Brookhaven National Laboratory

U.S. DEPARTMENT OF ENERGY

BNL Network Infrastructure

Mark Lukasczyk

Network Engineering

mlukasczyk@bnl.gov

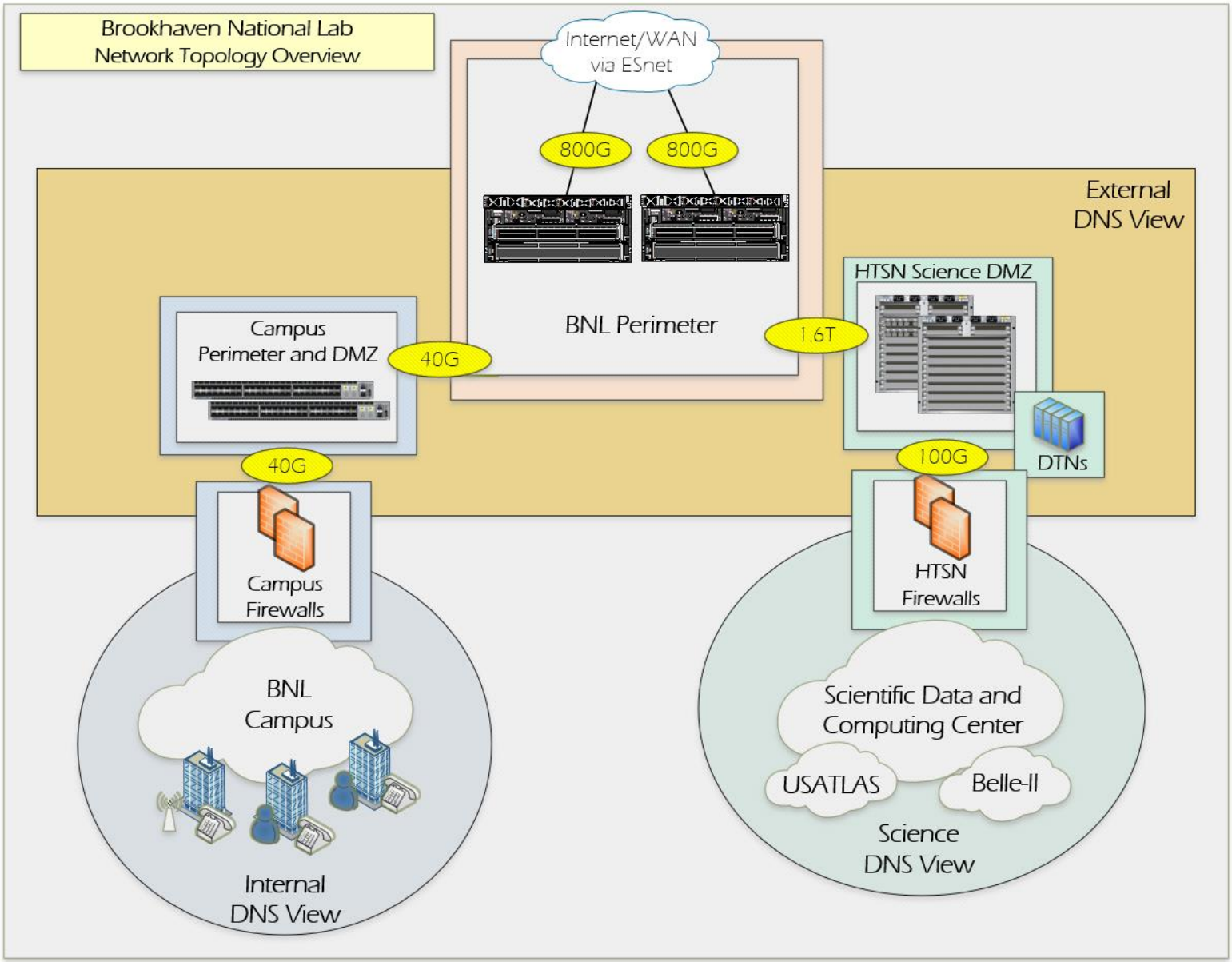
March 18, 2024

@BrookhavenLab

BNL Network Infrastructures

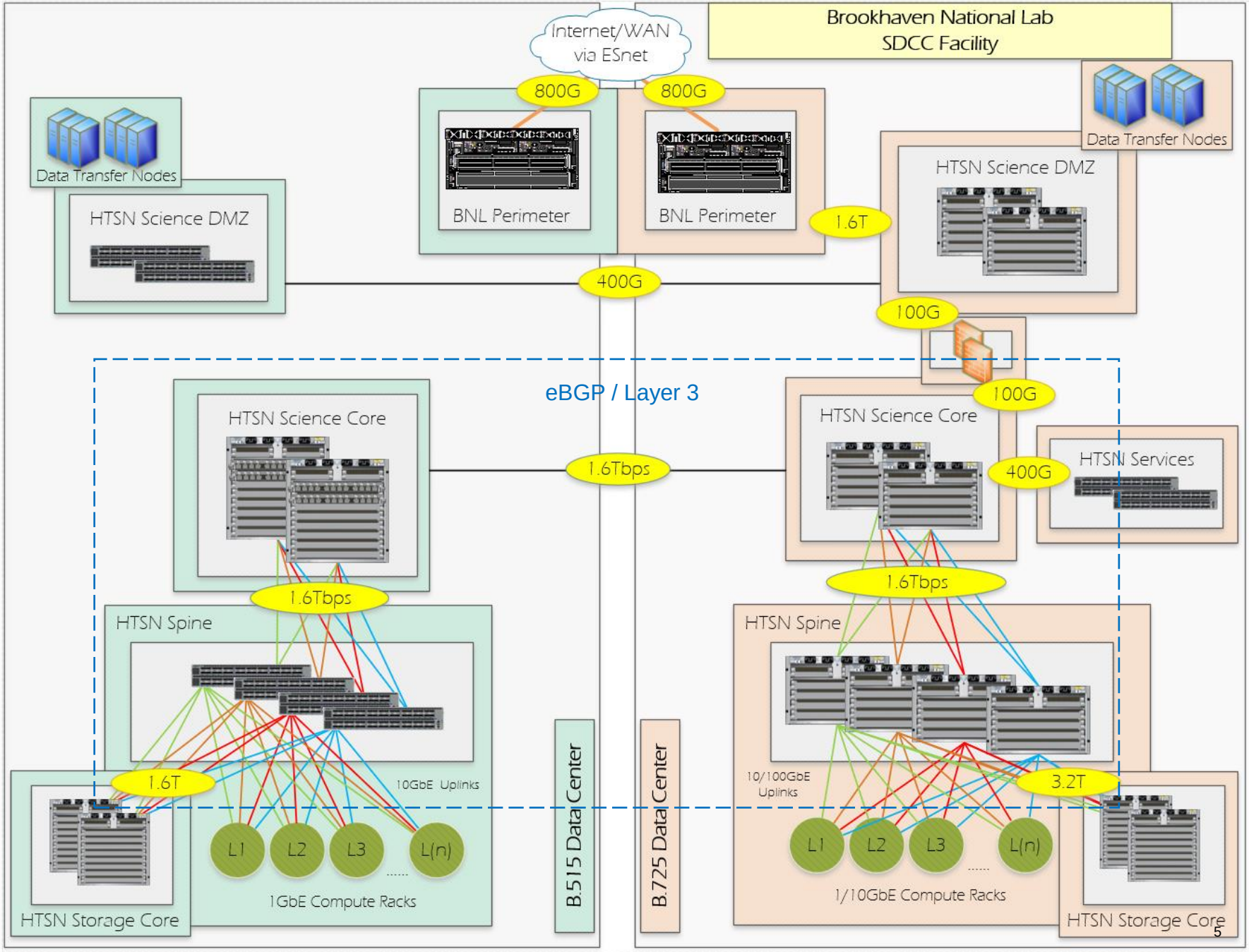
- Campus Network
  - Network foundation underlying the critical services within the enterprise such as workstations, phone service, security, safety and monitoring and enterprise computing.
- High Throughput Science Network (HTSN)
  - A centralized fault-tolerant, scalable, Terabit per second network infrastructure to support all collaborations compute and storage requirements.

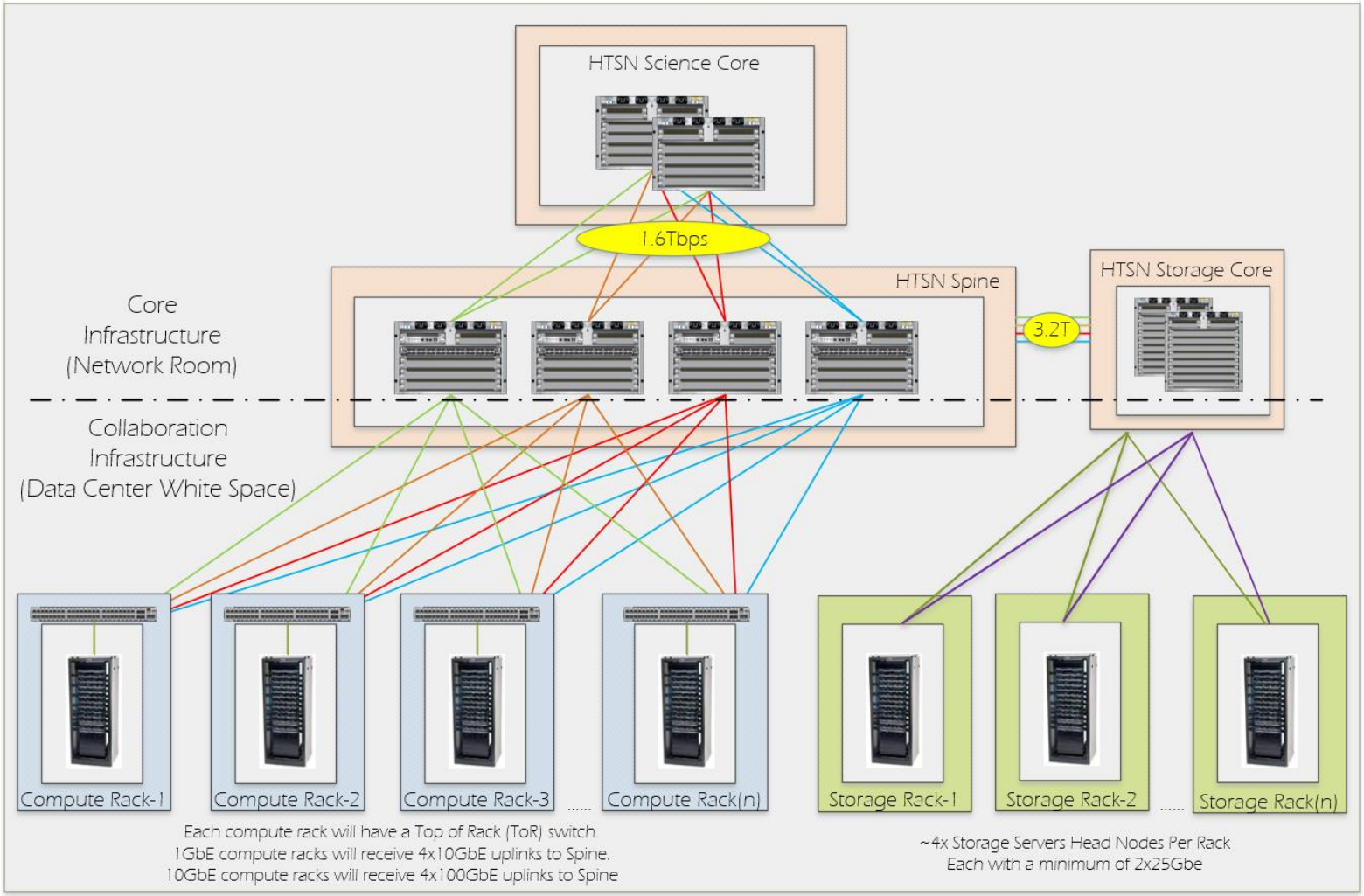
High Level Network Overview



BNL/ESnet Connectivity to USATLAS T2s

- ESnet's Connectivity to USATLAS T2 RENs:
  - AGLT2 & MWT2
    - OmniPoP Chicago at 800G (2x4x100G)
  - NET2
    - Northern Crossroads in Boston at 100G
  - SWT2
    - Great Plains Network (GPN) at 400G
    - Lonestar Education and Research Network (LEARN) at 100G
  - CERN
    - Aggregate 1Tbps for which 800G is for LHCOPN/LHCONE.
      - Note: BNL's primary LHCOPN circuit traverses a single 400GbE Trans-Atlantic circuit.


Brookhaven National Lab
SDCC Facility
Internet/WAN
via ESnet
800G
800G
Data Transfer Nodes
BNL Perimeter
BNL Perimeter
Data Transfer Nodes
HTSN Science DMZ
1.6T
HTSN Science DMZ
400G
100G
eBGP / Layer 3
100G
HTSN Science Core
HTSN Science Core
HTSN Services
1.6Tbps
400G
1.6Tbps
1.6Tbps
HTSN Spine
HTSN Spine
1.6T
10GbE Uplinks
10/100GbE
Uplinks
3.2T
L1
L2
L3
L(n)
L1
L2
L3
L(n)
1GbE Compute Racks
1/10GbE Compute Racks
HTSN Storage Core
HTSN Storage Core
B.515 Data Center
B.725 Data Center



Summary of the Network Architecture:

- Since eBGP is utilized, there is little to no vendor interoperability issues.
- The Spine and Science Core are comprised of Arista 7504R3 while the leaves (compute racks) are Arista 7280R3's or Nokia 7250's. Each compute rack utilizes either a /26 or /27 and a /64 (No VXLAN required).
- The Storage Core is an MLAGed pair of Arista 7508R3's which allows storage head nodes to redundantly connect to each Arista chassis in an active/active fashion. (Note: BNL doesn't utilize top of rack switches for storage connectivity since the head nodes can be spaced far apart).