EPJ Web of Conferences **295**, 07016 (2024) https://doi.org/10.1051/epjconf/202429507016
CHEP 2023

Evolving the LHCOPN and LHCONE networks to support HL-LHC computing requirements

Edoardo Martelli[1]

[1]CERN (Geneva, Switzerland), IT department CS group– email: edoardo.martelli@cern.ch

Abstract. The LHCOPN network, which links CERN to all the WLCG Tier 1s, and the LHCONE network, which connects WLCG Tier1s and Tier2s, have successfully provided the necessary bandwidth for the distribution of the data generated by the LHC experiments during first two runs of the LHC accelerator. We give here an overview of the most significant achievements and the current state of the two networks. It also explains how the two networks are evolving to support Run-3 and how they are preparing to meet the high demands foreseen for Run-4, notably by adopting new transmission technologies to increase the available bandwidth, introducing new software tools to improve the efficient utilization of all the links, as well as new monitoring capabilities to increase the understanding of the network traffic.

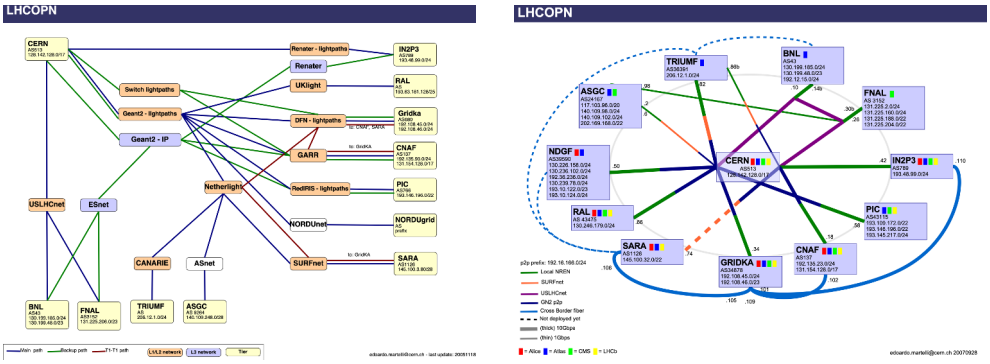
1 Introduction

The need to rely on data networks to distribute experimental data from CERN to collaborating institutes for data analyses was understood from the early stages of planning for the LHC accelerator and the associated detectors, even if the network technologies of that time were slow and in a very early stage of development. Fortunately, network technologies and capabilities evolved rapidly and the successes of the first two runs of the LHC, including the discovery of the Higgs Boson, owe much to the widespread availability of fast and reliable computer networks, including two dedicated infrastructures.
The first network dedicated to the Worldwide LHC Computing Grid (WLCG), LHCOPN [1], was fully operational during LHC Run-1. LHCONE [2] soon followed, providing the necessary bandwidth for the evolving computing models during LHC Run-2. The development and state of these networks in 2015 was described in a previous CHEP paper [3]; we describe here their evolution over the past eight years.

2 LHCOPN, the Tier0-Tier1 network

The LHC Optical Private Network, or LHCOPN, is a private network that connects the computing resources and storage elements at the WLCG Tier0 (CERN) to those at the Tier1s. It has a star topology: direct links from all the Tier1s to the Tier0.

From a network service provider point of view, LHCOPN is a Layer2 VPN: long distance Layer 2 (Ethernet) links that connect the border routers of two sites. The Layer 3 routing is implemented by the sites' routers using External BGP peerings which exchange the IP prefixes used by the servers.



(a) LHCOPN in 2005. **(b)** LHCOPN in 2007

Figure 1: Early LHCOPN topology diagrams

Conceived in 2005 and implemented the same year, LHCOPN initially connected eleven Tier1s: CA-TRIUMF, DE-KIT, ES-PIC, IT-INFN-CNAF, NDGF, NL-T1, TW-ASGC, UK-RAL, US-BNL, US-FNAL. The topology was a star and a partial mesh, with direct links between some pairs of Tier1s. Figure 1 shows two early topologies of LHCOPN.

In 2013 RU-KI and RU-JINR joined, followed by KR-KISTI in 2014. The membership remained stable until 2023, when, having stepped back from its Tier-1 role, TW-ASGC disconnected and two new sites, PL-NCBJ and CN-IHEP, connected to LHCOPN as they prepared to become new Tier1 sites.

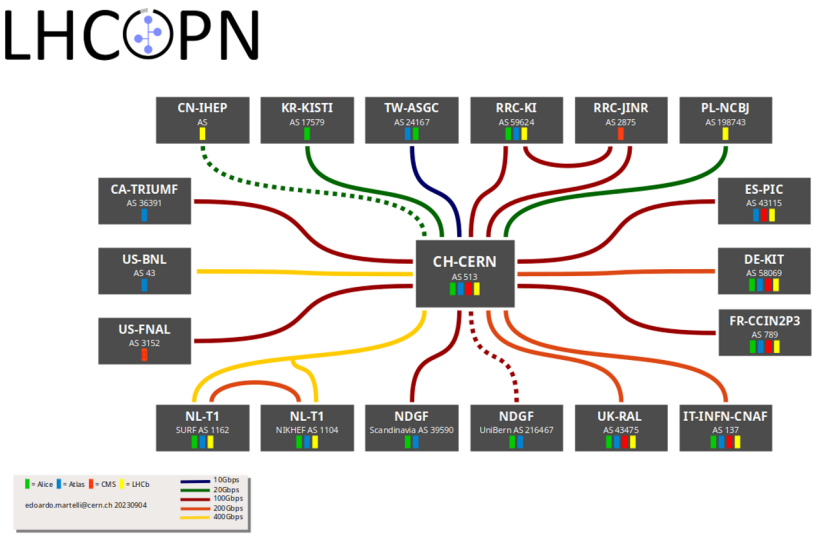


Figure 2: Actual LHCOPN map (2023).

Although membership remained relatively stable, the network evolved after LHCONE was introduced with the Tier1-Tier1 links gradually being replaced by LHCONE for their

direct and back-up connectivity functionalities (see section 3). The current state of the LHCOPN is depicted in figure 2.

Over the years the speed of the links has evolved from 1 Gbps to multiple 100 Gbps. Very recently NL-T1 (Nikhef) has connected with a 400 Gbps link. Today the aggregated bandwidth from all the Tier1s to the Tier0 exceeds 2.2 Tbps.

In preparation for the high bandwidth demand of HL-LHC (see section 4), new technologies are being explored, such as the use of shared spectrum on the service provider's dark-fibre infrastructure, directly from transmission devices at the premises of the WLCG sites. A pilot of this emerging technology is described in section 5.5

3 LHCONE, the Tier1-Tier2 network

Given the state of networks in the late 1990's, the intitial plans for the flow of LHC data were very hiearchical, with data flowing between the Tier-0 and regional Tier-1s and then between the Tier-1s and the Tier-2 centres in their region. Having seen the capability of networks during LHC Run-1, the Experiments started thinking of a less structured computing model, with Tier2 sites free to connect to any Tier1 in the world. As a consequence, in 2010 the LHCOPN community started exploring the possibility to create a new network to better connect Tier2s to any Tier1. Thus was born the LHC Open Network Environment (LHCONE), a private network that interconnects computing resources and storage elements at any particpating WLCG site, whether Tier0, Tier1 or Tier2s and regardless of location.

By 2012 a prototype of LHCONE spanning from Europe to North America was implemented. It soon went in production and it has been growing since, reaching now South America, Asia and Australia.

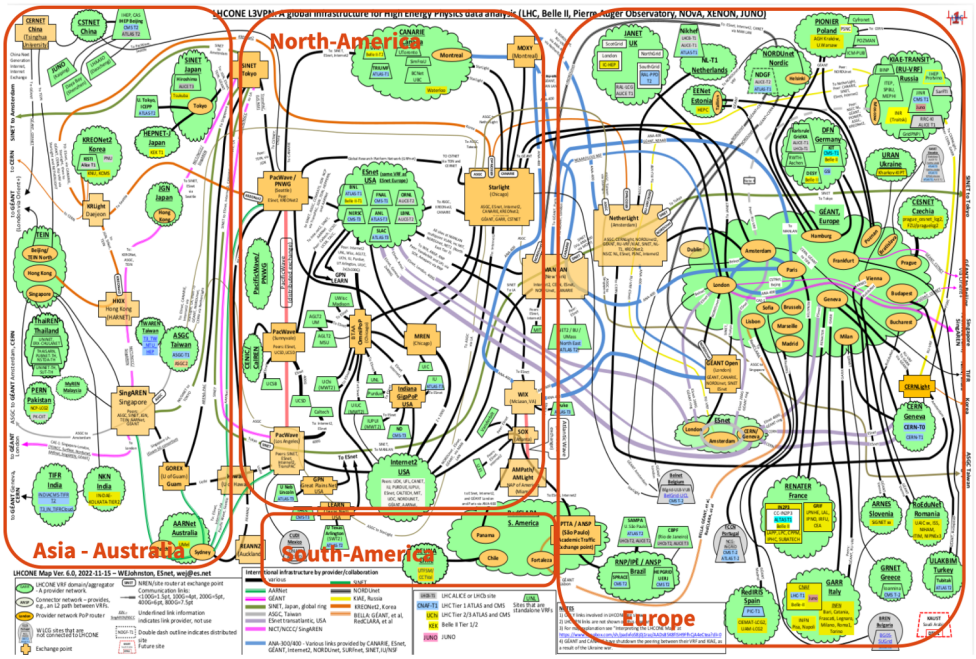


Figure 3: LHCONE network infrastructure (2023).

Unlike LHCOPN, a Layer 2 VPN, LHCONE is a Layer 3 VPN implemented by several network service providers (NSP). Each NSP provides an LHCONE instance as an overlay network on top of its own. LHCONE instances exchange routing information using the BGP protocol at interconnection points. At the edges, WLCG sites establish BGP peers with the closest LHCONE edge router and announce the IP prefixes used for their computing and storage resources. The NSP routers forward these announcements to all the other connected sites and to all the LHCONE instances of others NSPs to ensure global reachability. This is, of course, the same routing principle as the public Internet, but with the difference that only WLCG sites can connect to it.

This substantial difference, i.e. that only WLCG sites can connect, is what makes LHCONE very appealing: sites can trust the traffic that comes from the LHCONE upstreams and can connect these links directly into their data-centres, avoiding the need for expensive stateful security inspection tools at the network borders which would otherwise be needed to handle the high bandwidth data transfers. It is thanks to this trust that sites can afford very high-speed connections to LHCONE and achieve very high transfer throughput.

As of today, 31 Research and Education NSPs each provide an LHCONE instance; these 31 instances connect 117 sites in 5 continents. Figure 3 shows the complexity of the network infrastructure on top of which LHCONE is built.

Network information about the connected sites is stored in the CRIC (Computing Resource Information Catalog) database [4].

Permission to connect and use LHCONE is regulated by the LHCONE AUP (Acceptable Use Policy) [5], which has been defined and agreed by the LHCONE community.

Over the years, other High Energy Physics collaborations and experiments have joined LHCONE, mostly because many of their sites were already connected to LHCONE. As of today, the following non-LHC collaborations are part of LHCONE: BelleII, DUNE, JUNO, NOvA, the Pierre Auger Observatory and XENON (details in [5]).

This growth in the number of connected sites has, however, led to questions about the scalability of the LHCONE model as well as and raising doubts on the sustainability of the LHCONE security model. These issues are discussed in section 5.2

4 HL-LHC network requirements

The LHC has already started an upgrade process to implement improvements that will lead to a much higher luminosity (i.e. a higher number of collisions at every particle cycle), the High Luminosity LHC (HL-LHC) [6]. This higher luminosity will lead not only to a higher event rate for the experiments, but also to more complex events as the number of collisions at each bunch crossing increases, so adding more background clutter.

The LHC experiments estimate that the HL-LHC will generate a ten-fold increase in the amount of data compared to the current run (Run-3). Extraploating from current connection bandwidths and the current data rates of the the major LHC experiment observed during Run-3, Tier1s will need at least a 1 Tbps connection to the Tier0 at the beginning of Run-4 as well as an additional aggregated bandwidth of 1 Tbps for traffic to Tier2s [7]. LHCOPN and LHCONE are evolving to meet these requirements, not only by increasing the bandwidth capacity, but also by augmenting resiliency, monitoring capabilities and the ability to dynamically reconfigure themselves to better serve the most demanding traffic. Section 5 describes these improvements in detail.

The WLCG community has planned a series of data challenges to make sure that the complete system will be ready to support the foreseen computing load and associated network traffic. The first challenge took place in 2021 and demonstrated the possibility to reach 10%

of the HL-LHC needs. The next challenge is planned for 2024 and aims to demonstrate the ability to reach 25% of the foreseen capacity. [8]

5 Network Research and Development

The LHCOPN and LHCONE community has proposed and is developing several research and development projects [9]. These projects have different scopes, but mainly they aim to improve the visibility of network performance to end users, to reduce network costs, to better exploit all the available bandwidth and to reduce link idle time.

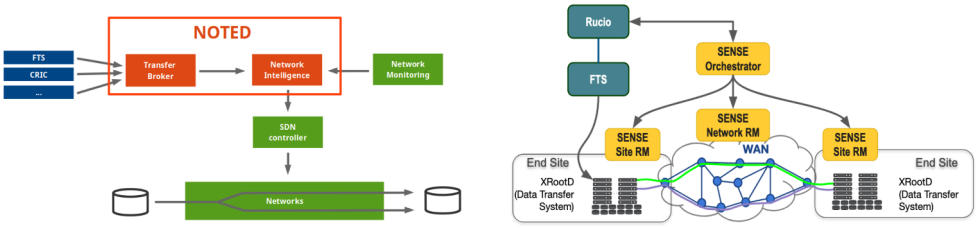
5.1 Software Defined Networking

SDN or Software Defined Networking is a paradigm in which an application can request that the network modify its behaviour in order to better serve the network activities of the application itself.

Two SDN tools, NOTED and SENSE, are being developed by the WLCG community with the aim of improving the performance of data transfers over heavily loaded networks

NOTED (for **N**etwork **O**ptimized **T**ransfer of **E**xperimental **D**ata) is an SDN application designed and developed at CERN. It aims to detect large data transfers and estimate their duration, so as to request network improvements where they are needed and when the duration of the transfer is long enough to justify the request. This can reduce the duration of a file transfer and avoid congestion on network links.

The NOTED architecture is described in figure 4a. The Transfer Broker component looks for large and long lasting data transfers generated by FTS [10]; the Network Intelligence component requests network re-configurations where they can reduce the duration of those transfers. NOTED is described in more detail in other papers presented at CHEP conferences: [11] and [12]



(a) NOTED architecture. **(b)** SENSE Rucio Science Workflow.

Figure 4: SDN projects

The **SENSE-Rucio Science Workflow** provides mechanisms for domain science workflows and middleware to identify priority data flows and implement wide area network traffic engineering. In this case the SENSE service provided by ESnet is used to improve the network Quality of Services for the data flows specified by the Rucio data management framework. This use case is described in [13]. This project also aims to reduce the duration of data transfers over congested networks.

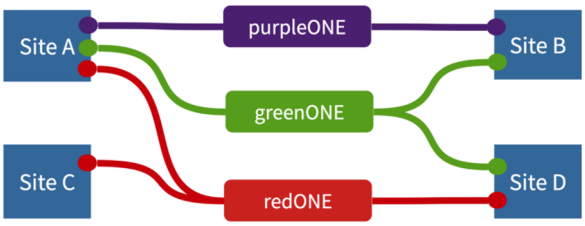


Figure 5: multiONE structure

5.2 MultiONE

MultiONE is an activity aiming to offer the advantages of LHCONE to other large science projects (such as SKAO, ITER, Rubin Observatory) that will start operating at the same time as HL-LHC. MultiONE can provide a secure network environment for sites serving multiple large-science collaborations without compromising the security of LHCONE sites; this is achieved by separating each collaboration in a dedicated virtual private network.
MultiONE is described in more detail in another paper presented at this CHEP conference: [14]

5.3 Flow and Packet Marking

Scientific Network Tags (scitags, https://www.scitags.org/) is an initiative promoting identification of science domains and their high-level activities at the network level. Two methods are proposed:
- Flow marking with UDP fireflies: issuing a specific UDP packet in a separate channel.
- Packet marking: adding a 14-bit code in the IPv6 flowlabel field.
Scitags will improve the understanding of the network utilization, the monitoring of the transfer applications and may allow advanced traffic engineering techniques.
Scitags is described in more detail in another paper presented to CHEP [15]

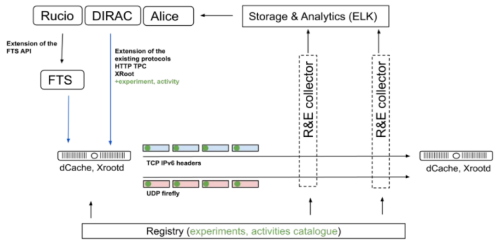


Figure 6: Science Tags.

5.4 Commercial Cloud Resources

LHC Experiments have been looking at the possibility to use commercial cloud resources to expand the computing capacity in an opportunistic way, i.e. in case of high demand and when affordable [16].

Having proved to be agile and economical, these resources will need excellent network connectivity to be fully exploited. One interesting possibility would be to have them reachable over LHCONE. For this reason LHC experiments and network service providers are working together to try to implement a virtual WLCG site in the cloud with full and secure LHCONE connectivity.

5.5 CERN-CNAF DCI

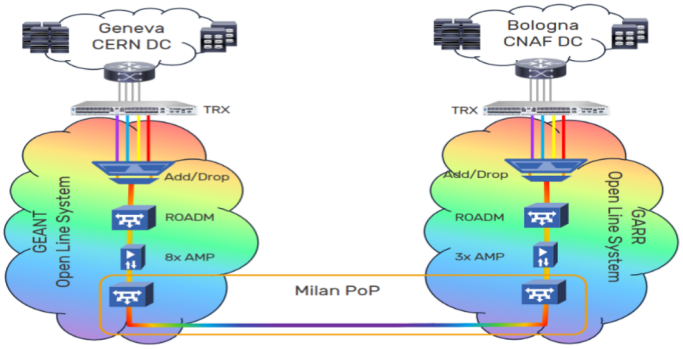


Figure 7: CERN-CNAF Data Centre Interconnect

INFN-CNAF, GARR, GEANT and CERN have setup a Data Centre Interconnect service prototype between CNAF and CERN. The service is implemented using long distance transmission devices injecting multiple 400Gbps wavelenghts on the dark fibres of GEANT and GARR. This DCI can easily grow to 1.6Tbps, more than the T0-T1 bandwidth required by HL-LHC
The innovations of this prototype are: a) the injection of long distance wavelength using transmission devices operated by the clients (CERN and INFN-CNAF); b) the use of a DWDM channel spanning over two dark fibres operated by two different network providers (GEANT and GARR). The setup is described in detail in [17].

Such transmission technologies will allow sites to get affordable and large network connections, with enough bandwidth to meet the HL-LHC requirements.

6 Conclusion

LHCOPN and LHCONE are two networks built by the WLCG community and the Research and Education network providers. They are an essential component of the computing infrastructure used by the High Energy Physics community to achieve impressive scientific results.

LHCOPN and LHCONE are evolving to prepare for the demanding computing requirements of HL-LHC. The community has so far been fortunate in that networks have been over provisioned relative to HEP needs. Competition for network bandwidth is increasing, though, given the growing number of large-scale global science collaborations; the research projects listed in this paper may help to increase the speed of data transfers and improve the efficient use of long-distance network links. They also introduce complexity and thus fragility, so their use in production has to be carefully evaluated.

Written in Geneva (Switzerland), in August 2023

References

[1] *LHCOPN web site*, https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome, (accessed: 10.08.2023)

[2] *LHCONE web site*, https://twiki.cern.ch/twiki/bin/view/LHCONE/WebHome, (accessed: 10.08.2023)

[3] E. Martelli, S. Stancu, *LHCOPN and LHCONE: Status and Future Evolution*, Journal of Physics: Conference Series, **664** (2015)

[4] C.D.T. at CERN, *CRIC computing resource information catalog. core service/site*, http://wlcg-cric.cern.ch/ (2020), accessed: 12.07.2023

[5] *LHCONE acceptable use policy* (2021), https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneAup, (accessed: 12.07.2023)

[6] *High luminosity LHC*, https://hilumilhc.web.cern.ch/, (accessed: 12.07.2023)

[7] S. Campana, *WLCG data challenges for HL-LHC - 2021 planning* (2021), https://zenodo.org/record/5532452, (accessed: 12.07.2023)

[8] W.C. Lassnig Mario, *WLCG DOMA news and status*, https://indico.cern.ch/event/1225114/contributions/5476124/attachments/2682805/4654196/WLCG%20DOMA%20News%20and%20Status%20-%20GDB%20072023.pdf, (accessed: 12.07.2023)

[9] M. Babik, S. McKee, *Network Capabilities for the HL-LHC Era*, EPJ Web of Conferences **245** (2020)

[10] E. Karavakis, A. Manzi, M.A. Rios, O. Keeble, C.G. Cabot, M. Simon, M. Patrascoiu, A. Angelogiannopoulos, *FTS improvements for LHC Run-3 and beyond*, EPJ Web of Conferences **245** (2020)

[11] J. Waczynska, E. Martelli, E. Karavakis, T. Cass, *NOTED: a framework to optimise network traffic via the analysis of data from File Transfer Services*, EPJ Web of Conferences **251**, 02049 (2021)

[12] C. Misa Moreira, E. Martelli, , T. Cass, *NOTED: An intelligent network controller to improve the throughput of large data transfers in File Transfer Services by handling dynamic circuits*, EPJ Web of Conferences (2023)

[13] J. Balcas, C. Guok, *Complete End-to-End Network Path Control for Scientific Communities with QoS Capabilities*, EPJ Web of Conferences (2023)

[14] C. Misa Moreira, E. Martelli, , T. Cass, *P4flow: A software-defined networking approach with programmable switches for accounting and forwarding IPv6 packets with user-defined flow label tags*, EPJ Web of Conferences (2023)

[15] S. McKee, M. Babik, T. Chown, A. Hanushevsky, T. Sullivan, B. Hoeft, J. Letts, D. Carder, G. Attebury, M. Lambert et al., *Identifying and Understanding Scientific Network Flows*, EPJ Web of Conferences (2023)

[16] F.H.B. Megino, M. Borodin, K. De, J. Elmsheuser, A. Di Girolamo, N. Hartmann, L. Heinrich, A. Klimentov, M. Lassnig, F. Lin et al., *Accelerating science: the usage of commercial clouds in ATLAS Distributed Computing*, EPJ Web of Conferences (2023)

[17] G. Vaugnin, *Deploying a Pilot Spectrum Connection Service over GARR/GEANT: Lessons Learned* (2023), https://indico.geant.org/event/2/contributions/210/attachments/144/356/TNC23_Vuagnin.pptx, (accessed: 08.08.2023)