

# U.S. ATLAS Tier-1 Network Status

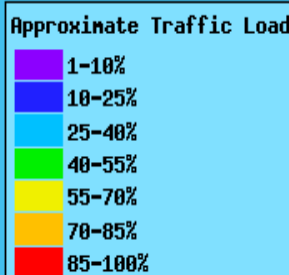
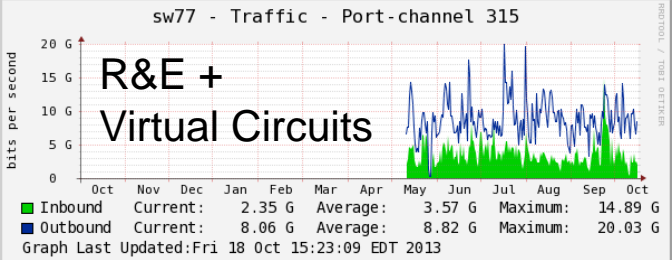
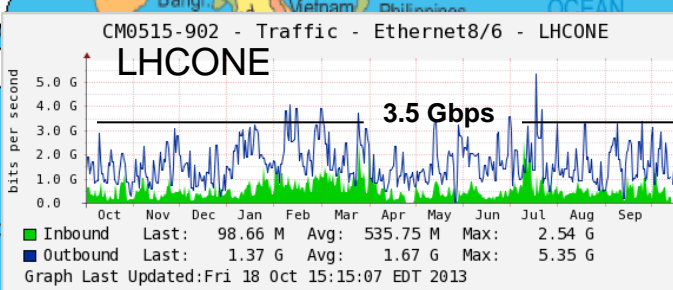
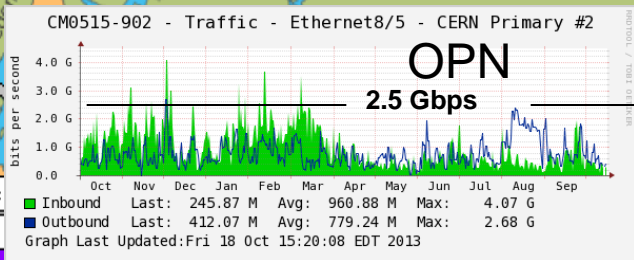
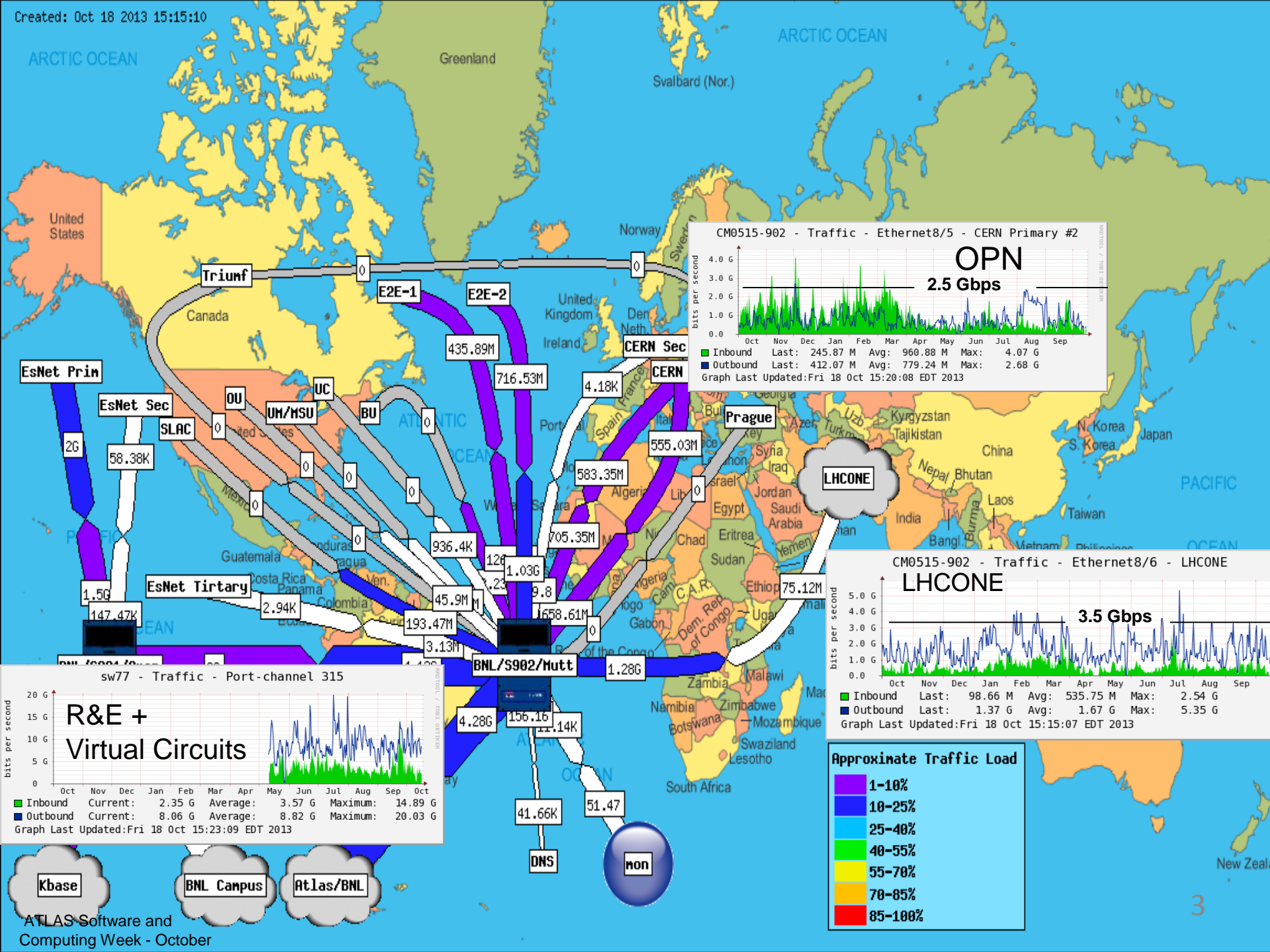
Michael Ernst  
mernst@bnl.gov



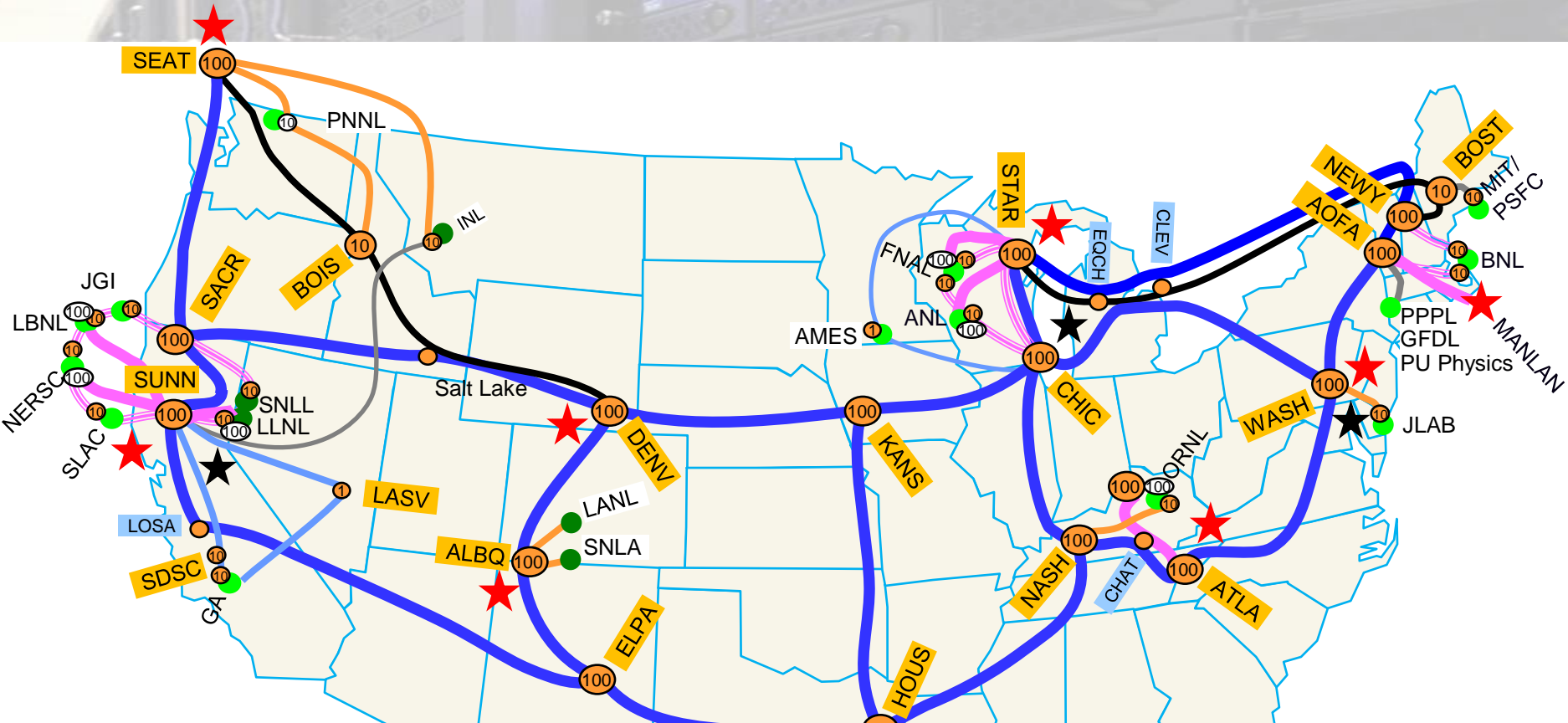
# Tier-1 Production Network Connectivity

- **As to the Tier-1: the maximum usable b/w is 70 Gbps**
  - 50 Gbps dedicated circuits/unshared plus 20 Gbps general IP service shared across all departments at BNL)
  - Currently available for Tier-0 ↔ Tier-1 and Tier-1 ↔ Tier-1: 17 Gbps via OPN/USLHCNet + 2\*10 Gbps ESnet/GEANT shared by researchers in US
  - One dedicated 10 Gbps circuit for LHCONE (LHC Open Network Environment, connecting the Tier-1 at MANLAN in New York)
  - DOE/ESnet “dark fiber project” has brought abundant physical fiber infrastructure into the lab
- **BNL connected to ESnet at 100G**
  - Have T1 facility connected to ESnet at 100G for R&D (ANA TA link)
  - In the process of moving BNL/T1 production environment to 100G





# ESnet5 March 2013



- SUNN** ESnet PoP/hub locations
- 100** ESnet managed 100G routers
- 10** ESnet managed 10G router
- 10** **100** Site managed routers
- LOSA** ESnet optical node locations (only some are shown)
- ESnet optical transport nodes (only some are shown)
- ★** commercial peering points
- ★** R&E network peering locations
- LBNL Major Office of Science (SC) sites
- LLNL Major non-SC DOE sites

- Routed IP 100 Gb/s
- Routed IP 4 X 10 Gb/s
- 3rd party 10Gb/s
- Express / metro 100 Gb/s
- Express / metro 10G
- Express multi path 10G
- Lab supplied links
- Other links
- Tail circuits

*Geography is  
only representational*



# BNL's LHCOPN Connectivity is provided by USLHCNet

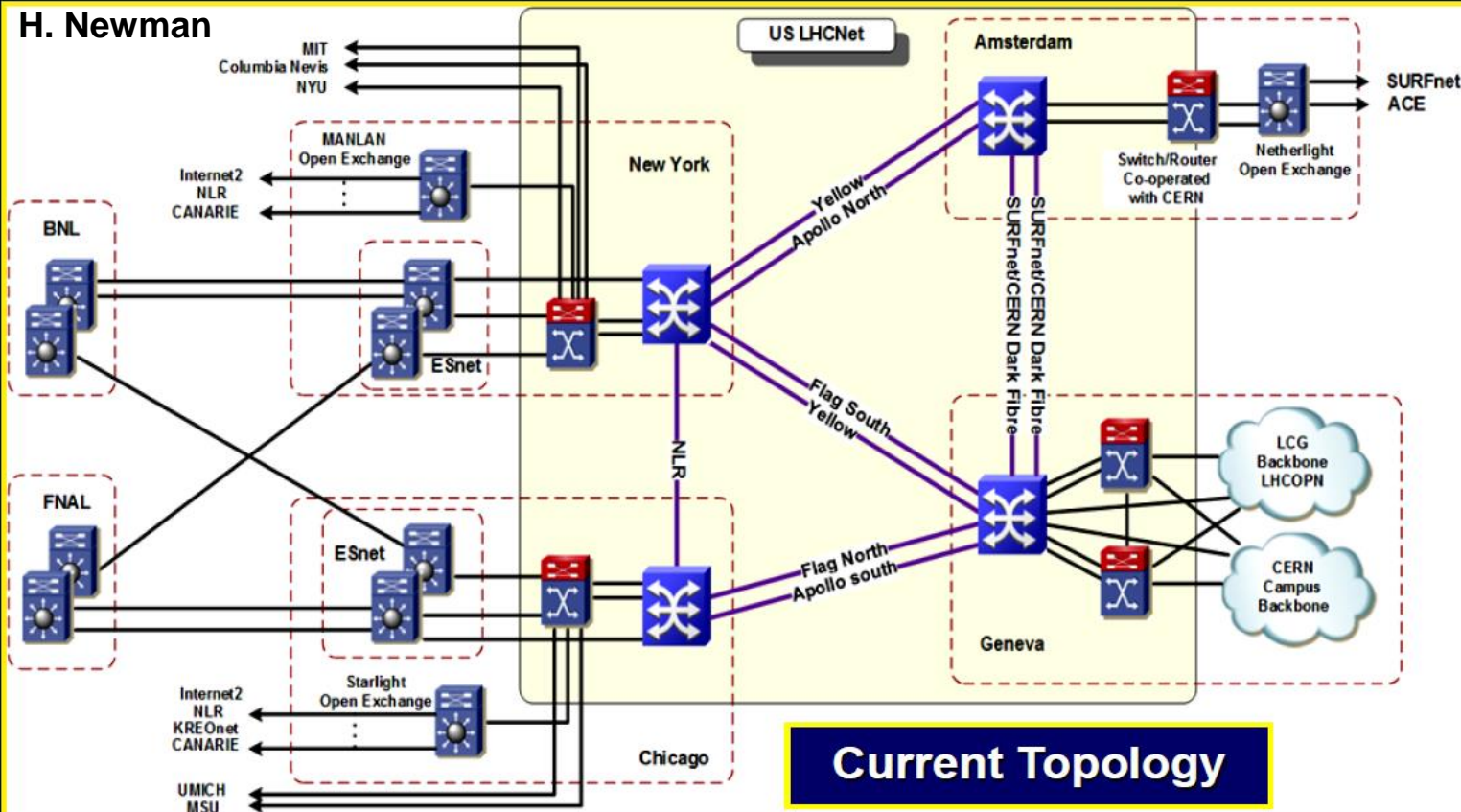


## US LHCNet Layout and Technology

Non-stop Operation; Circuit-oriented Services



H. Newman



**Core: Optical multiservice Switches [\*] that provide resilience**

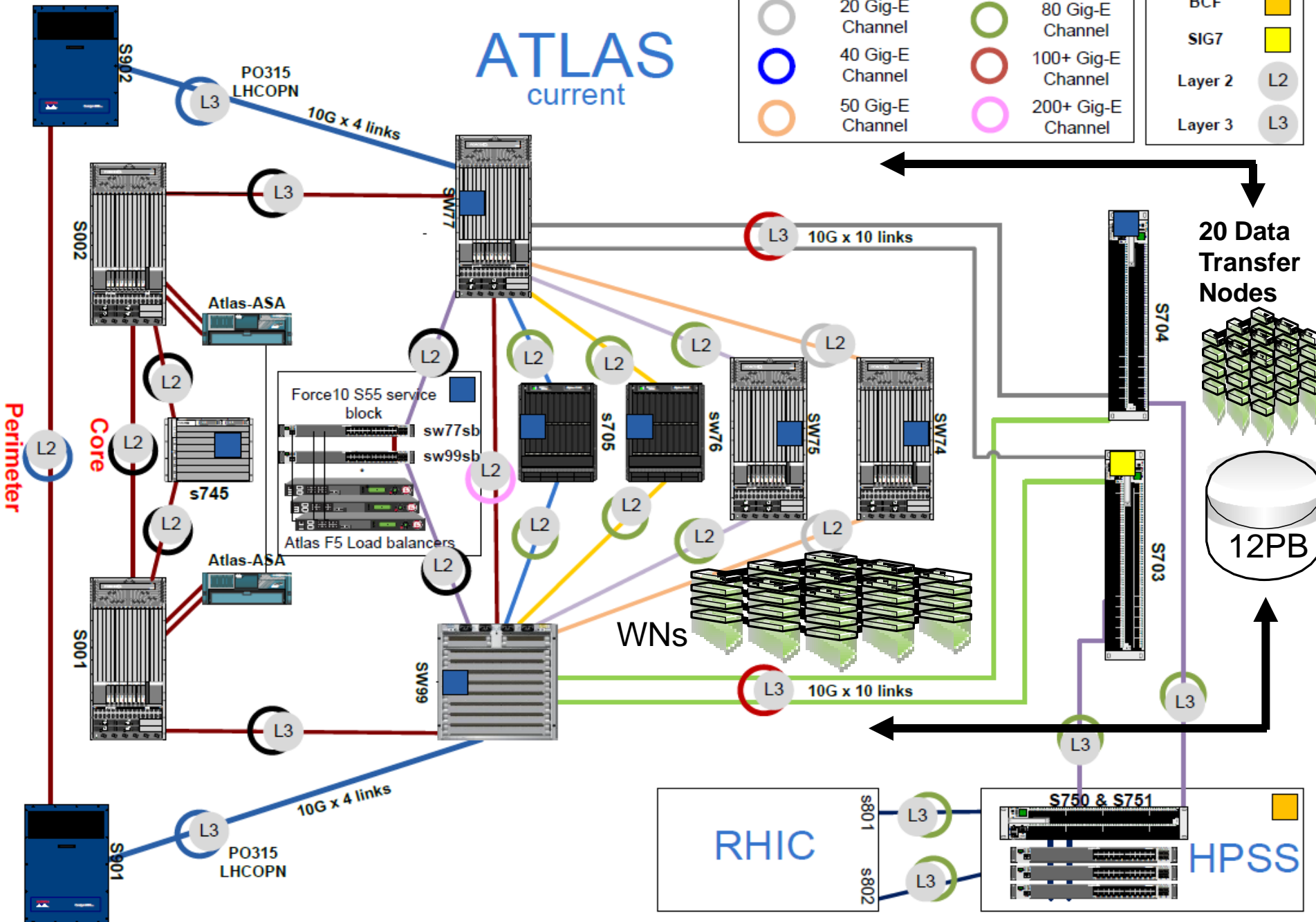
**Performance enhancing Standard Extensions: VCAT, LCAS**

**USLHCNet, ESNET, BNL & FNAL: Equipment and link redundancy**

**[\*] Dynamic circuit-oriented Carrier services with BW guarantees, with robust fallback at Layer 1: Hybrid optical network**

# ATLAS current

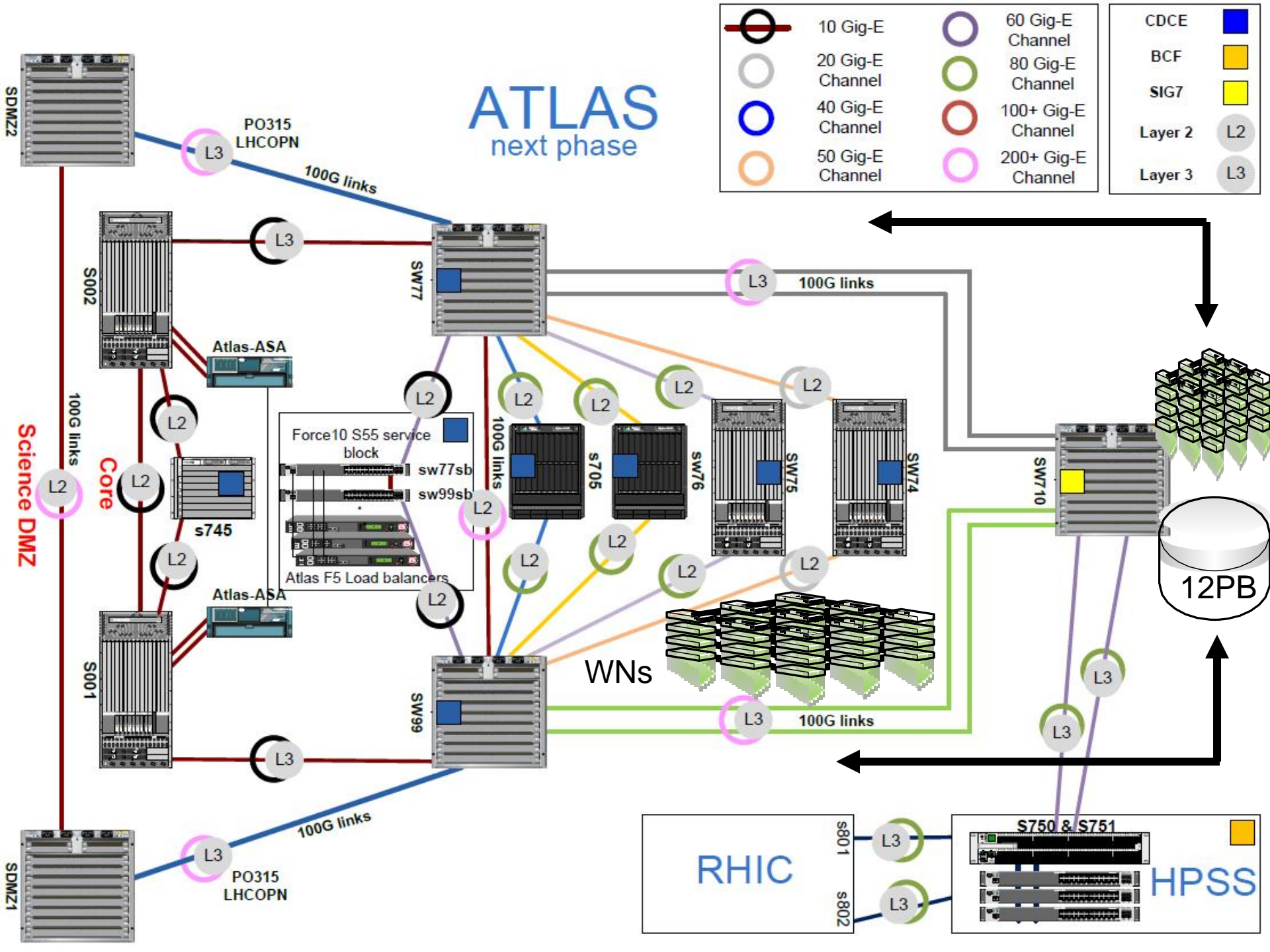
	10 Gig-E		60 Gig-E Channel	CDCE	
	20 Gig-E Channel		80 Gig-E Channel	BCF	
	40 Gig-E Channel		100+ Gig-E Channel	SIG7	
	50 Gig-E Channel		200+ Gig-E Channel	Layer 2	
				Layer 3	



RHIC

HPSS



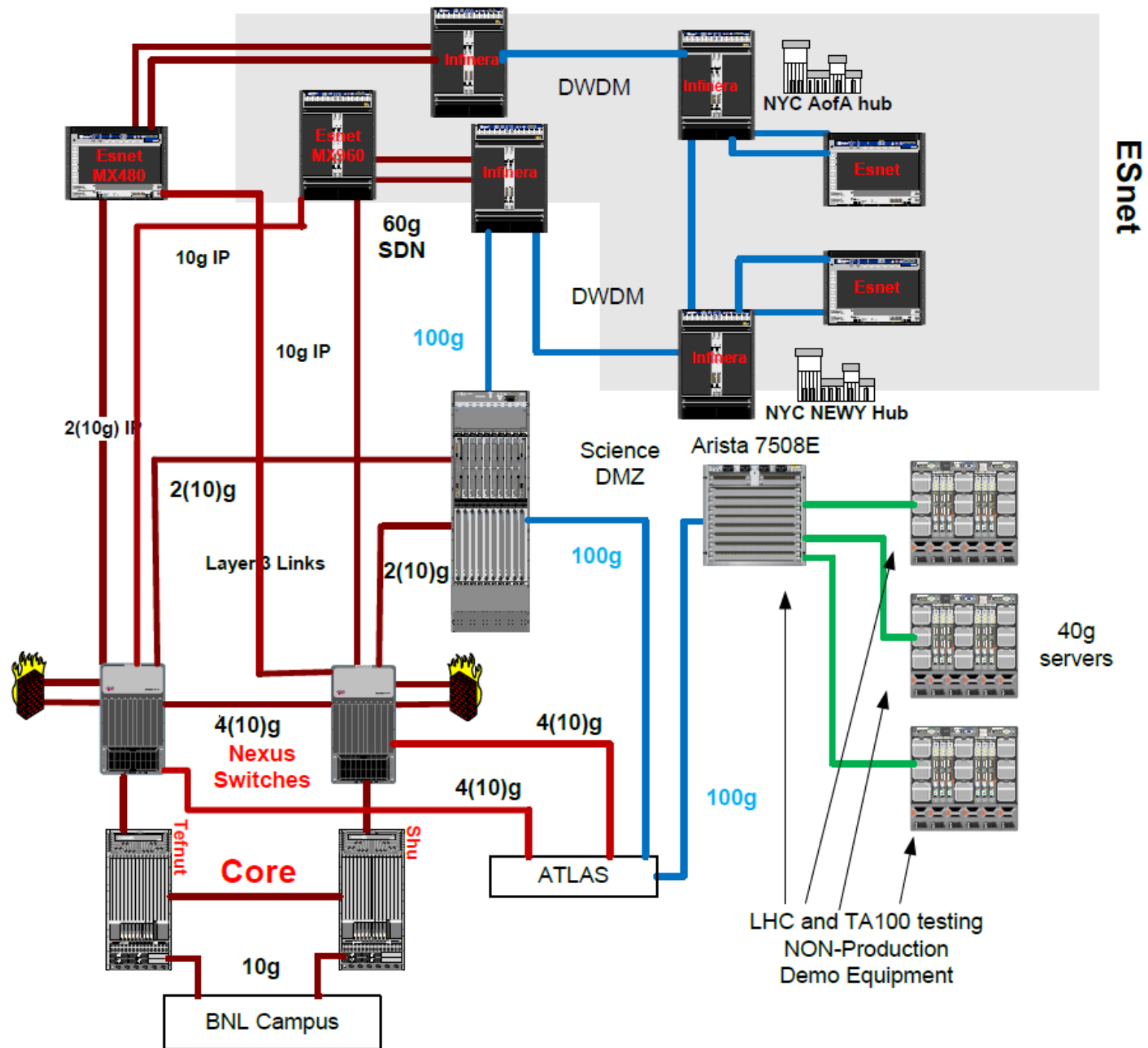


# BNL PDN and Science DMZ

(Current)

## Legend

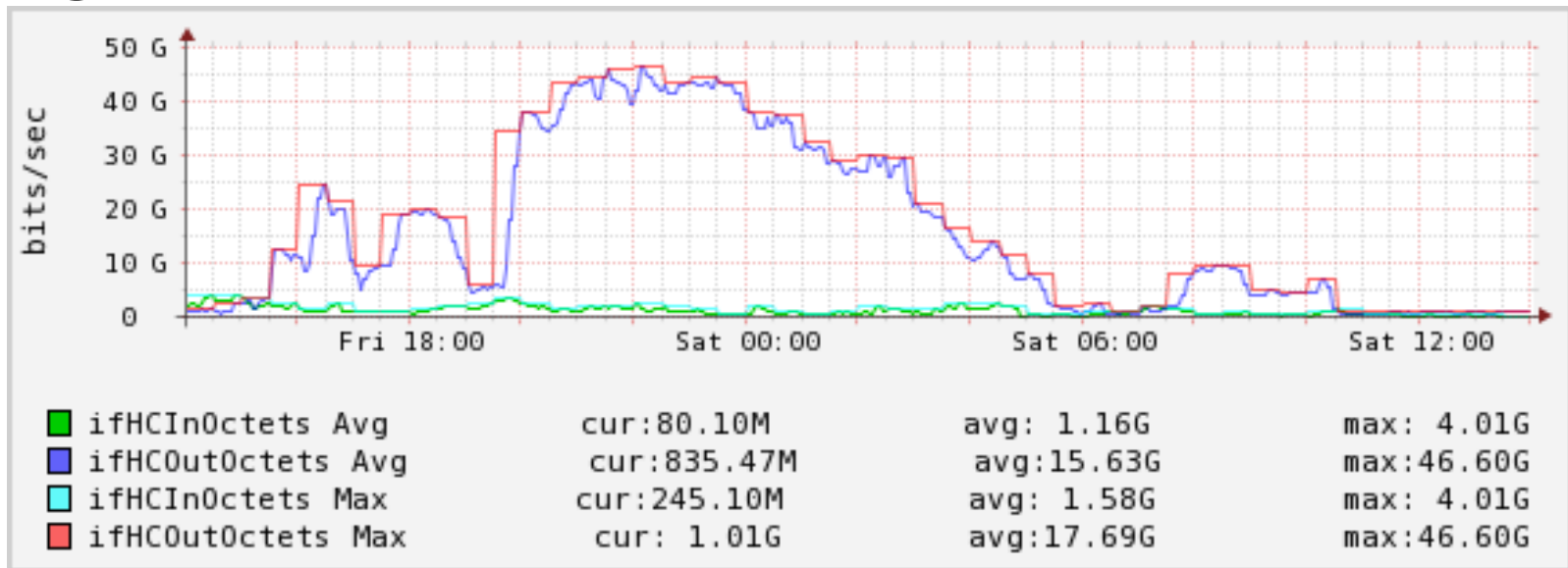
- 1Gig-E
- 10 Gig-E
- 40 Gig-E
- 100 Gig-E





# CERN/T1 -> BNL Transfer Performance via ANA 100G

- Regular ATLAS Production + Test Traffic

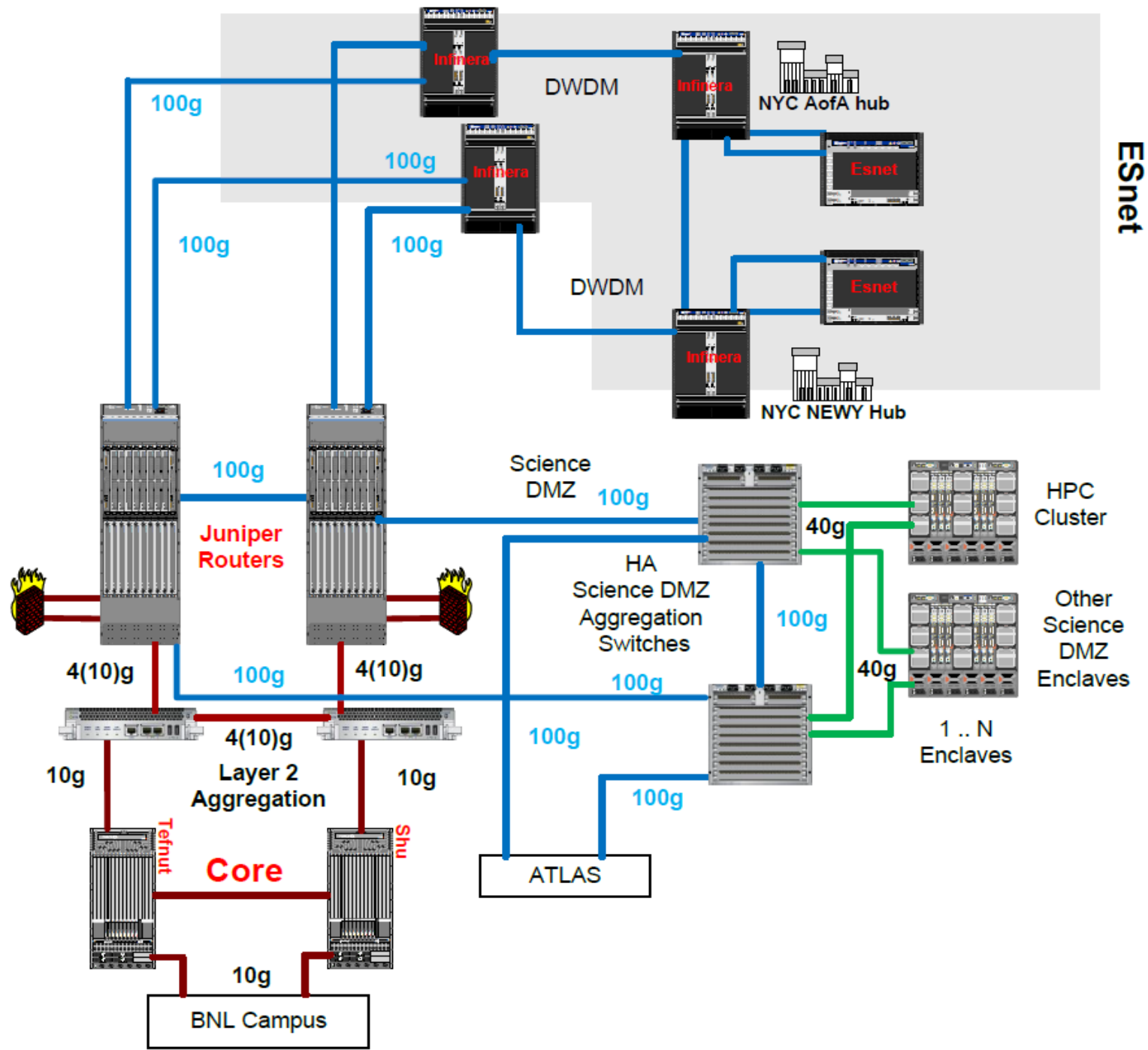


- Observations (all in the context of ATLAS)
  - Never exceeded ~50 Gbits/sec
  - CERN (ATLAS EOS) -> BNL limited at ~1.5 GB/s
    - Achieved >8 GB/s between 2 Hosts @ CERN and BNL
  - Each T1 (via OPN/CERN) -> BNL limited to ~0.5 GB/s

# BNL PDN and Science DMZ (Proposed Final)

## Legend

- 1Gig-E
- 10 Gig-E
- 40 Gig-E
- 100 Gig-E



# Evolving Tier-2 Networking

- All 5 US ATLAS Tier-2 (10 Sites) sites are currently connected at rate of at least 10 Gbps
  - This has proven not sufficient to efficiently utilize the resources at federated sites (CPU & Disk at different sites)
- US ATLAS Facilities have recognized the need to develop network infrastructure at sites
  - A comprehensive, forward looking plan exists
  - Additional Funding was provided by US ATLAS Mgmt & NSF
- Sites are in the process of upgrading their Connectivity to 100 Gbps
  - 6 Sites will have completed upgrade by end of April
  - All others will be done by the end of 2014

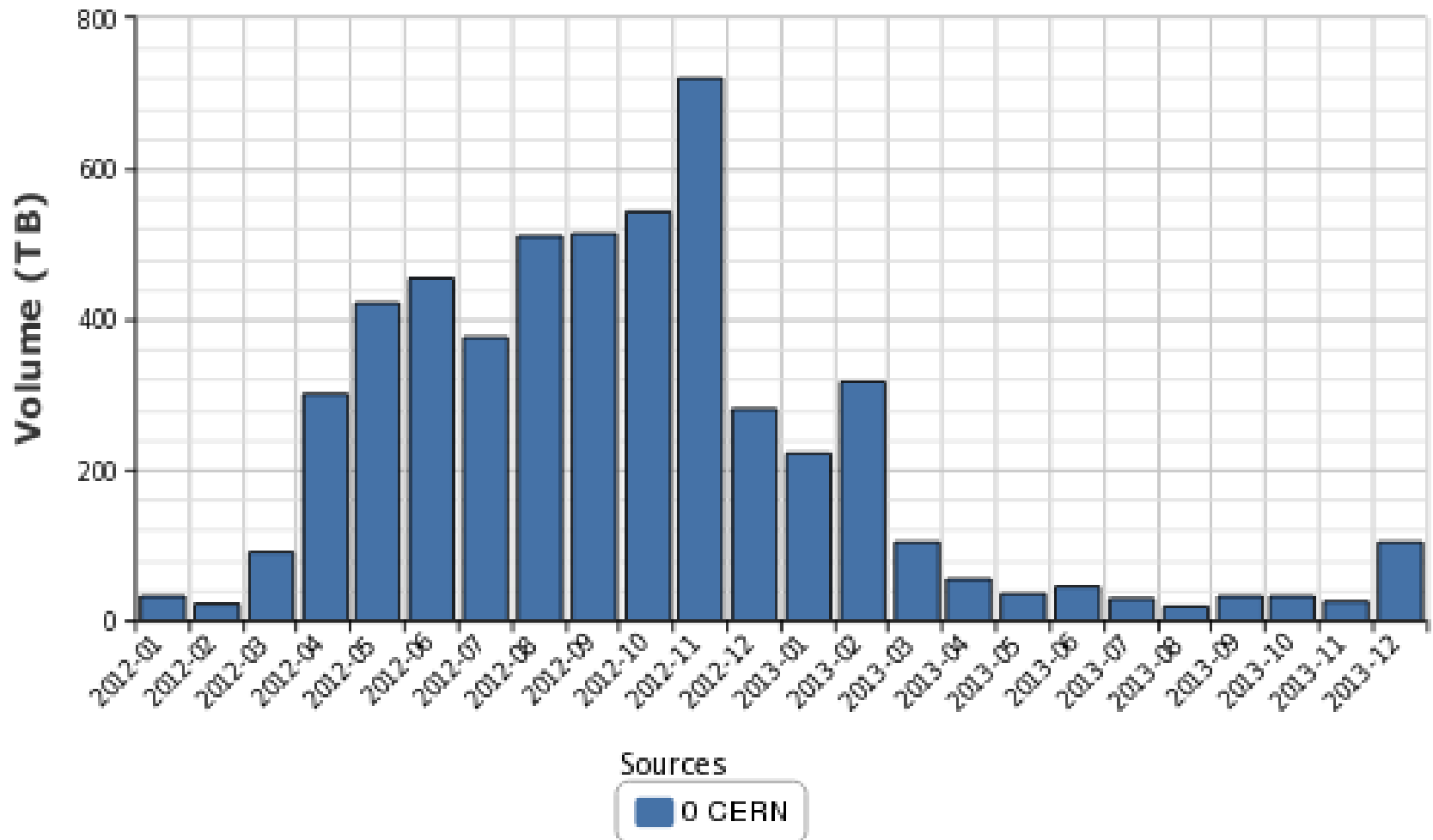


# From CERN to BNL



## Transfer Volume

2012-01-01 00:00 to 2013-12-31 00:00 UTC

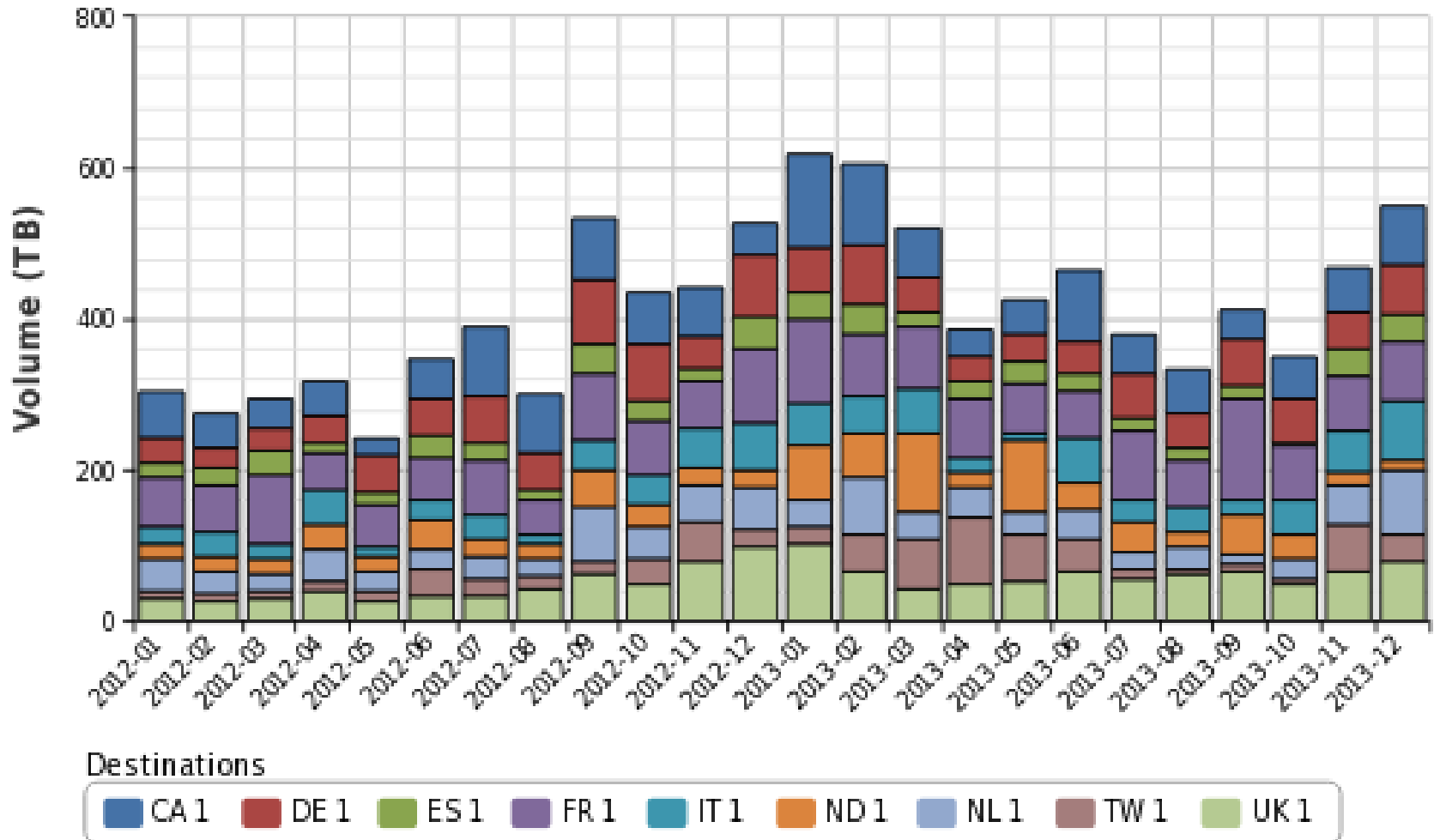


# From BNL to T1s



## Transfer Volume

2012-01-01 00:00 to 2013-12-31 00:00 UTC

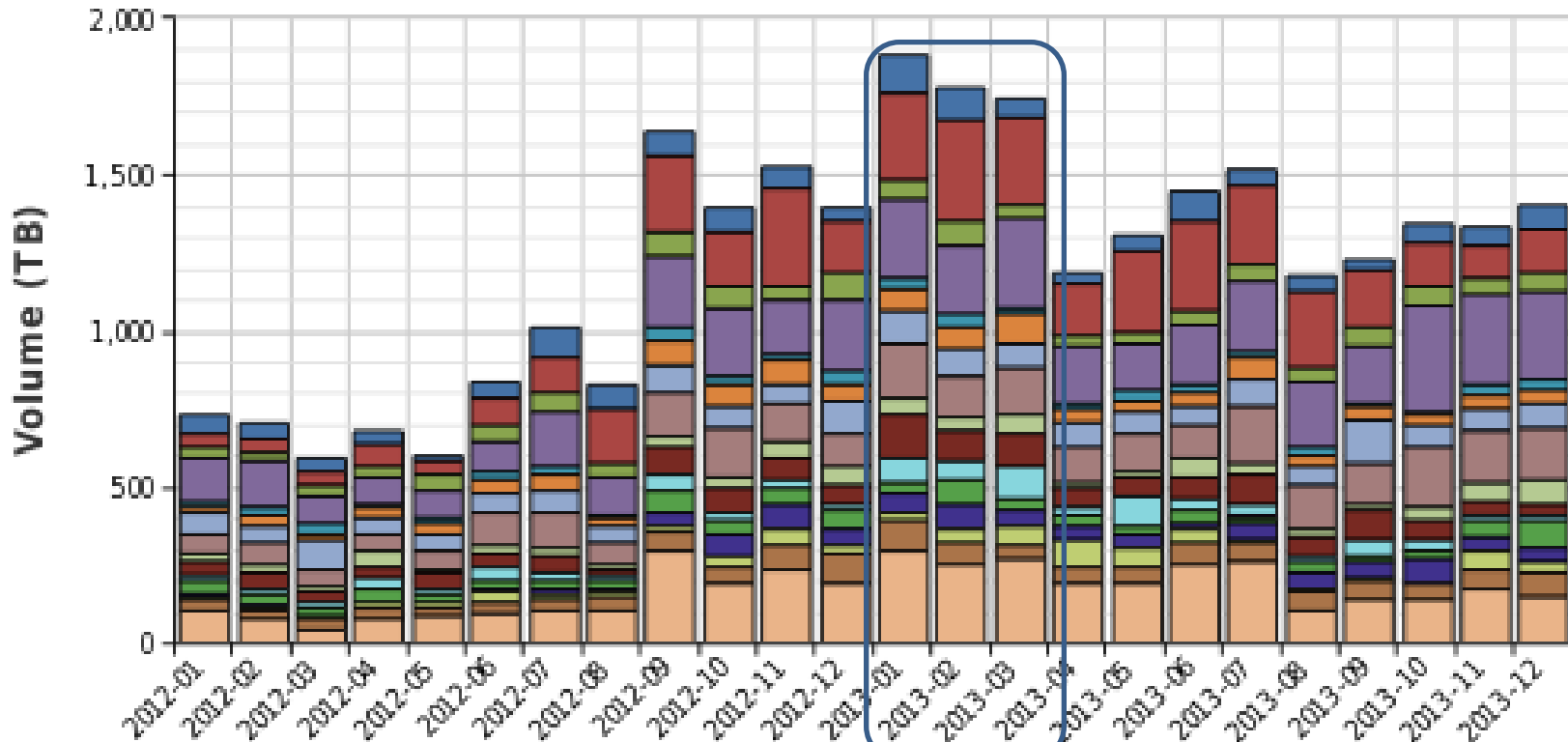


# From BNL to T1s and T2s



## Transfer Volume

2012-01-01 00:00 to 2013-12-31 00:00 UTC

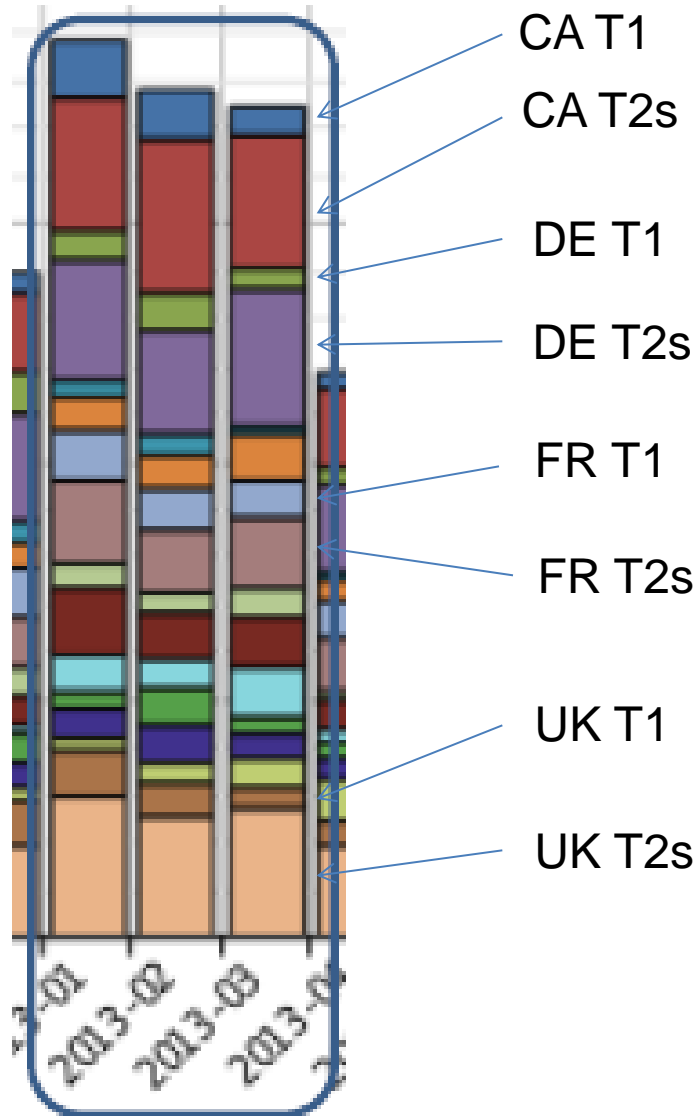


### Destinations





# T1s vs. T2s from BNL (2013 Winter Conference Preparations)



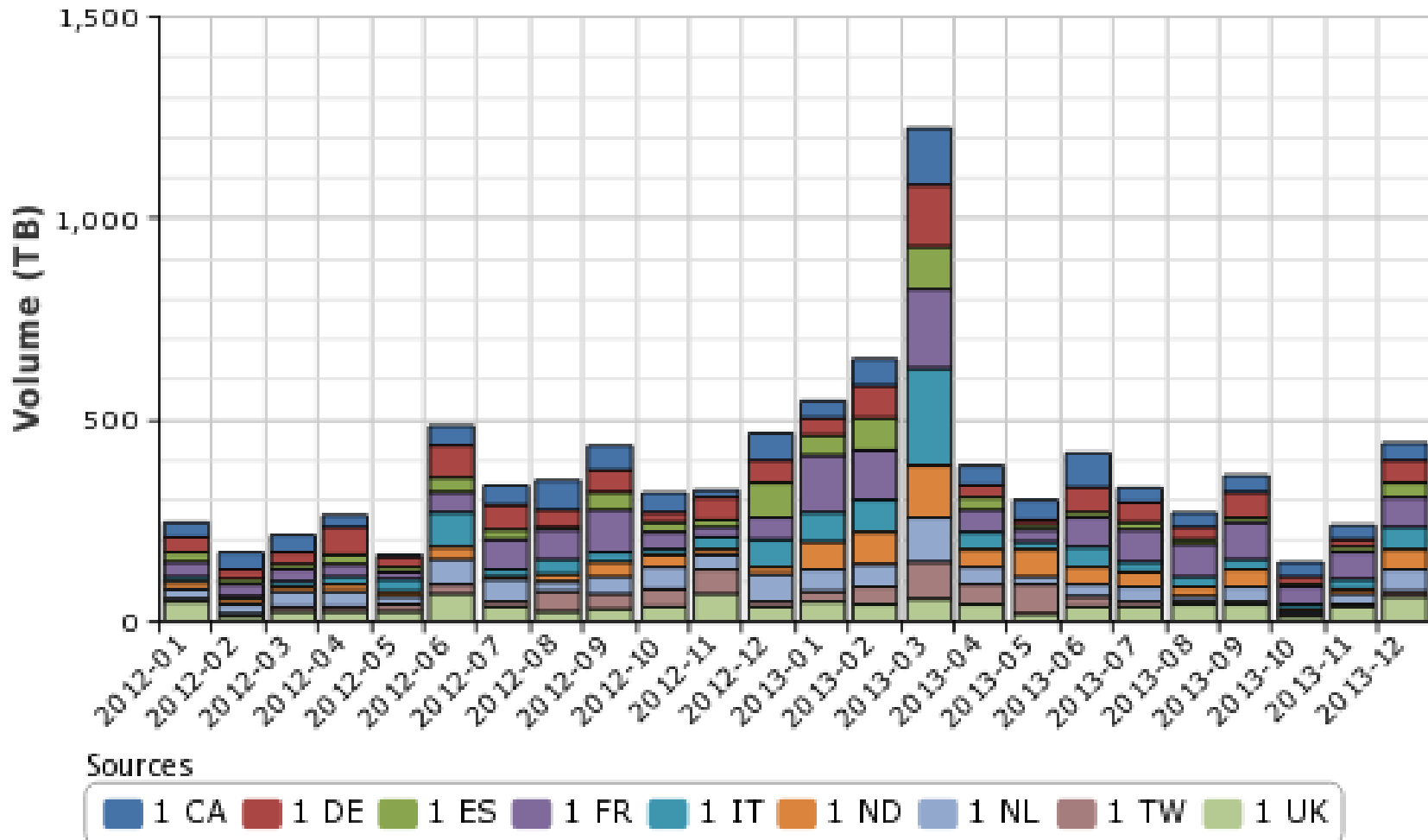
T2s in several regions are getting ~an order of magnitude more Data from BNL than the associated T1s

# From T1s to BNL



## Transfer Volume

2012-01-01 00:00 to 2013-12-31 00:00 UTC

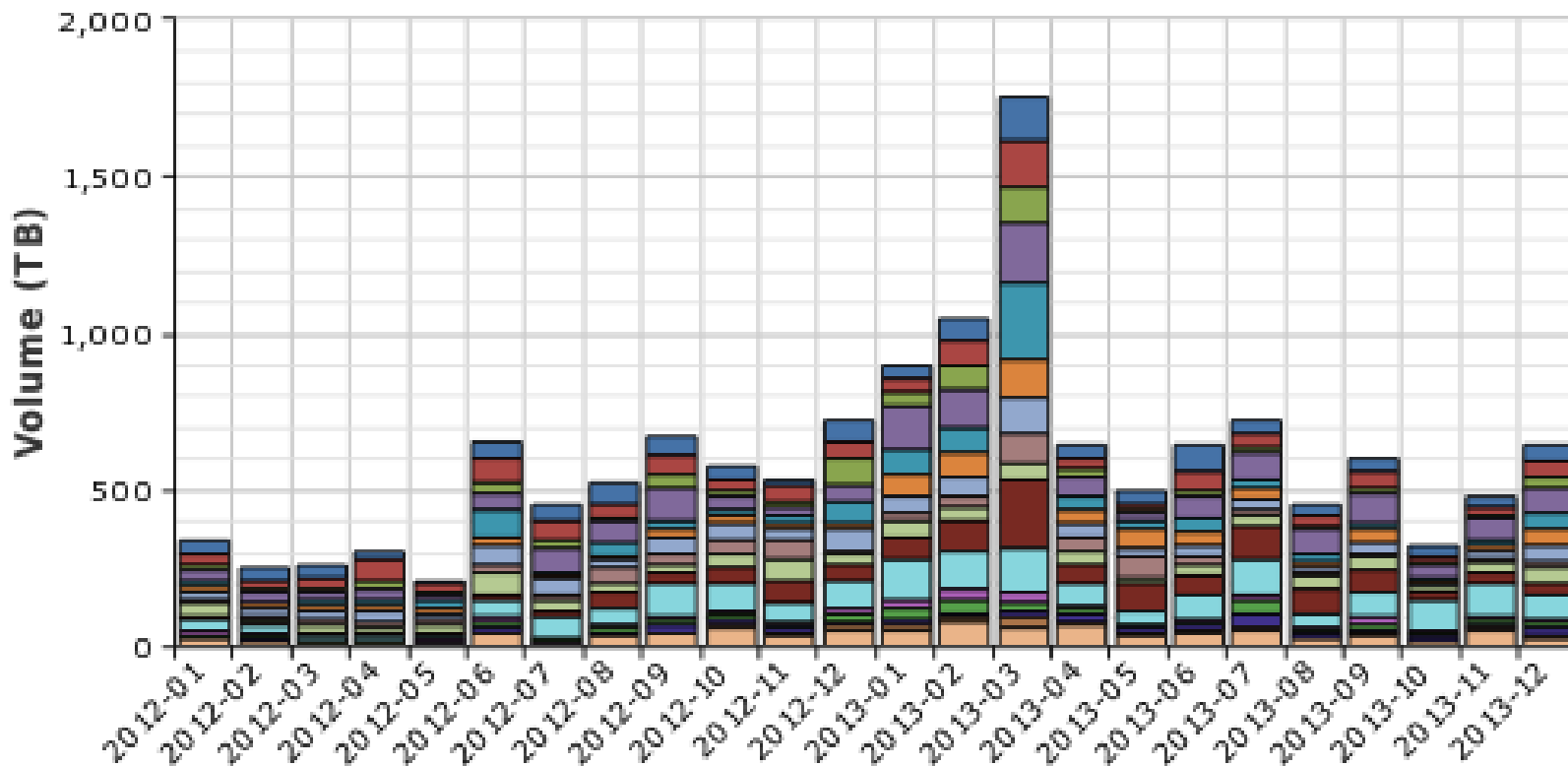


# From T1s and T2s to BNL



## Transfer Volume

2012-01-01 00:00 to 2013-12-31 00:00 UTC



### Sources



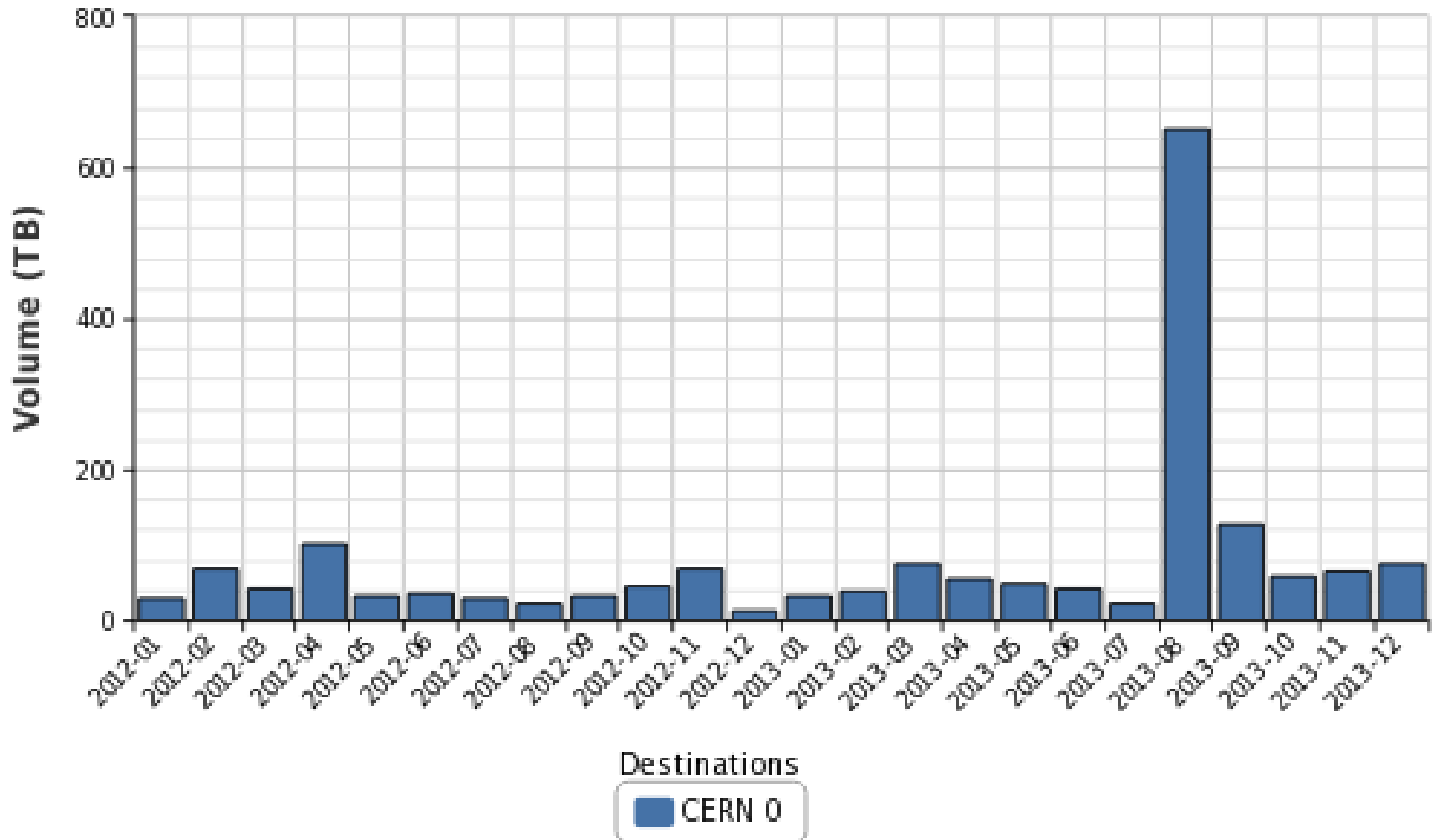


# From BNL to CERN

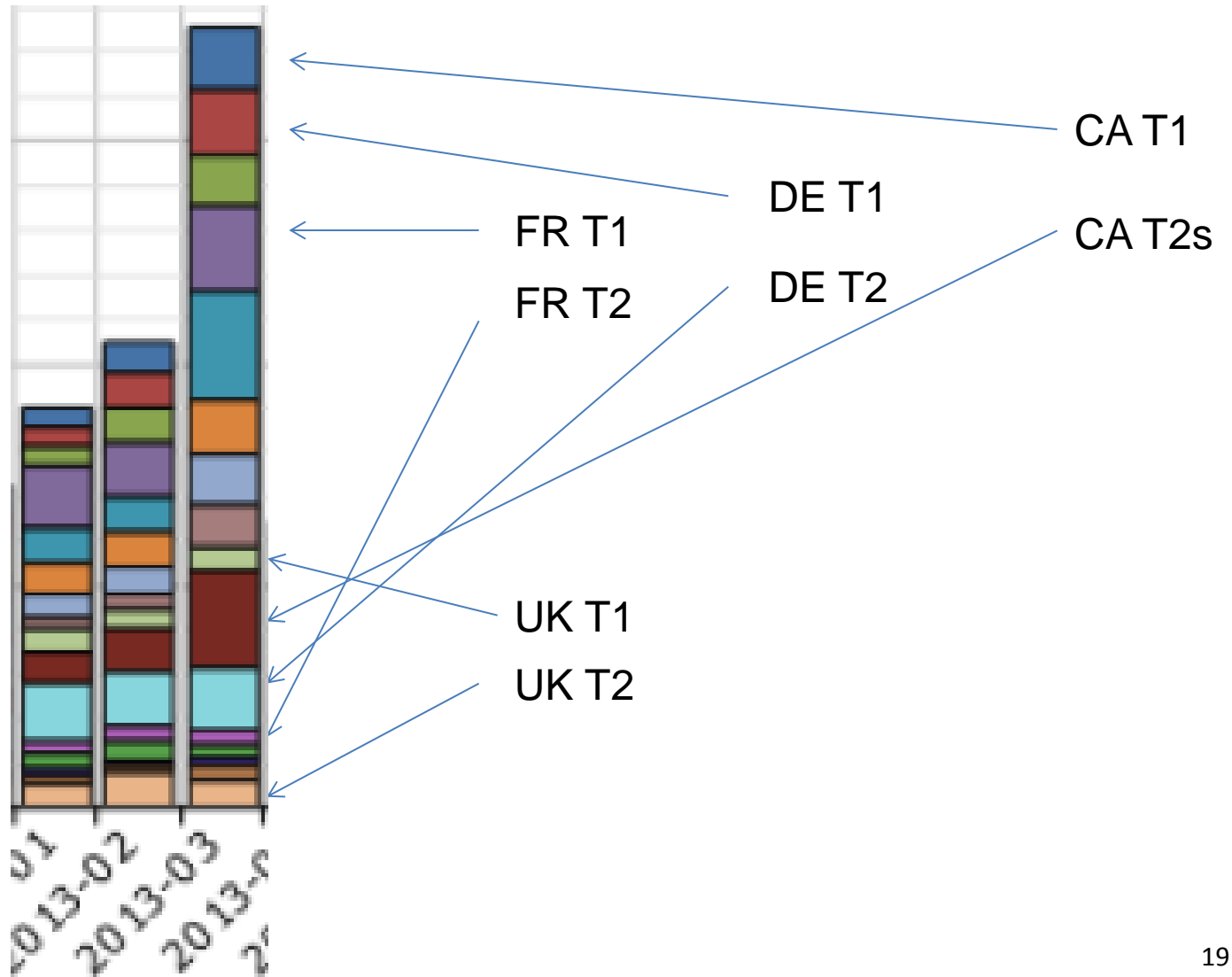


## Transfer Volume

2012-01-01 00:00 to 2013-12-31 00:00 UTC



# T1s vs. T2s to BNL

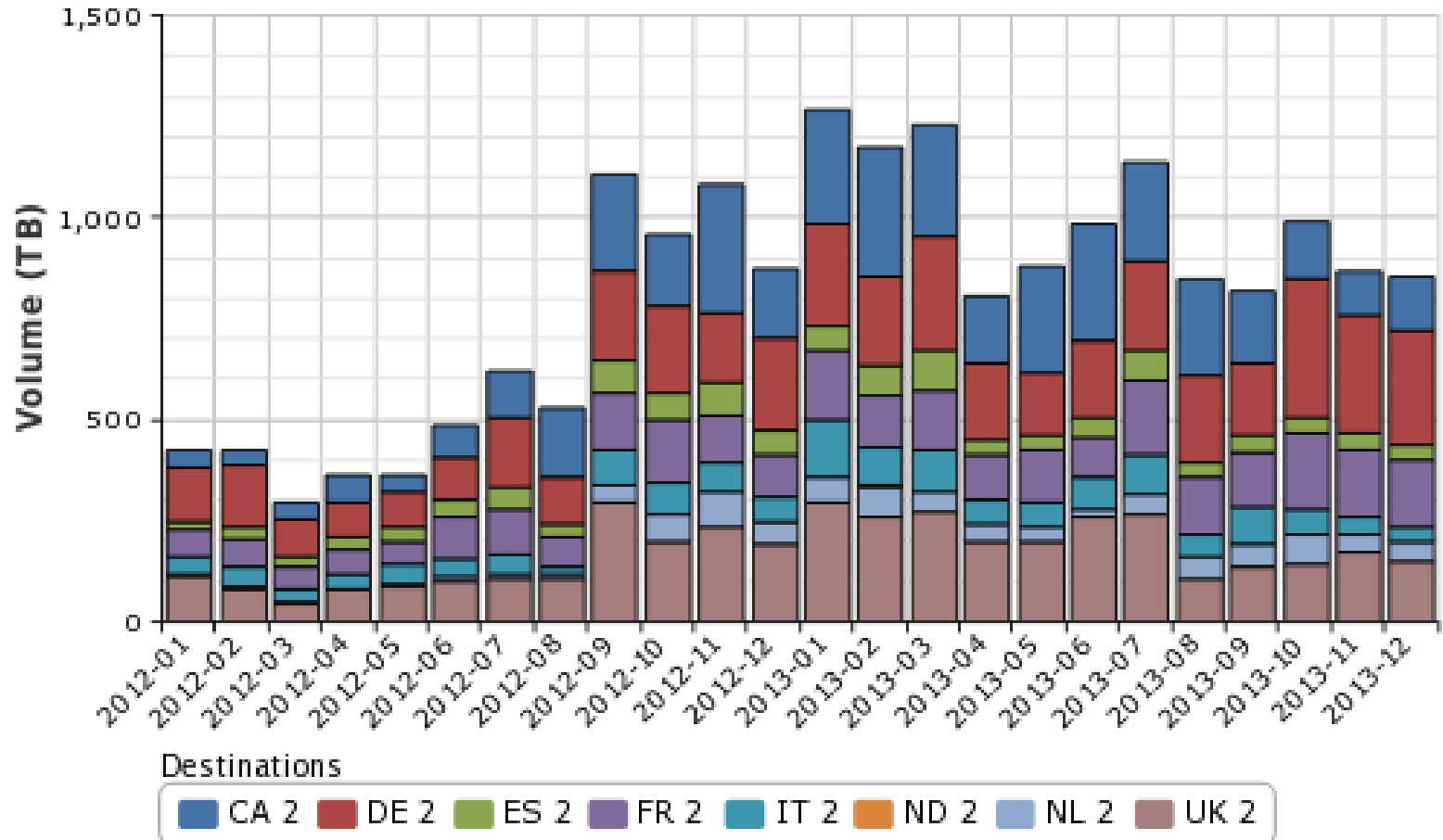


# From BNL to non-US T2s



## Transfer Volume

2012-01-01 00:00 to 2013-12-31 00:00 UTC

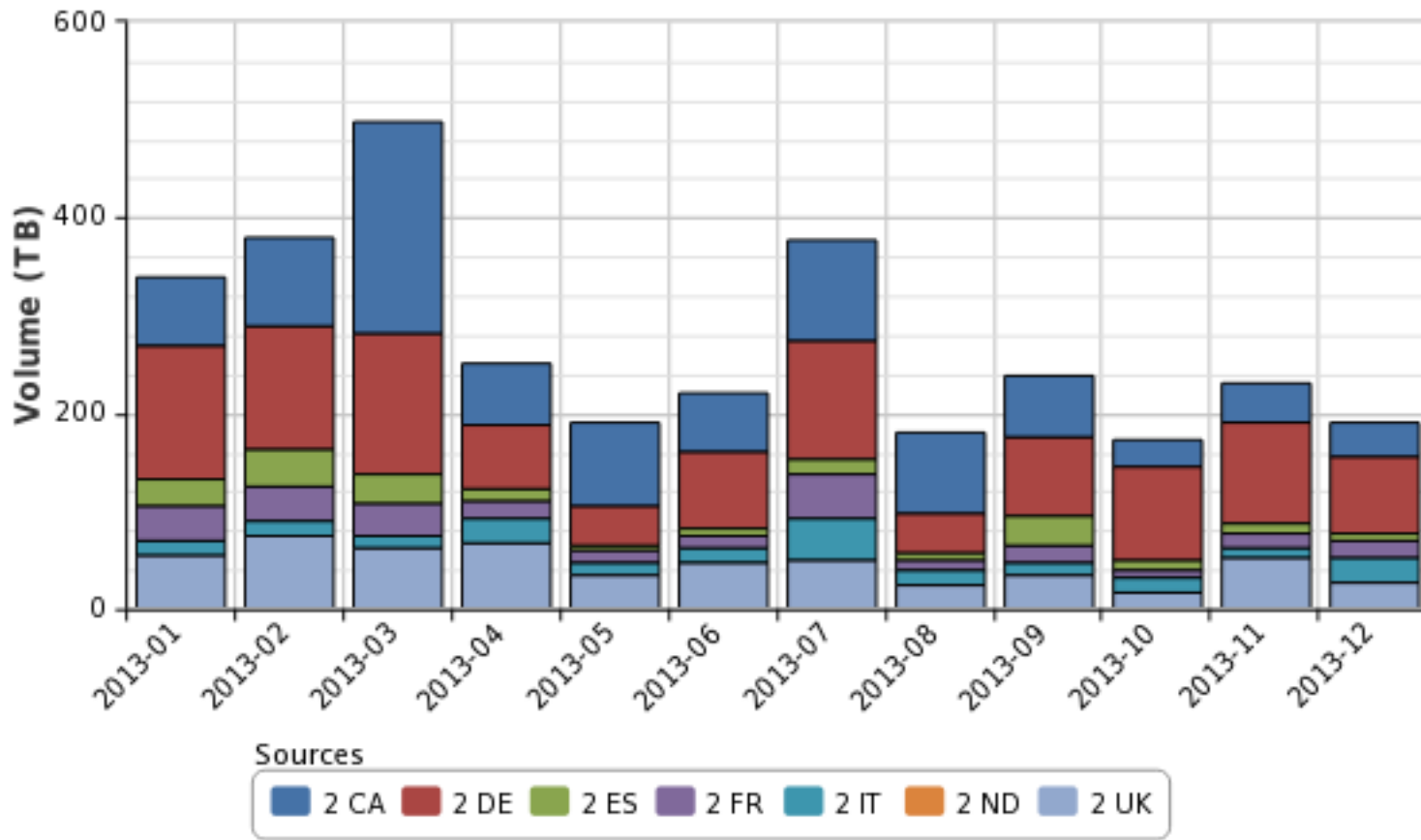


# From non-US T2s to BNL



## Transfer Volume

2013-01-01 00:00 to 2013-12-31 00:00 UTC



# Remote Access – A possible Game-Changer

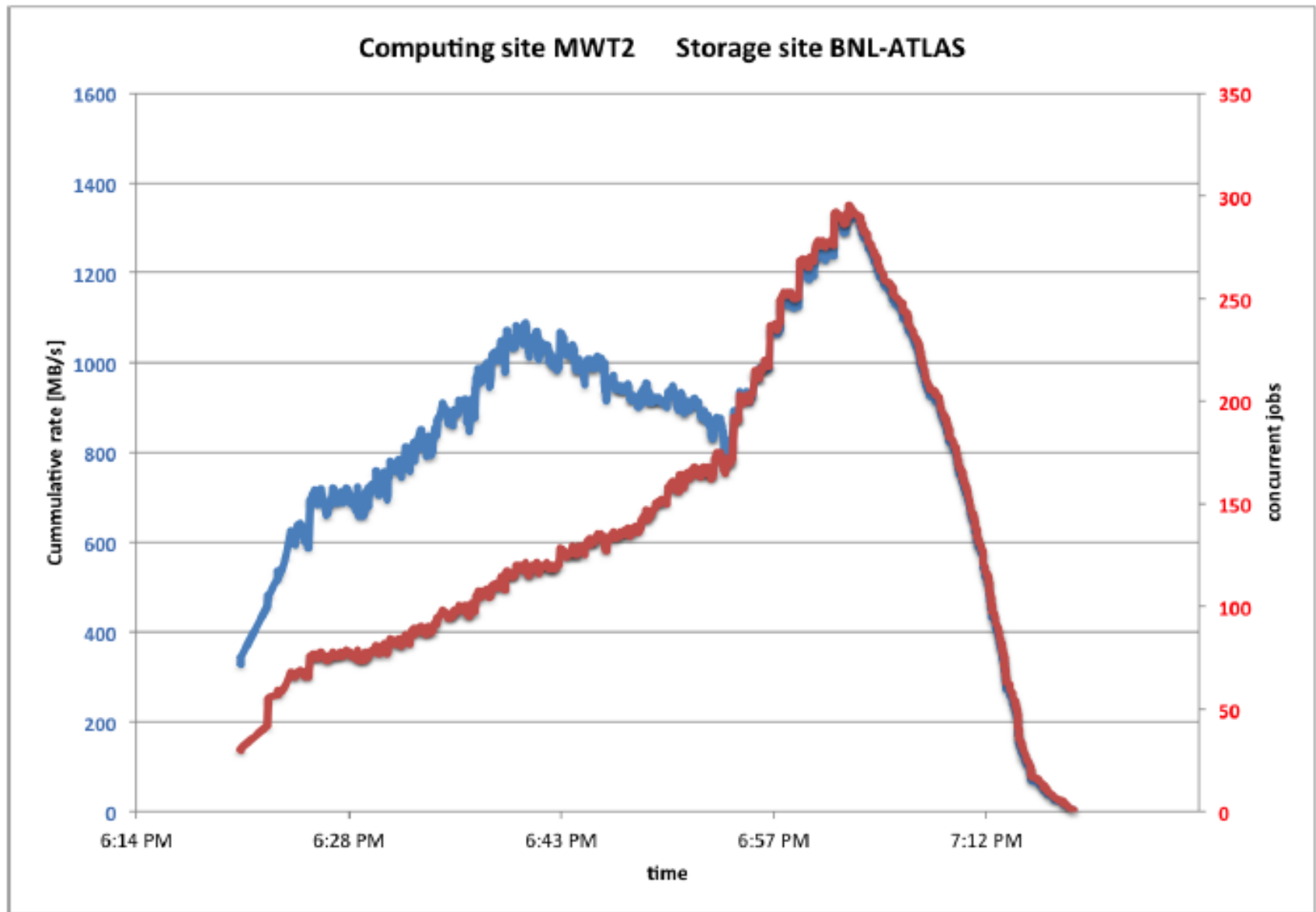
- Data access over the WAN at job runtime
  - Today tightly coupled with Federated Data Access
    - Automatic data discovery with XrootD redirector
    - Unpredictable network/storage bandwidth requirement
      - Possible issues include hotspots, campus network congestion, storage congestion, latency
  - Totally synchronous: time to completion within minutes/seconds (or less)



# Worldwide FAX Deployment



# Jobs Accessing Data Remotely w/ FAX



# Traffic Statistics - Observations

- Traffic Volume To/From BNL
  - From CERN to BNL: ~500 TB/month during ATLAS data taking
  - To BNL: 1,400 TB/month (Peak 1,900 TB/month)
  - From BNL: 1,900 TB/month (Peak 2.200 TB/month)
- T1 Traffic Volume To/From BNL via LHCOPN
  - To BNL: 400 TB/month (Peak 1,200 TB/month)
  - From BNL: 400 TB/month (Peak 600 TB/month)
  - BNL to T2 Volume during conference preparation order of magnitude higher than BNL to T1 Volume
- Traffic Volume From/To BNL via LHCONE and GIP
  - To BNL from non-US T2s: 200 TB/month (Peak 500 TB/month)
  - From BNL to non-US T2s: 1000 TB/month (~400MB/s)
  - Traffic clearly driven by analysis activities

# Trends

- Looking at 2012 and 2013 statistics data, from BNL's perspective Traffic statistics suggests BNL/T1 to T2 Traffic is dominating
  - Traffic to non-US T2s doubled to 1 PB/month in September 2012
    - ~Constant since then, potential to grow w/ new data
    - Largely driven by analysis
  - BNL Traffic Volume from/to T1s via LHCOPN staying fairly constant for 2 years at ~500 TB/month
    - Largely independent from data taking

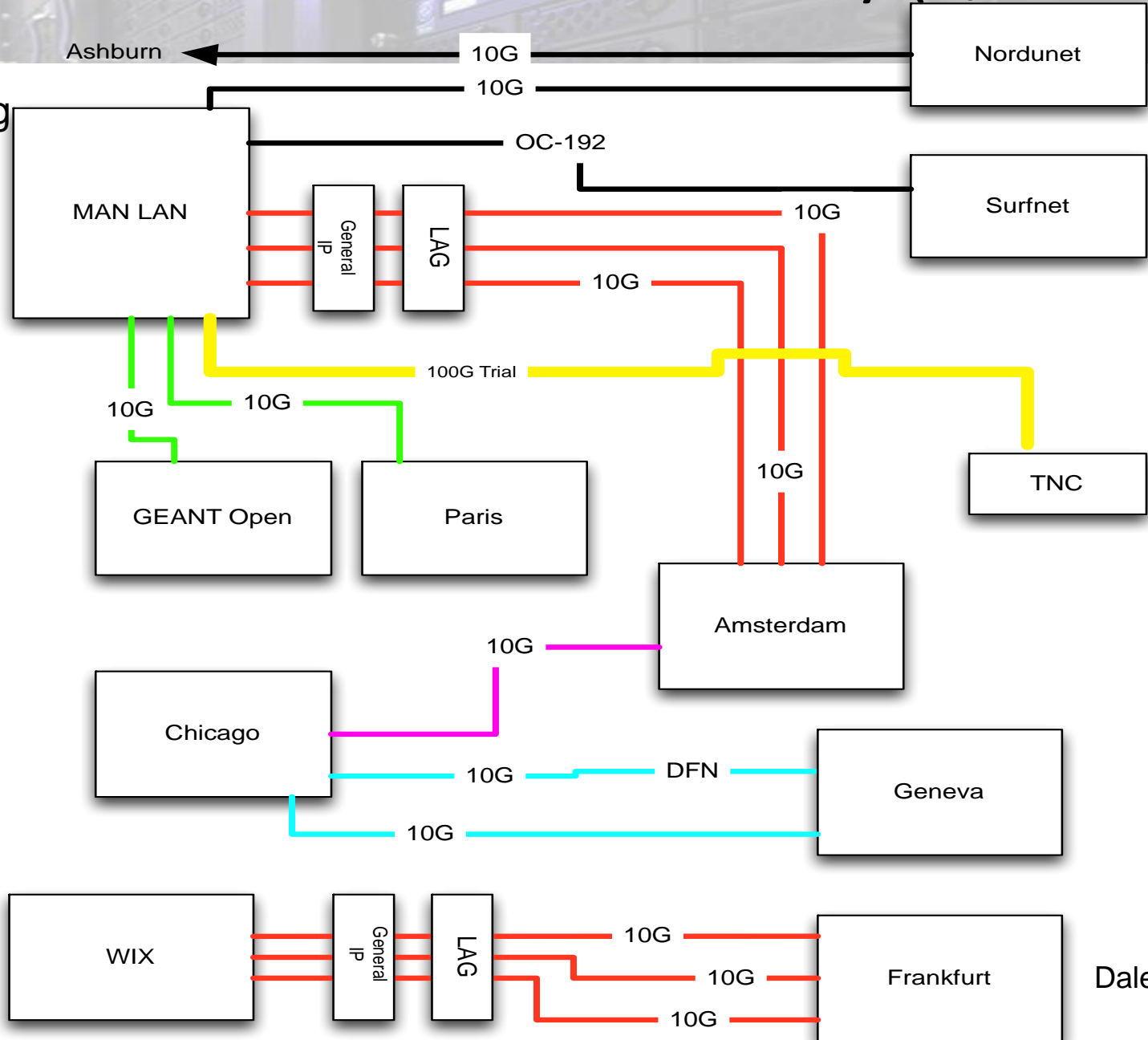
# Conclusion

- Rather than maintaining distinct networks the LHC Community should aim at unifying its network infrastructure
  - In ATLAS Tiers are becoming more and more meaningless
  - We are thinking about optimizing usage of CPU and Disk and we also need to think about optimizing usage of Network resources
  - Load-balanced links
  - Traffic prioritization, if necessary
- Traffic aggregation on fewer Links



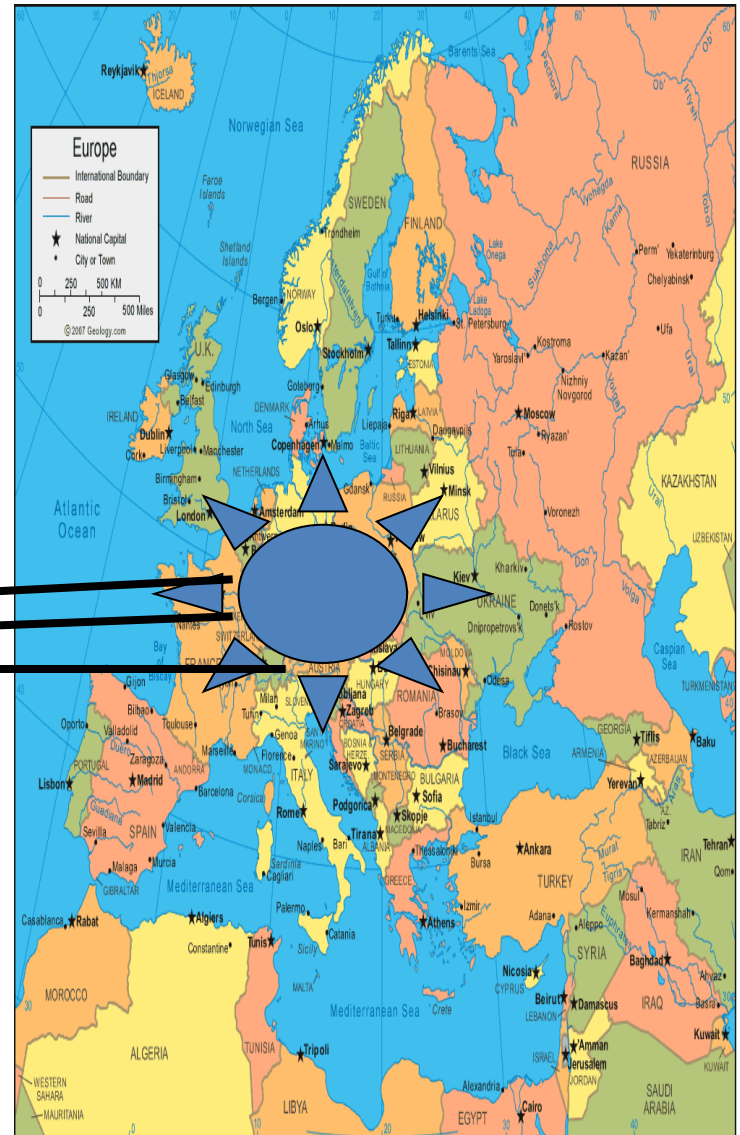
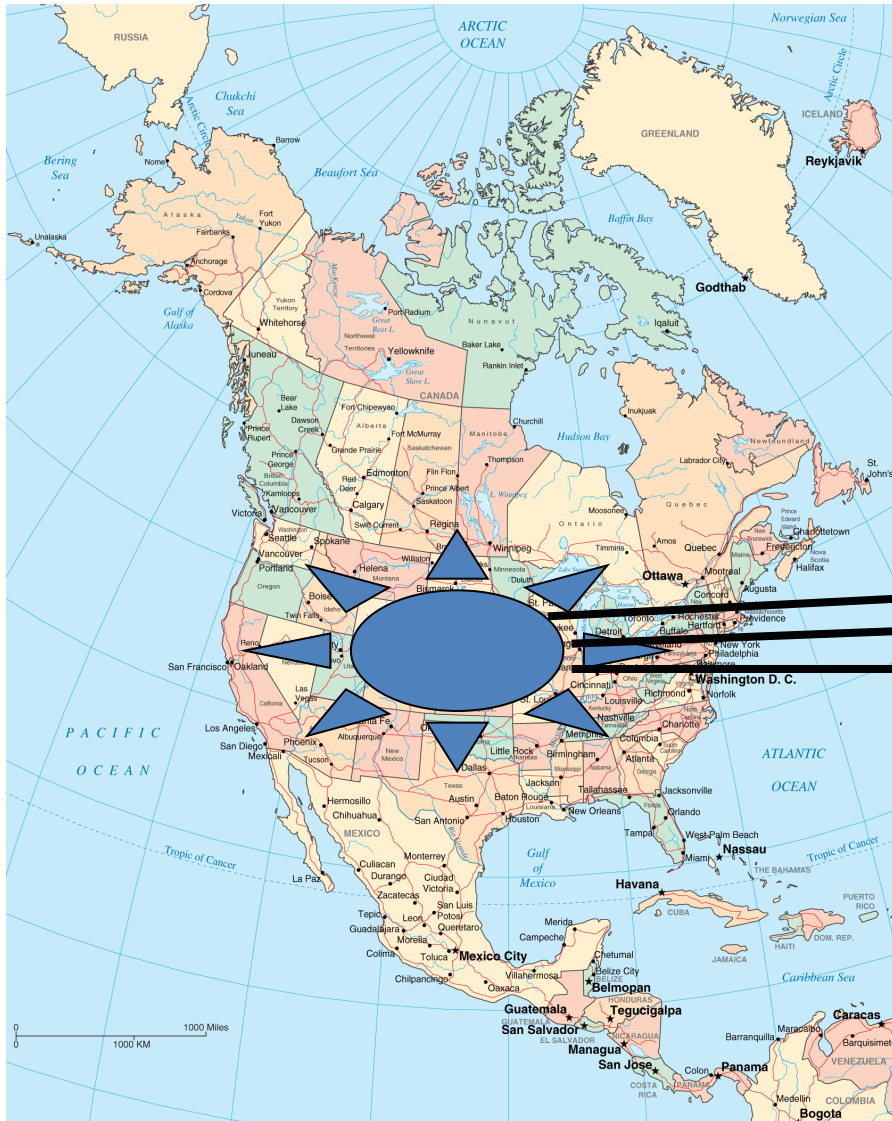
# R&E Transatlantic Connectivity (5/2013)

Not showing  
LHCOPN  
Circuits



Dale Finkelson

# Looking at it generically ...



# Concerns

- With the T1 and the T2s in the US upgrading now to 100G the global infrastructure needs to follow
- LHCONE Evolution
  - Currently the LHCONE side-by-side w/ general R&E infrastructure
  - Traffic segregated, but what is actually the benefit?
    - Is anyone looking at the flows for optimization, steering?
    - Is it really true that our ‘Elephant’ flows interfere w/ traffic from other science communities?
- P2P/Dynamic Circuit Infrastructure
  - Are the interface definitions and components mature enough to serve applications?
  - What would happen if the experiments started to use dynamic circuits extensively, in multi-domain environment?
  - Would there be sufficient infrastructure in the system?

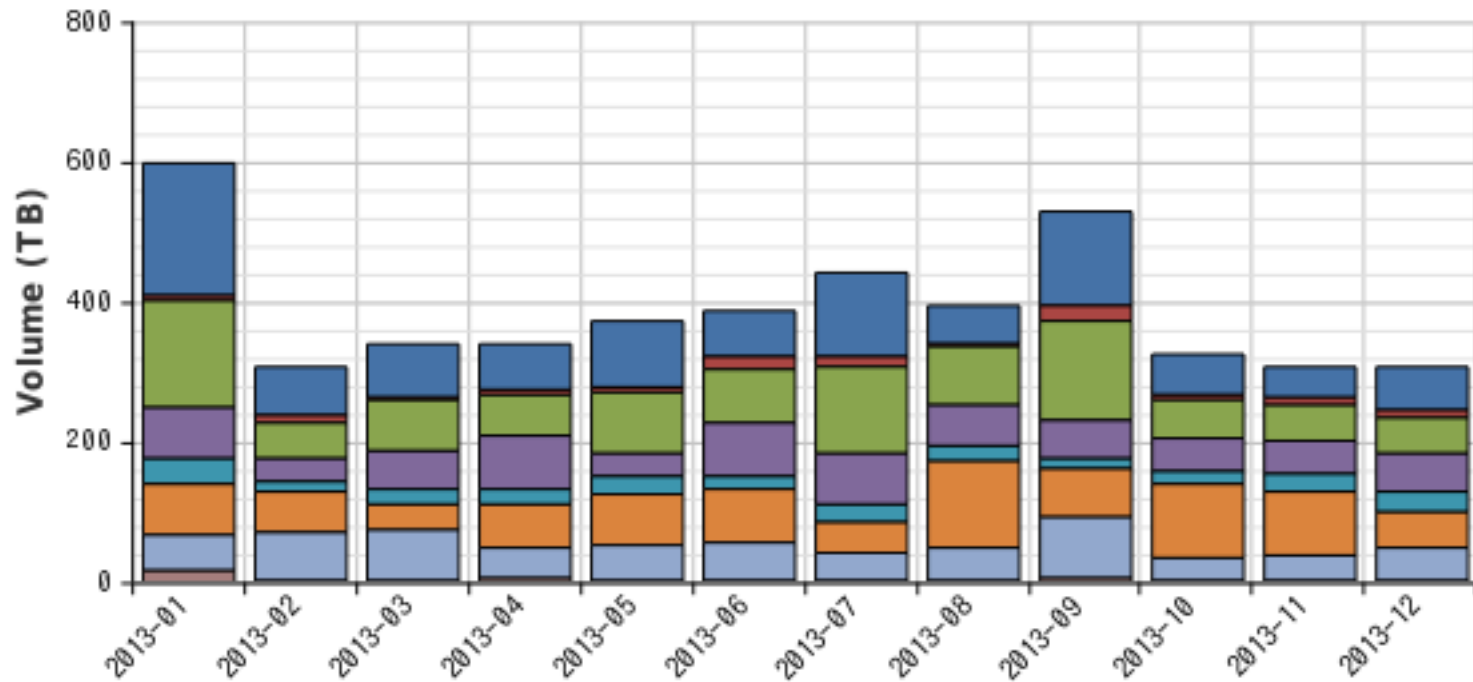
# Backup Material

# From US T2s to BNL



## Transfer Volume

2013-01-01 00:00 to 2013-12-31 00:00 UTC



### Sources



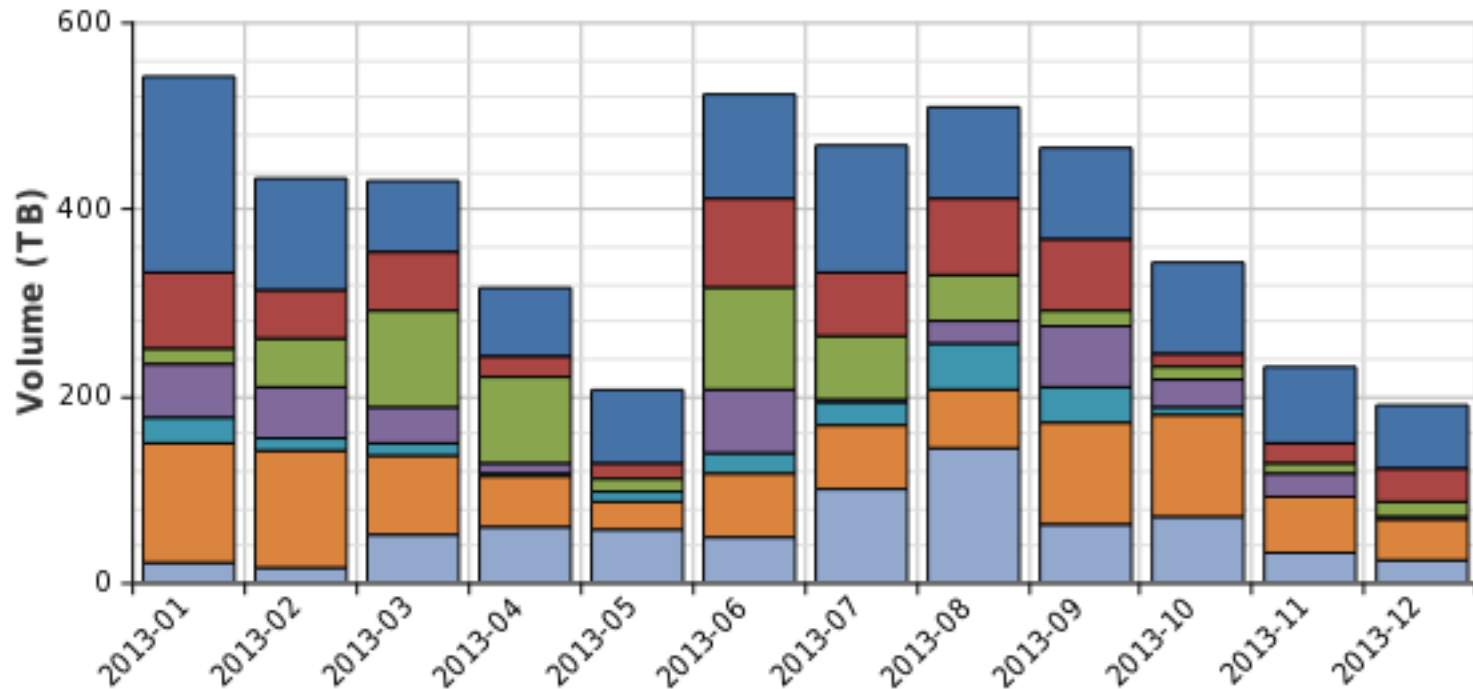


# From BNL to US T2s



## Transfer Volume

2013-01-01 00:00 to 2013-12-31 00:00 UTC



Destinations

