
- LHCONE enables high-volume data transport between T1s, T2s, and T3s.
- LHCONE separates LHC-related large flows from the general purpose routed infrastructures of R&E networks.
- LHCONE incorporates all viable national, regional and intercontinental ways of interconnecting Tier1s, Tier2s, and Tier 3s.
- LHCONE uses an open and resilient architecture that works on a global scale.
- LHCONE provides a secure environment for T1-T2, T2-T2, and T2-T3 data transport.
- LHCONE provides connectivity directly to T1s, T2s, and T3s, and to various aggregation networks, such as the European NRENs, GÉANT, and North American RONs, Internet2, ESnet, CANARIE, etc., that may provide the direct connections to the T1s, T2s, and T3s.
- LHCONE is designed for agility and expandability.
- LHCONE allows for coordinating and optimizing transoceanic data flows, ensuring the optimal use of transoceanic links using multiple providers by the LHC community.

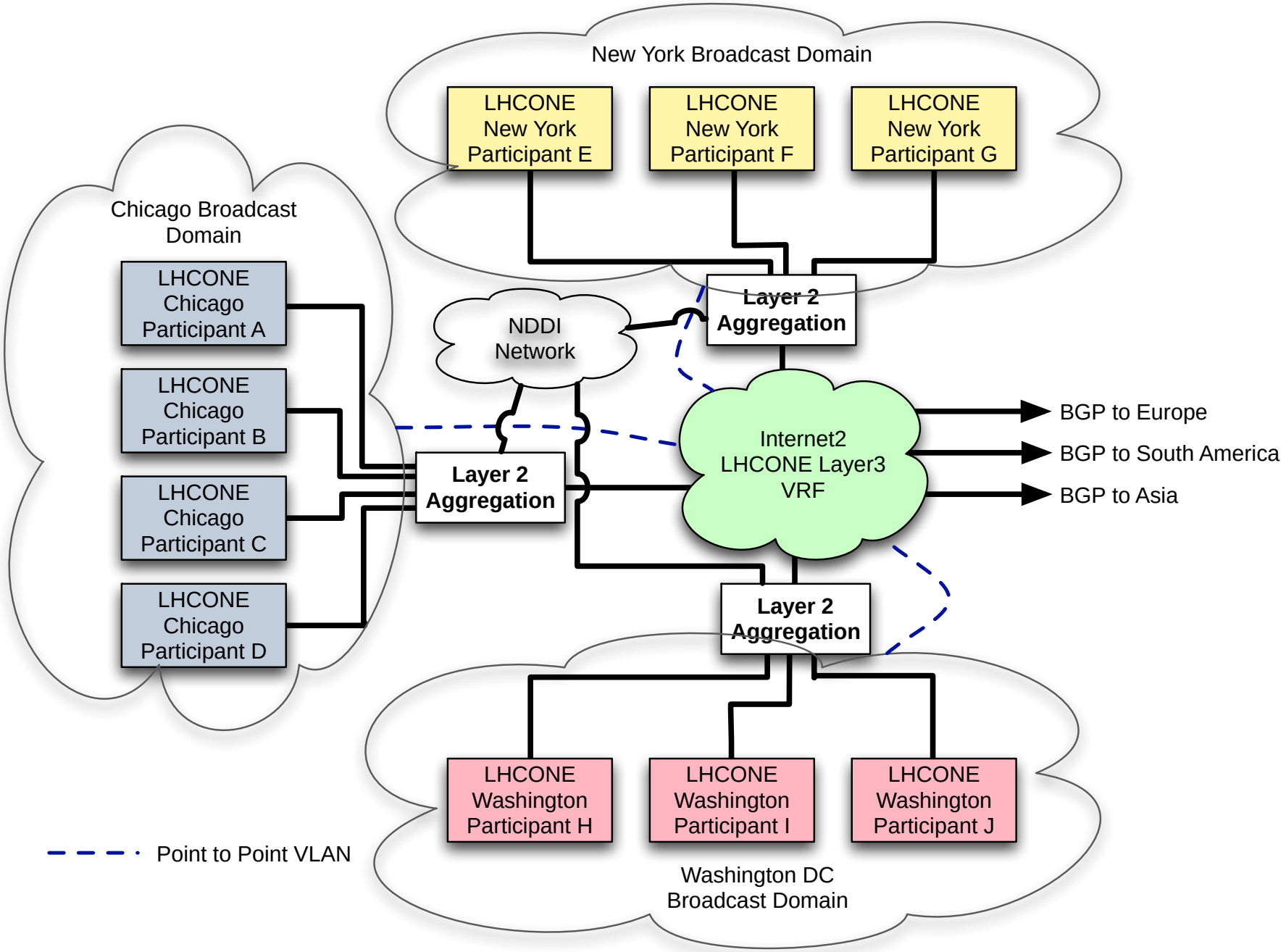
LHCONE Services

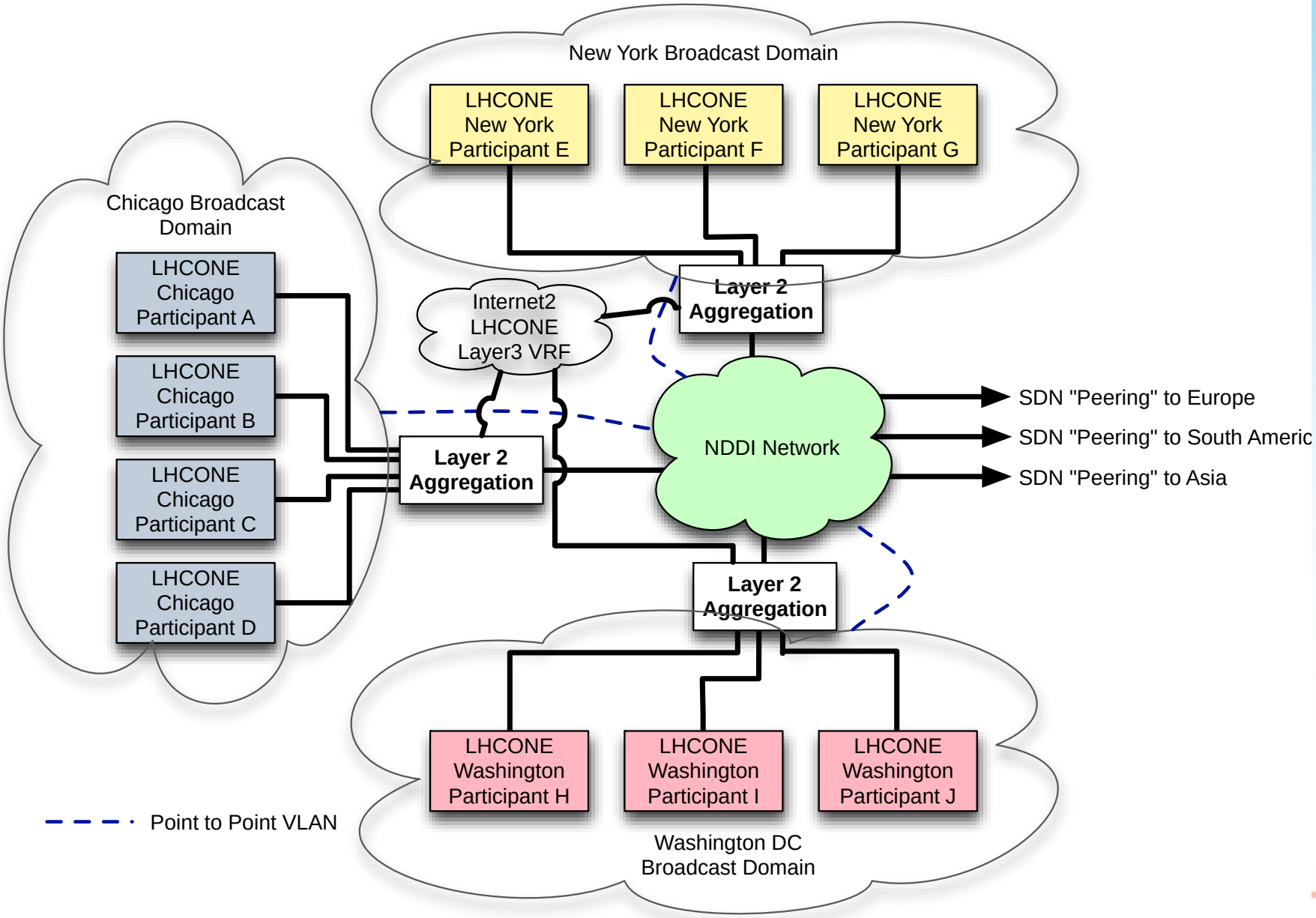
- Multipoint Service (TODAY)
- Point-to-Point Service (TOMORROW)
  - With bandwidth guarantees
  - Without bandwidth guarantees
- Diagnostic Service (TOMORROW)


LHCONE New York Participant E
LHCONE New York Participant F
LHCONE New York Participant G
LHCONE Chicago Participant A
LHCONE Chicago Participant B
LHCONE Chicago Participant C
LHCONE Chicago Participant D
Internet2 LHCONE Layer3 VRF
BGP to Europe
BGP to South America
BGP to Asia
LHCONE Washington Participant H
LHCONE Washington Participant I
LHCONE Washington Participant J


LHCONE New York Participant E
LHCONE New York Participant F
LHCONE New York Participant G
LHCONE Chicago Participant A
LHCONE Chicago Participant B
LHCONE Chicago Participant C
LHCONE Chicago Participant D
Layer 2 Aggregation
Layer 2 Aggregation
Layer 2 Aggregation
Internet2 LHCONE Layer3 VRF
BGP to Europe
BGP to South America
BGP to Asia
LHCONE Washington Participant H
LHCONE Washington Participant I
LHCONE Washington Participant J
Point to Point VLAN


New York Broadcast Domain
LHCONE New York Participant E
LHCONE New York Participant F
LHCONE New York Participant G
Chicago Broadcast Domain
LHCONE Chicago Participant A
LHCONE Chicago Participant B
LHCONE Chicago Participant C
LHCONE Chicago Participant D
Layer 2 Aggregation
Layer 2 Aggregation
Internet2 LHCONE Layer3 VRF
BGP to Europe
BGP to South America
BGP to Asia
Layer 2 Aggregation
LHCONE Washington Participant H
LHCONE Washington Participant I
LHCONE Washington Participant J
Washington DC Broadcast Domain
Point to Point VLAN


New York Broadcast Domain
LHCONE New York Participant E
LHCONE New York Participant F
LHCONE New York Participant G
Chicago Broadcast Domain
LHCONE Chicago Participant A
LHCONE Chicago Participant B
LHCONE Chicago Participant C
LHCONE Chicago Participant D
NDDI Network
Layer 2 Aggregation
Layer 2 Aggregation
Internet2 LHCONE Layer3 VRF
BGP to Europe
BGP to South America
BGP to Asia
Layer 2 Aggregation
LHCONE Washington Participant H
LHCONE Washington Participant I
LHCONE Washington Participant J
Washington DC Broadcast Domain
Point to Point VLAN

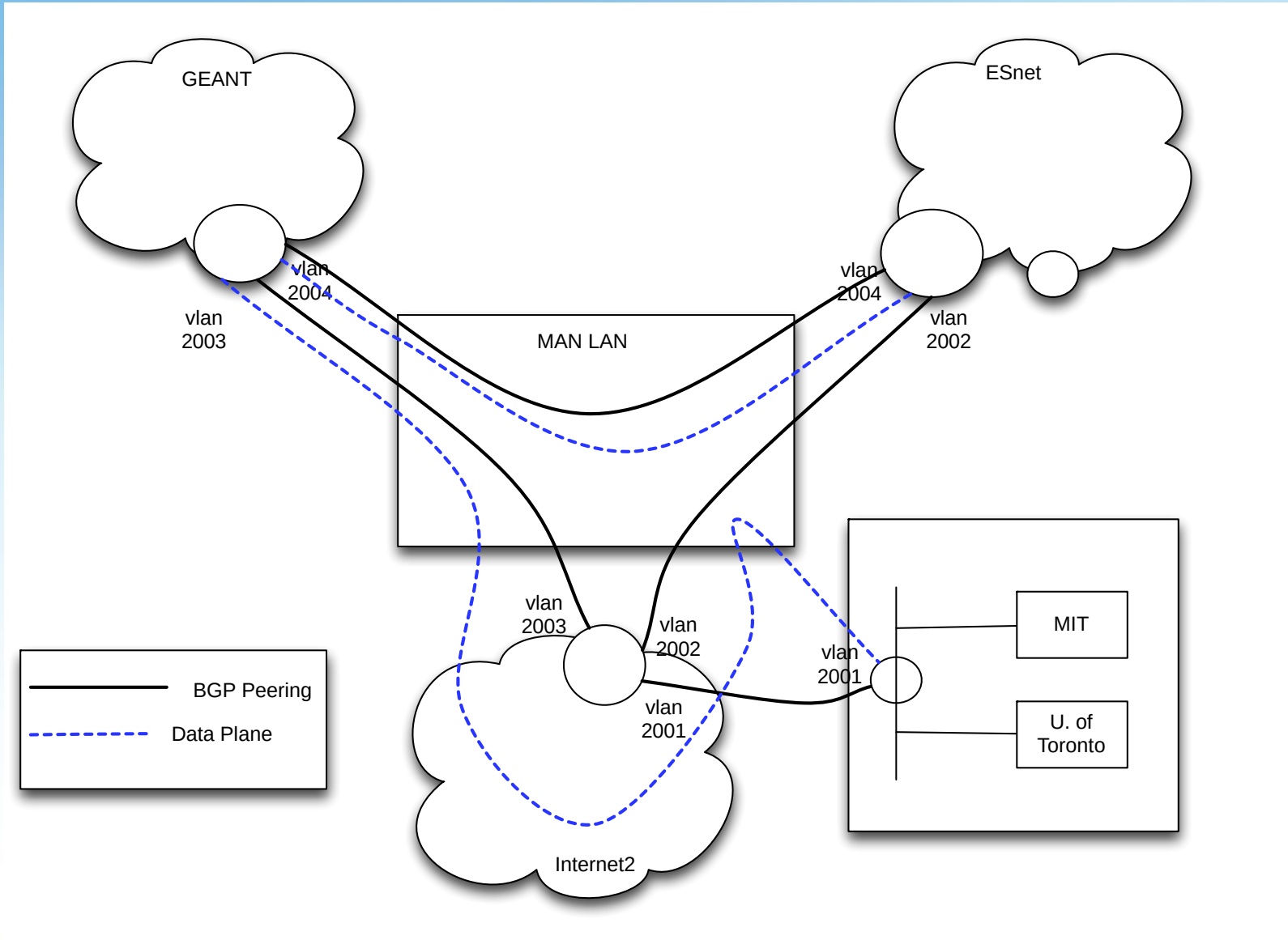

New York Broadcast Domain
LHCONE New York Participant E
LHCONE New York Participant F
LHCONE New York Participant G
Chicago Broadcast Domain
LHCONE Chicago Participant A
LHCONE Chicago Participant B
LHCONE Chicago Participant C
LHCONE Chicago Participant D
Internet2 LHCONE Layer3 VRF
Layer 2 Aggregation
Layer 2 Aggregation
NDDI Network
SDN "Peering" to Europe
SDN "Peering" to South Americ
SDN "Peering" to Asia
Layer 2 Aggregation
LHCONE Washington Participant H
LHCONE Washington Participant I
LHCONE Washington Participant J
Washington DC Broadcast Domain
Point to Point VLAN

Current Status

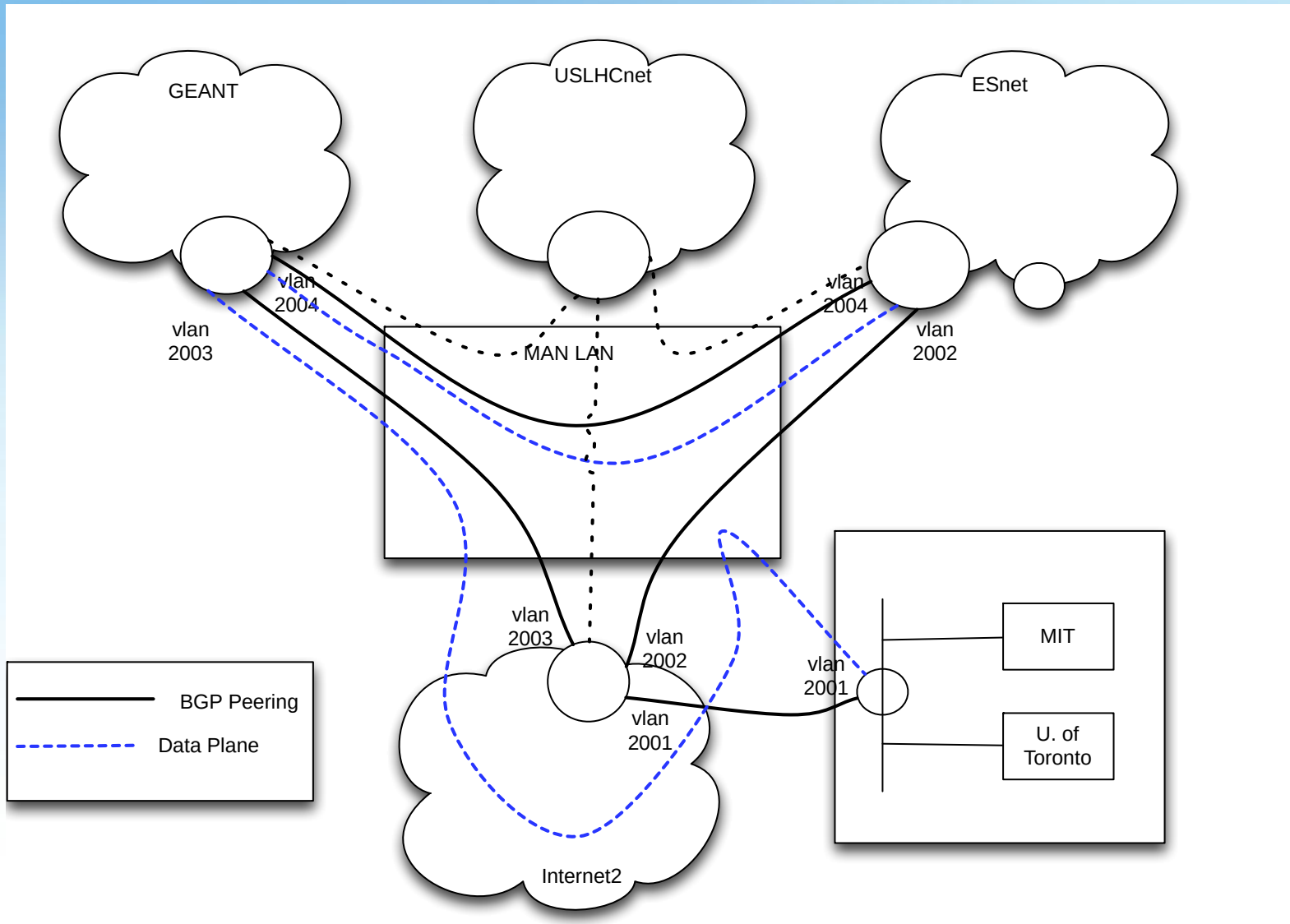
- The VRFs for Internet2, GEANT and ESnet are in place.
  - Internet2's VRF extends over its entire core.
  - GEANT's VRF extends over its entire core.
  - ESnet's VRF is partially implemented.
- The VLANs to establish peerings across MAN LAN have been established and sent out.
  - 2002 Internet2-GEANT
  - 2003 Internet2-ESnet
  - 2004 GEANT-ESnet
- Others can be set up as needed.
  - The next diagrams show the current status and adding USLHCnet.

INTERNET 2

VRF Solution: Current Status



If USLHCnet adds a VRF …



Current Status

- VLANs have not yet been assigned across WIX.
  - That will happen immediately after this discussion.
- As of Friday the 27th:
  - Peering mesh between GEANT, ESnet and Internet2 have all been set up.
  - There are probably not a lot of routes yet.
    - Those will grow as sites start to participate in the VRF's

INTERNET 2

Current Status - Policy

- We need to discuss what sort of policies will be applied:
  - There are BGP policies for traffic balancing.
  - There are 4 Trans-Atlantic paths
    - ACE/GEANT in NY and DC
    - USLHCnet
    - European circuit in Chicago
  - These are not all equal bandwidth.
    - There will need to be some traffic engineering between those paths.

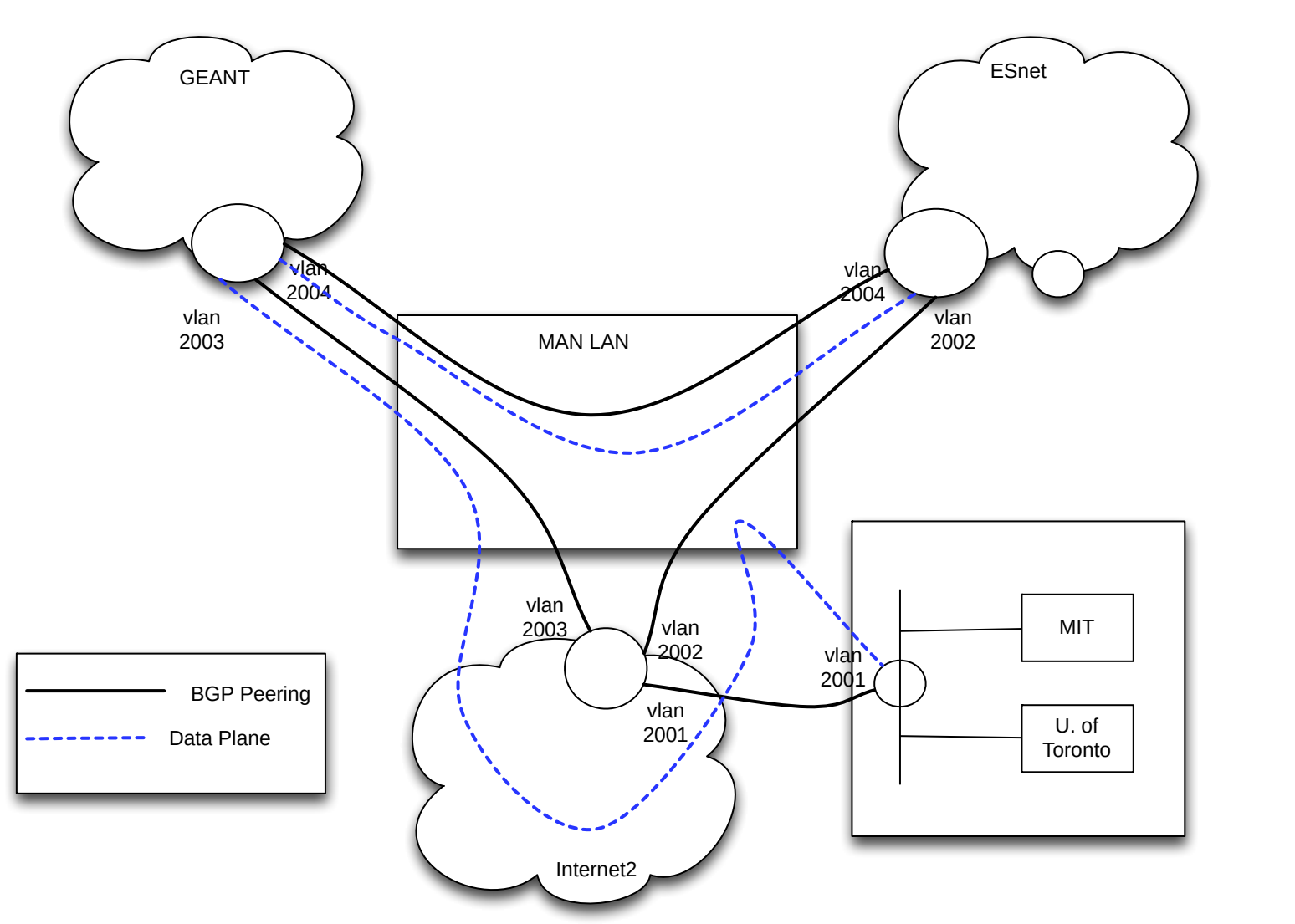
INTERNET 2

Current Status - Issue

- Would like the traffic passing through an exchange point to take the most direct path at the same time we minimize the number of peers with the VRFs.
- Looking at this diagram again, we see that the data path is not optimal.

INTERNET 2

Sub-optimal Data Plane

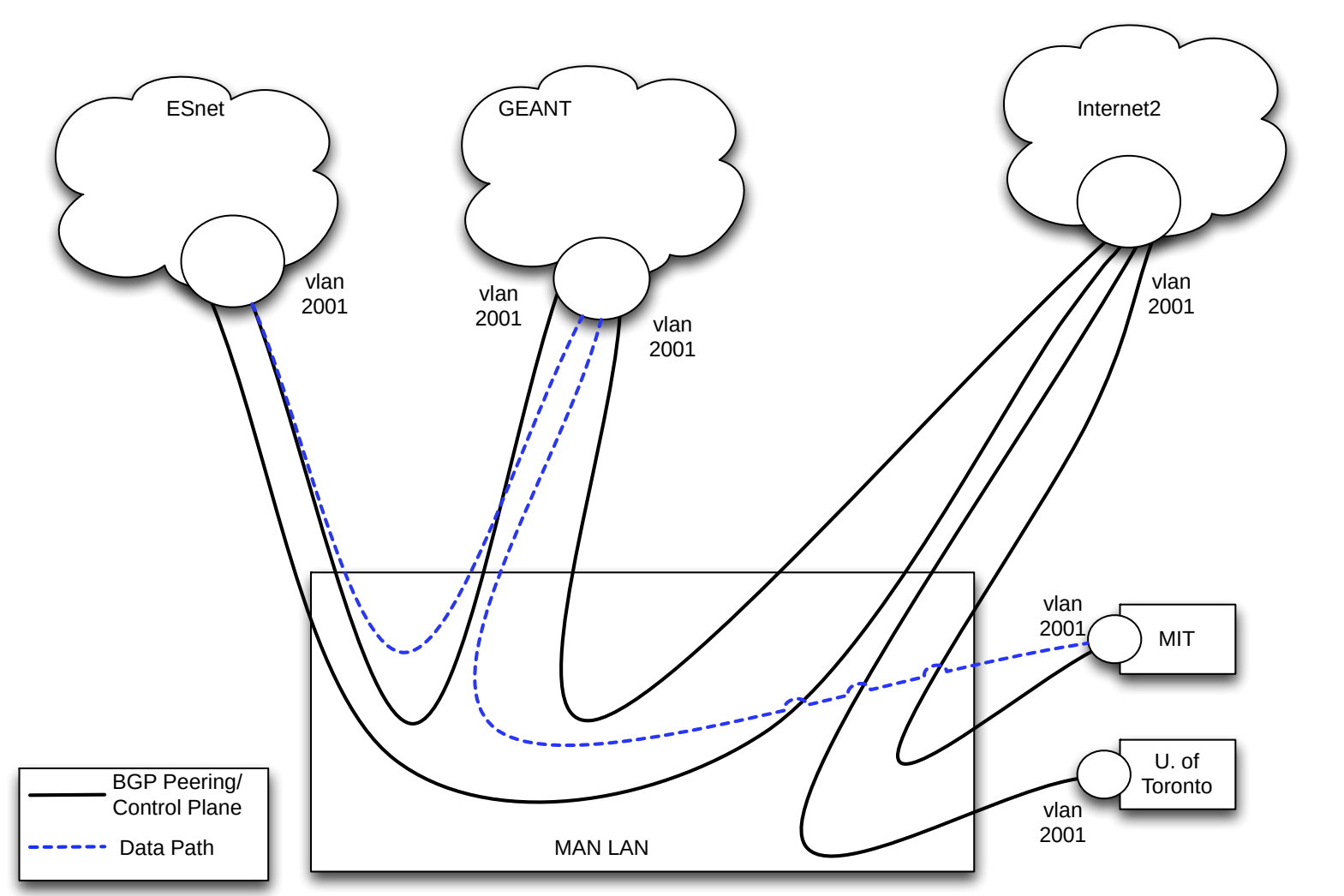


Solution

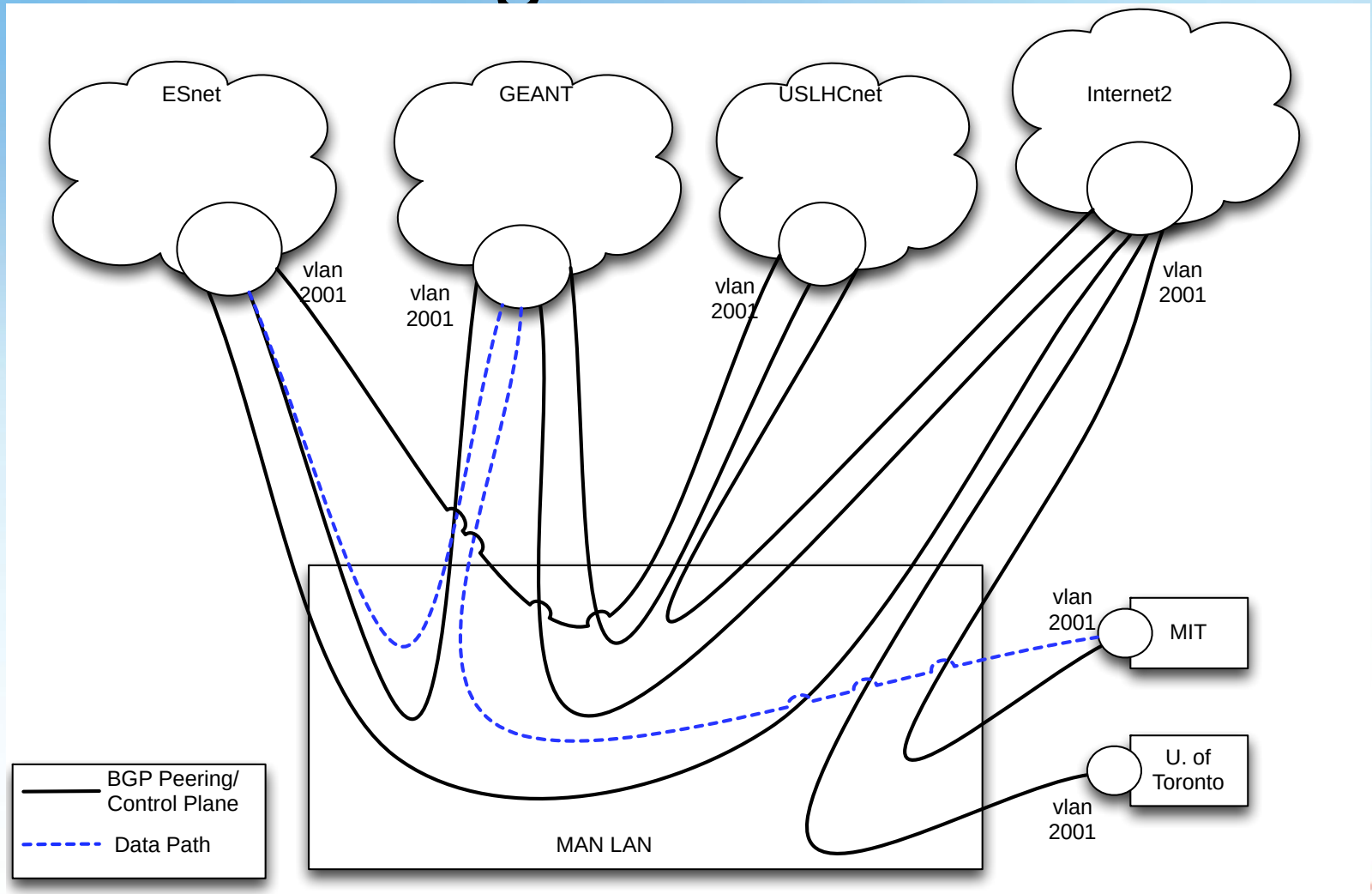
- To address this problem we are proposing separating the control plane (BGP) path from the data plane path.
- By setting up a shared VLAN within the exchange point, 3rd party next hop routing is enabled.
  - This will allow a VRF to better manage the traffic of direct connectors to an exchange point.
- It will also require an address block for each exchange that can be used for the LHCONE participants.
  - Can this be provided by CERN?

INTERNET 2

3rd Party Next Hop Routing



Adding USLHCnet



Next Step

- The shared VLAN in MAN LAN and WIX could be set up at any time.
- Seeking consensus from the LHCONE community prior to setting this up.

Discussion

- What did we get right?
- What did we get wrong?
- What else needs to be included?

INTERNET 2