CERN, LHCOPN and LHCONE networks update

HEPIX Fall 2023, Victoria - 18 October 2023
edoardo.martelli@cern.ch

Agenda

- CERN Networking

- LHCOPN

- LHCONE

- DC24 and Network R&D

CERN networking - update

Highlights

Campus Network:
- Completed campus upgrade of routers and switches
- Deploying virtual multi-domain service with BGP EVPNs
- Evolving WIFI users admission

Technical Network (LHC control network):
- Evaluating deployment of EVPN service to increase resiliency

New Prevessin Data Centre (PDC):
- October 2023: data-centre building, cooling, power, fibres ready
- November 2023 : installation of network equipment and first servers
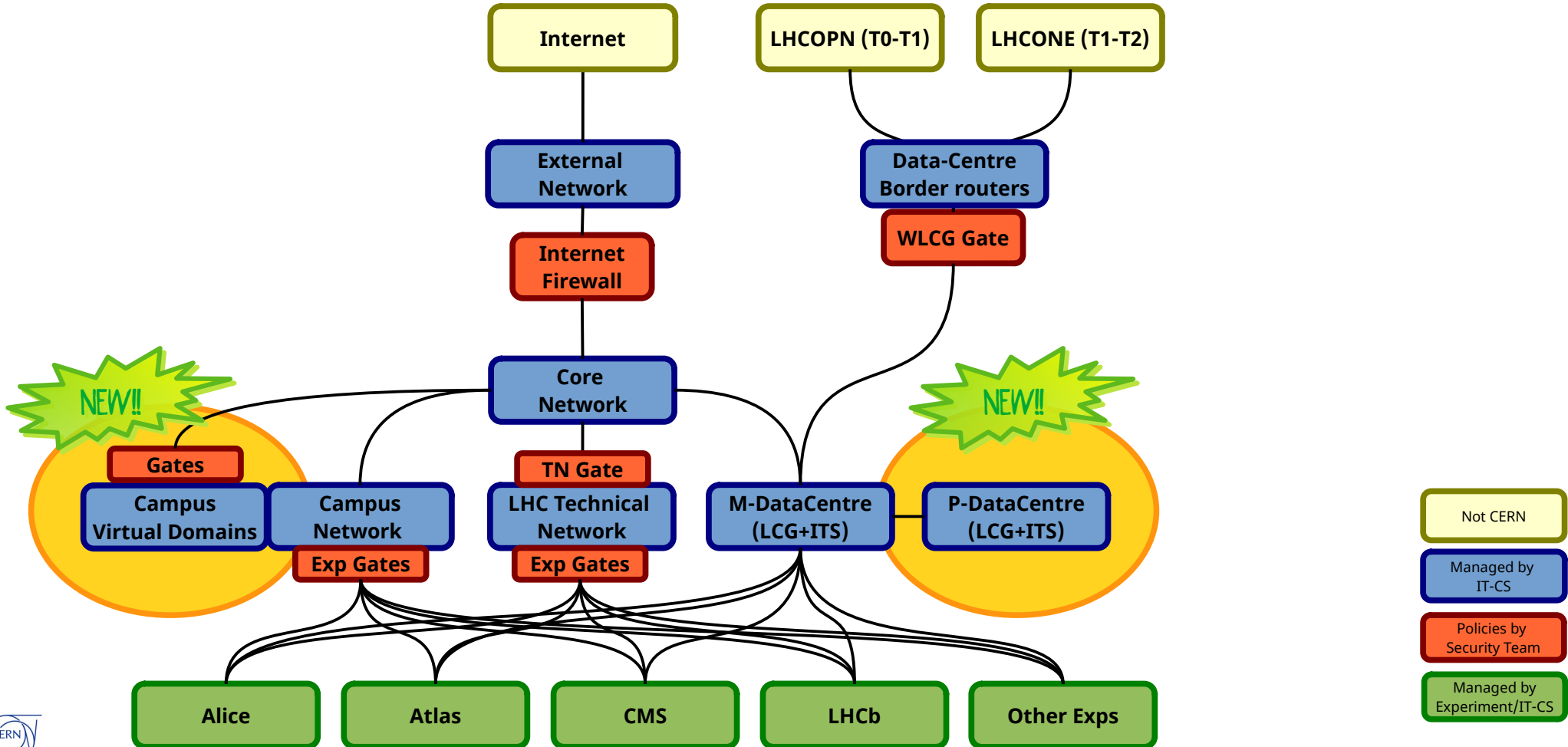- Q2 2024: ready for production

Telephony
- New GSM contract awarded to Swisscom (CH) and Orange (FR)
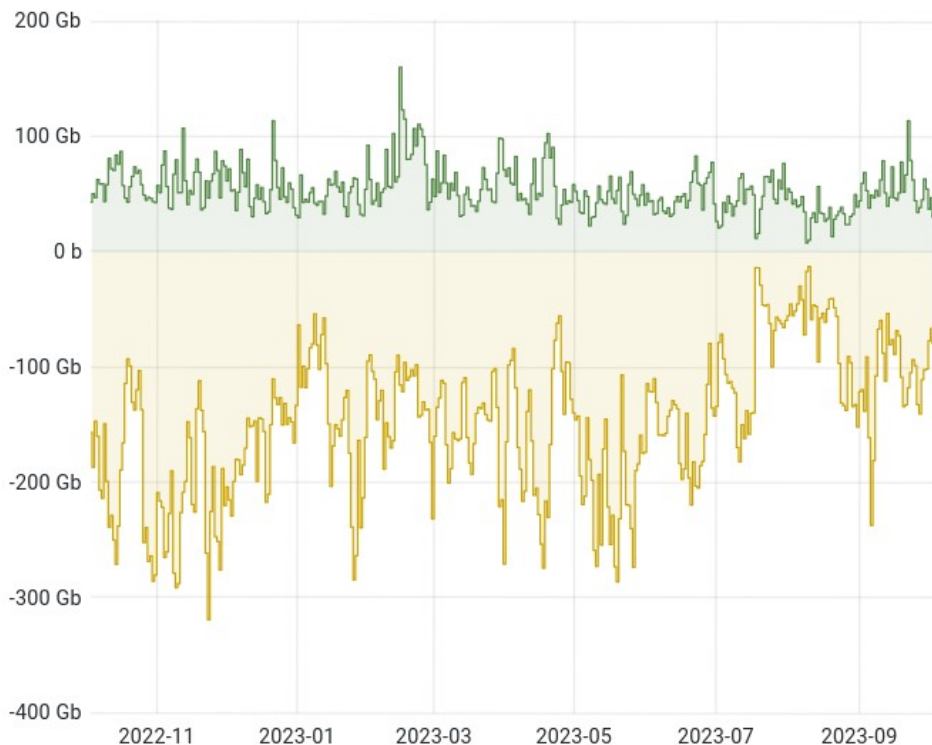- PABX phase out in final phase

PDC: Prevessin data-centre

CERN Network Domains

CERN total external traffic

LHCOPN+LHCONE+Internet, last 12 months

	min	max	avg
Internet IPv4 In	0 b	13.7 Gb	3.78 Gb
Internet IPv4 Out	0 b	107 Gb	20.8 Gb
Internet IPv6 In	0 b	24.2 Gb	7.57 Gb
Internet IPv6 Out	0 b	38.5 Gb	9.71 Gb
Internet Total In	0 b	28.1 Gb	11.3 Gb
Internet Total Out	0 b	124 Gb	30.6 Gb
WLCG IPv4+IPv6 In	0 b	141 Gb	42.1 Gb
WLCG IPv4+IPv6 Out	0 b	271 Gb	118 Gb
Total In	7.49 Gb	160 Gb	53.4 Gb
Total Out	13.1 Gb	319 Gb	148 Gb

Numbers:

Sent out ~583 PB in the last 12 months
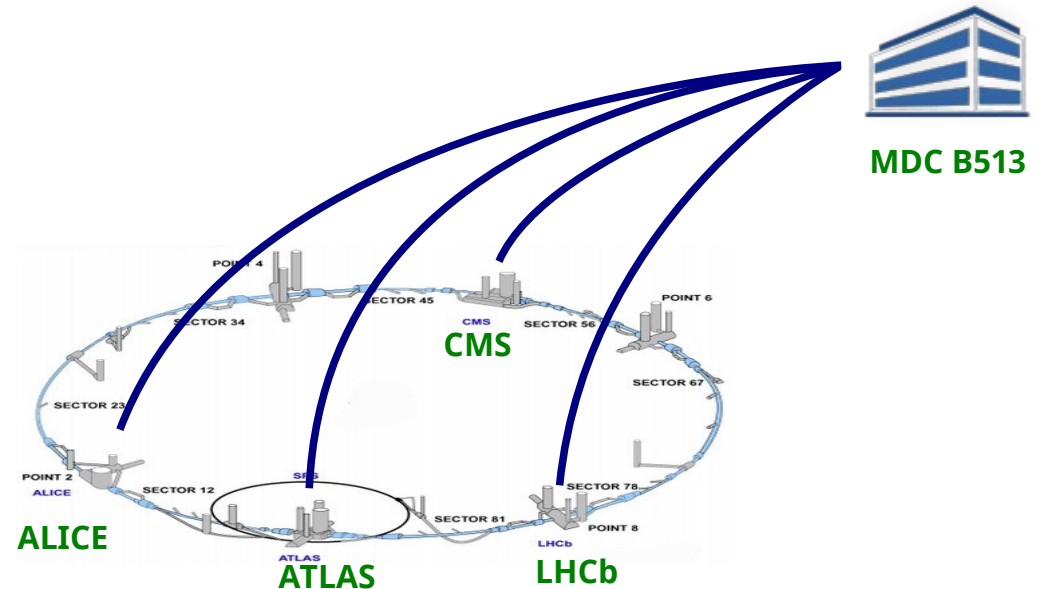
+9% compared to previous year (536PB)

Ref: https://monit-grafana.cern.ch/d/cScW82Tnz/00-overview?orgId=14&var-source=long_term&var-bin=1d&from=now-1y&to=now

Experiments' DAQ lines to IT data-centre

Capacity in place for Run3:
- ALICE: 3.2 Tbps
- LHCb: 400 Gbps
- CMS: 400Gbps
- ATLAS: 200Gbps

MDC B513

CMS

ALICE

ATLAS

LHCb

External Network: some numbers

- LHCONE capacity: 1.2 Tbps
- LHCOPN capacity: 2.1 Tbps
- Internet capacity: 1 Tbps
- Statefull Firewall capacity: 0.6 Tbps
- HTAR (firewall bypass) capacity: 0.6 Tbps *(waiting for new routers with more interfaces on Core routers)*
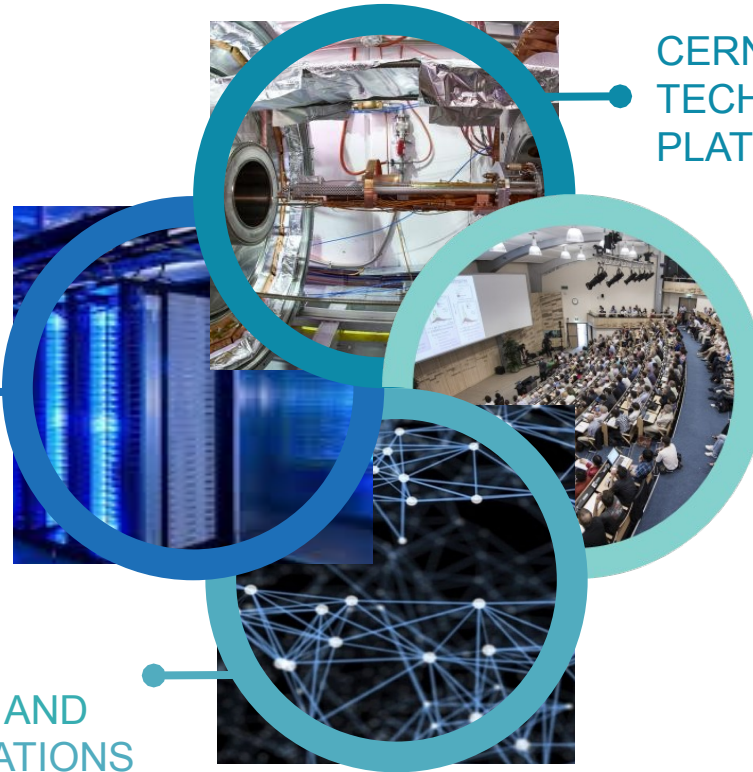
CERN Quantum Technology Initiative

Phase 2 (2024-2028) just approved

CERN QUANTUM
TECHNOLOGY
PLATFORMS

HYBRID QUANTUM
COMPUTING AND
ALGORITHMS

COLLABORATION
FOR IMPACT

QUANTUM
NETWORKS AND
COMMUNICATIONS

QTI2 - Quantum Communications

Set up the CERN Quantum Networks Hub (physical layer)

Quantum Key (QKD), Quantum Communication
- set up a test environment, to gain experience with required network
  capabilities and equipment
- use of White Rabbit for key synchronization
- interconnect with other NRENs Quantum networks

Optical Time and Frequency Distribution
- identify experiments needing it, reach them with optical connections
- connect to metrology institutes in Europe. GEANT is planning an
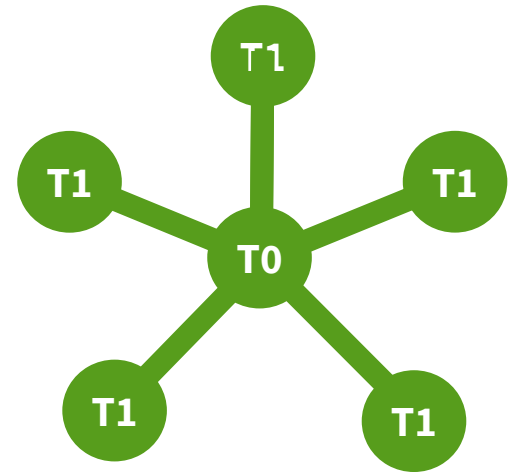 European network and is interested in a collaboration

LHCOPN

LHCOPN

Private network connecting Tier0 and Tier1s

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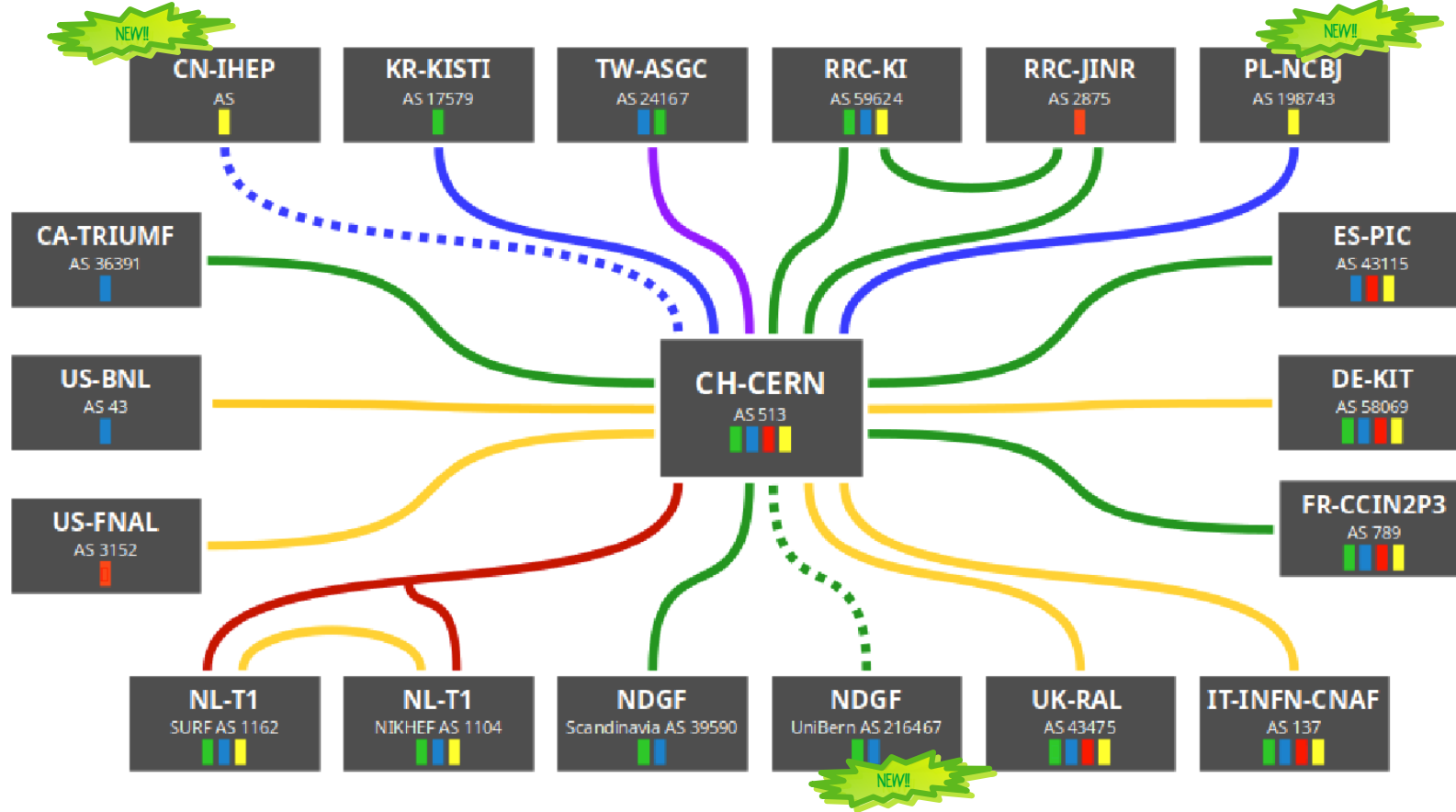- Can connect directly to Science-DMZ,
  bypass perimeter firewalls

Advanced routing:
- BGP communities for traffic engineering
- load balancing

Numbers

- 18 sites for 15 Tier1s + 1 Tier0
- PL-NCBJ just joined, CN-IHEP and NDFG-LHEP in the process to connect
- 15 countries in 3 continents
- 2.1 Tbps to the Tier0

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

IHEP (CN): new LHCb Tier1

IHEP LHCb Tier-2 has started the procedure to become LHCb Tier-1

- CSTNet is the network service provider for IHEP International links

- All domestic connections will be upgraded from 10G to 100G

- New international connections will be deployed to improve the bandwidth between China and Europe

- LHCOPN: new link to CERN via CSTNet and GEANT being deployed
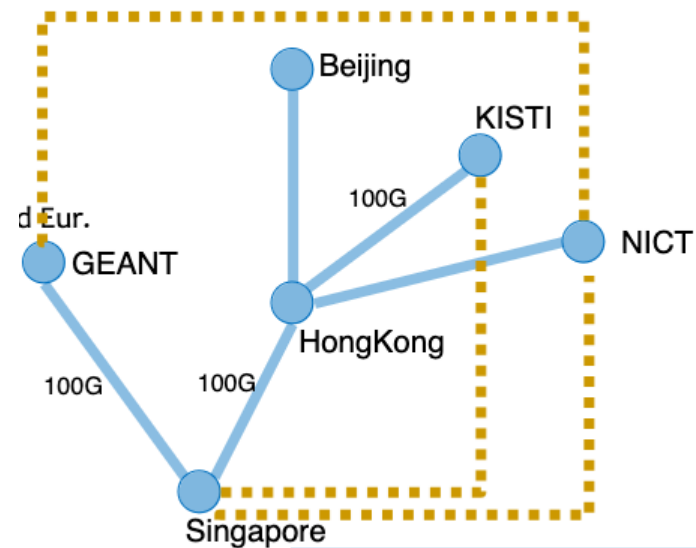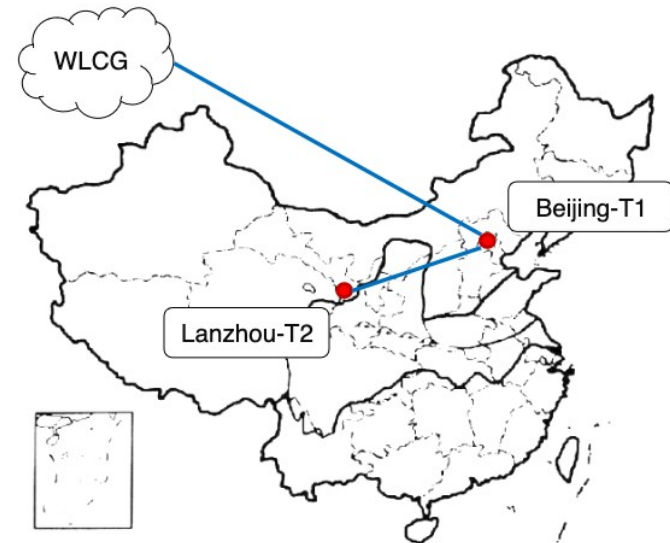
NCBJ (PL) - new LHCb ⊤

NCBJ, National Centre for Nuclear Research in Warsaw has started the procedure to become a LHCb Tier1.

It hosts the Świerk Computing Centre (CIŚ)
- Computing: 1.4 PFLOPS, 36000 cores, 200 TB RAM
- Disk storage: 26 PB (Lustre, Isilon, Netapp, dCache)
- Tape storage: TSM4500, 16 PB (uncompressed)

Network resources:
- 100 Gbps link to PIONIER (academic internet, GEANT)
- 20 Gbps dedicated VLAN to LHCONE
- 2x10 Gbps dedicated link for LHCOPN connectivity, plus 20G backup VLAN over shared PIONIER network to CERN
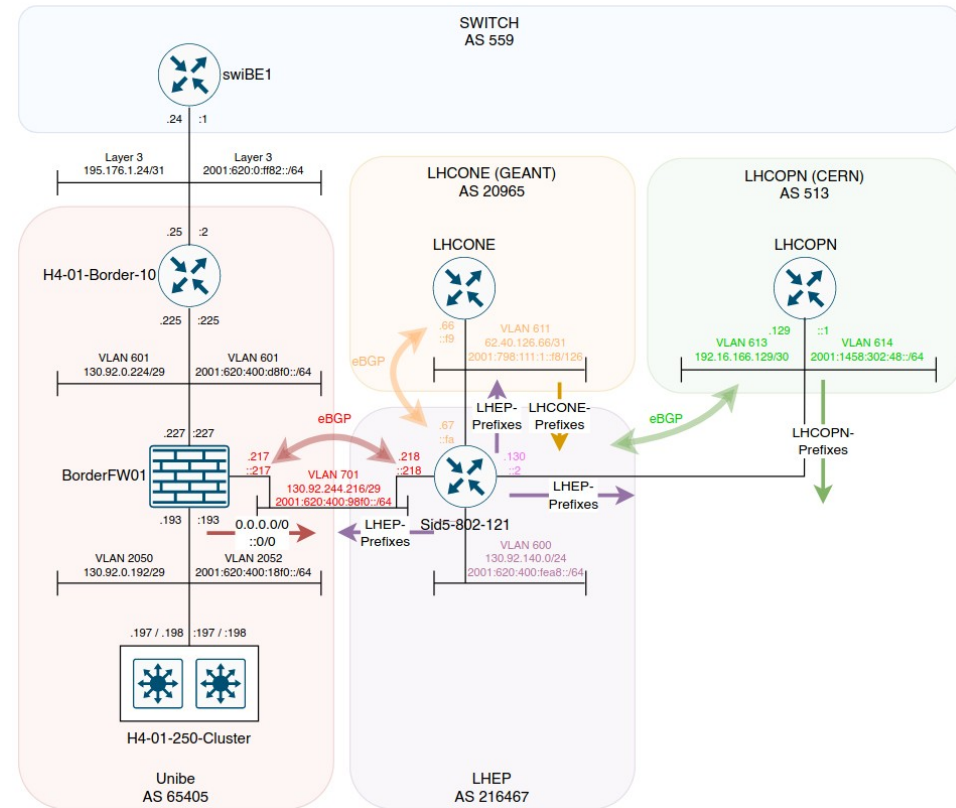
LHEP (CH) new NDGF Tier1 site

LHEP at University of BERN has joined the NDGF distributed Tier1.

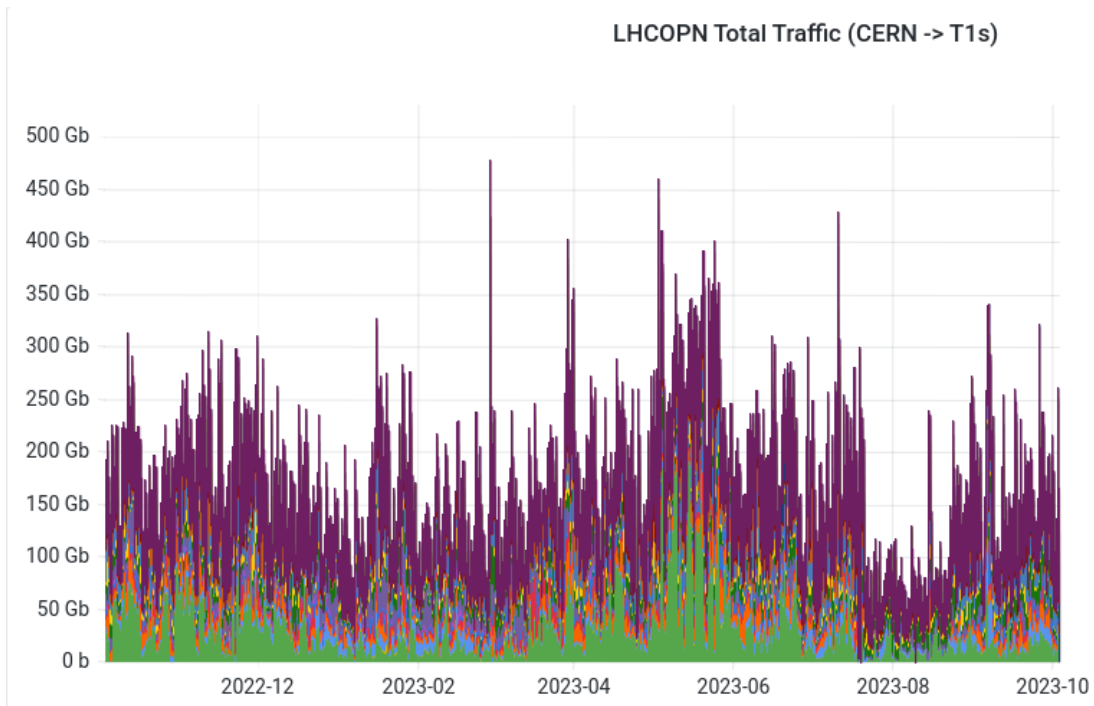LHEP will be connected to CERN with a 100G LHCOPN link

The physical connection is provided by SWITCH (Swiss NREN) and it is already in place.

LHCOPN routing will be configured in November 2023

LHCOPN Traffic – last 12 months

LHCOPN Total Traffic (CERN -> T1s)

	Mean	Max
Outgoing DE-KIT	53.0 Gb	251 Gb
Outgoing KR-KISTI	198 Mb	9.39 Gb
Outgoing RU-T1	6.60 Gb	42.6 Gb
Outgoing FR-IN2P3	11.5 Gb	89.4 Gb
Outgoing NDGF	6.87 Gb	83.3 Gb
Outgoing NL-T1	6.12 Gb	104 Gb
Outgoing TW-ASGC	1.14 Gb	9.69 Gb
Outgoing IT-INFN-CNAF	12.3 Gb	130 Gb
Outgoing UK-RAL	9.44 Gb	40.5 Gb
Outgoing CA-TRIUMF	5.96 Gb	75.7 Gb
Outgoing US-BNL	12.7 Gb	98.8 Gb
Outgoing US-FNAL	6.76 Gb	81.3 Gb
Outgoing ES-PIC	4.43 Gb	81.8 Gb
Outgoing-PL-NCBJ	246 Mb	18.9 Gb
Total	137 Gb	479 Gb

Numbers:

Moved ~540 PB in the last 12 months

+18% compared to previous year (457PB)

Peak at ~479Gbps

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

LHCOPN: Long-term growth
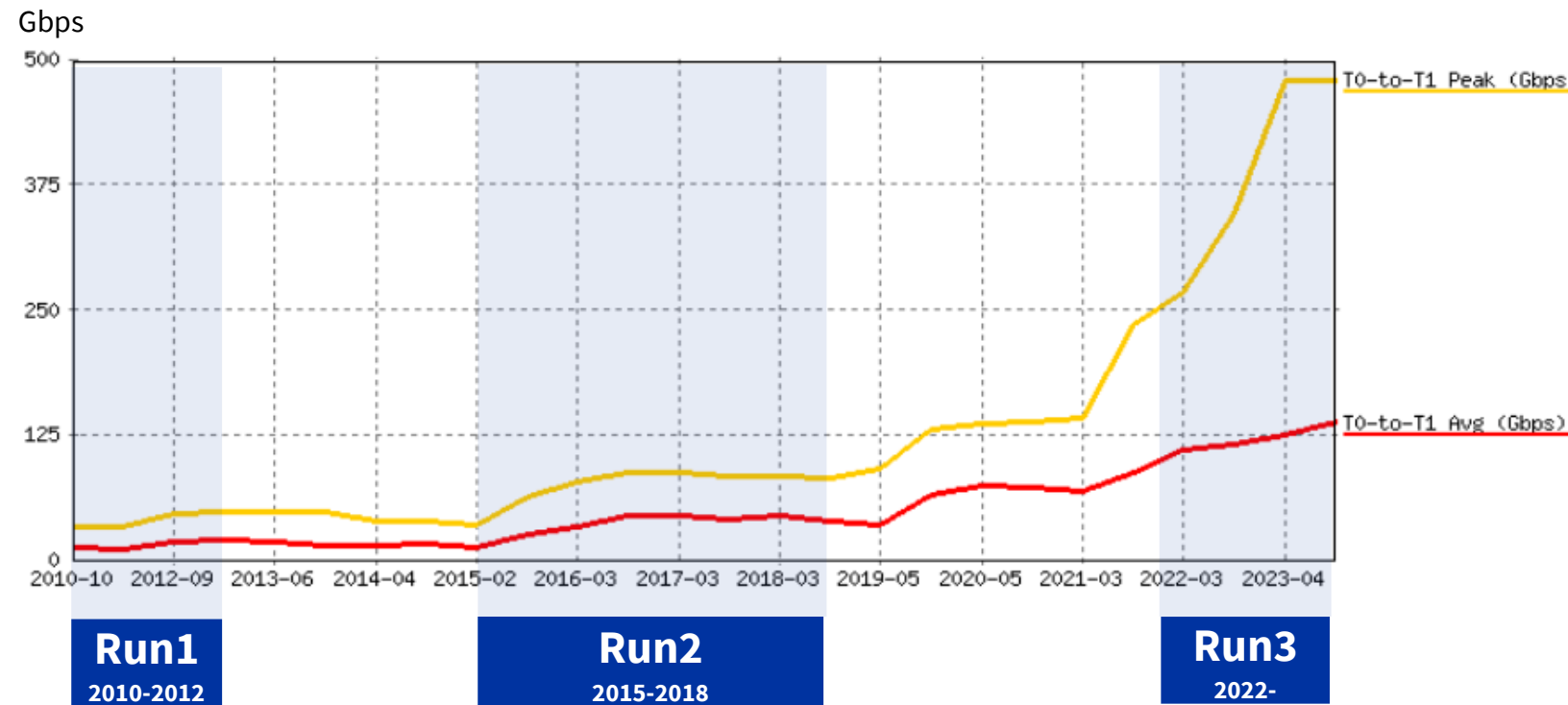
Gbps

Run1: 2010-12
LS1:   2013-14
Run2: 2015-18
LS2:   2019-21
Run3: 2022-25

Y-Axis: Gbps - Average bandwidth of previous 12 months

Ref: https://twiki.cern.ch/twiki/bin/view/LHCOPN/LhcopnStats

CERN-CNAF Data-Centre Interconnect
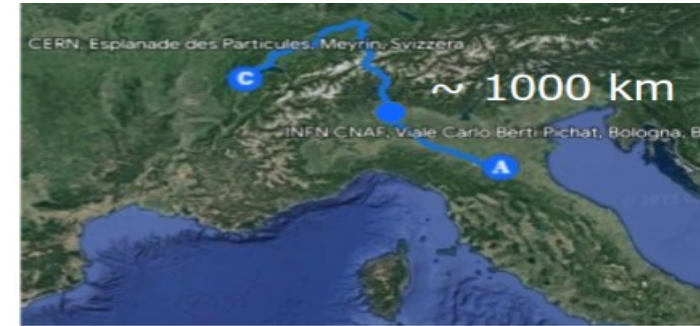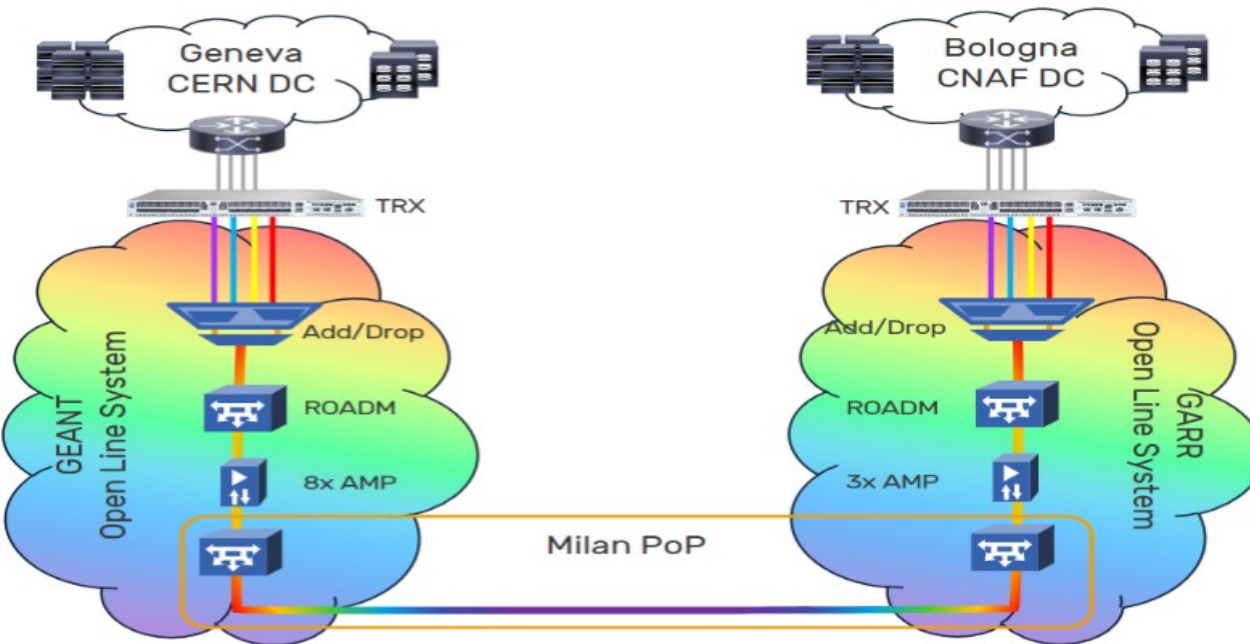
1.6 Tbps

4 carriers
- DP-16QAM
- 27% SD-FEC
- 69 Gbaud

- Commissioning phase. Connected 2x 100Gbps client interfaces, 400G coming soon
- It will be used for IT-INFN-CNAF's LHCOPN link
- Plan to grow to 1.2 Tbps by HL-LHC

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
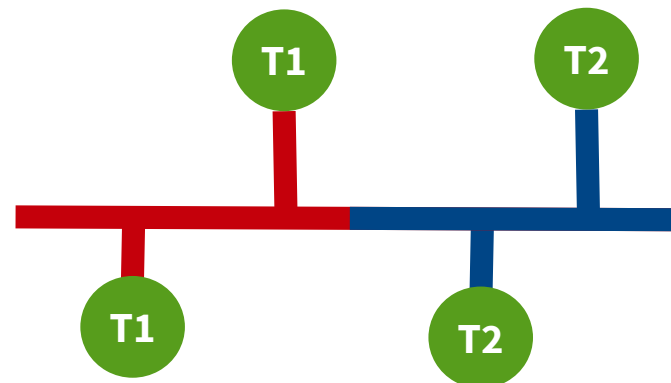
Open to other HEP collaborations

DUNE joins LHCONE

At the last LHCONE meeting in Prague (April 2023), DUNE has formally requested permission to join LHCONE

The LHCONE community expressed its approval to allow DUNE to join LHCONE

The request and decision was presented to the WLCG Management Board of September 2023. The Board had no objections and endorsed the decision

LHCONE L3VPN – latest news

News

- CERN has upgraded its LHCONE connections to 400G with ESnet and 2x 400G with GEANT

- New sites:
    - Lawrence Berkeley National Laboratory (ESnet)
    - University of Massachusetts – Amherst (ESnet)

- KIFU (Hungarian NREN) joins LHCONE

Traffic statistics:

- continue increase

- first peak above 1Tbps seen in GEANT

LHCONE L3VPN: A global infrastructure for High Energy Physics data analysis (LHC, Belle II, Pierre Auger Observatory, NOvA, XENON, JUNO)

LHCONE map

LHCONE Map Ver. 6.0, 2022-11-15 — WEJohnston, ESnet, wej@es.net

LHCONE monitoring

- perfSONAR 5 is out and being deployed Some bugs have pushed new release (latest 5.0.5)
- perfSONAR 5 now uses Elasticsearch and Grafana
- LHCONE 100G mesh: data is now shown correctly , but results are not great. Work in progress

LHCONE Looking Glass

Running looking-glass to analyses the routing tables of the VRFs

Implemented on a CERN router. Now peering with these VRFs:
- ASGC AS24167
- CANARIE AS6509
- CERNlight AS20641
- ESnet AS293
- KREOnet AS17579
- GEANT AS20965 (Geneva and Frankfurt routers)
- NORDUnet AS2603
- RU-VRF AS57484

The looking glass is accessible at http://lhcone-lg.cern.ch/

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 netwrk information related t oLHCOPN 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)

CRIC NetSite table

https://wlcg-cric.cern.ch/core/netsite/list/

- NOC contact email
- URLs to monitoring and other info pages
- LHCONE AUP acknowledgment
- LHCONE and LHCOPN participation
- Network bandwidth: WAN, LHCOPN, LHCONE (-1 = not connected)
- LHCONE providers

~140 entries

RC Site	NetSite	NOC	monit URL	info URL	AUP	LHCONE active	LHCOPN active	WAN	LHCONE	LHCOPN	providers
AGLT2	✏ ▦ US-AGLT2 Michigan State University	aglt2-noc@umich.edu			✓	✓	✗	100	100	-1	ESnet
AGLT2	✏ ▦ US-AGLT2 University of Michigan	aglt2-noc@umich.edu			✓	✓	✗	80	80	-1	ESnet
ANLASC	✏ ▦ US-ANL	noc@anl.gov			✓	✓	✗	0	100	-1	ESnet
ARNES	✏ ▦ SL-ARNES-NREN				✓	✓	✗	0	9	-1	ARNES
ARNES	✏ ▦ SL-IJS-Ljubljana				✓	✓	✗	0	9	-1	ARNES
ARNES	✏ ▦ SL-IZUM-Maribor				✓	✓	✗	0	9	-1	ARNES
Australia-ATLAS	✏ ▦ AU-Australia-ATLAS				✓	✓	✗	20	10	-1	AARNET
BEgrid-ULB-VUB	✏ ▦ BE-ULB-VUB				✗	✗	✗	20	-1	-1	

CRIC NetworkRoute table

https://wlcg-cric.cern.ch/core/networkroute/list/

One entry per set of prefixes sharing common routing policies

- set of v4 and v6 prefixes
- "More specific" flag (in case of disaggregated prefixes)
- Autonomous System number (ASN)
- collaborations using these prefixes
- LHCOPN and LHCONE bandwidth specific to the prefixes of the record
- monitoring URL specifics to the prefixes of the record

~150 entries

RC Site	NetworkRoute	NetSite	ASN	monit URL	MS	Subnets	LHCONE limit	LHCOPN limit	collaborations
AGLT2	✎ ⊞ AGLT2_LHCONE_RT	US-AGLT2 Michigan State University	229		✖	2001:48a8:68f7::/48, 2001:48a8:68f7::/50, 2001:48a8:68f7:4000::/50, 2001:48a8:68f7:8000::/50, 2001:48a8:68f7:c000::/50	100	-1	US-ATLAS, WLCG
AGLT2	✎ ⊞ AGLT2_MSU	US-AGLT2 Michigan State University	229		✖	192.41.236.0/23, 192.41.238.0/28	100	-1	US-ATLAS, WLCG
AGLT2	✎ ⊞ AGLT2_MSU IPv6	US-AGLT2 Michigan State University	237		✖	2001:48a8:68f7:8001::/64	100	-1	US-ATLAS, WLCG

WLCG Data challenges & Network R&D

WLCG Data Challenge 2024

Next data challenge (DC24)

- Planned for 12-23 February 2024

- Target of 25% of HL-LHC requirements

- Network providers are increasing their network capacity for LHCONE

- Defined the list of projects that will be tested during DC24

- Preparation Workshop at CERN 9-10 of November at CERN

DC24 projects

List of the projects on networking:
- Packet marking
- Packet pacing, BBR performances
- perfSONAR for network alarms and debugging
- Site Network monitoring of in/out bandwidth
- Use of Jumbo frames to improve performances
- NOTED: FTS driven SDN
- Rucio SENSE
- ALTO FTS Rucio

multiONE

LHCONE already very large, it could become risky to include other large science projects

Better to implement multiple VPNs, one for each collaboration:

- Each site joins only the VPNs it is collaborating with, to reduce the exposure of their data-centre

But it's difficult to separate the traffic for sites member of multiple collaborations. Work in Progress. A new proposal will be made tomorrow at the LHCONE meeting
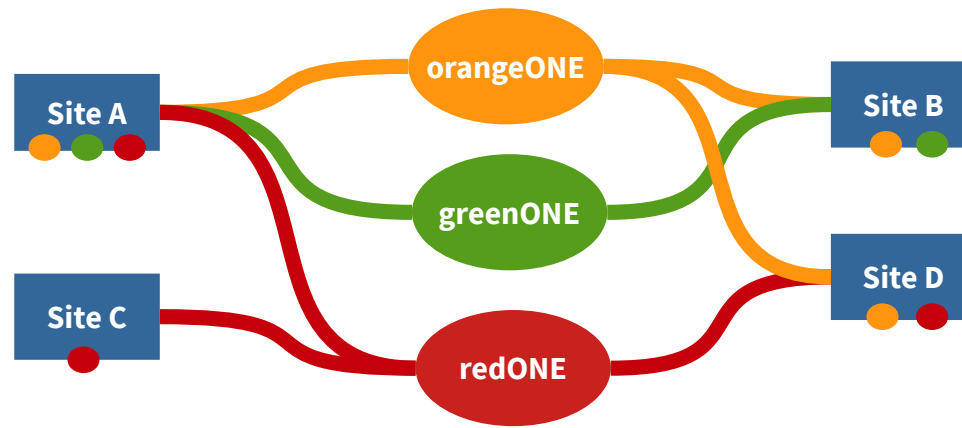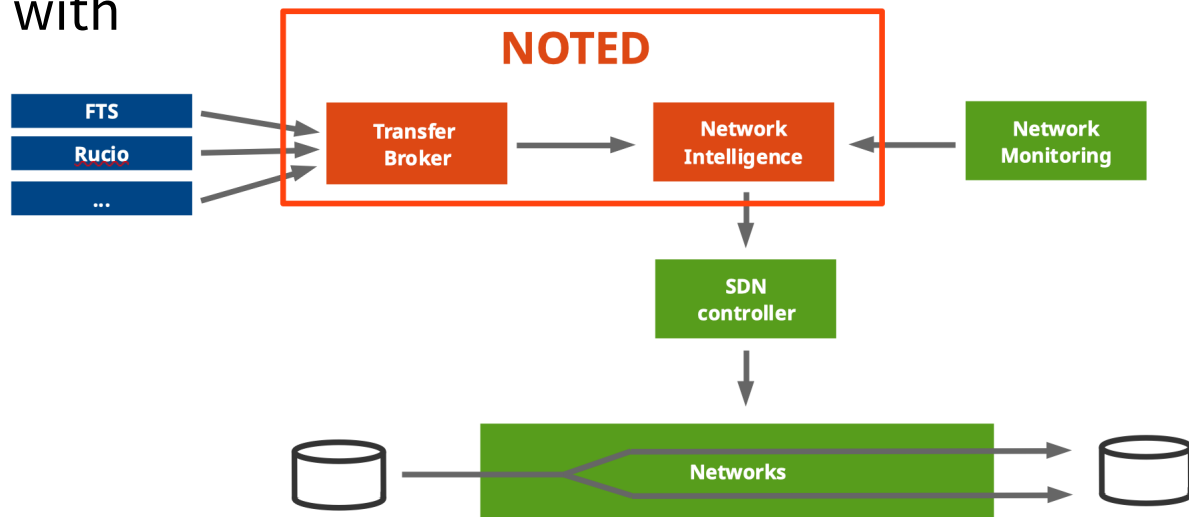
NOTED SDN

NOTED is a framework that can detect large FTS data transfers and trigger network optimization actions to speed up the execution of the transfers

Already tested with production transfers:
- CERN-PIC with LHCOPN-LHCONE load balancing
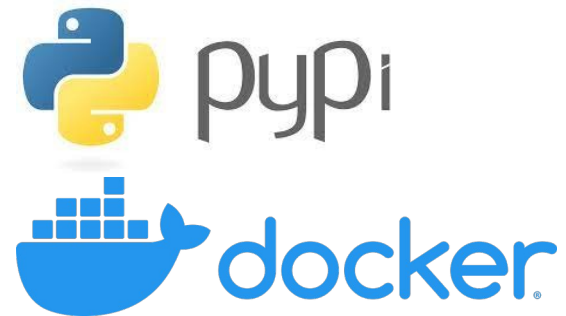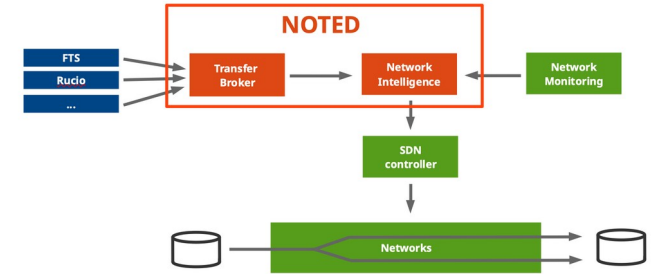- CERN-TRIUMF and KIT-TRIUMF with the activation of dynamic circuits

NOTED status

Version 2 released:

- rewritten in Python

- improved efficiency and stability

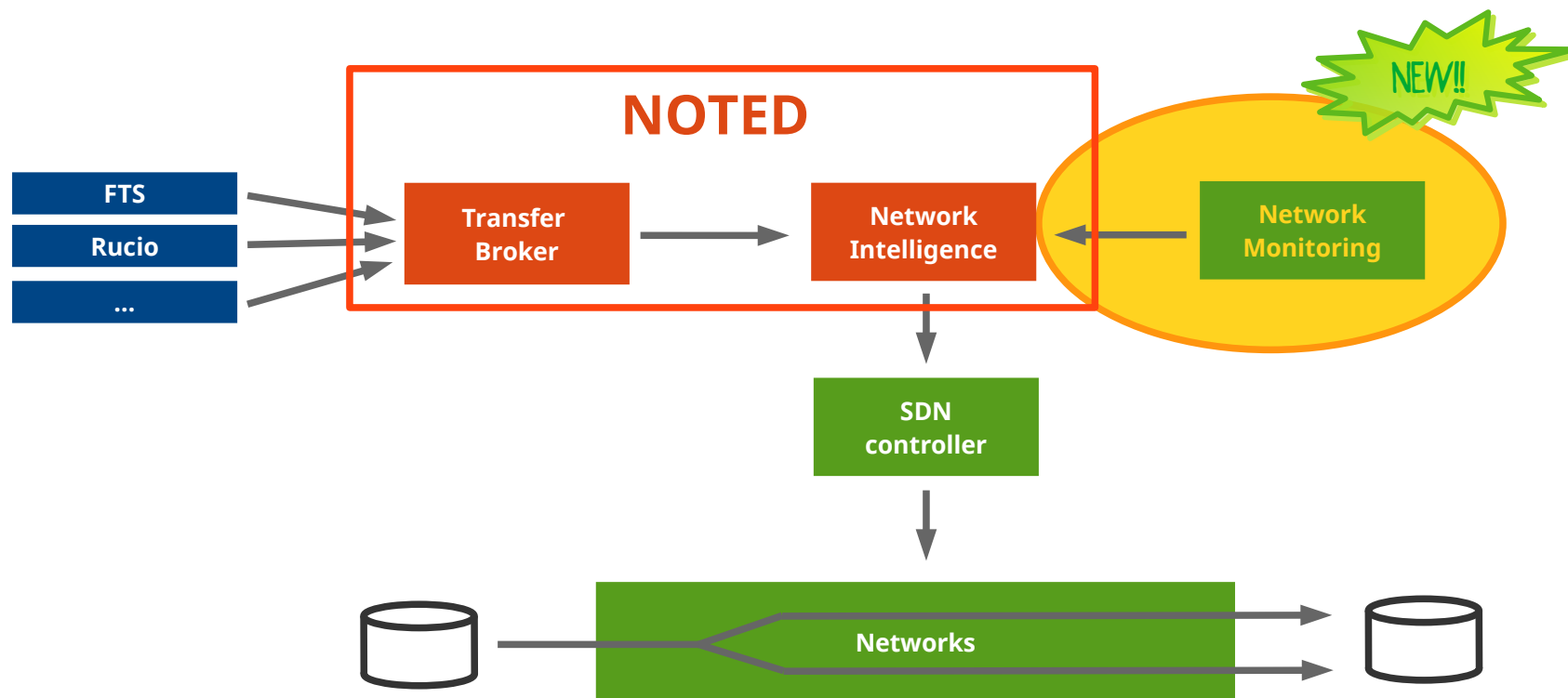- easier configuration

- open source (GPL v3)

Package distribution:

- available at https://pypi.org/project/noted-dev/

- also as docker container:

      https://hub.docker.com/r/carmenmisa/noted-docker

NOTED with network monitoring trigger

Integration with network monitoring: NOTED can be triggered by link saturation alarms

Research Network Technology WG

The RNTWG is currently working on two main actiities: packet marking and packet
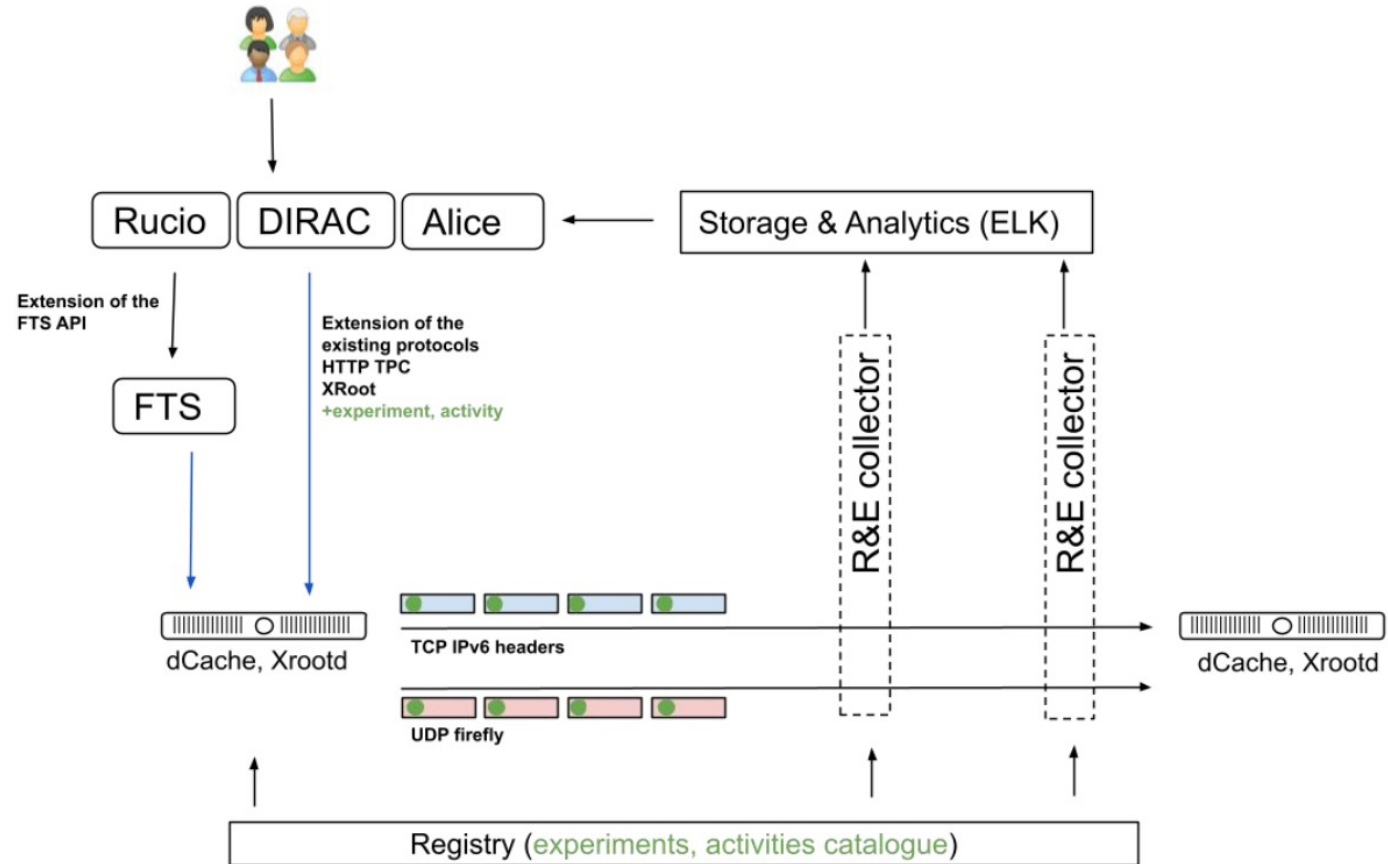   pacing

Packet marking:
- significant progress on network traffic visibility. Two options
   - IPv6 flowlabel tagging
   - Firefly flow marking
- both will be tested during DC24

scitags.org

Packet pacing:
- pacing of TCP packets to avoid buffer exhaustion and packet drops
- considering new TCP congestion control protocols like BBRv3

How scitags works

AutoGOLE and SENSE

AutoGOLE: Infrastructure which provides "end-to-end" network services in a fully automated manner

Open-source software framework based on:

- Network Service Interface (NSI): multidomain network provisioning

- SENSE: end-system provisioning and realtime integration with network services

Persistent Infrastructure, somewhere in between production and a testbed

AutoGOLE, NSI and SENSE work together to provide the mechanisms for complete end-to-end services that include network and attached End Systems DTNs

Using SENSE to move CMS data in Rucio


Project led by UCSD and Caltech

The increased requirements of the HL-LHC requires to use any resource in the most efficient way, including networks

Objectives of the project:

#1 Make Rucio capable to schedule transfers on the network and prioritize them

#2 Predetermined transfer speed and quality of service (time to completion)

Demonstrated:

- SENSE can build VPNs between pairs of XrootD servers in charge of FTS transfers requested by Rucio

- QoS can be provisioned in the network to prioritize the traffic in the VPN
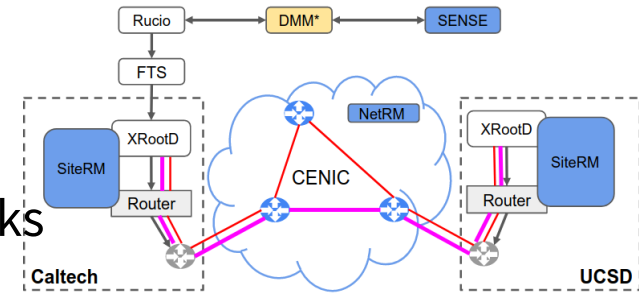
Conclusions

Summary

- CERN networks: developing for Run4
- LHCOPN: Three new Tier1s connected
- LHCONE continue to grow. DUNE has joined it
- LHCONE R&D: several projects on going
- WLCG DC24: several network projects proposed
- Network R&D essentials to be ready for Run4 data deluge

Questions?

edoardo.martelli@cern.ch