

Thoughts on Future LHCOPN

Some ideas

Artur Barczyk, Vancouver, 31/08/09

Problem definition

- The LHCOPN model proved so far to be very successful!
 - Simple design
 - Addresses primarily raw data distribution
 - Provides sufficient resiliency (c.f. STEP09)
 - Security
 - Solid operational model
- Covers Tier0-Tier1 and to some extent Tier1-Tier1 requirements
 - T1-T1 is using
 - CBF, i.e. not fully interconnected
 - Scavenger on T0-T1 links

Problem definition

- Current LHCOPN adequately addresses problem of raw data distribution, but (for good reason) leaves out the aspect of data distribution down to analysis level
- Tier1-Tier2 data transfers
 - Fermilab estimates:
 - ~11 Gbps Tier2 related bandwidth
 - ~40-45% to non-US Tier2s
 - Assume similar numbers for other Tier1s?
- Tier2-Tier2 estimates less known
 - Varies by experiment, different data distribution models
- Tier2 connectivity via GPN
 - Security, access through firewall
 - Potential congestion effects

Problem definition

- Estimate Tier2 transfer rates (CMS)
- Dataset size: 30-50 TB x N
- Typical Tier2 storage: 200 TB
 - “bigger” ones (will) have 400 TB and more
- At 1 Gbps (sustained, no transfer errors):
 - 2.8 days to transfer a single dataset
- OR: to transfer a dataset within a day, need >3Gbps sustained rate
- Each dataset will be re-processed and re-distributed several times per year
 - Possibly 6 times/year, at least in the beginning
- Tier1s will need to cope with several such transfers in parallel
- For efficient operation of a Tier2, 1 Gbps bandwidth will be just sufficient
 - Below 1G connectivity, a Tier2 will be hardly functional
- Safe to assume an increase in access bandwidth of Tier2s (min 1G, 10G norm)

Dedicated Tier2 connections?

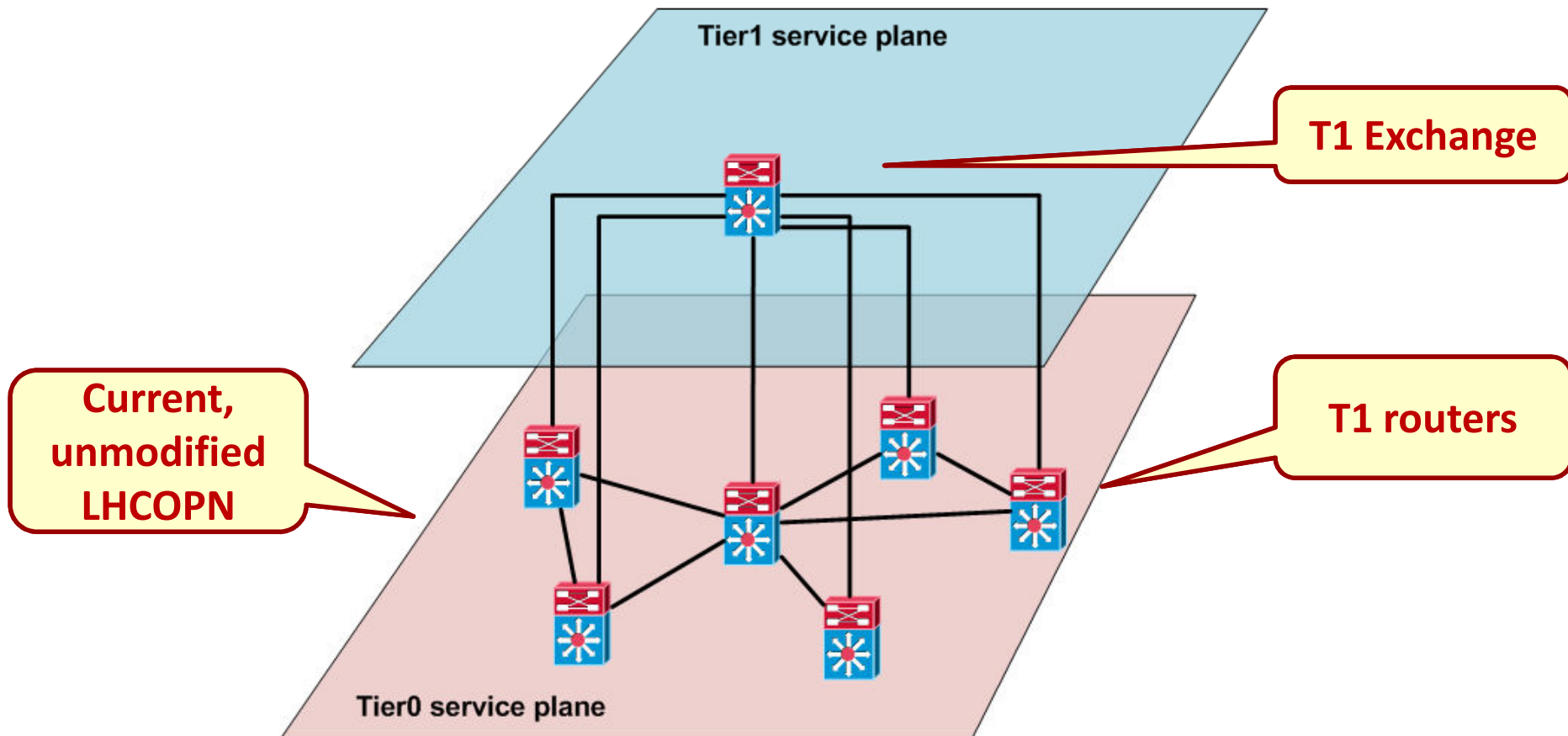
- US Atlas has mandated all US Tier2s to have a dedicated link to their Tier1 (BNL)
- Static links being deployed
- What about Europe and other places?
- Often a cost factor, depending on region

What's ahead?

- LHCOPN “core” (Tier0, Tier1s):
 - Extend capacity for Tier1-Tier1?
 - Extended deployment of Cross-Border Fibre?
 - Possibly good strategy in Europe
 - Raises capacity
 - Raises path diversity, i.e. resilience
 - But:
 - poor scalability
 - Raises complexity
 - Dedicated Tier1 Exchange Point (T1XP)?
 - One? Where?
- Extended LHCOPN (include Tier2s)
 - Guarantee performance necessary for LHC data operation
 - Provide intrinsic security bypassing the GPN

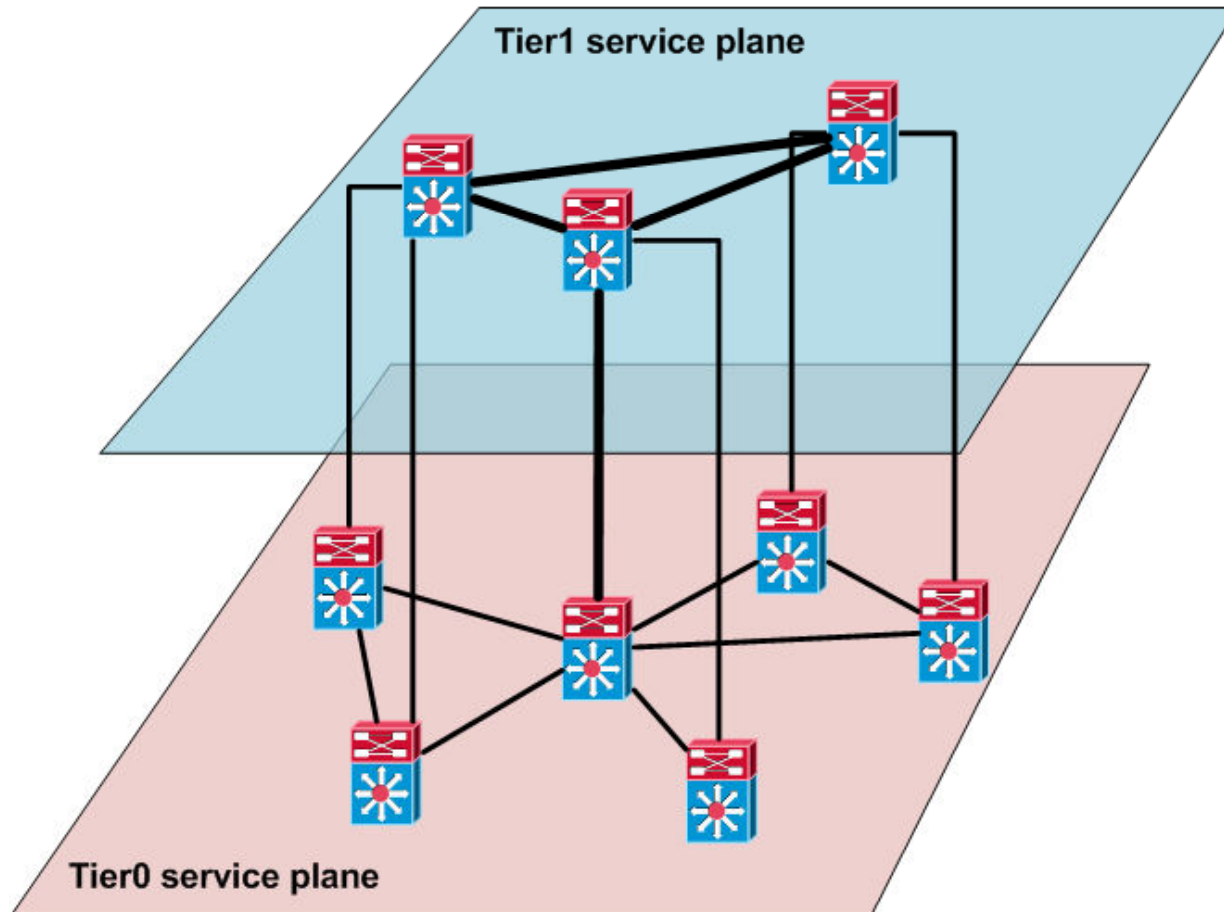
Proposal 1 (Tier1-Tier1)

- Add T1XP router, (centrally) at CERN only



Proposal 1' (Tier1-Tier1)

- Add T1XP routers: distributed T1XP, interconnected with dedicated lightpaths, e.g. CERN, Netherlight, Starlight

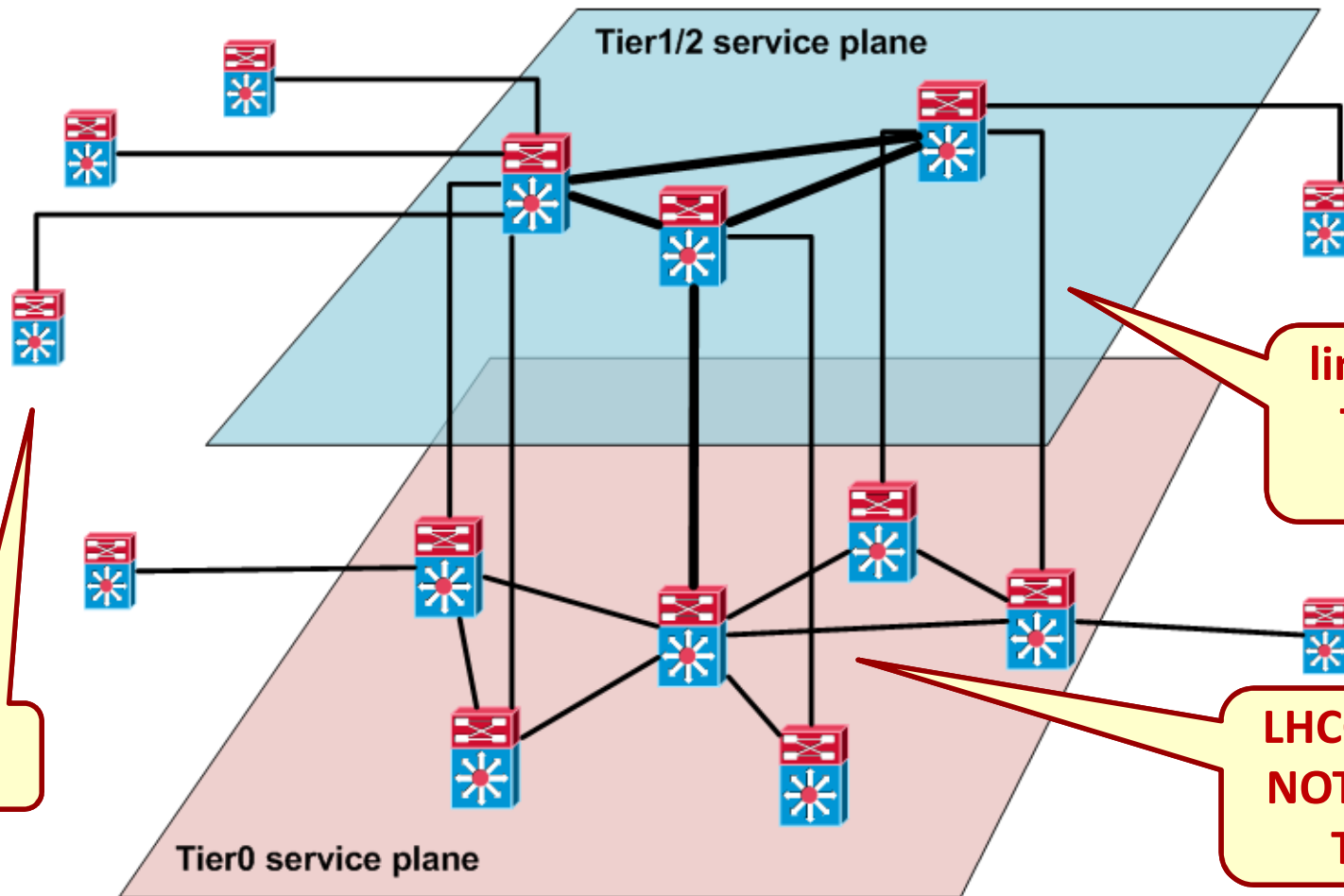


Proposal 2 (Tier1-Tier2)

- ***LHCOPNcore***: no modifications to infrastructure
 - apart from possible additions, and/or T1XP
- ***LHCOPNedge***: extend OPN to reach Tier2 sites
 - From Tier1s
 - From T1XP (T1T2XP)
 - At least to start with, “nailed down” circuits, possibly later with dynamic lightpaths

Proposal 2 (Tier1-Tier2)

- (building on 1')



T2s

links used for
Tier1s and
Tier2s

LHCOPN links
NOT used for
Tier2s!

Tier0 service plane

Tier1/2 service plane

Dynamic Layer 2 connections

- Admittedly in pilot/development stage
 - Internet2: DCN, pilot
 - ESnet: SDN (OSCARS), in production
 - GEANT: AutoBAHN, in development
 - SURFnet: DRAC, in production
 - EU project: PHOSPHORUS, in development
 - + other projects (Japan, Korea, ...)
- Reservation of end-to-end fixed bandwidth Layer2 connections
 - end-system to end-system
 - Edge-to-edge
- Uses “middle-ware” (user/application agents)
 - Set up network path
 - Configure end-system (optional)

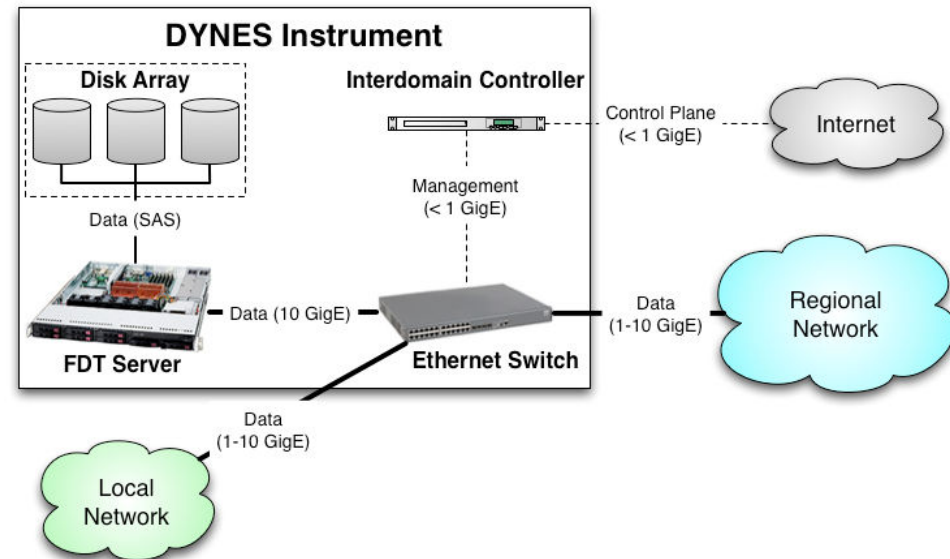
“Don’t touch my border router!”

- Middleware-driven configuration of network devices is needed for end-to-end provisioning
- This can or cannot include the existing border router and/or LAN devices
 - Fermilab, University of Nebraska, Caltech use LambdaStation to automatically switch data paths within the campus
 - TeraPaths in Atlas
- E.g. Internet2 & Caltech & UMICH & Vanderbilt University submitted a proposal to NSF to fund infrastructure at Tier2 and 3 sites in US
- Includes a small but powerful “border switch”, to be installed in addition the to the existing GNP devices

DYNES proposal (US)

- If successful, each participating site (40 identified) will install a small switch with a connection to the regional network, dedicated to DCN
- A server will be installed running the Inter-Domain Controller software as well as PerfSONAR monitoring tools

Tier 2/3 Hardware Configuration



Final Thoughts

- Current LHCOPN (core) does what it was designed for - shouldn't be modified (extend – not redesign)
- We should care about the “broader picture”, data distribution and movement does not stop at Tier1s – important for LHC experiments' operation
- Tier1-Tier1 and Tier1-Tier2 data movements will need to be addressed, probably sooner rather than later
- Adding edge support to the existing network will invite Tier2s to use the provided infrastructure, instead of looking later for quick-fixes