

LHCOPN-LHCONE meeting #42

Umeå (SE)

summary report

4th - 5th June 2019 – v1.1

edoardo.martelli@cern.ch



Venue

Hosted by Umeå University in Umeå, Sweden



Participants

- 22 Participants
- 16 Institutes
- 2 Collaborations
- 5 Research Networks

Day 1: updates and operations

LHCOPN update

- **LHCOPN traffic volume:** increased till the end of Run2
- **CH-CERN:** completed replacement of legacy routers with new Juniper
- **IT-INFN-CNAF:** second 100G link deployed as backup link. Will be turned in load balancing configuration with the primary because already hitting the ceiling
- **DE-KIT:** plan to deploy 100G backup link
- **NL-T1:** deployed 100G primary link
- **ES-PIC** and **UK-RAL:** will deploy 100G link for Run3
- **NDGF** will upgrade to 2x100G as soon as network hardware available in Geneva (currently 4x10G)

Slides: <https://indico.cern.ch/event/772031/contributions/3357233/attachments/1855272/3047009/LHCOPNE-20190604-Umea-LHCOPN-update.pdf>

CERN Tier0 update

- Another record luminosity year for LHC
- IT infrastructure well coped with data taking and analyses
- Long Shutdown 2 (LS2) just started. Run3 will start in 2021. High Luminosity LHC will be ready in 2026 and will bring 5x luminosity
- HEP community has been discussing the future of accelerators. CERN has proposed the FCC (circular) and the CLIC (linear) colliders
- IT data-centre network being upgraded. Wigner extension being dismantled and servers being moved to temporary containers at LHCb premises. Preveessin Computer Centre (PCC) waiting the green light of the Finance Committee

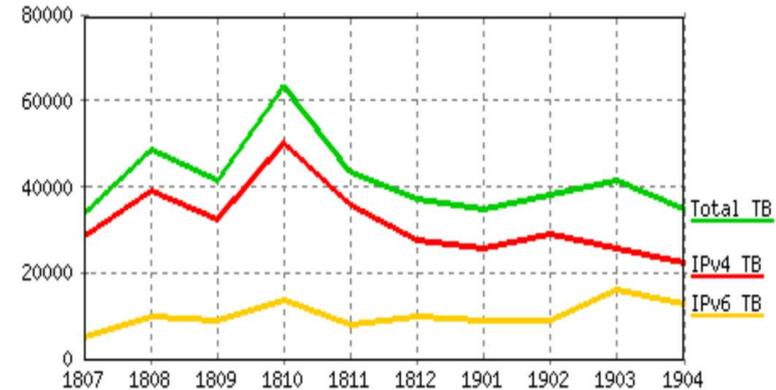
Slides: <https://indico.cern.ch/event/772031/contributions/3357239/attachments/1855273/3047011/LHCOPNE-20190604-Umea-CERN-update.pdf>

IPv6 adoption in LHCOPN and LHCONE



LHCOPN: all Tier1s connected with IPv6. IPv6 transfers happening among all site except 2

WLCG: 40% of the sites didn't meet the deadline of end-2018 to deploy dual-stack storage and perfSONAR. IPv6 traffic is almost 40% of the total.



LHCOPN+LHCONE traffic seen on CERN border routers

Looking for areas where to deploy IPv6 only services

Graph: <https://twiki.cern.ch/twiki/bin/view/LHCOPN/LHCOPNEv4v6Traffic>

Slides: <https://indico.cern.ch/event/772031/contributions/3372804/attachments/1855550/3047553/ipv6-at-LHCOPNONE-update.pdf>

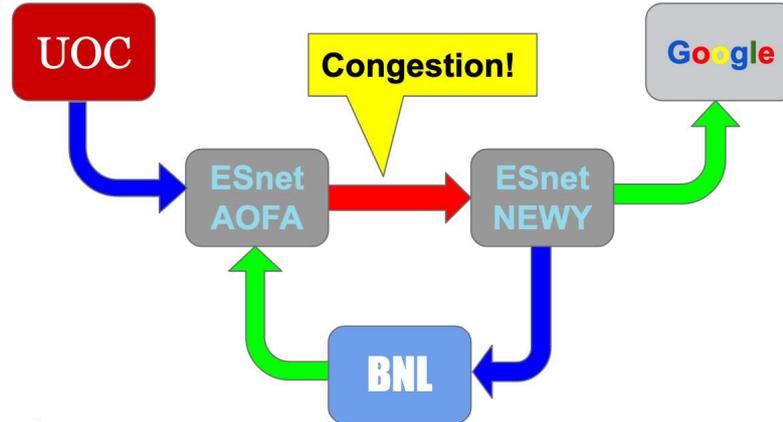
LHCONE L3VPN status - update

- Huge increase in IPv6 traffic
- Remarkable increase in traffic for sites that upgraded capacity
- Still some work to do for Intra-Asia, Asia to Europe and Asia to North America routes stability
- The availability of new inter-continental routes will likely help
- Awaiting for new joiners (Chile expected soon)

Slides https://indico.cern.ch/event/772031/contributions/3416857/attachments/1855601/3047668/2019-6-4_ECapone_LHCONE_L3VPN.pptx

LHCONE operations update

- ESnet LHCONE traffic has increased 80% in the past year. The ESnet LHCONE traffic counts for 54% of ESnet total volume.
- Presented analyses of an event that led to an ESnet core link saturation



Slides: https://indico.cern.ch/event/772031/contributions/3416864/attachments/1855526/3047498/Umea_LHCONE-OperationsUpdate.pptx.pdf

STFC and RAL Tier1 - update

Connection to LHCONE

- Peering and IP address allocation expected from JISC in coming days
- Start Tier-1 joining LHCONE (perfSONAR) later in August

Tier-1 is still seeing unacceptably high outbound packet loss over IPV6.
Investigation is on-going

Upgrade of LHCOPN link:

- Planning to move to a single 100Gb/s link for Run3
- Resilience will come from LHCONE

Slides: <https://indico.cern.ch/event/772031/contributions/3449119/attachments/1855535/3047636/STFC-LHCONE-STATUS-IPC-2019-06-04.pdf>

perfSONAR and monitoring update

Four projects around the core OSG Network Area:

1. SAND (NSF) project for analytics
2. HEPiX NFV WG
3. perfSONAR project
4. WLCG Network Throughput WG

PerSONAR news:

- 4.1.6 is the latest version. 4.2.0 will add preemptive scheduling & gridftp testing
- Campaign to update all perfSONARs to CC7 and 4.1 still on-going
- All meshes now test throughput and traces via both IPv4 and IPv6
- OCRE Cloud testing (<https://github.com/cern-it-efp/OCRE-Testsuite/>)

New LHCONE mesh proposal:

- Our proposal is to replace the current mesh with one that does uni-directional tests from set of selected sites to set of R&E endpoints

Testing of 100G perfSONAR servers on going. Reached ~80Gbps between CERN and NLT1 via LHCOPN

Slide: https://indico.cern.ch/event/772031/contributions/3360614/attachments/1855592/3047650/LHCOPN_LHCONE_perfSONAR_Update_2019spring.pdf



BelleII update



- Phase 3 beam run has started on 11th March 2019
- Current resource usage:
 - 9PB available on the main Storage Elements (7 PB used)
 - 62 Sites used in the last 12 Months
 - Up to 27k Running jobs
- Data Challenge 2019 have shown improvement thanks to new SINET 100G link to Amsterdam and new 40G IN2P3 LHCOPN link. GEANT DTNs were used to compare results

Slides: <https://indico.cern.ch/event/772031/contributions/3360616/attachments/1855510/3047812/BELLEII-UPDATE-LHCONE-2019-04-June-v7.0.pdf>

GEANT network upgrade

Challenges:

- Exponential traffic increase but flat budget
- Need for programability and reduce vendor lock-in

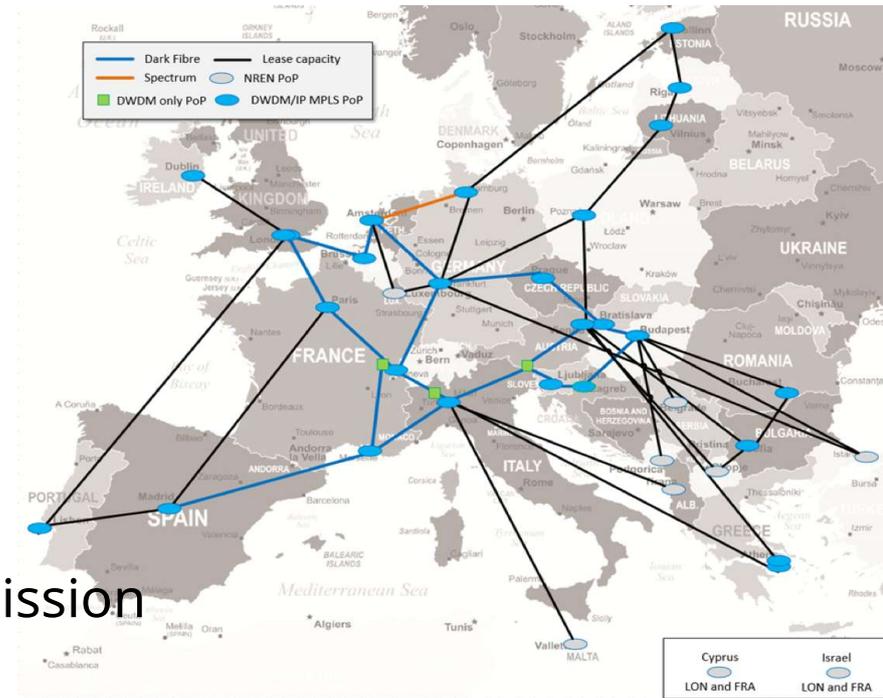
GN4-3N project:

- Funding cycle 2019-2022
- Budget at least 16M€, out of total 4-year GÉANT project budget of 128M€

Aiming to more dark fibres, cheaper transmission boxes, cheaper routers for small PoPs.

The upgrade will be completed by 2021.

Slides: https://indico.cern.ch/event/772031/contributions/3416871/attachments/1855784/3048048/2019-6-4_ECapone_GEANT_updates.pdf



ESnet6 - update

Working on ESnet6, the next-generation ESnet network

ESnet6 project, next steps:

- Finish final design review
- Finish fibre and colocation procurement by the summer 2019
- Complete optical vendor selection by October 2019
- ESnet6 optical layer will be upgraded in 2020 (during LS2)
- ESnet6 packet layer will be upgraded in 2021 (during Run3)



Slides https://indico.cern.ch/event/772031/contributions/3416872/attachments/1855529/3047501/ESnet6_Overview.pdf

TRANSPAC and NSF projects update

International Networks at Indiana University: provisions and operates international circuits to support US science collaborations

Current projects:

- TransPAC4: Pragmatic Application-Driven International Networking
- NEAAR: Networks for European, American, and African Research
- NetSage: An Open, Privacy-Aware, Network Measurement, Analysis, and Visualization Service
- EPOC: Engagement and Performance Operations Center

Slides: https://indico.cern.ch/event/772031/contributions/3416875/attachments/1855781/3048036/LHCONE_INIU_overview_.pdf

Chasing unwanted packets in LHCONE

- DE-KIT and ESnet have run an investigation on unwanted packets coming from LHCONE connections
- LHCONE ingress filtering/control has improved dramatically since measurement began in Q1 2018
- Growth in Asia has likely contributed to a certain number of exceptions
- Routing table inconsistency may also be a source of false positives
- Exchanges and transit providers are being added to LHCONE, the community needs to provide better guidance to these providers

Slides https://indico.cern.ch/event/772031/contributions/3360612/attachments/1855532/3047503/LHCONE_Edge_Filtering_Policy_and_Practice_Umea_1.pdf

LHCONE VRF hacking

Exploiting the SURFnet peering status with NORDUnet, a SURFnet engineer demonstrated the possibility to inject a network prefix from the Internet into the LHCONE routing table, creating a potential security issue

LHCONE Network Providers will fix this weakness

Slides:

https://indico.cern.ch/event/772031/contributions/3452643/attachments/1856017/3048497/20190604-LHCONE-hacking_myself_into_LHCONE.pptx

WLCG workshop and DOMA activities

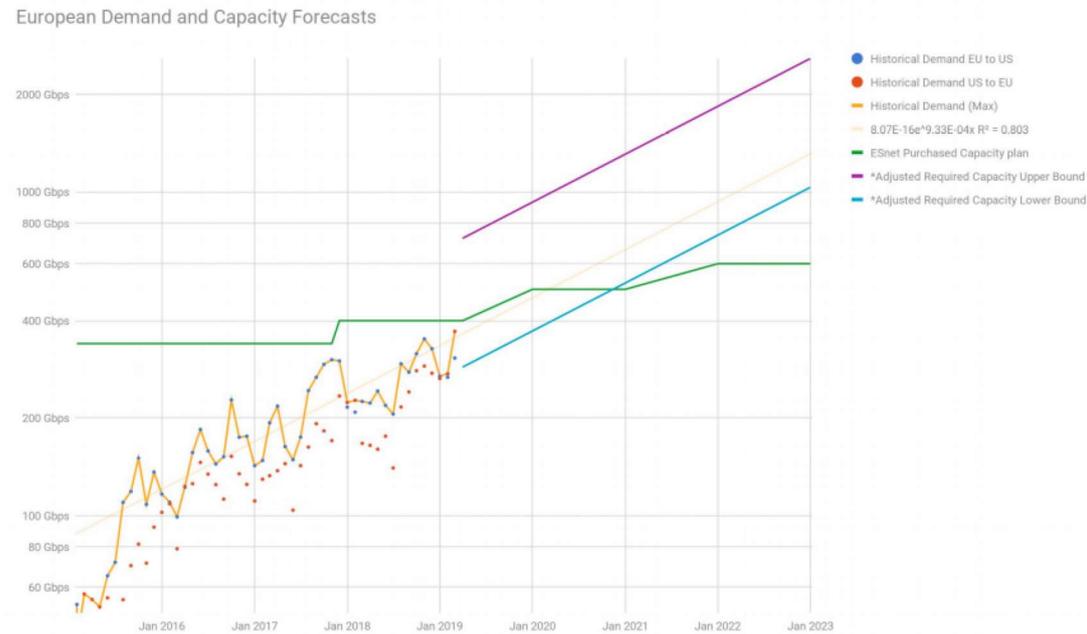
- Can no longer assume increases in performance/capacity
- ATLAS and CMS Run 4 requirements are driving a lot of the activities
- Other experiments and communities have requirements at least comparable scale
- Profit better from shared efforts and investments
- Funding agencies have finally started recognizing the importance of sustainable SW development for big science
- Use of HPC centers and clouds will increase
- Use of GPUs, other accelerators and machine learning will become more significant
- Authentication and authorization becoming easier
- The organization, management and access of big data will shift toward data lakes
- Sites will be able to choose between several ways to make their service deployment and operations more lightweight

Slides: <https://indico.cern.ch/event/772031/contributions/3373083/attachments/1856018/3048706/WLCG-Report-IPC-2019-06.pdf>

LHCONE traffic analyses: ATLAS and CMS

- ESnet Transatlantic (TA) traffic continues to grow. LHCONE is the main contributor
- TA traffic increased 50% over the last year
- TA traffic growth changes continuously: several slow down in the last 5 years, except LHCONE

Transatlantic Forecasting



Slides: https://indico.cern.ch/event/772031/contributions/3360613/attachments/1855600/3047667/ESnet_LHCONE_traffic_insights_-_Umea_meeting_2019_.pdf

LHCONE Looking Glass – Follow up

More peering have been added to the LHCONE looking glass:

- ASGC AS24167
- CERNlight AS20641
- ESnet AS293
- KREOnet AS17579
- GEANT AS20965 (Geneva and Frankfurt routers)
- NORDUnet AS2603
- RU-VRF AS57484

SINET, CANARIE, TEIN will follow soon

The looking glass is accessible here: <http://lhcone-lg.cern.ch/>

Slides: <https://indico.cern.ch/event/772031/contributions/3357234/attachments/1853899/3047014/LHCOPNE-20190604-Umea-LHCONE-looking-glass.pdf>

LHCONE routing table analyses

NORDUnet has updated the analyses of the routing tables of most of the LHCONE VRFs.

- IPv4 reachability has improved thanks to increased mesh of peerings
- IPv6 reachability is still fragmented and needs to be improved (action)

Slides: https://indico.cern.ch/event/772031/contributions/3428968/attachments/1855890/3048260/LHCone_routing_digging_2019.pdf

MTU size recommendation

The working working has produced a twiki page to document the MTU issue:

<https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneMTU>

MTU size recommendations for LHCONE networks

- In order to avoid issues with Jumbo frames, these recommendations on MTU size are given:
- The link layer MTU must be set to the maximum supported
- IPv4 and IPv6 MTU must be set to 9000 Bytes

The Path MTU Discovery protocols need these ICMP packets to be allowed:

- IPv4: ICMP Fragmentation Needed - Type 3, Code 4
- IPv6: ICMPv6 Packet Too Big - Type 2 (Value 0)

It's important that these packets have a routable IP address as sources, because unroutable addresses may be dropped by antispoofing filters.

Slides: https://indico.cern.ch/event/772031/contributions/3360615/attachments/1855275/3048624/LHCONE-MTU-Status-Umea-2019.pptx_1.pdf

LHCONE AUP review

Discusses whether is necessary to review the LHCONE AUP

Still concern on upcoming big data science project: add them to LHCONE or push them to create their own VPN?

- Network providers can easily implement any VPN
- Sites may have problems in selecting the right VPN to use

Proposed to test a separated VPN with an upcoming project. DUNE could be a candidate

Day 2: R&D

How to expose network status information

Experiments would like to get information about the status of the network to better debug transfer issues

Ideally, Network Operators could send network status information to a centralized repository

The information can be overwhelming. Concern on how the Experiment can understand it

In addition, Network Operators may have political reason not to share such information

Whole LHCONE usage statistics

Aiming to producing a graph with the whole LHCONE traffic

Discussed the implementation: all the VRF operators should send/generate a number with the amount of input traffic from all their client interfaces every 5 minutes

A prototype using the Grafana infrastructure at CERN will be tested. GEANT and NORDUnet will produce the initial data

NOTED project - update

- Network activity in the contest of **WLCG DOMA**
- CERN has hired a technical student to write a Transfer Broker
- Access to FTS transfer information in Elasticsearch and InfluxDB
- Need to enrich transfers records with IP prefixes of storage element
- Working on storing the IP prefixes of storage nodes in AGIS and CRIC
- The project welcomes input and suggestions

Slides: https://indico.cern.ch/event/772031/contributions/3428931/attachments/1856338/3049170/presentation_NOTED.pdf

LHCONE P2P and AutoGOLE activities - update

LHCONE P2P:

- DTN nodes placed in US and EU
- Experiments of Univ. Amsterdam to integrate DTNs to FileSender and OwnCloud

AutoGOLE:

- Dynamic ANA planned for this year, expecting scaling up to other regions
- Expanding AutoGOLE with connectivity to DTNs
- MEICAN has been chosen for the dynamic provisioning of circuits on a multi domain network. MEICAN software development by RNP (Brazil)

Slides: https://indico.cern.ch/event/772031/contributions/3448045/attachments/1856468/3049418/LHCOPN-LHCONE_Umea_-_Gerben_van_Malenstein.pdf

R&D 100G and DTNs

Overview of projects on DTN e data transfer services, especially on those based on servers capable to stream data from storage at 100Gbps

Slides: https://indico.cern.ch/event/772031/contributions/3449066/attachments/1855584/3047622/DTN_Presentation_LHC_Networking_Joe_Mambretti.pdf

LHCONE for HPC

Growing interest of LHC Experiments in HPC resources. Many presentations on the subject at the last HSF/OSG/WLCG workshop

Those infrastructure could be connected to LHCONE, possibly with DTNs, but only when dedicated to WLCG.

The community will investigate possible solutions in the coming meetings

Slides: <https://indico.cern.ch/event/725706/contributions/3118915/attachments/1744145/2823038/DataTransfer.pdf>

HEPiX NFV Working Group

- Explored several existing SDN/NFV approaches and use cases
- In Cloud Native Networking focusing primarily on Tungsten Fabric, OVS, but also looking into commercial projects. CERN and Nikhef leading pilot projects for deploying Tungsten Fabric
- SENSE and BigData Express leading projects in programmable networks and data transfers
- Surveying R&E plans for higher-level services
- Introductory tutorial and material to help sites establish their testbed is ready
- Finalizing white-paper process. Whitepaper draft:
<https://docs.google.com/document/d/1w7XUPxE23DJXn--j-M3KvXlfXHUnYgsVUhBpKFjyjUQ/edit?usp=sharing>

Slides:https://indico.cern.ch/event/772031/contributions/3335971/attachments/1856728/3049975/HEPiX_Network_Functions_Virtualisation_Working_Group_Update_June_2019_F2F.pdf

Conclusions

Summary

LHCOPN:

- most Tier1s connected at 100Gbps to CERN. The others will upgrade for Run3

LHCONE:

- Traffic grown ~50% in the last year. IPv6 traffic more than 30% of the total
- The sites upgrading to 100Gbps see immediate traffic grows
- Global reachability improved for IPv4. IPv6 still lacking behind
- AUP review: it's necessary to test separated VRFs for other science

Monitoring:

- perfSONAR moving to 100Gbps probes
- Looking glass for LHCONE has 7 peerings. More will follow

R&D:

- Several projects based on DTNs
- NOTED will implement a Transfer Broker for WLCG
- Work needed to connecting HPC centres

HEPiX NFV WG:

- working on white paper to help sites deploy NFV and SDN

Actions for next meeting

- LHCONE unwanted packets: follow up with GEANT and TEIN the packets generated by the Thai VRF
- Produce a better LHCONE connectivity guide for VRFs and sites
- Follow up on malicious injection of prefixes into LHCONE
- Looking glass: add more peerings, set addpath option
- MTU: send recommendations to mailing lists. Add list of jumbo capable perfSOANR servers
- AUP: test separated VRF for another collaboration (DUNE?)
- Implement prototype of LHCONE total stats
- Progress on NOTED Transfer Broker

Following Meetings

Next meeting: 13th and 14th of January 2020 at CERN

<https://indico.cern.ch/event/828520/>

Following meeting co-located with ISGC. Possible date: 8-9 of March 2020

Meeting in Fall 2020 could be co-located with NORDUnet conference. Meeting in Spring 2021 with HEPiX in US. TBC

References

Meeting agenda and presentations:

<https://indico.cern.ch/event/772031/>

Questions?

edoardo.martelli@cern.ch