unroutable LHCONE traffic

Bruno Hoeft / KIT
Michael O‘Connor / ESnet
Richard Cziva / Esnet
Samuel Ambroj Perez / KIT
Magnus Bergroth / NORDUnet
LHC → VPNs

- LHCOpticalPrivatNetwork → a CERN centric star VirtualPrivatNetwork
  - VPN with 15 participants only, mutual agreements between connected partners, no formal rules
LHCOpenNetworkEnvironment – MAP

- LHCONE → a still growing international distributed VPN
  - VPN with 104 endsites connected through 26 Virtual Routing and Forwarding implementations at the connecting Network Service Providers (NSPs)
NSP Packet Filtering Requirements

All LHCONE Traffic is subject to the following conditions:

- Traffic injected into the LHCONE must only be originated from addresses within an LHCONE routable prefix
- Only address ranges present in the LHCONE routing table should be transported on the network

Objective: In order to maintain route symmetry and access control, each NSP will implement policy and packet filters to manage their connected customer address prefix ranges.

- Ensures that a return route exists in the LHCONE network
- Blocks spoofed packets (Similar to BCP 38)

https://twiki.cern.ch/twiki/pub/LHCONE/LhcOneVRF/LHCONEconnectionguide-1.2.pdf
NSP Border Gateway Protocol (BGP) Import Policy

Prefix Lists will be negotiated between connecting institutions and their NSP within the constraints imposed by the LHCONE Acceptable Use Policy (AUP).

LHCONE NSPs have agreed to configure:

1. BGP import filters
2. Source address packet filters

End sites are encouraged to implement source address filters at their edge in order to eliminate their own unroutable LHCONE packets. NSPs will generally discard non-compliant packets without informing the site.

Connecting institutions/sites will not add prefixes to the LHCONE routing table without direct cooperation with their NSP.

AUP: https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneAup
The Investigation

ESnet
- Three months of ESnet netflow IPv4 & IPv6 sampling from Feb. 2019 - April 2019 for the following connected sites and peers

<table>
<thead>
<tr>
<th>aarnet</th>
<th>duke</th>
<th>ornl</th>
<th>transpac</th>
</tr>
</thead>
<tbody>
<tr>
<td>aglt2</td>
<td>fllr</td>
<td>mit</td>
<td>uchicago</td>
</tr>
<tr>
<td>anl</td>
<td>fnal</td>
<td>net2</td>
<td>ucsb</td>
</tr>
<tr>
<td>ansp</td>
<td>geant</td>
<td>nordunet</td>
<td>ucsd</td>
</tr>
<tr>
<td>asgc</td>
<td>ind-gpop</td>
<td>ou</td>
<td>uiuc</td>
</tr>
<tr>
<td>bnl</td>
<td>internet2</td>
<td>pnnl</td>
<td>unl</td>
</tr>
<tr>
<td>caltech</td>
<td>JGN</td>
<td>rnp</td>
<td>uta</td>
</tr>
<tr>
<td>canet</td>
<td>kiae</td>
<td>sinet</td>
<td>uwmadison</td>
</tr>
<tr>
<td>cern</td>
<td>kreonet</td>
<td>slac</td>
<td>vanderbilt</td>
</tr>
<tr>
<td>cernlight</td>
<td>cern</td>
<td>tacc</td>
<td></td>
</tr>
</tbody>
</table>

ESnet counted:
- All LHCONE ingress packets
- Unroutable source packets
- Packets with non-lhcone/missing origin ASN
* corrected for netflow sampling rate (1000)
Very good results!!!
A small number of U Toronto hosts are transmitting large amounts of data on LHCONE. ESnet had routes for U Toronto in the past.
Ingress Packet Filters based on LHCONE routing table

<table>
<thead>
<tr>
<th>Total packets (during 4 weeks)</th>
<th>IPv4</th>
<th>IPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.044.471</td>
<td>1.376.338</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Packet/day</th>
<th>IPv4</th>
<th>IPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>37.302</td>
<td>35.290</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Privat IP destination</th>
<th>IPv4</th>
<th>IPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10.0.0.0/16, 172.16.0.0/12, 192.168.0.0/16</td>
<td>fe00::/16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of sources</th>
<th>IPv4</th>
<th>IPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>44 + private</td>
<td>33 + private</td>
</tr>
</tbody>
</table>

**Total LHCONE traffic → 0,2% / none ICMP → approx. 57%**

- priv
- AARNIEC-RoEduNet
- AIT-TH
- ASGC-NET
- Associação Rede Nacional (BR)
- CANARIE
- Bcnet Vancouver
- CAS-PRG-6TCZ
- CAS-TCZ
- CERNET
- CHINANET-FJ
- CIECCHQ-CN
- CIEMAT
- CZ-RELCOM-19930901
- DFN WIN-IPV6
- ERNET-IN
- ES-REDIRIS-20010521
- Fundação Carlos Chagas Filho (BR)
- FR-RENATER
- FR-IN2P3-LAL-ORSAY
- FR-IN2P3-LLR-PALaiseau
- FR-CEA-SCLAY-GRILLES
- FR-IN2P3-LPNHE-PARIS
- FR-IN2P3-LPC-CLERMONT-AUBIER
- FR-IN2P3-LAPP-ANNECY
- FR-IN2P3-CPPM-MARSEILLE
- GEANT
- GARR-P-P
- GARRB-NET
- GARRX-NET
- GR-GRNET-19991208
- GZIN
- IANA – reserved
- IHEP-IPV6
- Imperial College London
- Indiana University
- INFNNET-LNF
- Inst. Nat. de Physique Nucléaire
- IT-GARR-20011004
- JINR-NET
- Kasetsart University, Thailand
- KREONet-KR
- NL-GEANT-20020131
- PNPI
- REDIRIS
- RoEduNet-IPV6-NET-1
- RRC-KIAE-Moscow
- RU-ROSNIROS-20180806
- RWTH Aachen
- SAVECOM SAVECOM-NET
- SINEJ-JPNC
- SUIT-TH
- T-NCU.EDU.TW-NET
- T-NSYSU.EDU.TW-NET
- T-NTHU.EDU.TW-NET
- TANET
- TANET-8 T-HCRC.EDU.TW-NET
- TANET-BMETA
- TANET-BNETS Taiwan
- TANET-NET
- TANET-NET Taiwan
- TANET Taiwan
- TEIN2-CERNET
- THAINE-T-H
- TFN-NET TAI-SHIN-NET
- TRIUMF
- UAM
- UK-GEANT-20020131
- UNI Michigan
- UNIVERSIDADE DE SAO PAULO
- UNI of Nebraska-Lincoln
- University of Toronto
- IJS-IPV6-NET - Ljubljana
- VANDERBILT
- WIN-IP

---

Bruno Hoeft / Michael O’Connor / Richard Cziva / Samuel Ambroj Perez -- CHEP 2019 – Adelaide – 7th Nov. 19
unroutable IPv4 packets

- Reduced to 30 different sites only by
  - Combining private address areas
  - Removed sites with less than 1000 packets (per month)
  - Pull different subnets of one source/site together

total: 1.010.889 + 33582 = 1.044.471
packts per day: 19.707

ICMP approx. 43% = 459.508
none ICMP = 584.963
Reduced by FR-Renater and GARR (ICMP only)

632 different sources
min: 3 average: 53 max: 2910

16 different sources
min: 12 average: 6.834 max: 41.790

67 different sources
min: 2 average: 1.195 max: 31.734

33 different sources
min: 2 average: 12.907 max: 418.228

18 different sources
min: 5 average: 3.559 max: 22.647

18 different sources
min: 5 average: 3.559 max: 22.647

2 different sources

Reduced by Geant and FR-Renater and GARR (ICMP only)
**IPv6 Filter:** unroutable LHCONE packet

- filter Privat address space (link local)
- FR-Renater → ICMP packet of network devices
- IT-GARR → ICMP packet of network devices and the main injector (2 Perfsonar Server) are became part of LHCONE

<table>
<thead>
<tr>
<th>Source</th>
<th>IPv6 Address</th>
<th>Min.</th>
<th>Average</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEBO:7AB1:1</td>
<td>66005</td>
<td>1 source min: 66.005</td>
<td>208.693</td>
<td>1 source min: 208.693</td>
</tr>
<tr>
<td>RWTH AACHEN:2400:3960:0:1</td>
<td>1026</td>
<td>914626</td>
<td>11 different sources min: 4 average: 5.050 max: 23.712</td>
<td></td>
</tr>
<tr>
<td>IMPERIAL COLLEGE LONDON:2A0C:58004:1</td>
<td>728</td>
<td>55554</td>
<td>6 different sources min: 10 average: 152.437 max: 749.085</td>
<td></td>
</tr>
<tr>
<td>UNIVERSIDADE DE SAO PAULO</td>
<td>693</td>
<td>4640</td>
<td>2862</td>
<td></td>
</tr>
<tr>
<td>CERNET 2001:25:20:1</td>
<td>7681</td>
<td>37187</td>
<td>352512</td>
<td></td>
</tr>
<tr>
<td>DFN WIN-IPV6 2001:1:38:6:1</td>
<td>208.693</td>
<td>1 source min: 208.693</td>
<td>208.693</td>
<td>1 source min: 208.693</td>
</tr>
<tr>
<td>CERNET 2001:25:20:1</td>
<td>7681</td>
<td>37187</td>
<td>352512</td>
<td></td>
</tr>
<tr>
<td>CIEMAT</td>
<td>2001:720:400:1:1</td>
<td>2862</td>
<td>11 different sources min: 4 average: 5.050 max: 23.712</td>
<td></td>
</tr>
<tr>
<td>ROEDUNET-IPV6-NET-1</td>
<td>2001:830:10:1</td>
<td>2862</td>
<td>11 different sources min: 4 average: 5.050 max: 23.712</td>
<td></td>
</tr>
<tr>
<td>INFN-Padova</td>
<td>2001:1720:400:1:1</td>
<td>2862</td>
<td>11 different sources min: 4 average: 5.050 max: 23.712</td>
<td></td>
</tr>
<tr>
<td>INFN-Pisa</td>
<td>2001:1720:400:1:1</td>
<td>2862</td>
<td>11 different sources min: 4 average: 5.050 max: 23.712</td>
<td></td>
</tr>
</tbody>
</table>
Project @ DE-KIT: unroutable packet -- web portal

- automate unroutable packet information gathering
  - → store into an organized and structured database (kibana)
  - visualise the data (packet beat / elastic search / Kibana / grafana) with different levels
    - abstract overview
    - and zoomable into detailed view (up to source/dest. of a single packet)
      this data shall be available for the LHCONE connected sites (but not for the world),
      one idea:
    - community securing the data
    - restrict access --> personal authentication enabled (via eduGAIN)

project just started,
working on first results by the end of this year (2019)
Worldmap of misrouted packets captured during last 30 days
LHCONE regional different routing tables

Routing-tables of collected and analysed NSPs
Conclusion / actions

- Number of unroutable LHCONE packets reduced from 2.4 million to 1 million found 1.4 million packets marked as “false positive“ or could be sorted out
- LHCONE ingress filtering/control has improved dramatically since measurement began in Q1 2018!
  But still Detection will be beneficial:
  - Regularly scheduled monitoring?
  - Periodic NSP self run audits?

- Growth in Asia has likely contributed to a small number of exceptions.
  - Work towards Prevention
    - NSPs initiative at Edge Site filter configuration
      - RPF too strict? Rather opt for ACLs
      - Templated policy & filter configuration

- Routing table inconsistency may also be a source of false positives.

Bruno Hoeft / Michael O’Connor / Richard Cziva / Samuel Ambroj Perez  --  CHEP 2019 – Adelaide – 7th Nov. 19
Questions
Suggestions
Discussion
Backup Slids
Recent presentations

- unroutable LHCONE packets:
  - LHCOPN/ONE Meeting March 06, 2018
    https://indico.cern.ch/event/681168/contributions/2848474/attachments/1611723/2559528/LHCONE-Filter-Policy-Practice.pdf
  - LHCOPN/ONE Meeting Oct. 30, 2018
  - LHCOPN/ONE Meeting June 04, 2019
    https://indico.cern.ch/event/772031/contributions/3360612/attachments/1855532/3047503/LHCONE_Edge_Filtering_Policy_and_Practice_Umea_1.pdf

- regional LHCONE routing table differences:
  - LHCOPN/ONE Meeting Oct. 30, 2018
    https://indico.cern.ch/event/725706/contributions/3149436/attachments/1744301/2823417/LHCone_routing_digging.pdf
  - LHCOPN/ONE Meeting June 04, 2019
An exchange is like an NSP:
- BGP import filtering
- Packet filtering
- Community based BGP filtering

An exchange is like a site:
- Require the full LHCONE table via a transit NSP
- Packet filters are configured and require maintenance

Indiana GigaPOP is a current ESnet example. SOX is planned to be the second and will connect UFL, FSU and others.

- Will L3 Exchange Fabrics implement and maintain LHCONE specific services?
- Should there be an LHCONE defined role for these network organizations?
- Are they permitted to attach new sites?
private

grouping different CIDR of one organization

removing sites with less than 1000 unroutable packets over four weeks
Routing table differences e.g. Géant (IPv6)

No change between Oct. 2018 and June 2019

Missing:
- AARNET
  - Via Internet2
  - ESnet
- AMPRO
  - via ESnet