

LHCOPN routing wg

Seattle - 14th November 2005

edoardo.martelli@cern.ch

Network dedicated to LHC data (while connected end hosts might not be dedicated to LHC).

Primary and backup T0-T1 connections: 10G lightpaths between the T0 and T1 routers.

Pre production backup: path across IP networks.

Direct eBGP peering between T0's and T1's routers to exchange only the LHCOPN prefixes.

Optional T1-T1 connectivity: via T0 or direct lightpath.

LHCOPN architecture – IP addressing

- Every T1 and the T0 must allocate publicly routable IP address space to the machines that need to be reached over the T0-T1 links (the "LHC prefixes", as defined in the Introduction).
- LHC prefixes should be aggregated into a single CIDR block for every T1; if this is not possible, only a very small number of CIDR blocks per T1 would be acceptable.
- LHC prefixes should be dedicated to the LHC network traffic.
- If a T1 can't dedicate a prefix but need to use its whole address space, then it will be its own responsibility to inject in the LHCOPN only the traffic destined to the LHC-OPN resources.
- LHC prefixes can be carved as a CIDR block from a T1's existing allocations or obtained as new allocation from the appropriate RIR through already established channels.
- LHC prefixes cannot be from RFC1918 and related (like RFC3330) addresses.
- T0 will allocate /30 prefixes for the addressing of the T0-T1 links, i.e. the links that connect to CERN up to the first BGP speaker in the path.
- Every T1 and T2 interested in exchanging traffic directly with the T0 is required to provide the T0 with the list of its LHC prefixes. T0 will maintain a global list of all LHC prefixes and inform T1s and T2s about any changes.

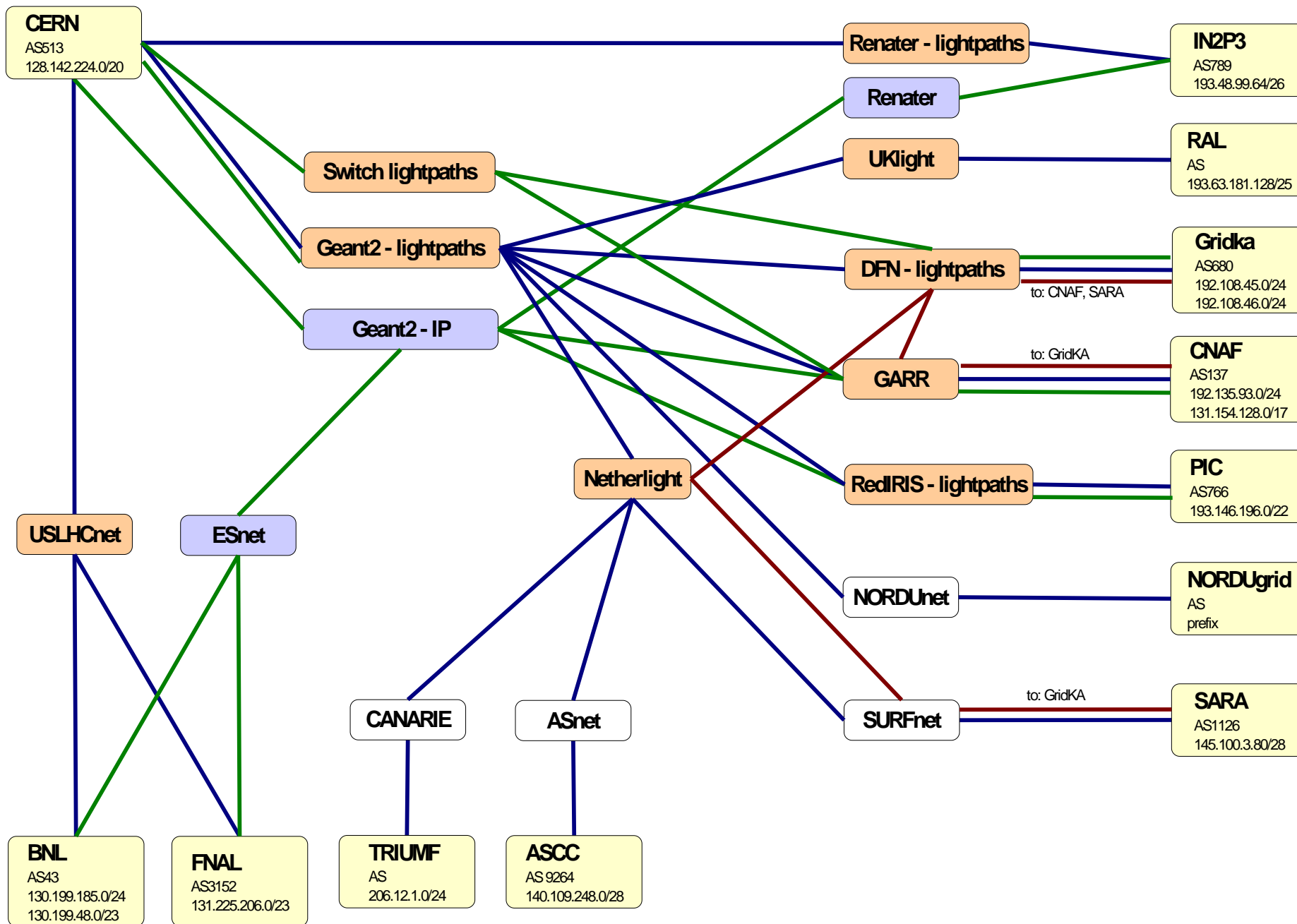
External BGP peerings will be established between T0 and each T1. More precisely, the T1 peer is the BGP speaker directly connected to the T0 on behalf of a specific T1, e.g. an NREN connecting a T1 or the BGPcapable gigabit Ethernet switch of the T1.

Guidelines for the BGP configuration:

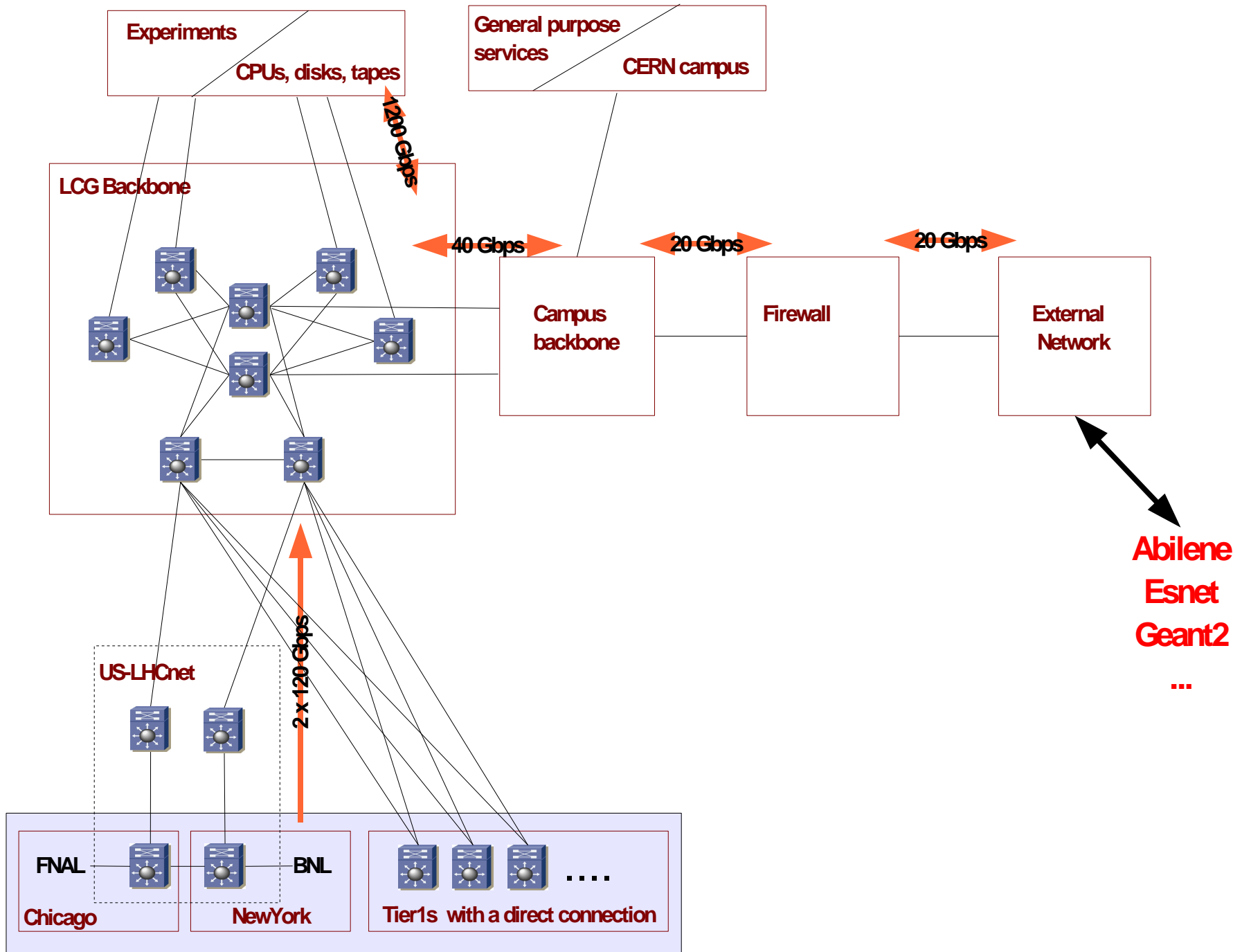
- T0 will use the CERN Autonomous System number (AS513).
- T1s will use the AS number of the entity that provides the LHC prefixes to them or the AS number of their standard upstream NREN.
- Every T1 will announce its own LHC prefixes to T0.
- T0 will announce its LHC prefixes to every peering T1.
- T0 will accept only the LHC prefixes related to a specific T1, i.e. the T1's own LHC prefixes, plus LHC prefixes of any T1 or T2 for which that T1 is willing to provide transit for.
- T0 will re-announce to all the T1s all the LHC prefixes received in BGP. Nevertheless, since T1s are encouraged to establish direct connectivity among themselves, they can filter out unnecessary LHC prefixes according to each individual T1-T1 routing policy.
- T1 will accept T0's prefixes, plus, if desired, some selected T1's prefixes (see previous bullet).

- T0 and T1s should announce their LHC prefixes to their upstream continental research networks (GÉANT2, Abilene, ESnet) in order to allow connectivity towards the T2s. Special care must be taken by each Tier to not leak out BGP prefixes that belong to other Tiers.
- Usage of static routes is not advisable.
- No default route must be used in T1-T0 routing.
- It is the responsibility of every Tier to make sure that any of its own machines within the LHC prefix ranges can reach any essential service (for instance the DNS system).

LHCOPN Overview



Tier0 - CERN



Technical contact: Edoardo Martelli, extip@cern.ch

Operations contact: noc@cern.ch +41227674011 (24x7)

AS number: AS513

LHC prefixes: 128.142.224.0/20

Notes:

- Force10 E1200 routers. LANPHY (SX, LX, EX) and WANPHY interfaces available.

Technical contact: You Lin Chang llchang@gate.sinica.edu

Operations contact:

Primary link: Concatenad lightpath: Taipeh-Amsterdam (10G?,
Provider?) - Amsterdam-Geneva (10G, GN2).- WANPHY?

Backup link:

AS number: AS9264

LHC prefixes: 140.109.248.0/28

Notes:

Technical contact: John Bigrow big@bnl.gov

Operations contact:

Primary link: Concatenated lightpath: Brookhaven-NYManlan (10G? Esnet), NYManlan-Geneva (10G, USLHCnet). LANPHY?

Backup link:

AS number: AS43

LHC prefixes: 130.199.185.0/24 - 130.199.48.0.23

Notes:

Technical contact: Marco.Marletta@garr.it

Luca.DellAgnello@cnafe.infn.it

Operations contact:

Primary link: Concatenated Lightpath: Bologna-Milano (10G, GARR), Milano-Geneva (10G, GN2). 10G-baseSX??

Backup link: Concatenated lightpath: Bologna-Milano (?), Milano-Zurich-Geneva (10G, SWITCH). LANPHY?

AS number: AS137

LHC prefixes: 131.154.128.0/17 - 192.135.23.0/24

Notes:

- Direct link with Gridka
- Peering with the GARR router in Milano
- Policy Based Routing used inside GARR network

Technical contact: Vyto Grigaliunas vyto@fnal.gov, Andrey Bobyshew bobyshew@fnal.gov

Operations contact: helpdesk@fnal.gov +1-630-840-2345 (8x5)

Primary link: Concatenated lightpath: Batavia-Chicago (10G, FNAL), Chicago-Geneva (10G, USLHCnet). 10GbaseLX.

Backup link: L3 path over ESnet and GN2

AS number: AS3152

LHC prefixes: 131.225.204.0/22

Notes:

- Policy Based Routing used inside FNAL network

Technical contact: Bruno.Hoeft@iwr.fzk.de

Operations contact:

Primary link: Concatenated lightpath: Karlsruhe-Frankfurt(10G, DFN), Frankfurt-Geneva (10G, GN2). 10GbaseLX.

Backup link: Concatenated lighpath provided bt DFN and SWITCH.

Interface?

AS number: AS34878

LHC prefixes: 192.108.45.0/24 192.108.46.0/24

Notes:

- Direct link with GARR and SARA
- Policy Based Routing used inside FZK network

Technical contact: Jerome.Bernier@in2p3.fr - Simon Frank
simon@renater.fr

Operations contact: noc@in2p3.fr +41 33 78 93 08 80

Primary link: Direct lightpath: Lyon-Geneva (10G, Renater).
10GbaseSX??

Backup link: L3 path over Renater and Geant

AS number: AS789

LHC prefixes: 193.48.99.64/26

Notes:

Technical contact: Per Nihlen

Operations contact:

Primary link: Concatenated lightpath

Backup link:

AS number:

LHC prefixes:

Notes:

Technical contact: Maripaz de Andres mpandres@pic.es

Operations contact:

Primary link: Concatenated lightpath: Barcelona-Madrid (10G, Rediris), Madrid-Geneva (10G,GN2). Interface?

Backup link: MPLS tunnel over Rediris + GN2 (MPLS or IP or lightpath?)

AS number: AS766

LHC prefixes: 193.146.196.0/22

Notes:

Technical contact:

Operations contact:

Primary link: Concatenated lightpath? Interface?

Backup link:

AS number:

LHC prefixes: 193.63.181.128/25

Notes:

Technical contact: Pieter de Boer pieter@sara.nl -

Hanno.Pet@sara.nl

Operations contact: nic@sara.nl, nrg@sara.nl +312-5028000

Primary link: Direct lightpath: Amsterdam-Geneva (10G,
Netherlight). WANPHY

Backup link:

AS number: AS1126

LHC prefixes: 145.100.3.80/28

Notes:

Technical contact: Steven.McDonald@triumf.ca

Operations contact:

Primary link: Concatenated lightpath: Vancouver-Amsterdam (10G, ?), Amsterdam-Geneva (10G, GN2). Interface?

Backup link:

AS number:

LHC prefixes: 206.12.1.0/24

Notes:

CERN's Force10s have already been installed. One is currently used for Service challenge 3.

T0-T1 links will be immediately used for Service Challenges as soon as they are available. CNAF will be connected next Friday; FNAL, IN2P3, SARA will soon follow.