EPJ Web of Conferences **245**, 07018 (2020) https://doi.org/10.1051/epjconf/202024507018
CHEP 2019

Unroutable Traffic: Maintaining Trust and Integrity of the LHC Open Network Environment

Bruno Hoeft[1], *Samuel* Ambroj Pérez[1], *Magnus* Bergroth[2], *Michael* O'Connor[3], *Richard* Cziva[3]

[1]Karlsruhe Inst. of Tech. (KIT), Herm. von Helmholtz Pl. 1, 76344 Eggenstein-Leopoldsh., Germany
[2]NORDUnet, Tulegatan 11, 2tr. SE-113 53 Stockholm Sweden
[3]ESnet, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, US

Abstract. This paper explores the methods and results confirming the baseline assumption that LHCONE[1] (Large Hadron Collider Open Network Environment) traffic is science traffic. The LHCONE is a network conceived to support globally distributed collaborative science. LHCONE connects thousands of researchers to Large Hadron Collider (LHC) datasets at hundreds of universities and labs performing analysis within the global collaboration on high-energy physics. It is "Open" to all levels of the LHC as well as a short list of approved non-LHC science collaborations. LHCONE satisfies the need for a high performance global data transfer network of supporting scientific analysis. Even though LHCONE is a closed virtual private network, packets from non-LHCONE sites were found within the network on multiple occasions. This paper describes the findings, discusses the reasons and proposes some ideas on how to prevent "unroutable LHCONE packets" in order to maintain trust and integrity within the network.

1 Introduction

LHCONE is a virtual private network (VPN) connecting numerous universities and research facilities of the High Energy Physics (HEP) community around the globe. It has spanned over America, Europe, Asia and Australia joined lately, while the first steps to include African sites are under preparation. As of 2020, there are more than 130 endsites connected to LHCONE and close to 30 involved Network Service Providers (NSPs) imple- menting the connecting "Virtual Routing and Forwarding" (VRF) basis. The individual link capacity of the VPN varies from 1 Gbps to multiple 100 Gbps. The investigation into the packet "quality" of LHCONE was initiated when the first none LHCONE-compliant packets were found in Germany and it was discovered that they were injected to LHCONE from quite a high number of sites.

Packets are defined as unroutable packets, if they are found in the LHCONE VPN with source or destination address outside the LHCONE address spaces. All sites connected to the LHCONE VPN are announcing their respective address spaces with the Border Gateway Protocol (BGP). Through BGP each site has a routing table including all addresses of the LHCONE VPN. For example, an ingress packet at the LHCONE interface of a LHCONE site with a source address outside of the LHCONE VPN address space allows no return packet to the source through the LHCONE VPN, which is marking this packet therefore as "unroutable LHCONE packet". Even though the source address might be a valid internet address, if it is not within the LHCONE routing table the packet is considered as unrouteable.

LHCONE is dedicated to LHC science and the National Research and Education Networks (NRENs). They actively support LHCONE in the scientific community and play an important role in supporting the Globally Distributed Computing Model essential to LHC experiments. LHCONE was conceived to provide a network dedicated to analyse and connect the compute resources at LHC Tier centers to research performed at universities and other collaborating institutions worldwide. The need to connect research institutions to the relatively small group of dedicated NRENs is an important advantage to span geographies and achieve scale in the LHC analysis community. However the charters of many NRENs prevent them from providing general Internet service to Universities within their respective national borders. The LHCONE service is not a general Internet service, it is exclusively a scientific transport network and sites are required to limit access specifically to the group, department etc. affiliated with the LHC collaboration. This focus on science and the controlled access features in the LHCONE design have enabled direct NREN connectivity to large numbers of educational compute centers that would not otherwise be permitted by NREN Acceptable Use Policies (AUPs). The closed LHCONE VPN implemented a trust relation between the connected sites. Packets injected from unlegitible sources are jeopardising this trust (/security) frame work.

2 LHCONE Policy

If all connected sites and their NSPs would comply with the requirements defined at LHCONE connection guide[3] and the AUP[2] unroutable LHCONE packets should not exist at all.

2.1 LHCONE Policy

The LHCONE connection guide is describing the regulation and policies the LHCONE connected site and the NSP should keep to. The traffic injected into LHCONE must only originate from addresses within the LHCONE routable prefix. Only address ranges present in the LHCONE routing table should be transported on the network. In order to maintain route symmetry and access control, each NSP has to implement policy and packet filters to manage their connected customer address prefix ranges. This ensures that a return route exists within the LHCONE network and that spoofed packets (similar to BCP 38 [4]) are prevented.

2.2 Acceptable Use Policy

The AUP defines the different collaborations participating in LHCONE. These are the high-energy physics experiments Belle-2, NOvA, Pierre Auger Observatory, U.S.ATLAS, U.S.CMS, WLCG, XENON. Nodes which are currently and primarily used to distribute, store, process and analyze the data generated by HEP sites can be part of LHCONE. The same applies to routers and switches for routing such data as well as perfsonar probes and corresponding management infrastructure used for LHCONE. The AUP states that only HEP sites of one of the participating collaborations are eligible to be connected to LHCONE. Moreover it defines the roles and responsibilities of sites and providers (NSPs). The sites have to acknowledge this AUP and must be responsible for forming their own security policy regarding traffic arriving from LHCONE. LHCONE providers (NSPs) have to make sure that only sites which have acknowledged this AUP become connected to the LHCONE L3VPN. They also have to implement BGP filtering based on LHCOPN BGP communities, which implies BGP import filters and source address packet filtering. The end sites are encouraged as well to implement (egress) source address filters at their edge in order to eliminate their own unroutable LHCONE packets.

2.3 NSP required actions

The NSPs are supposed to set up the LHCONE connection with their customers and take care that the site establishes the requirements/regulations of the aforementioned AUP. NSPs should generally discard non-compliant packets and inform the sites accordingly. It would be beneficial if the NSP would regularly check the LHCONE submitted packets and if necessary assist with reconfiguration of the site.

3 The Investigation and their Results

Two LHCONE collaborators analyzed the LHCONE traffic. One of them was NSP ESnet[5] and the other one the WLCG German Tier-1 site DE-KIT/GidKa located at Karlsruhe Institute of Technology[6] (KIT) (further named as DE-KIT). Both sites have been inspecting LHCONE traffic, however, with different methods.

3.1 ESnet

The US Department of Energy network "Energy Sciences Network" (ESnet) is a founding member of the LHCONE network collaboration and is very well positioned to observe a large percentage of global LHCONE data flows. ESnet supports two US LHC Tier 1 collaborating institutions, BNL and FNAL for ATLAS and CMS respectively. The Cern Laboratory Tier 1 compute center is also directly connected. ESnet direct site connections include US National Labs, LHC collaborating universities in the US as well as CERN. This represents about 35% of all LHCONE collaborating compute and analysis centers.

ESnet has built a rich peering environment connecting to approximately 54% of 26 total collaborating LHCONE network service providers, predominantly NRENs.

ESnet processed three months of LHCONE netflow[7] data (Feb. 2019 - Apr. 2019), with a sampling rate of 1000 packets, on these 38 directly connected AS's, with packets from a total of 95 unique LHCONE AS's observed. These networks provided over 46,000 unique LHCONE speakers with a combined output of 3.5 trillion packets for the three month period.

3.2 DE-KIT

At DE-KIT two Cisco Nexus border routers were deployed for the German LHC-Tier-1 DE-KIT. Each border router had a separate connection with the LHCONE VRF of the NSP Deutsches ForschungsNetz (DFN). Both were configured with an ingress filter matching the source address of each incoming packet with the LHCONE routing table. All unmatched packets are regarded as LHCONE noncompliant. The headers of the noncompliant packets were sent through a mirror port of the router and to linux server were the information is written to a logfile unsampled. That logfile was previously analyzed manually and then transferred to a spreadsheet for visualization. Today applications display each noncompliant packet. The current deployment of a chain of several applications: packet beat, elastic search, kibana, grafana will be discussed in a separate chapter.

During four weeks in March/April 2019 at DE-KIT 2.3 mil. unroutable packets were found. In relation to the total traffic only 0.2%, this is not very much, however, too much to be ignored. It is quite a high number of 77 different independent source subnets. When dividing the sources into IPv4 and IPv6 protocols there are 44 IPv4 and 33 IPv6 sources with only a very few overlapping. ICMP packets were often responses to packets that originated from our own site. Yet we were concerned that more than half of the packets were non ICMP packets (57%). Combining the subnets address spaces belonging to one institution or autonomous system (as) and removing all sources with less than 1,000 recorded packets reduced the number of IPv4 sources to 30 and IPv6 sources to 15 sites. At the protocol IPv4 6 sources with packets above 50,000 each were further analyzed. 109,340 packets from Géant

were ICMP packets only and issued from the Géant routing gears. For example these are ICMP packets containing information like destination not reachable or frame (MTU size) too big. The same applied for 61,566 packets of the Italien NSP GARR and for 61,366 packets of the French NSP RENATER. The filters at the border routers were adjusted so that these packets from known NSPs were no longer considered as unroutable packets and therefore not counted in the logfile.

80,093 unroutable LHCONE packets were discovered with a source address of the Tier-3 TW-NCUHEP the HEP group of the "National Central University of Taiwan. And 52,650 packets of the State Scientific Center of the Russian Federation the Institute for Theoretical and Experimental Physics in Moscow. 425,931 packets were injected by the Tier-3 site T3-TH-CHULA of the Chulalongkom University in Thailand. All three sites and their connecting NSPs were addressed and notified about the unroutable LHCONE packets, but none of them responded. TW-NCUHEP still injects packets to LHCONE but only at a much lower sending scale. T3-TH-CHULA disappeared and no packets have been recorded ever since. Only ITEP in Moscow is still injecting packets into LHCONE at unchanged rate until today. This requires further investigation and actions.

Looking at the 33 IPv6 sources those with less than 1.000 packages were dismissed equal to what was done with the IPv4 sources. This reduced the number to 15. From these 15 sources only 4 sources with more than 50,000 unroutable packets were to be analyzed. One of them was GARR, the Italian NSP with 914,626 packets. Several sources were the LHCONE trespassing router injecting ICMP packets. Two sources injecting a high packet rate were two Perfsonar servers, a bandwidth and a latency Perfsonar server of the Tier-2 INFN-ROMA1-CMS in Rome. After including the CIDR in LHCONE all injected packets by GARR are clarified. The injected packets by the French NSP RENATER were only ICMP packets of the LHCONE trespassing routing equipment and could therefore be ignored. 52,512 packets were injected by the Tier-2 "RO-02-NIPNE" of the Romanian National Institute for Physics and Nuclear Engineering in Bucharest and 208,693 packets were injected by the Tier-2 at the Universidad De Sao Paulo. The last two sites were approached both the site itself and the connecting NSP as well, without any response. Further investigations are going on.

Another discovered topic was that quite a high rate on private source addresses were discovered. At IPv4 it was distributed over three private subnet areas: 10.0.0.0/8, 172.16.0.0/12, 196.168.0.0/16. At IPv6 only private source addresses "link local" were discovered at the LHCONE incoming interface. There were only two possibilities:

- Ignore the packets with private source addresses and deploy a rejecting filter at the ingress interface
- Investigate and analyze at the NSP the source of the private address packets.

Up to now our connecting NSP could not get convinced to invest in the second option.

4 Visualizing "Unroutable Packets"

For reducing the time and human resource consuming gathering of the "unroutable LHCONE packets" data, the process was automated and equipped with a graphical data representation including a geolocated packet presentation. The graphical representation is still in a first more experimental state. The final version will be internet accessible and requires out of the sensitivity of the data edugain authentication. Here is a brief summary of the current setup. The two border routers streamed non-compliant LHCONE packets to a server where packetbeat ran two independent processes. Each process received only packages from one border router in one NIC, which facilitated the distinction between border routers. Packetbeat was configured to collect and report statistics on network flows. A flow is a group of packets sent over the same time period with common properties, such as source and destination address and protocol. The setup also enabled the ICMP protocol, which added common

protocol-specific options. This information gathered by packetbeat was sent to another server where logstash ran. Logstash applied different filters on the incoming packages to add information about the geographical location of IP addresses and ASN information. This information was obtained from the GeoLite2 databases. On the same server, there was an elasticsearch instance which received the output of logstash. Elasticsearch is a full text search server based on Lucene. Each document is stored as a JSON document and it is possible to contact it via a Restful API. In our analysis Kibana and Grafana have been used to facilitate the visualization of the data as shown in figure 1.

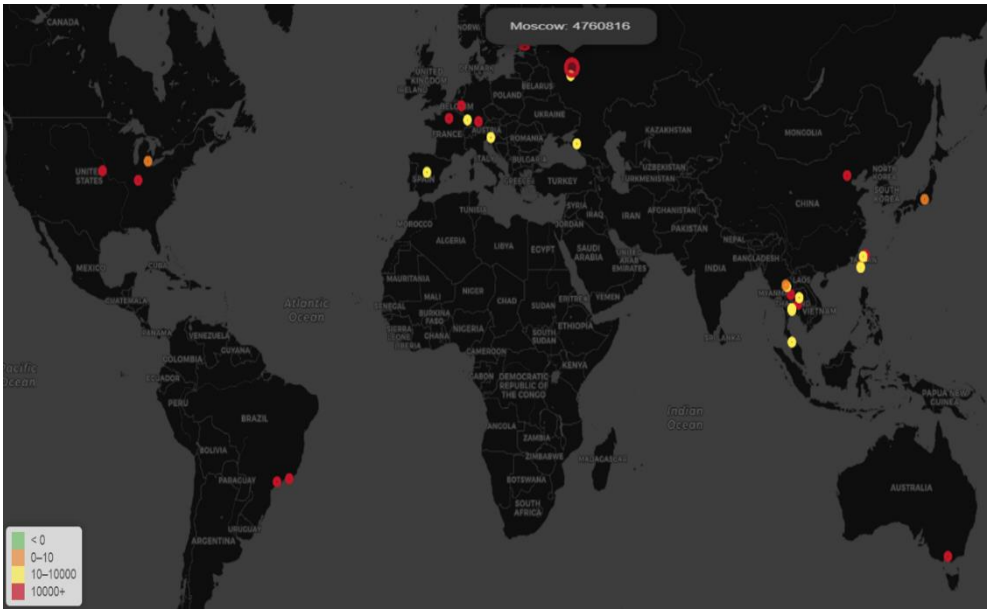

Figure 1. Geolocation representation of the unroutable LHCONE packet sources

5 LHCONE regional differences at the BGP routing table

Detecting unroutable LHCONE packets is predicated on a fully consistent routing table, any missing route prefixes in the LHCONE table used by the monitoring network will potentially cause legitimate sources to be labeled as unroutable.

LHCONE was conceived and initially implemented by NRENs in Europe and North America. GÈANT's collaboration of European national networks, NORDUnet serving the Nordic countries and ESnet, Internet2 and Canarie in North America. These networks already maintained well established general Internet peering relationships that were expanded to form a fully meshed and consistent LHCONE routing core. Since each of these networks peered with all of the other founding members, there was no need for "transit networks", as is necessary to fully connect the general Internet.

As LHCONE grew, the full mesh could no longer be maintained and the need for transit networks became obvious as in 2018 when RUNNET, the NSP connecting Federal State Institution Russian Scientific Center Kurchatovsky institute (AS57484), contacted their LHCONE upstream provider, NORDUnet with a problem reaching TEIN's Asian down-stream networks. During this period of growth, Asian NRENs began establishing BGP peerings directly with the established LHCONE NRENs. However, these European and North American networks were not all present in Asia and maintaining a full mesh was no longer possible. The result was that without recognized transit networks, inconsistency crept into the LCHONE routing tables across the globe and newer LHCONE compute centers were

not able to reach all of the other members. Since these issues were first brought to light, progress has been made establishing transit networks in Asia and the US, but this also requires monitoring and has not completely eliminated the routing reachability issues as reported[8] by Magnus Bergroth of NORDUnet at LHCONE meeting.

6 Regular Monitoring and follow up of discovered unroutable packets

It will still be necessary to observe packets transmitted through LHCONE to identify new "unroutable LHCONE packet" sources. Follow it up, communicate with the connected site and make them aware of the issue. On the other hand communicate with the connecting NSP and visualize the issue and ask them to redirect the traffic or at least stop sending the traffic to LHCONE. During the effort of the last month assumed over four weeks aggregated traffic of 2.4 million packets could be reduced to only a little bit over 1 mil. packets. Several sources were found which can be ignored: e.g. packets submitted by the routing equipment of NSPs (packet not deliverable, packet to big). One other source of a high rate of tcp packets could get omitted after the NSP included the site in the LHCONE routing table. One other site stopped after notifying submitting packets to LHCONE. All this was only possible with a detailed monitoring and analysing of the traffic.

7 Conclusion

Maintaining the integrity of the LHCONE AUP is critical to the deployment of LHCONE as a controlled access purpose built network service able to support the globally distributed computing model relied on by the LHC collaborating institutions. Relaxing the established access controls would render LHCONE indistinguishable from a general Internet transit network, which is absolutely not the intent of the LHC collaboration and could not possibly justify the necessary funding to support it.

Because LHCONE is distinct from the general public Internet, NSPs have been able to offer connectivity directly to University compute centers resulting in a more tightly integrated network that avoids local bottlenecks encountered when transiting multiple unaffiliated commercial and regional networks. The best example of this is the inclusion of US universities, the two US Tier 1 centers and Cern the Tier 0, directly on ESnet. Connecting general Internet services directly to Universities is not permitted by ESnet's own AUP. Regular and consistent monitoring of unroutable packets in the LHCONE network is essential to maintaining access control and ultimately the integrity of the established LHCONE AUP.

LHCONE access control is managed at the edge between NSP's and the compute centers, we have observed that as that perimeter expands to include new NSPs the potential for mistakes increases. Not all LHCONE NSPs attend our meetings or are as committed to staying current. Fortunately we have been able to demonstrate that monitoring from as few as two network perimeters can detect edge filtering configuration errors in remote networks across the globe.

The unroutable LHCONE packets found have been discussed during several LHCONE meetings. On one hand NSPs have been taking care about it and reduced them by better controlling. On the other hand the issues have been reduced by following up the cases and talking to the sites/NSPs. This resulted at ESnet in a decreasing rate from 2018 to 2019 and at DE-KIT in a stagnating rate.

There are still far too many unroutable LHCONE packets left to be neglected. Additionally, new sources injecting unroutable packets need to be approached and either the LHCONE site deployment or the packet routing adjusted. This requires permanent monitoring to make sure that unroutable packets are recorded.

The current monitoring system has to be developed further. One important task is automating the update between the LHCONE routing table and the monitoring system.

Another task is to increase the visibility of the monitoring system by installing an eduGAIN controlled open monitoring frontend visible to all collaborating partners.

There are plans to write an AUP Appendix mandating the LHCONE monitoring and offer the team a stronger position in approaching sites and NSPs.

References

[1] LHCONE - Large Hadron Collider Open Network Environment http://lhcone.web.cern.ch/

[2] LHCONE AUP - Acceptable Use Policy https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneAup

[3] Connection Guide for new LHCONE sites https://twiki.cern.ch/twiki/pub/LHCONE/LhcOneVRF/LHCONEconnectionguide-1.2.pdf

[4] Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing https://tools.ietf.org/html/bcp38

[5] ESnet - Energy Science Network https://www.es.net/engineering-services/the-network/

[6] Karlsruhe Institute of Technology, Research and Education Institute of Germany at State of Baden-Würtemberg: http://www.kit.edu

[7] Cisco Systems NetFlow Services Export Version 9 https://www.ietf.org/rfc/rfc3954.txt

[8] Digging in the LHCONE routing table (Magnus Bergroth / NORDUnet) https://indico.cern.ch/event/772031/contributions/3428968/attachments/1855890/3048260/LHCone_routing_digging_2019.pdf https://indico.cern.ch/event/725706/contributions/3149436/attachments/1744301/2823417/LHCone_routing_digging.pdf