WLCG networking

19nd September 2024

edoardo.martelli@cern.ch

Network connectivity

Computing and storage resources at any WLCG site:

- Must be reachable over the **Internet**

- Should be connected to **LHCONE** (highly recommended for Tier1s)

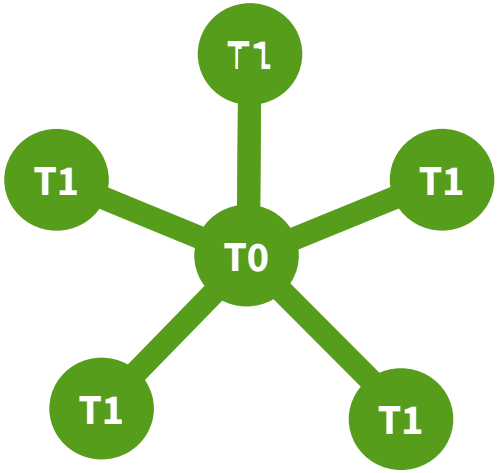
- Tier1s must have a direct link to CERN (**LHCOPN**)

LHCOPN

LHCOPN

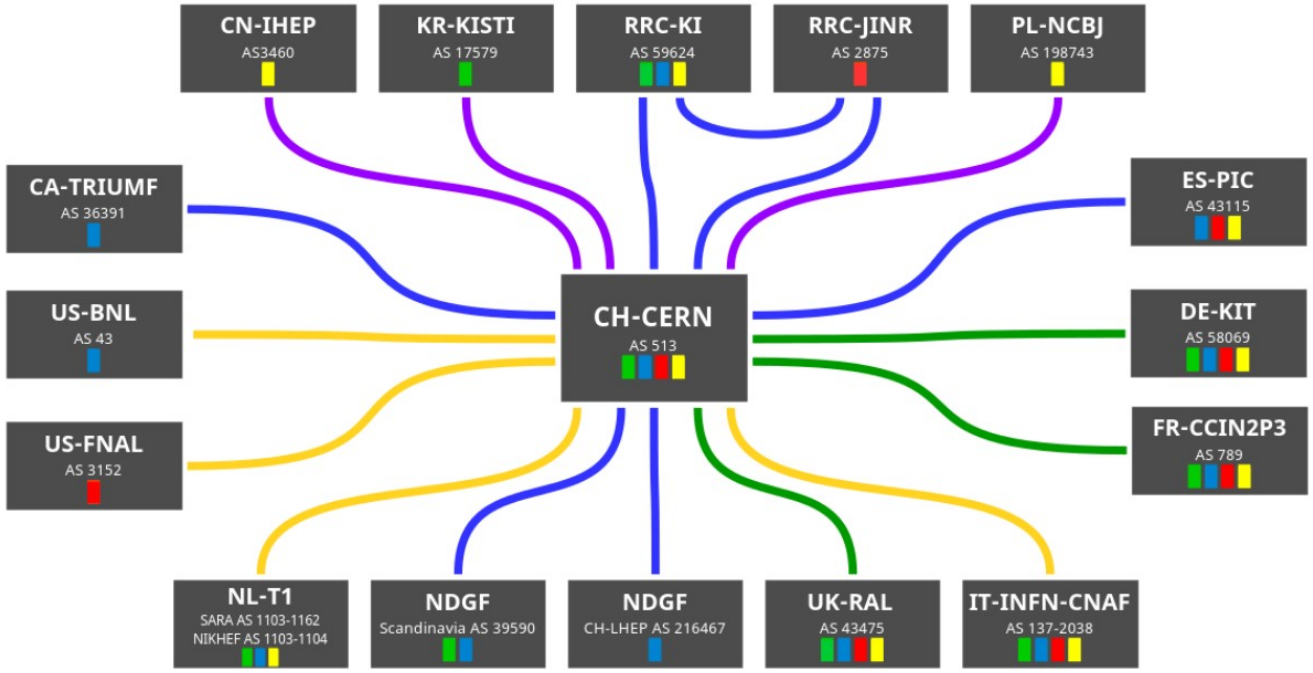
Private network connecting Tier0 and Tier1s

Secure:
- Dedicated to LHC data transfers
- Only declared IP prefixes can exchange traffic
- Can connect directly to Science-DMZ,
  bypass perimeter firewalls



Advanced routing:
- BGP communities for traffic engineering

LHCOPN

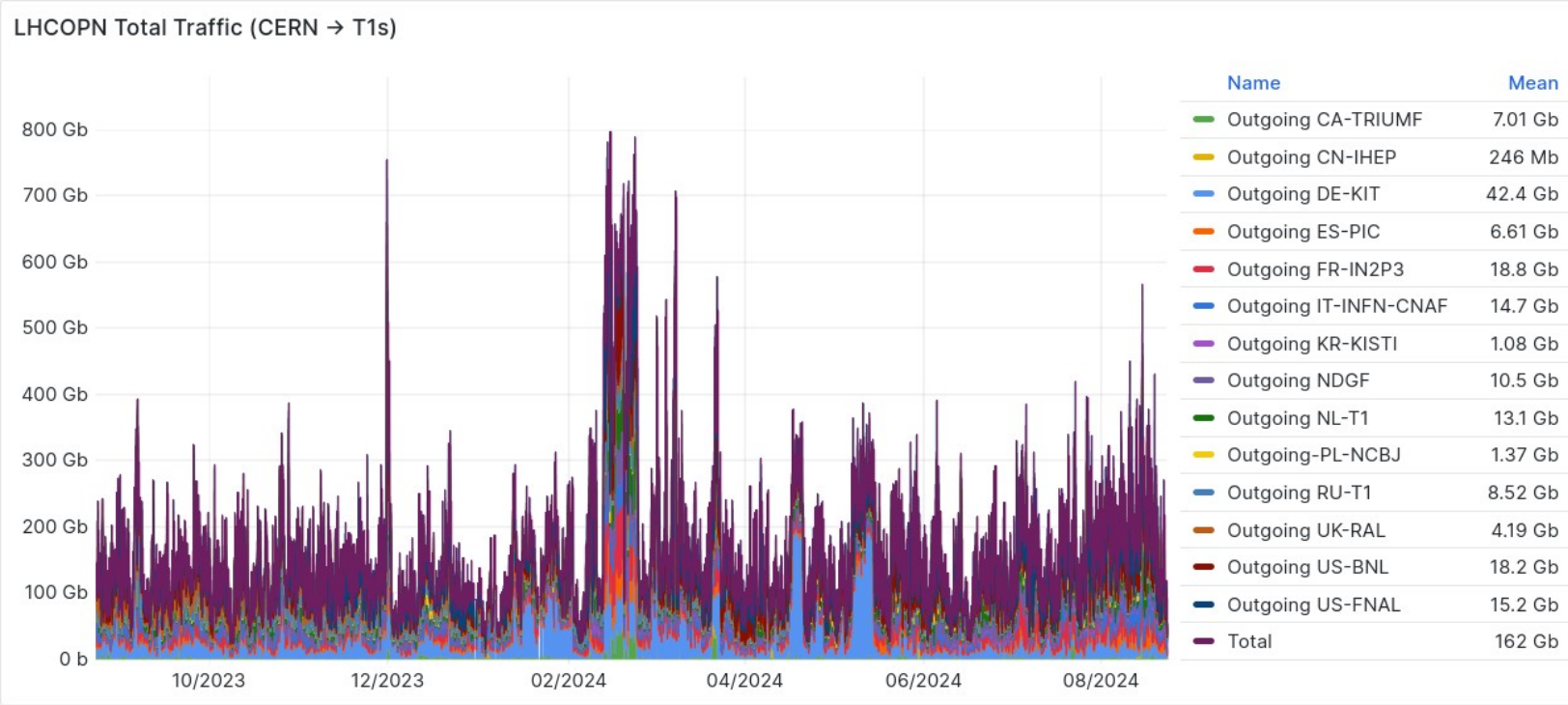


https://twiki.cern.ch/twiki/bin/view/LHCOPN/OverallNetworkMaps

Numbers

- 17 sites for 15 Tier1s + 1 Tier0
- 14 countries in 3 continents
- 2.86 Tbps to the Tier0
- CN-IHEP, PL-NCBJ and NDFG-LHEP last additions
- TW-ASGC has left

LHCOPN Traffic – last 12 months



Numbers:

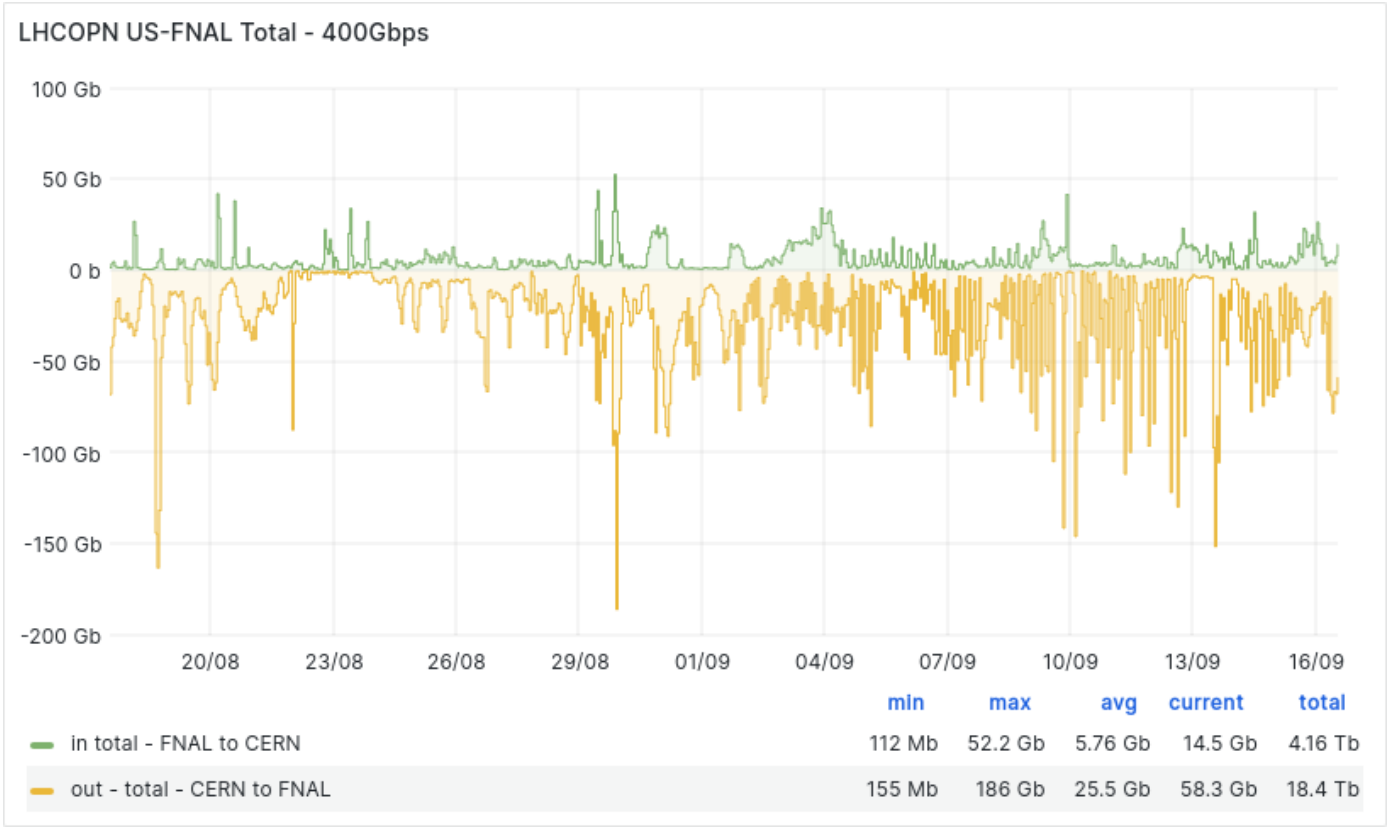
Moved ~638 PB in the last 12 months

+18% compared to previous year (540PB)

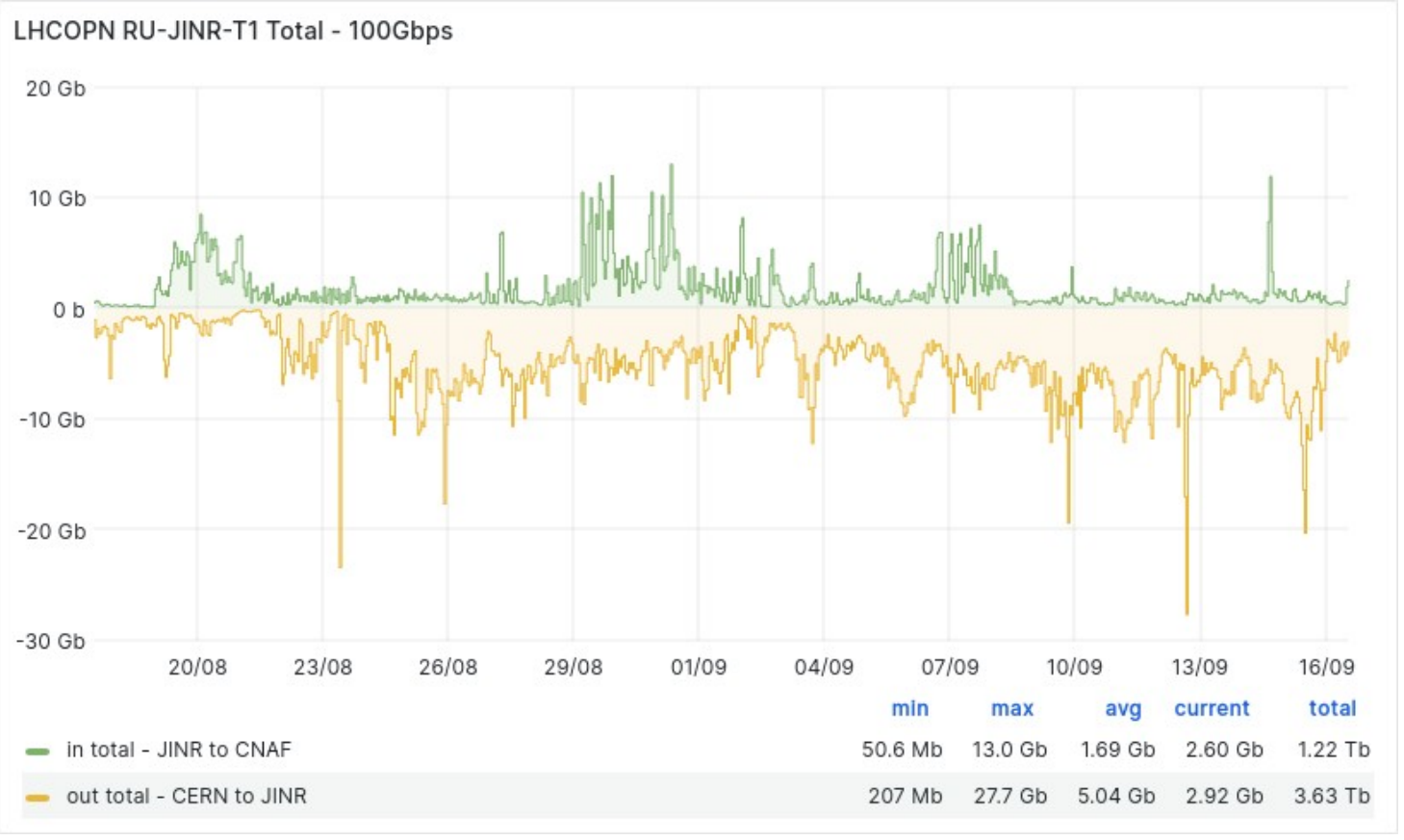
Peak at ~800Gbps during DC24

Ref: https://monit-grafana-open.cern.ch/d/HreVOyc7z/all-lhcopn-traffic?orgId=16

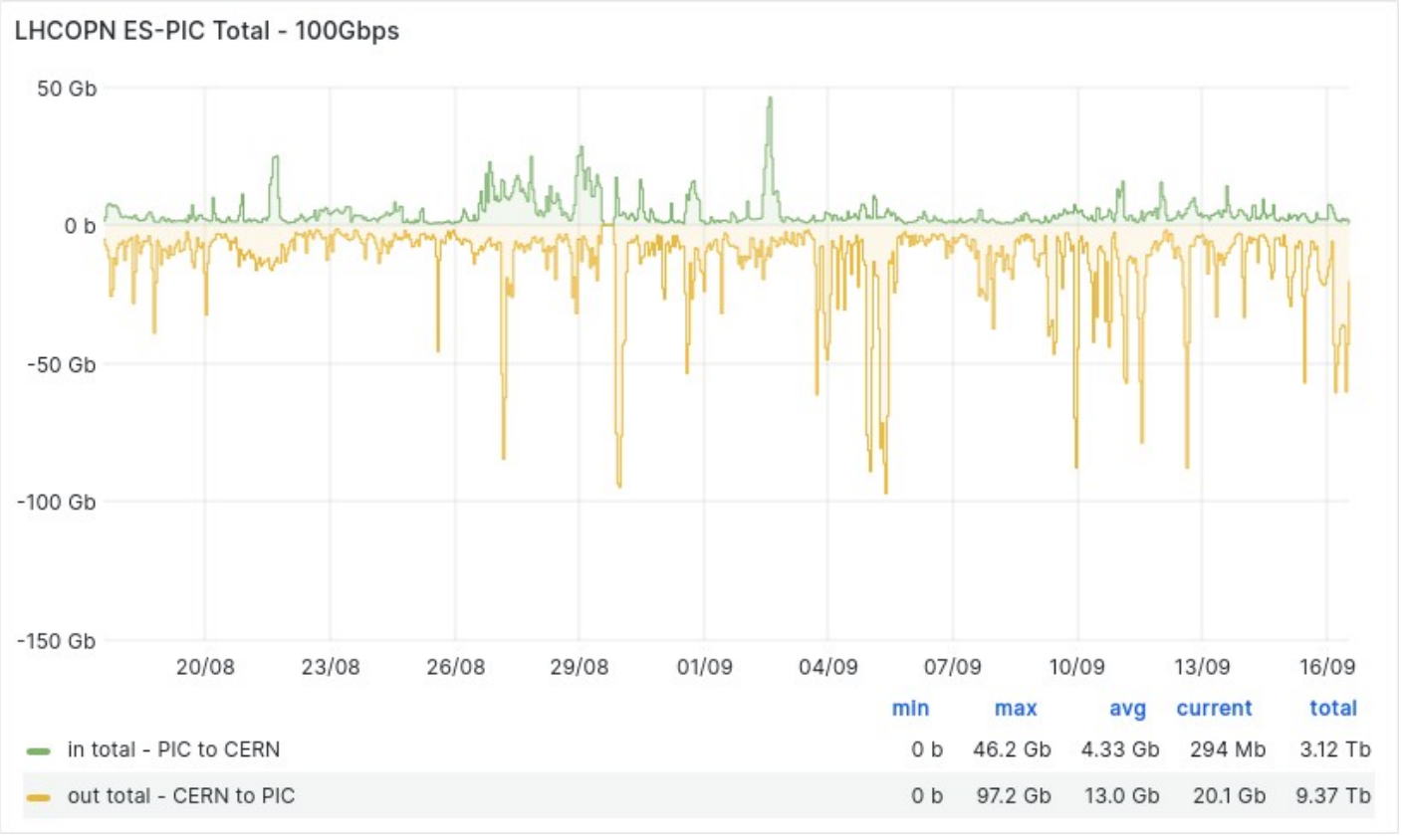
FNAL USA - CMS


LHCOPN US-FNAL Total - 400Gbps
100 Gb
50 Gb
0 b
-50 Gb
-100 Gb
-150 Gb
-200 Gb
20/08
23/08
26/08
29/08
01/09
04/09
07/09
10/09
13/09
16/09
min
max
avg
current
total
in total - FNAL to CERN
112 Mb
52.2 Gb
5.76 Gb
14.5 Gb
4.16 Tb
out - total - CERN to FNAL
155 Mb
186 Gb
25.5 Gb
58.3 Gb
18.4 Tb

JINR Russia - CMS



PIC Spain – ATLAS, CMS, LHCb



LHCONE

LHCONE L3VPN service

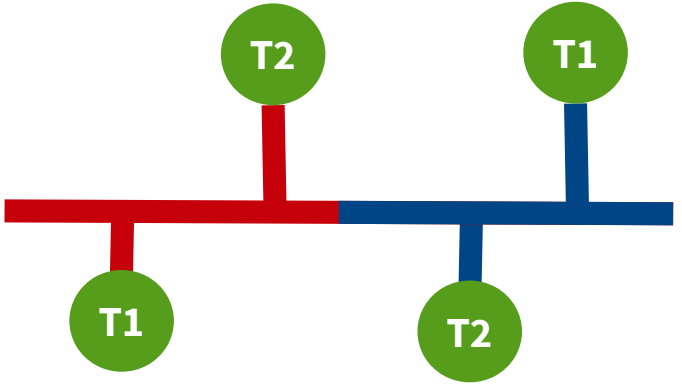
Private network connecting Tier1s and Tier2s

Secure:

- Dedicated to LHC data transfers
- Only declared IP prefixes can exchange traffic
- Can connect directly to Science-DMZ, bypass perimeter firewalls

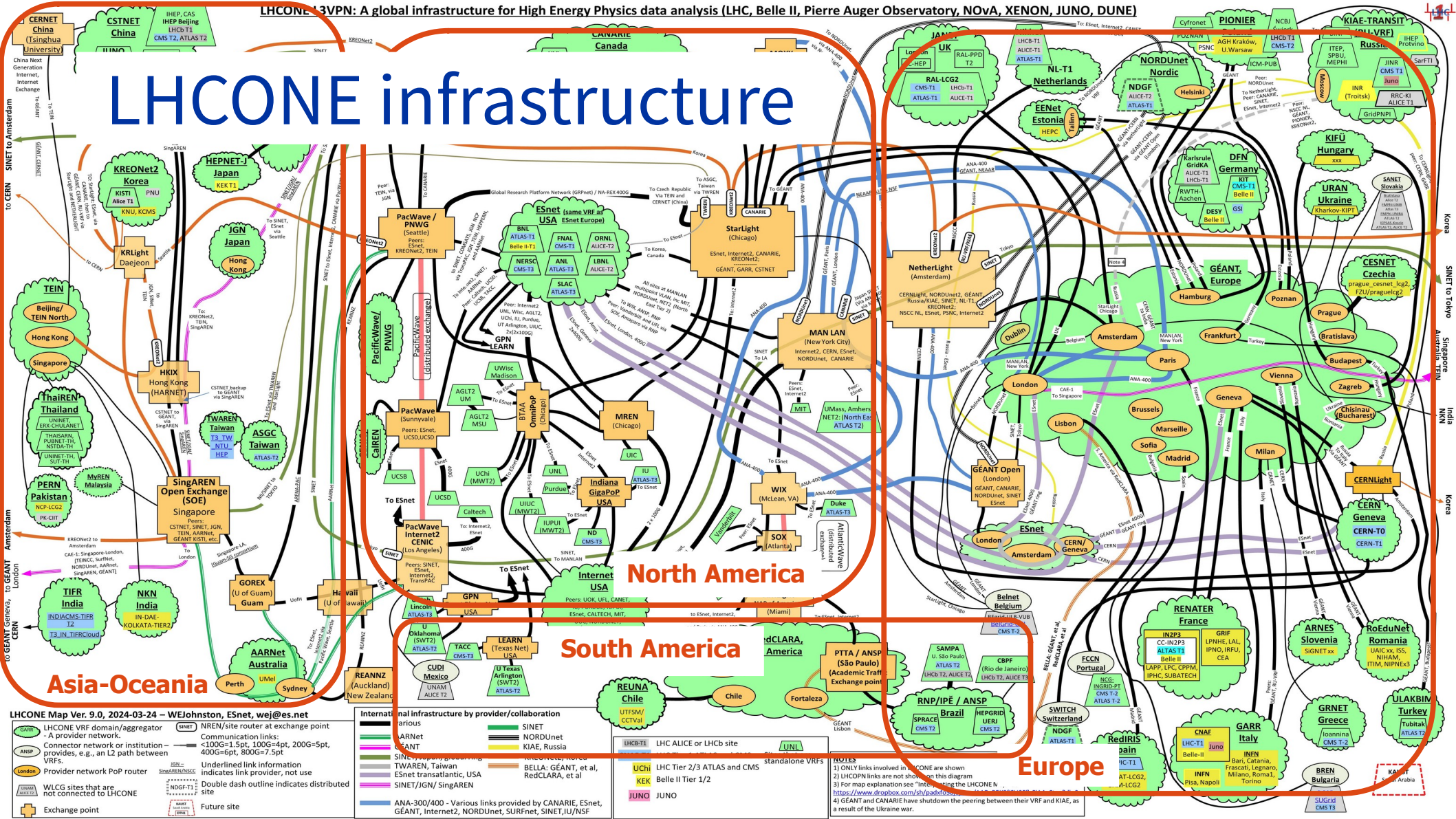
Advanced routing:

- Multi domain L3 VPN
- BGP communities for traffic engineering



Open to other HEP collaborations

LHCONE infrastructure



LHCONE status

- VRFs: 30 national and international Research Networks

- Connected sites: ~110 in Europe, North and South America,Asia, Australia

- Trans-Atlantic connectivity provided by ESnet, GEANT, Internet2, RedCLARA, NORDUnet, CANARIE and SURF

- Trans-Pacific connectivity provided by KREOnet, SINET, TransPAC

- Interconnections at Open Exchange Points including NetherLight, StarLight, MANLAN, WIX, CERNlight, Hong Kong, Singapore and others

Network information in CRIC

CRIC (Computing Resources Information Catalogue) is the database used by WLCG to document the available resources. It is used also to store network information related to LHCOPN and LHCONE

Easily accessible

- Netsite: https://wlcg-cric.cern.ch/core/netsite/list/ (login required)
- NetworkRoute: https://wlcg-cric.cern.ch/core/networkroute/list/ (login required)
- Json view: https://wlcg-cric.cern.ch/api/core/rcsite/query/?json (no login)

How to connect

Prerequisites

A WLCG facility must have:

- **Public IPv4 and IPv6 addresses**
- **A public Autonomous System number (ASN, a network identifier assigned by RIPE)**
- **A BGP capable router**

All of them can be provided by the local National Research & Education Network (NREN) provider

Physical connection

The WLCG facility has to:

- agree with the local NREN about how and where to connect to closest Point of Presence (PoP)
- LHCOPN: request a direct link to CERN
- LHCONE: connect to the closest LHCONE access router

LHCONE Acceptable Use Policy (AUP)

Use of LHCONE must be restricted to WLCG related traffic

IP addresses announced to LHCONE:

- should be assigned only to WLCG servers
- cannot be assigned to generic campus devices (desktop and portable computers, wireless devices, printers, VOIP phones....)

https://twiki.cern.ch/twiki/bin/view/LHCONE/LhcOneAup

LHCONE Routing setup

- A BGP peering is established between the WLCG site router and the closest LHCONE access router

- The site router announces only the IP subnets used for WLCG servers

- The site router accepts all the prefixes announced by the LHCONE access router

- The site router must ensure traffic symmetry: injects only packets sourced by the announced subnets

LHCOPN Routing setup

- A BGP peering is established between the WLCG site router and the CERN LHCOPN border router

- The WLCG site router announces only the IP subnets used for WLCG servers

- The WLCG site router accepts all the prefixes announced by the CERN router

- The WLCg site router must ensure traffic symmetry: injects only packets sourced by the announced subnets

- LHCOPN can be used to reach all the other Tier1s (transit via CERN)

AMRES Serbian NREN

AMRES should be the main contact to provide LHCOPN and LHCONE connectivity

LHCONE:

- AMRES can implement an LHCONE instance for all Serbian sites and then connect it to the LHCONE backbone provided by GEANT

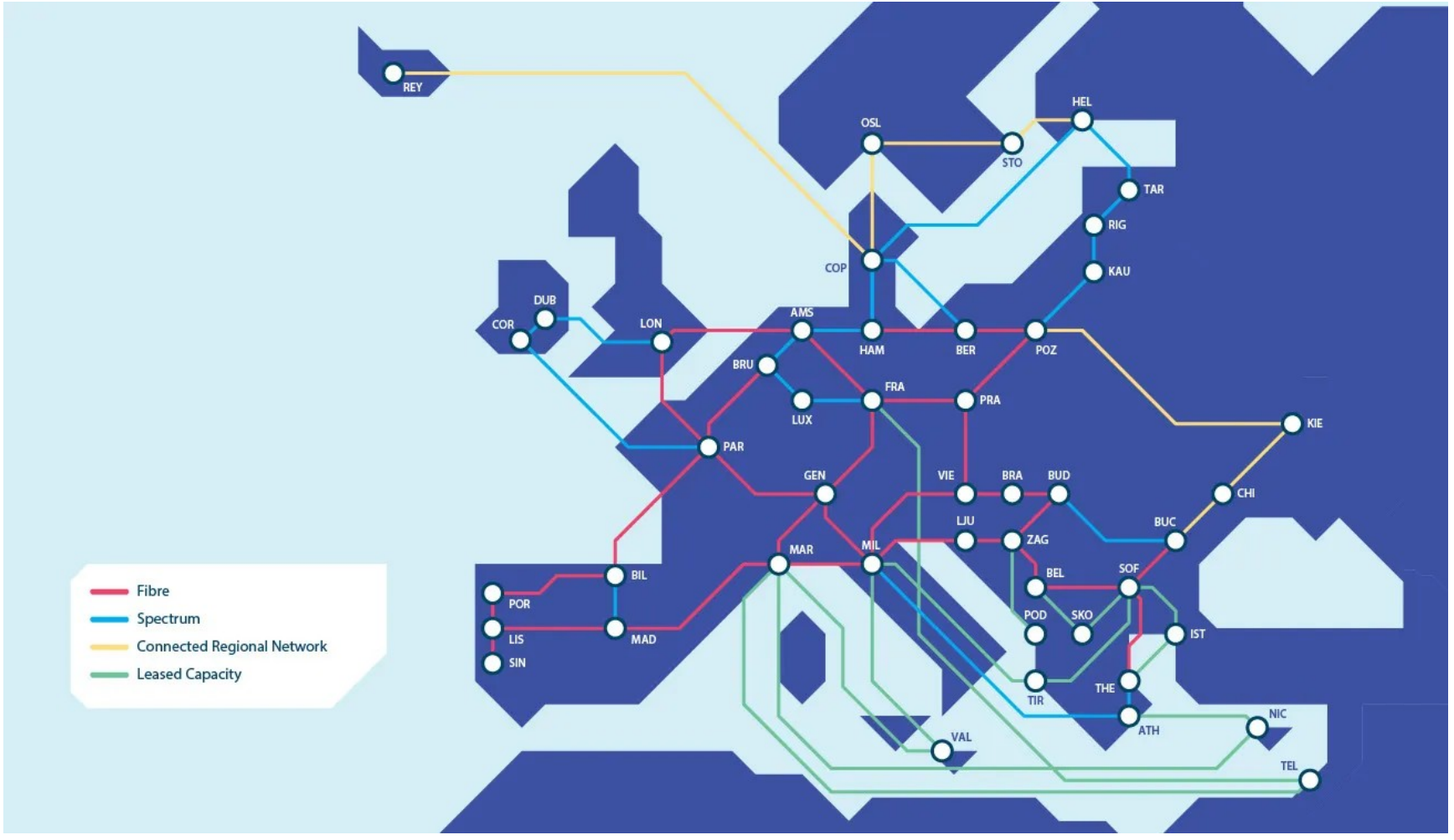
or

- AMRES can provide a L2 connection from the data-centre to the closest GEANT LHCONE access router

LHCOPN:

- AMRES should work with GEANT to provide the direct link form the Tier1 data-centre to CERN

GEANT


REY
OSL
HEL
STO
TAR
RIG
KAU
COP
DUB
COR
LON
AMS
HAM
BER
POZ
BRU
FRA
PRA
LUX
KIE
PAR
GEN
VIE
BRA
BUD
CHI
LJU
ZAG
BUC
MIL
MAR
BIL
BEL
SOF
POR
POD
SKO
IST
LIS
MAD
SIN
THE
TIR
ATH
NIC
VAL
TEL
Fibre
Spectrum
Connected Regional Network
Leased Capacity

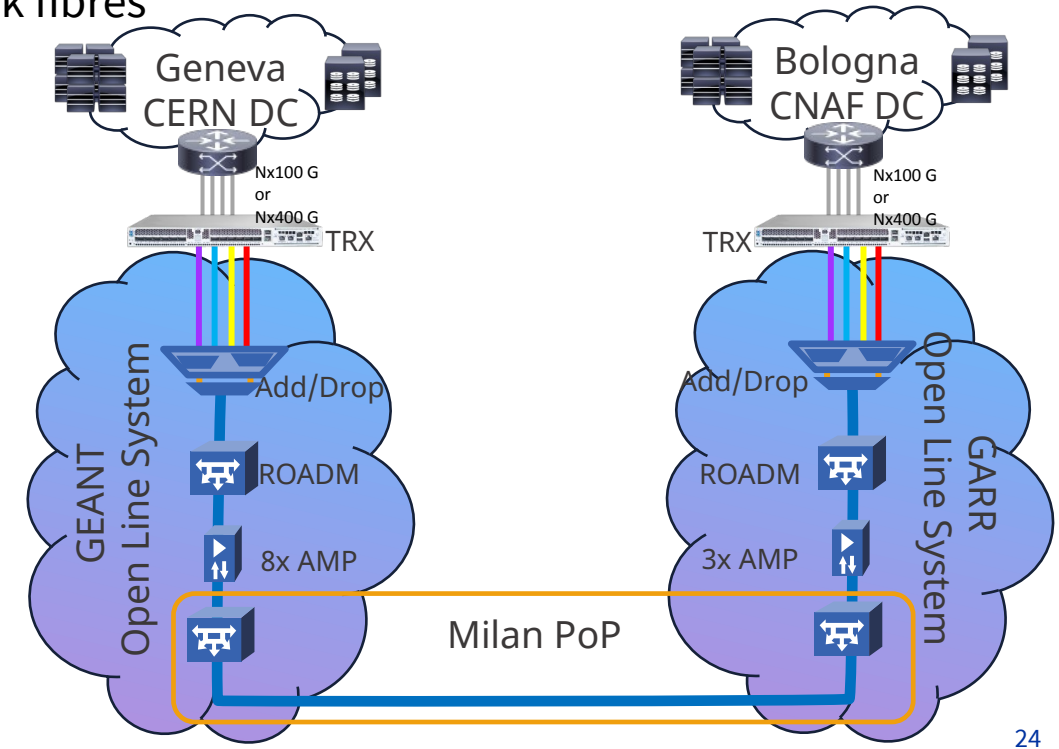
CNAF-CERN DCI

The LHCOPN link of IT-INFN-CNAF Tier1 is implemented using shared spectrum

- spectrum sharing over GEANT and GARR dark fibres

- 4x100Gbps links between CERN and CNAF used for DC24 and now in production

- cost effective technique to get >1Tbps LHCOPN connections already today



WLCG Data Challenges

Data Challenges for HL-LHC

WLCG has planned for a series of data challenges to prepare for HL-LHC data taking

- Demonstrate readiness for the expected HL-LHC data rates with:
  - Increasing volume/rates
  - Increase complexity (e.g. additional technology)
- A data challenge roughly every two years

2021: 10% of HL-LHC requirements *(480Gbps minimal – 960Gbps flexible)*
2024: 25% of HL-LHC requirements *(1.2Tbps minimal – 2.4Tbps flexible)*
2026/7: 50% of HL-LHC requirements *(date and % to be confirmed)*
2028/9: 100% of HL-LHC requirements *(date and % to be confirmed)*
2030: start of HL-LHC (Run4) *(4.8Tbps minimal – 9.6Tbps flexible)*

HL-LHC network requirements

ATLAS & CMS T0 to T1 per experiment

- 350PB raw data per year; average of 50GB/s or 400Gbps during LHC running time
- Another 100Gbps estimated for prompt reconstruction data tiers (AOD, other derived output)
- estimated 1Tbps for CMS and ATLAS summed

ALICE & LHCb T0 Export

- 100 Gbps per experiment, estimated from Run-3 rates

Minimal Model

- Sum (ATLAS,ALICE,CMS,LHCb)*2(for bursts)*2(safety-margin) =

  4.8Tbps expected HL-LHC bandwidth

Flexible Model

- Experiments may need to reprocess and reconstruct the collected data during the year
- This requires doubling the bandwidth of the Minimal model:

  9.6Tbps expected HL-LHC bandwidth

Overall network requirements for HL-LHC

Each Major Tier1s:

1 Tbps to the Tier0 (LHCOPN)

1 Tbps to the Tier2s (aggregated, LHCONE)

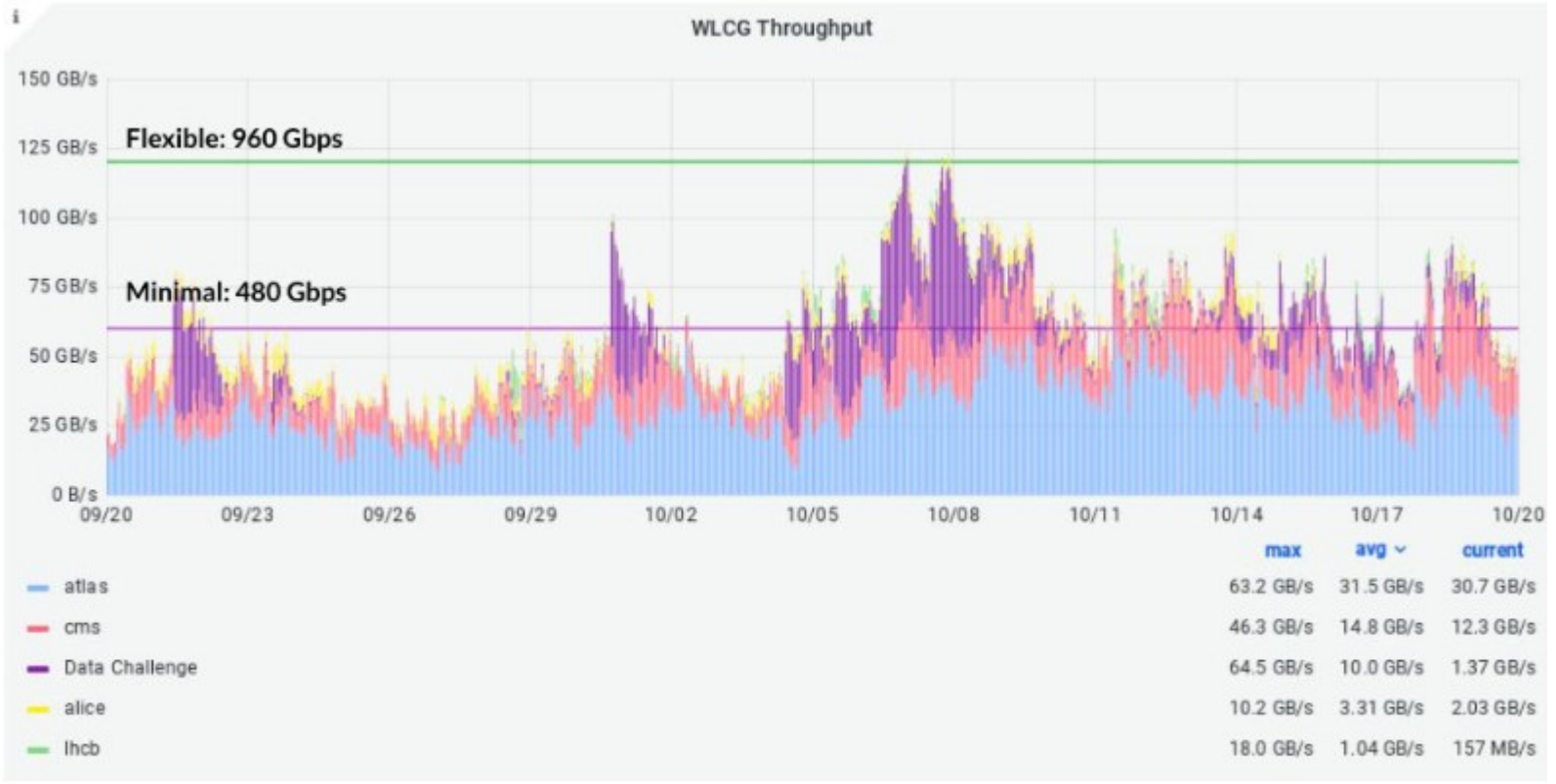
Each Major Tier2s:

>400 Gbps (LHCONE)

Over provisioning main not always be an option on transoceanic routes. More efficient technology may be needed

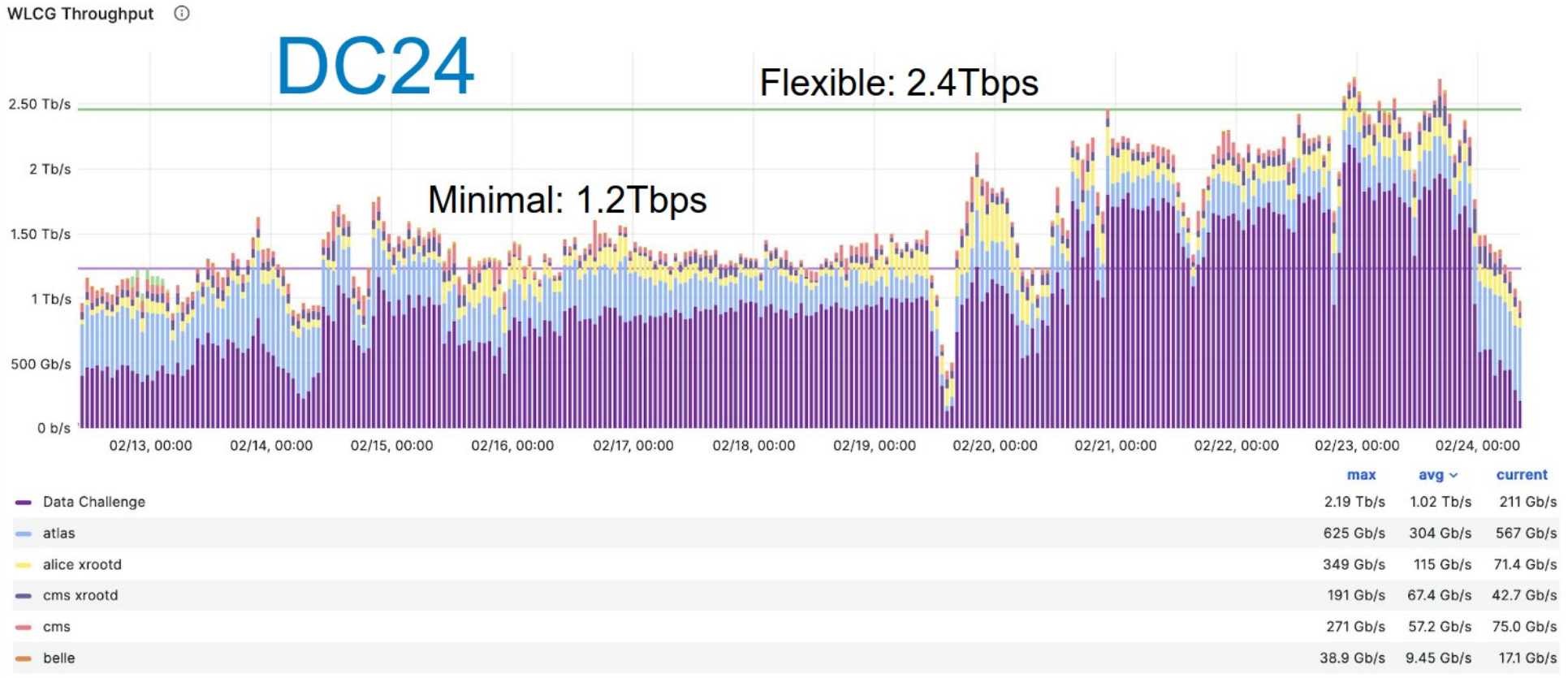
DC21

Mostly production transfers. Flexible model reached for a short time



DC24

Achieved full throughput of minimal model. Flexible model achieved only for short time.



DC24 results: applications

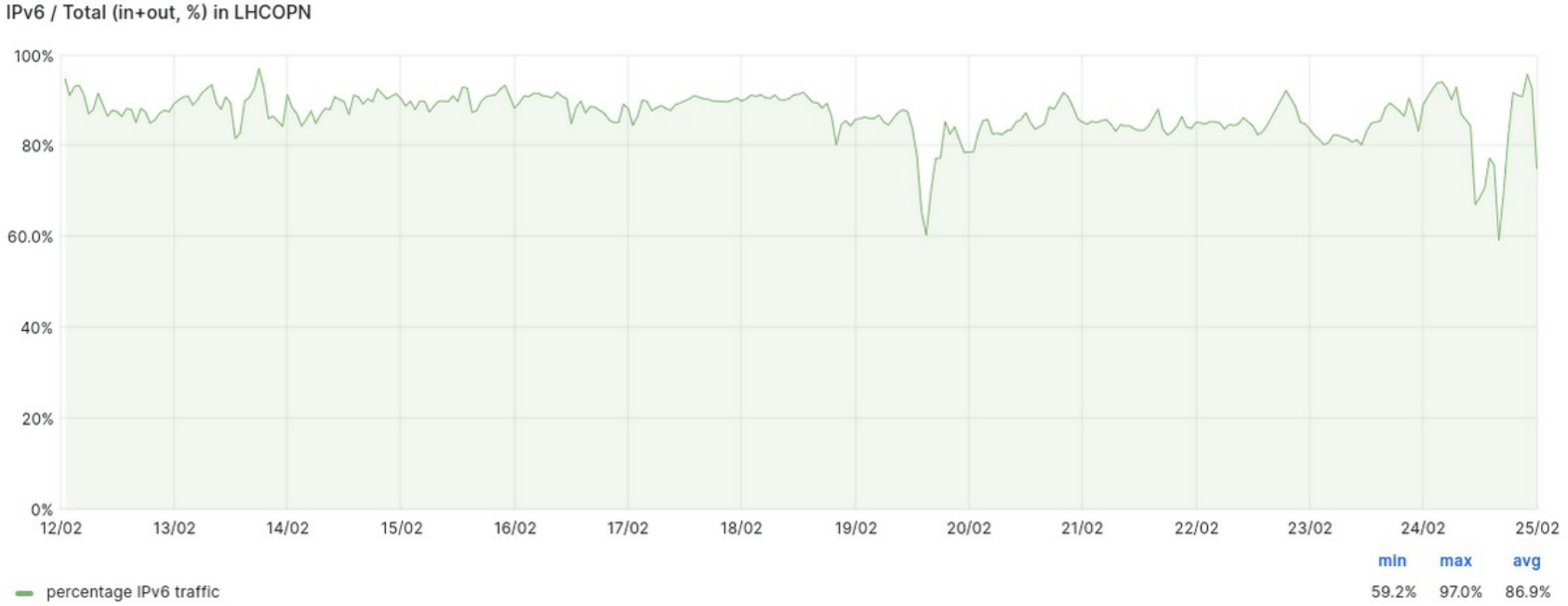
- Ran 12-23 February 2024
- Real data moved disk-to-disk using production applications: ad-hoc transfers on top of production traffic
- Applications pushed to uncharted levels showed unexpected limitations
- Using tokens for authorization added instabilities, but testing at such scale it was a necessary step to take

DC24 results: Networking

- **Global Research & Educations (REN) networks demonstrated more than sufficient capacity and reliability during DC24 and were NOT a bottleneck for any of the experiments.** Unexpected undersea cable cuts properly backed-up by alternative connectivity
- Some sites did identify local network bottlenecks or non-optimal architectures
- Various network technologies (NOTED, SENSE, BBR, perfSONAR, SciTags, Spectrum sharing…) were successfully tested during DC24 and showed promising results, now working to put them into production.
- Efforts to better monitor networks has become part of WLCG operational toolkits
- Regular mini-challenges will be run to track progress and prepare for DC26/7
- Storage infrastructure and middleware are being improved and the networks need to keep pace. Bandwidth upgrades may be needed for DC26/7

DC24 Network measurements: IPv6

- IPv6 Traffic in LHCOPN: 86.9% of the total
- Some sites testing IPv6 because of IPv4 scarcity



Resources

References

LHCOPN: https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome

LHCONE: https://twiki.cern.ch/twiki/bin/view/LHCONE/WebHome

LHCOPN/ONE CERN monitoring:
https://monit-grafana-open.cern.ch/d/000000523/home?orgId=16

Meetings

LHCOPN and LHCONE meetings:

- Every six months
- Upciming meetings:
    - 9-11 of October 2024 in Beijing CN
    - April 2025 in UK

Questions?

edoardo.martelli@cern.ch