ESnet
ENERGY SCIENCES NETWORK

LHCONE Implementation in America and Example Site Configurations

William Johnston
Senior Scientist and Advisor
ESnet
Lawrence Berkeley National Laboratory

Asia Tier Center Forum
KISTI, Daejeon, South Korea
September 22-24, 2015

U.S. DEPARTMENT OF ENERGY
Office of Science

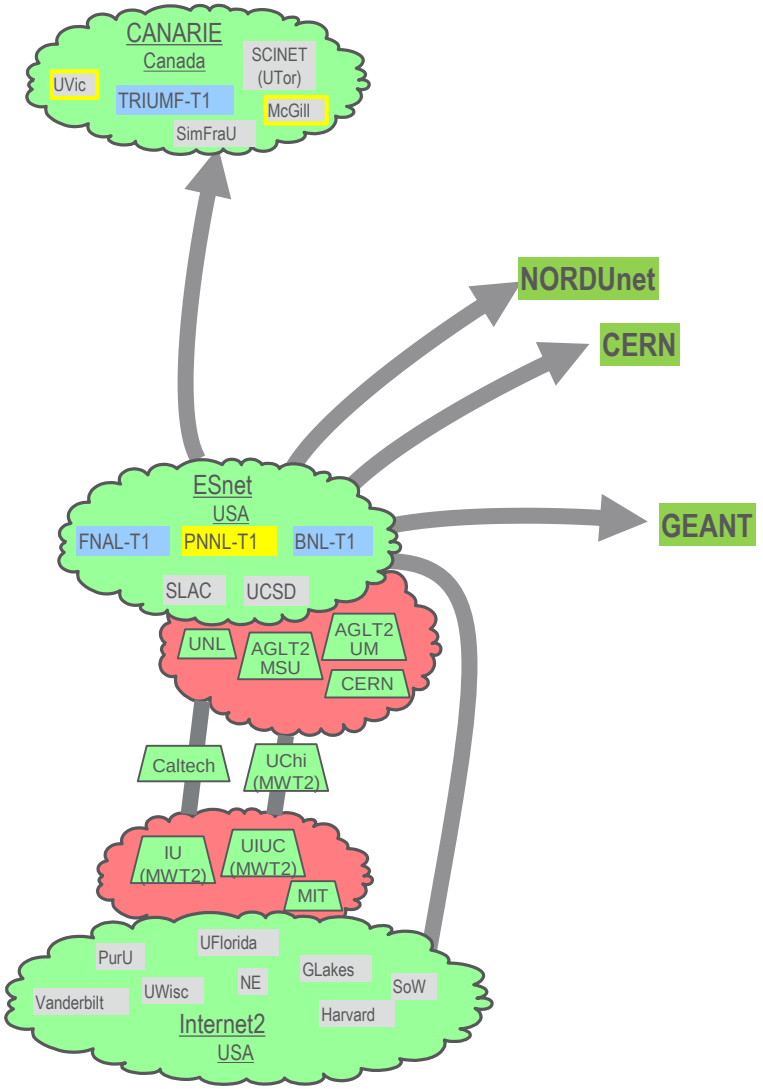
BERKELEY LAB

LHCONE in America

- Networking in the U.S. is structured somewhat differently than it is in Europe.
- The U.S. has many independent networks that have informal arrangements with each other to provide a coherent R&E national network
- The potential for fragmentation was recognized in the mid-1990s and NSF funded the construction of an initial set of “GigaPoPs” (gigabit point-of-presence) that were open exchange points where networks could interconnect in a policy-free environment
- Today the GigaPoPs or “exchange points” are an integral part of the U.S. R&E Internet
  - StarLight, MAN LAN, WIX, SOX, PNG, and many others
- The GigaPoPs play a key role in the U.S. LHCONE environment by providing the inter-VRF connections

LHCONE in America

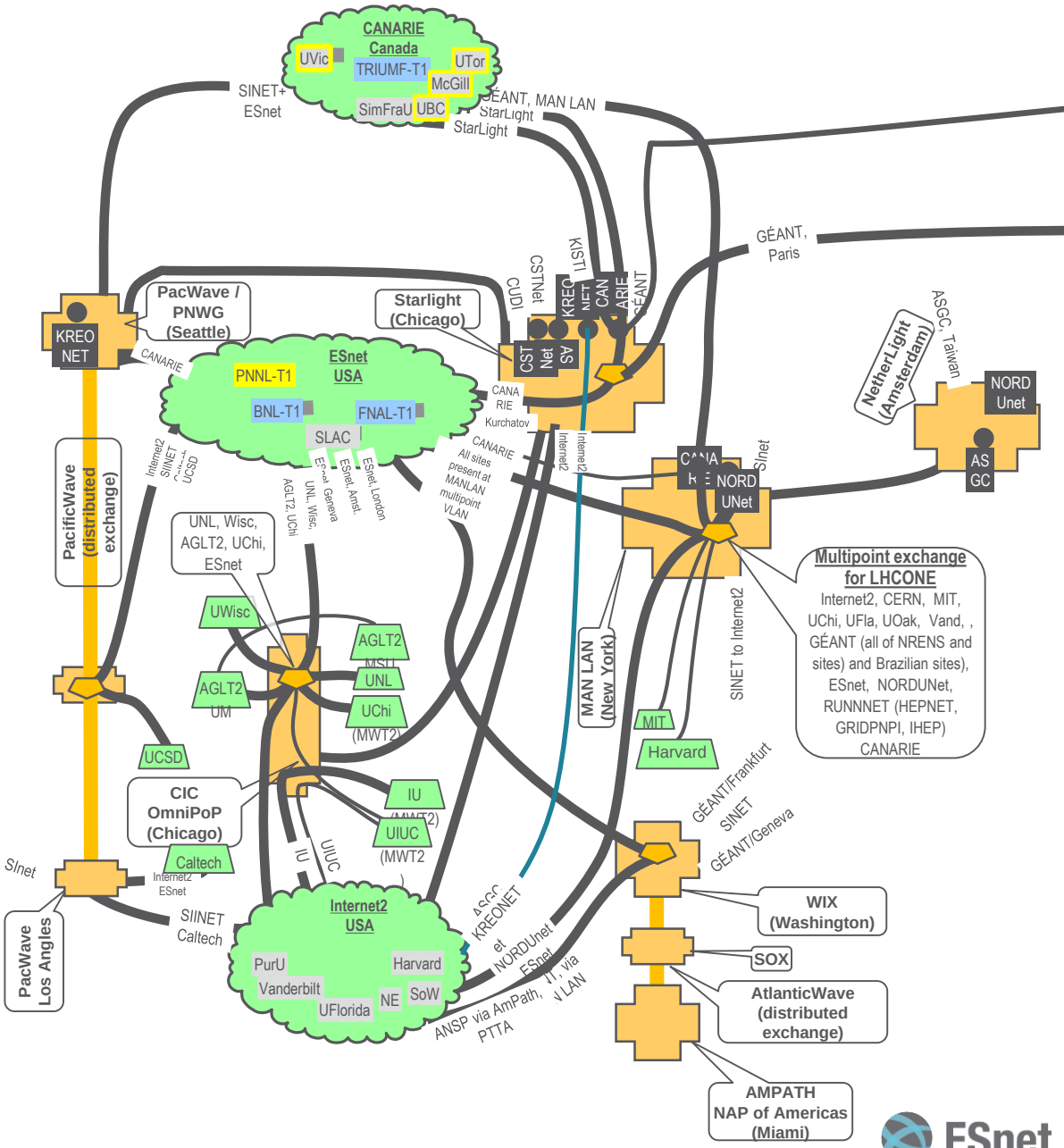
- Although LHCONE is frequently represented just by the connections between NSP and site VRFs, reality is quite different



ESnet

LHCONE in America

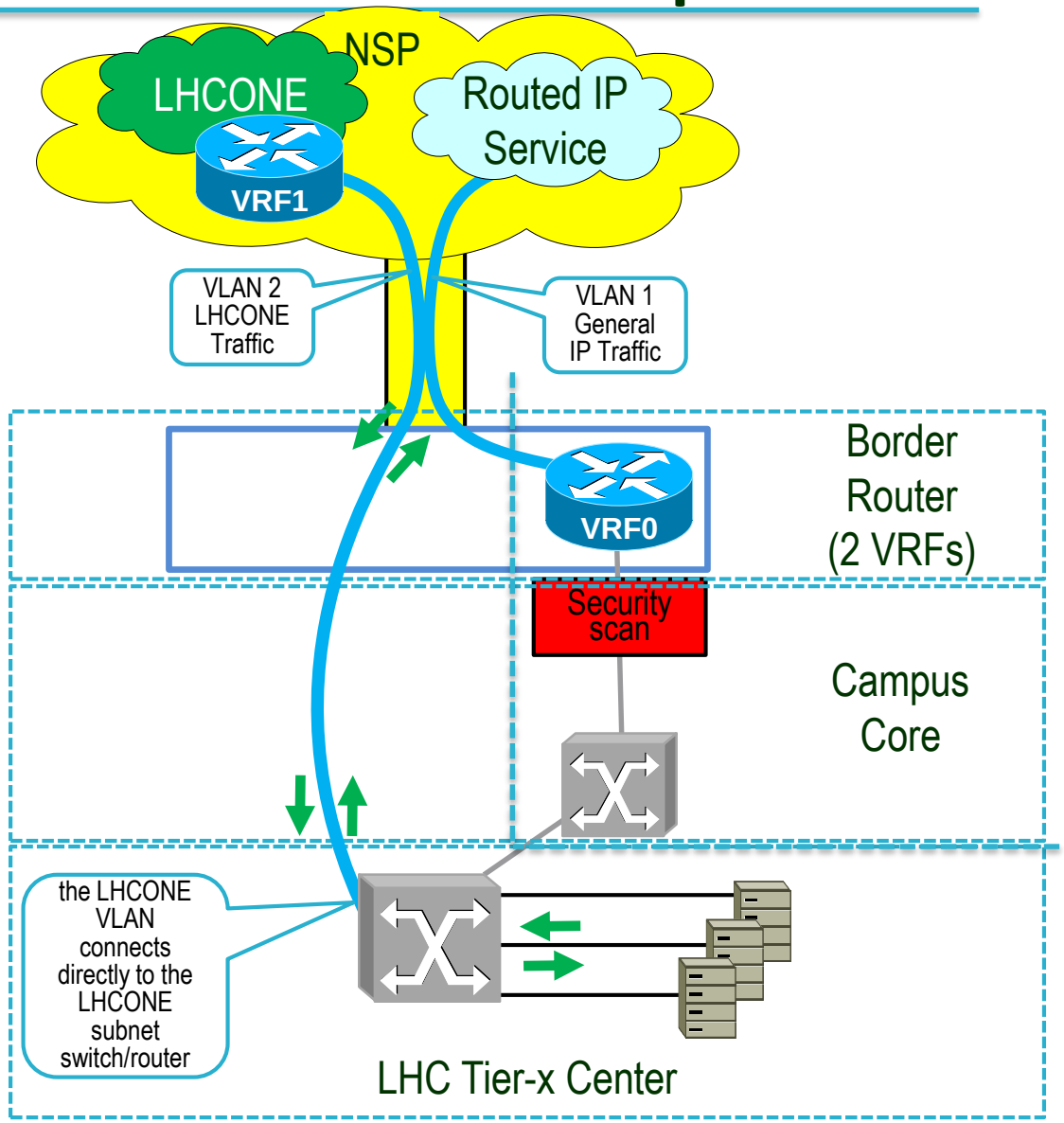
- The reality is that virtually ***all of these bilateral connections are mediated through an exchange point***
- Some of the exchange points are distributed in that cross domain connections can be made at different physical locations
- Some of the exchange points set up multipoint exchanges for LHCONE VRF cross connections (via either a dedicated physical switch or a virtual switch)



ESnet

LHCONE Basic Site Architecture Example-I

- There are a number of possible configurations and relationships between sites and the Network Service Provider runs the LHCONE VRF
- Sites must provide a separate subnet (and address space) for LHC resources
  - Separate routing instance for LHCONE
  - LHCONE subnet address block published to CERN
  - Direct access by LHCONE (and LHCOPN) to site LHC resources
- Very much like a ScienceDMZ
  - See fasterdata.es.net

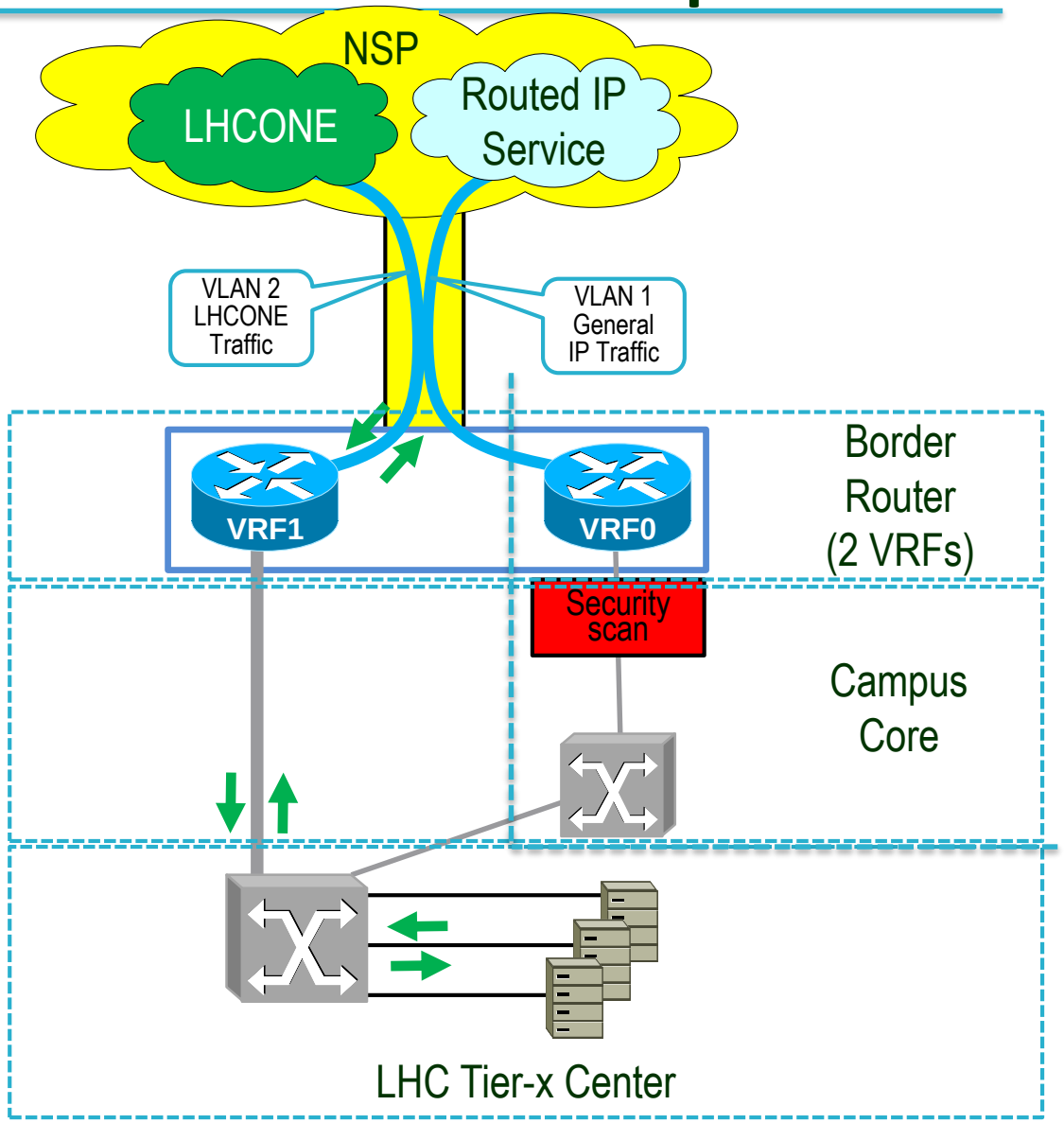


Thanks to Shawn McKee, U. Mich.

ESnet

LHCONE Basic Site Architecture Example-II

- Most big US sites, as well as others, run their own VRF because they want to have control over the LHCONE routing and have the ability to easily impose policy on the incoming traffic
- Note the for performance reasons the LHCONE subnet is frequently outside the site firewall
  - This is “reasonable” because the LHCONE sites agree to a common security policy and are all in the same science community (LHC)

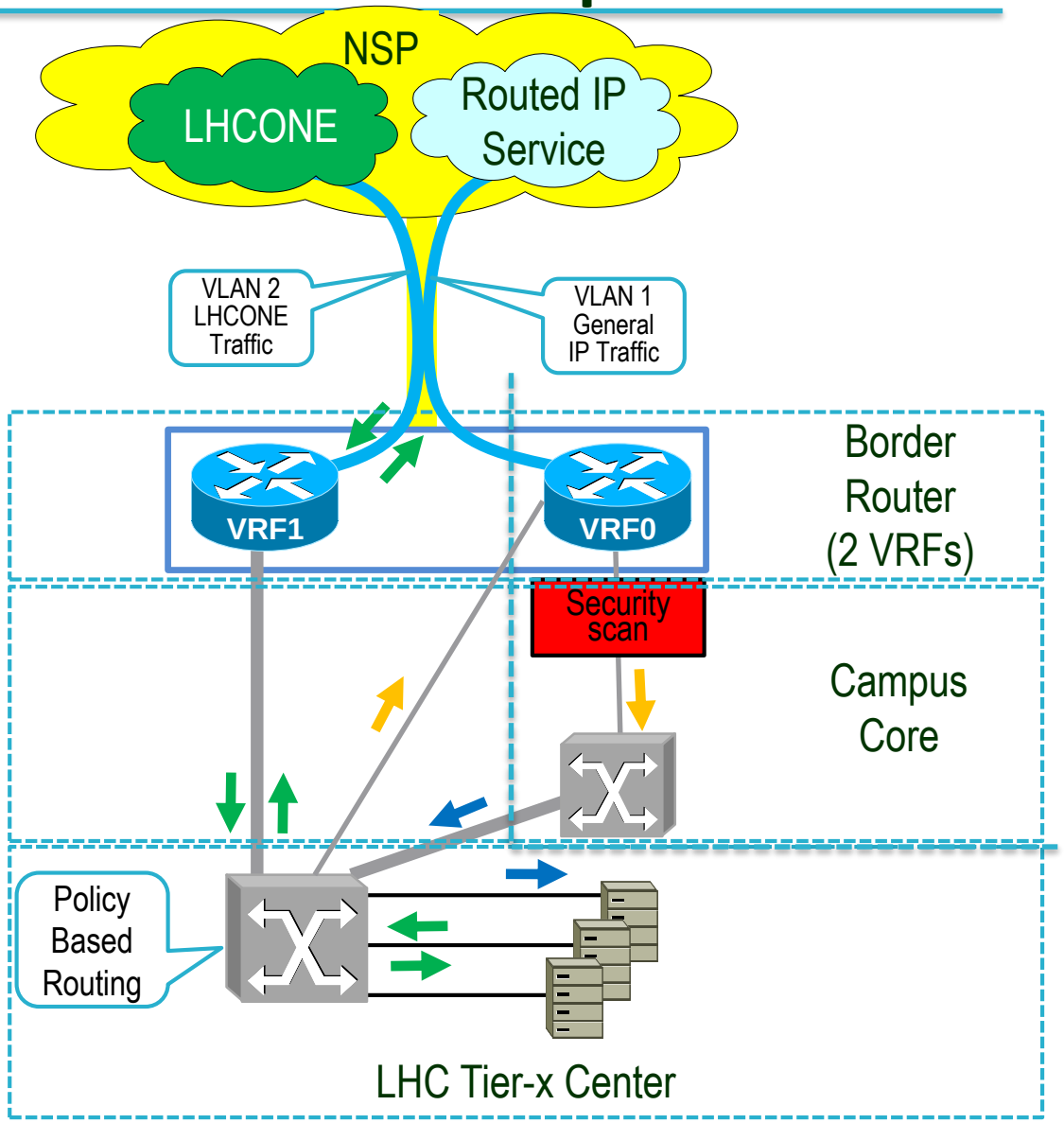


Thanks to Shawn McKee, U. Mich.

ESnet

LHCONE Site Architecture Example-IIa

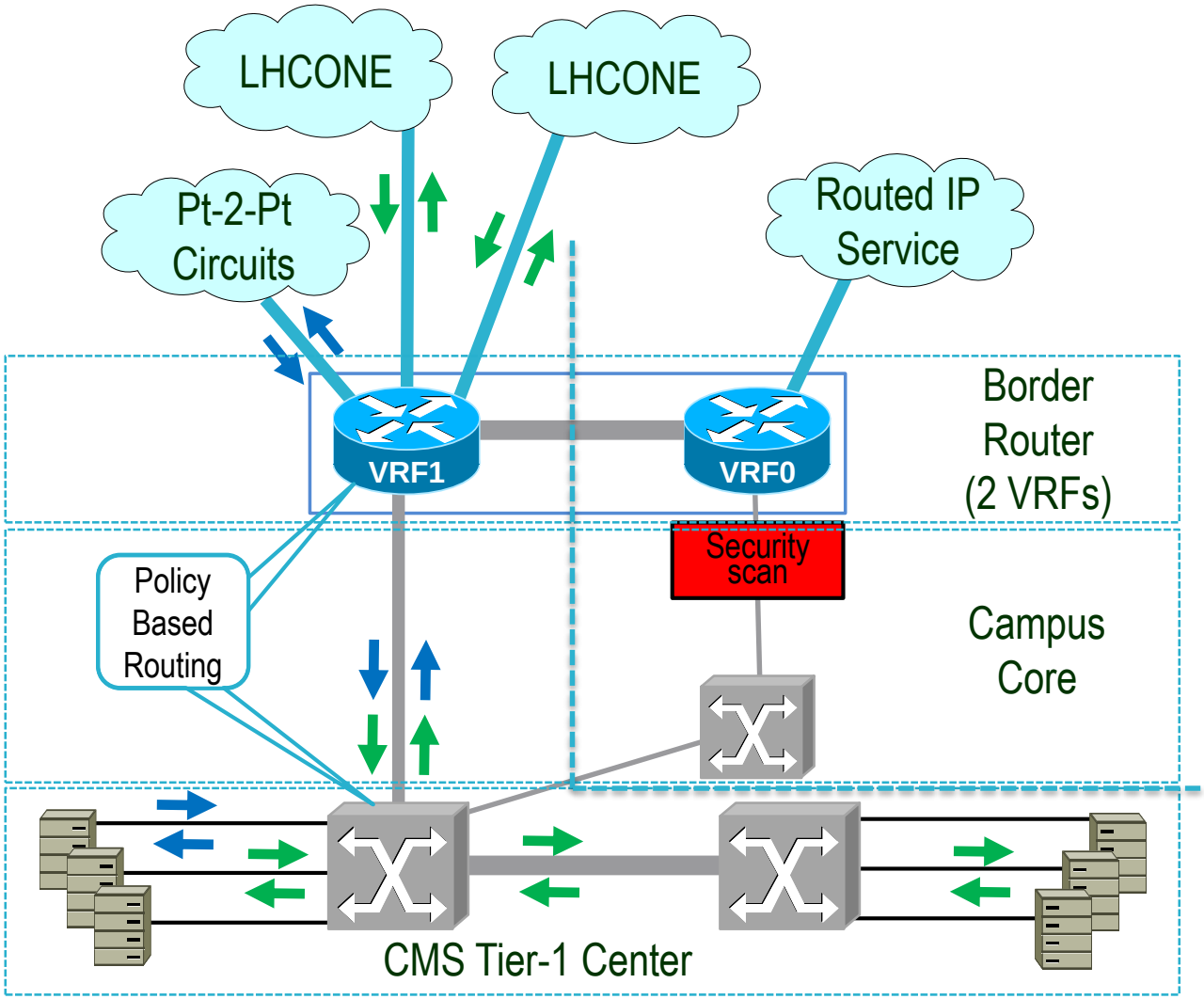
- A variant of the basic architecture is that
  - a policy is implemented that ***lets data from the campus move directly to the LHC systems***
  - however, ***traffic to campus from the LHC systems is security checked as through it were external traffic***
- Sometimes called a "diode" approach, this allows for high-speed data loading of LHC servers from campus, but screens all LHC Tier-x traffic because this environment is open to non-campus users



ESnet

LHCONE Large Site Architecture Example

- Policy Based Routing allows for site-specific decisions on what LHCONE sites will have direct access to the site LHC resources
  - E.g. have they agreed to the LHCONE security policy?
- This is a simplified diagram because such a site will probably have redundant border routers and redundant fiber to their upstream provider



Thanks to Phil DeMar, Fermilab and Bruno Hoeft, KIT

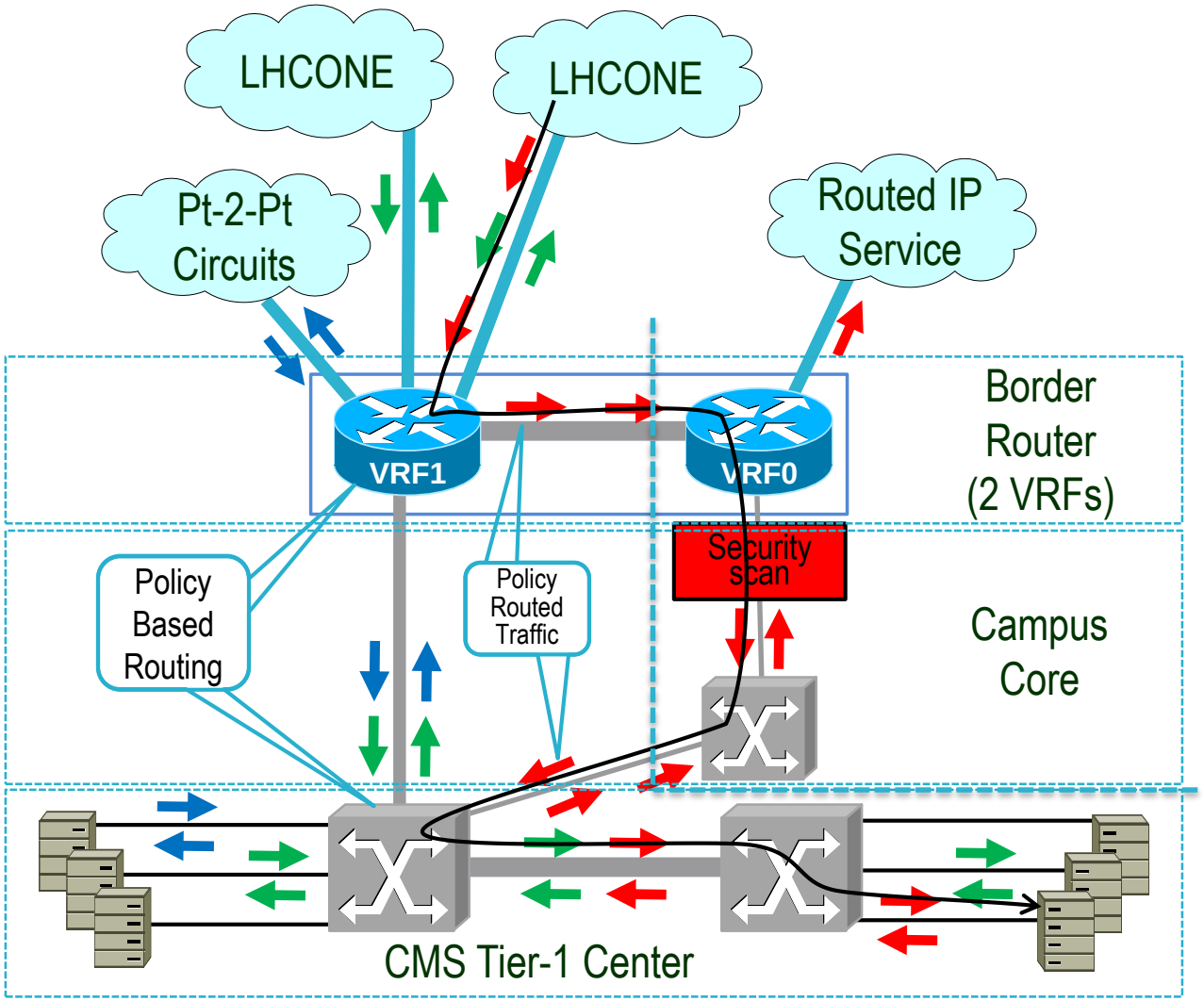
ESnet

LHCONE Site Architecture Example

- PBR is just one way to manage the incoming traffic, ACLs could be used, and in the (near) future, OpenFlow traffic management will likely be used

  But the net result is mostly the same:

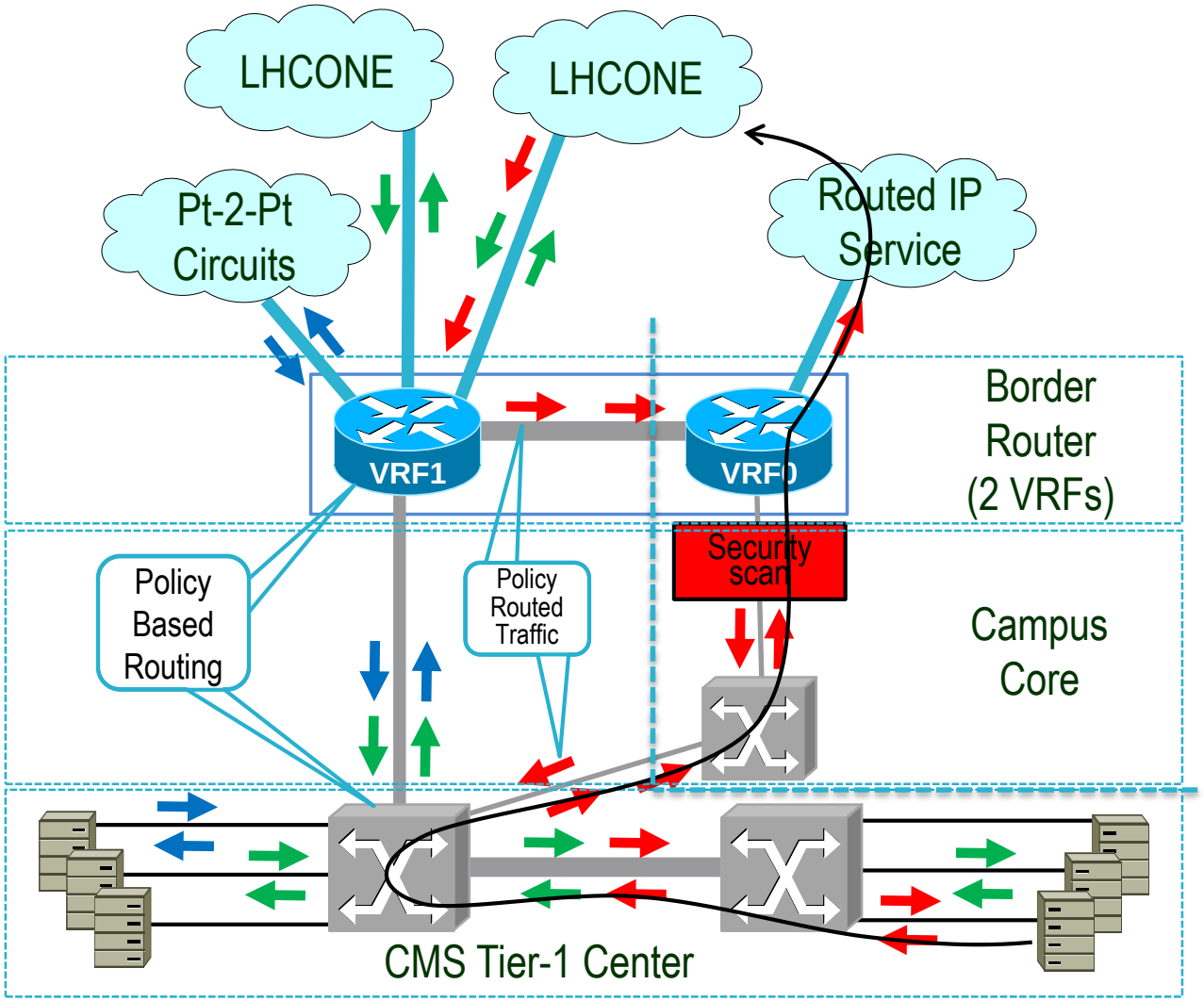
- *Non-conforming traffic is treated differently*
  - In this example, it is routed through the general, external access router and security screening



ESnet

LHCONE Site Architecture Example

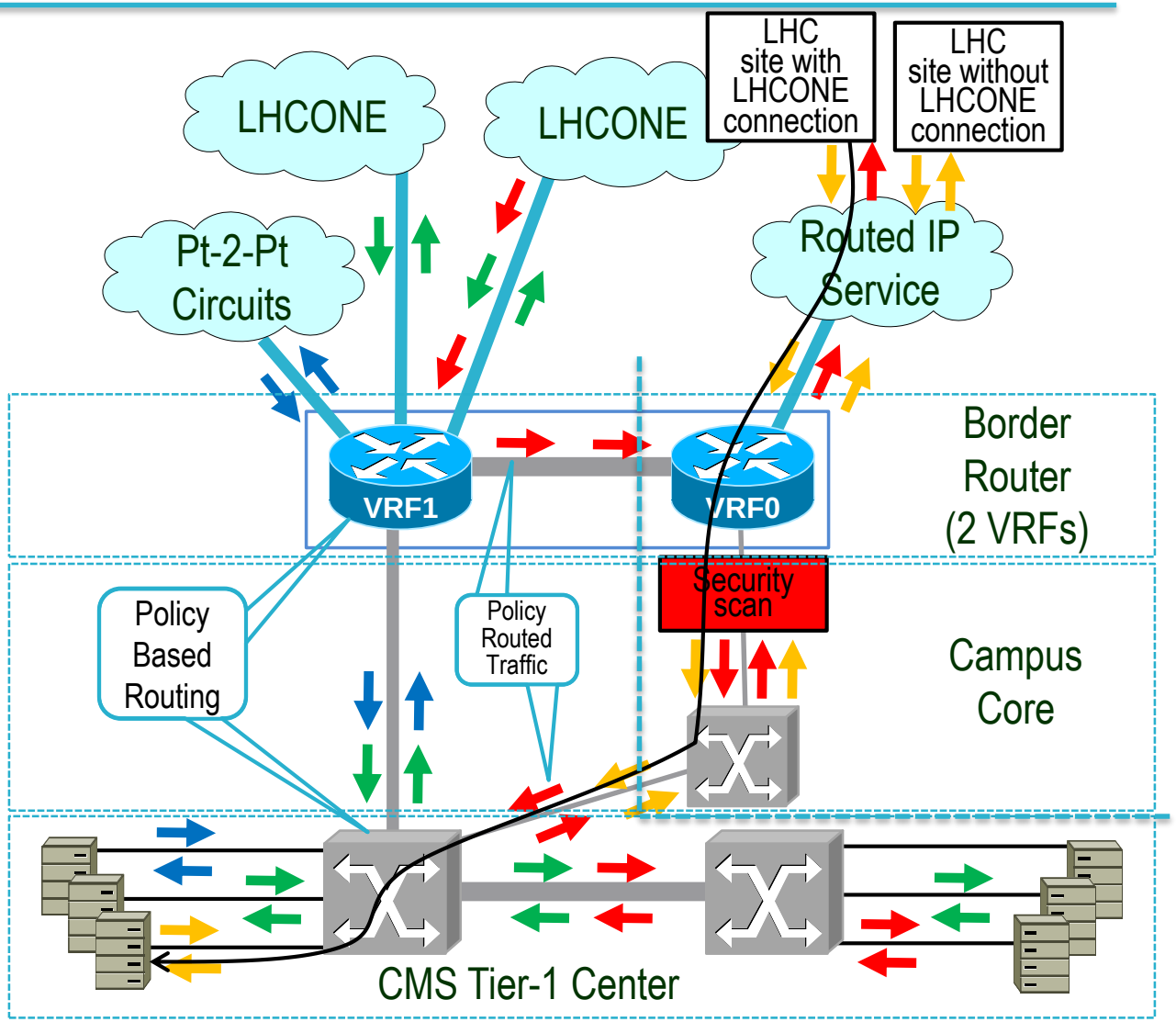
- In this example the non-conforming traffic is returned via the general Internet



ESnet

LHCONE Site Architecture Example

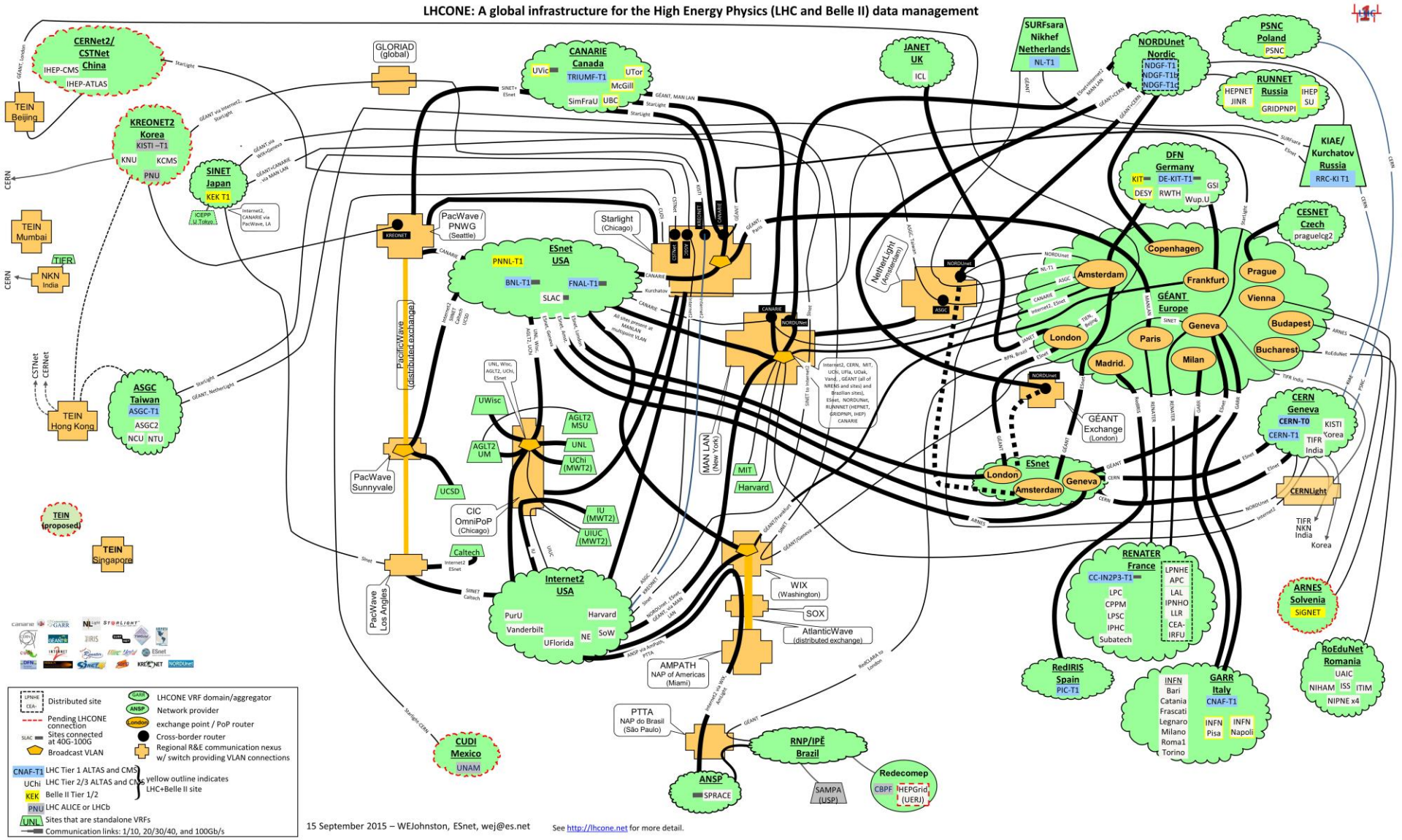
- If another site treats this center's traffic as non-conforming then that traffic is returned over the general Internet and the site must be able to route that traffic to the LHC resources internally (as illustrated here)
- If there is an LHC Tier-2 site that is not connected via LHCONE, then that traffic just looks like ordinary Internet traffic that happens to be accessing LHC resources



- This traffic must also be able to reach the LHCONE resources

END

- END


LHCONE: A global infrastructure for the High Energy Physics (LHC and Belle II) data management
CERNet2/ CSTNet China
IHEP-CMS
IHEP-ATLAS
TEIN Beijing
KREONET2 Korea
KISTI –T1
KNU
KCMS
PNU
SINET Japan
KEK T1
TEIN Mumbai
TIFR
NKN India
TEIN Hong Kong
ASGC Taiwan
ASGC-T1
ASGC2
NCU
NTU
TEIN (proposed)
TEIN Singapore
GLORIAD (global)
CANARIE Canada
UVic
TRIUMF-T1
UTor
McGill
SimFraU
UBC
JANET UK
ICL
SURFsara Nikhef Netherlands
NL-T1
NORDUnet Nordic
NDGF-T1
NDGF-T1b
NDGF-T1c
PSNC Poland
PSNC
RUNNET Russia
HEPNET
JINR
IHEP
SU
GRIDPNPI
KIAE/ Kurchatov Russia
RRC-KI T1
DFN Germany
KIT
DE-KIT-T1
GSI
DESY
RWTH
Wup.U.
CESNET Czech
praguelcg2
PacWave / PNWG (Seattle)
Starlight (Chicago)
NetherLight (Amsterdam)
PacificWave (distributed exchange)
ESnet USA
PNNL-T1
BNL-T1
FNAL-T1
SLAC
GÉANT Europe
Copenhagen
Amsterdam
Frankfurt
Prague
Vienna
Geneva
Budapest
London
Paris
Madrid
Milan
Bucharest
GÉANT Exchange (London)
ESnet
London
Amsterdam
Geneva
CERN Geneva
CERN-T0
CERN-T1
KISTI Korea
TIFR India
CERNLight
TIFR
NKN
India
Korea
ARNES Solvenia
SIGNET
RoEduNet Romania
UAIC
NIHAM
ISS
ITIM
NIPNE x4
RENATER France
CC-IN2P3-T1
LPNHE
APC
LAL
IPNHO
LLR
CEA-
IRFU
LPC
CPPM
LPSC
IPHC
Subatech
RedIRIS Spain
PIC-T1
GARR Italy
CNAF-T1
INFN Bari
Catania
Frascati
Legnaro
Milano
Roma1
Torino
INFN Pisa
INFN Napoli
PacWave Sunnyvale
PacWave Los Angles
UWisc
AGLT2 MSU
AGLT2 UM
UNL
UChi (MWT2)
UCSD
CIC OmniPoP (Chicago)
IU (MWT2)
UIUC (MWT2)
Caltech
MAN LAN (New York)
MIT
Harvard
Internet2 USA
PurU
Vanderbilt
UFlorida
Harvard
SoW
NE
WIX (Washington)
SOX
AtlanticWave (distributed exchange)
AMPATH NAP of Americas (Miami)
PTTA NAP do Brasil (São Paulo)
RNP/IPÊ Brazil
SAMPA (USP)
Redecomep
CBPF
HEPGrid (UERJ)
ANSP
SPRACE
CUDI Mexico
UNAM
Distributed site
Pending LHCONE connection
Sites connected at 40G-100G
Broadcast VLAN
LHC Tier 1 ALTAS and CMS
LHC Tier 2/3 ALTAS and CMS
Belle II Tier 1/2
LHC ALICE or LHCb
Sites that are standalone VRFs
Communication links: 1/10, 20/30/40, and 100Gb/s
LHCONE VRF domain/aggregator
Network provider
exchange point / PoP router
Cross-border router
Regional R&E communication nexus w/ switch providing VLAN connections
yellow outline indicates LHC+Belle II site
15 September 2015 – WEJohnston, ESnet, wej@es.net
See http://lhcone.net for more detail.

LHCONE: A global infrastructure for the High Energy Physics (LHC and Belle II) data management