



LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS
partículas e tecnologia

Computing and Networks in High Energy Physics

Advanced and Distributed Computing Group



WHAT

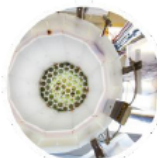
LIP IS

The reference institution for experimental particle physics in Portugal and the Portuguese reference partner of CERN as well as other international scientific infrastructures

Has centres in:
Lisbon, Coimbra, Braga



Experimental particle and astroparticle physics



Development of new instruments and methods

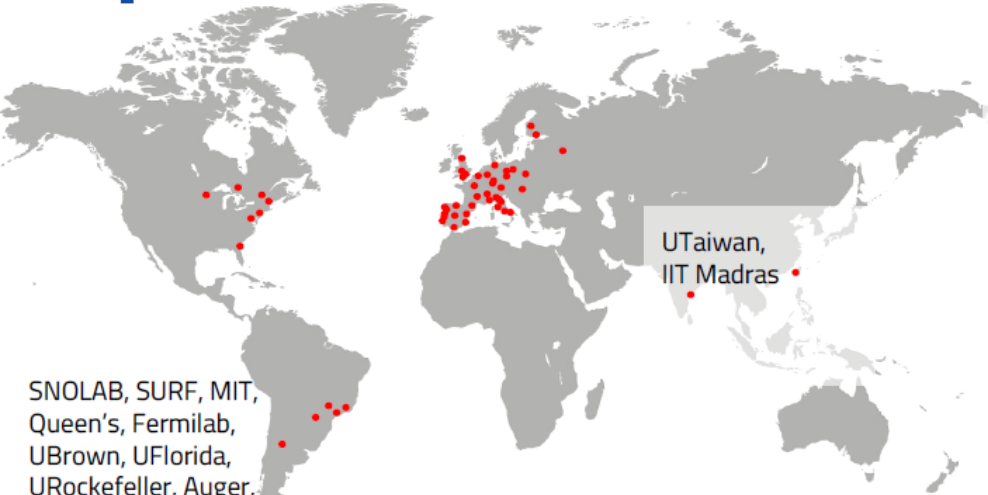


Scientific computing



Knowledge transfer, education and outreach

To whom we are connected



SNOLAB, SURF, MIT,
Queen's, Fermilab,
UBrown, UFlorida,
URockefeller, Auger,
CBPF, SBF, USP,
USC, UCampinas,
EERJ

UTaiwan,
IIT Madras

IST, FCUL, ULisboa, FCTUC, UM,
CTN, UA, ICNAS, LNEC, Ciência
Viva, IBEB, INESC-ID,
INESC-TECH, UBI, UÉvora, SPF,
ISEC/IPC, LIBPhys, BioSI,
CCMAR, ISEC, UPorto, IMM, IGC,
PORBIODATA, FCT-FCCN

CERN, ESA, EGI, DESY, HIP
Helsinki, MEPhi, Imperial
College, USurrey, UOxford,
TUDresden, LMU Munich,
HephyVienna, TUDortmund,
IPPP, LPC, TUDelft, GSI,
Humboldt, KIT, CEA,
CESNET,
Clermont-Ferrand,
CYFRONET, PSNC,
Utrecht



CSIC, IFCA, UPV, CESGA, BIFI, UAM, PIC, Lifewatch ESFRI,
UGranada, USC/IGFAE, INFN, INAF, UFerrara, UTorino,
UPadova, UPisa, UUdine, PoliMilano, PoliBari, LLR

LIP activities

High Energy Physics and related areas

High Energy Physics

Trying to answer many open questions about our universe

Many open questions ...

- What is the origin of mass? why some particles are very heavy while others have no mass at all ?
- Is the 125 GeV Higgs boson the fundamental scalar predicted by SM, or is it explained by an extend theory (why 125 GeV) ?
- What are dark Matter and dark energy?
- Why is there far more matter than antimatter in the universe?
- How does the quark-gluon plasma give rise to the particles that constitute the matter of our Universe?
- and many more ...

Search for new phenomena ...

- Gravity and extra dimensions ?
- Supersymmetric particles ?
- New fundamental interactions ?
- New generations of quarks/leptons ?
- Leptoquarks ?
- Something else completely new ?

Large Hadron Collider

Circumference: ~ 27 Km
Depth: up to 175 meters
Electrical power: 120 MW (600 GWh per year)

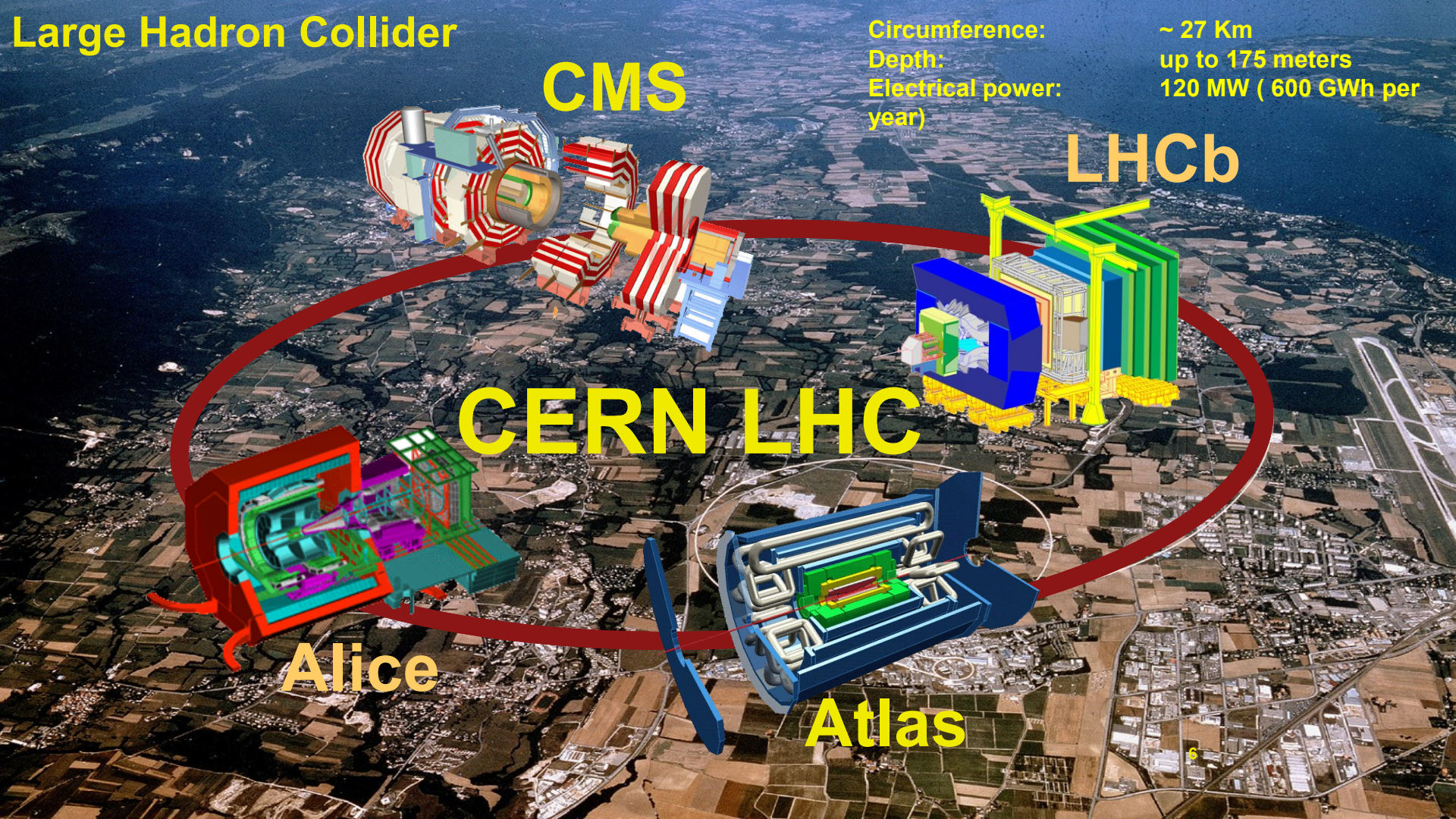
CMS

LHCb

CERN LHC

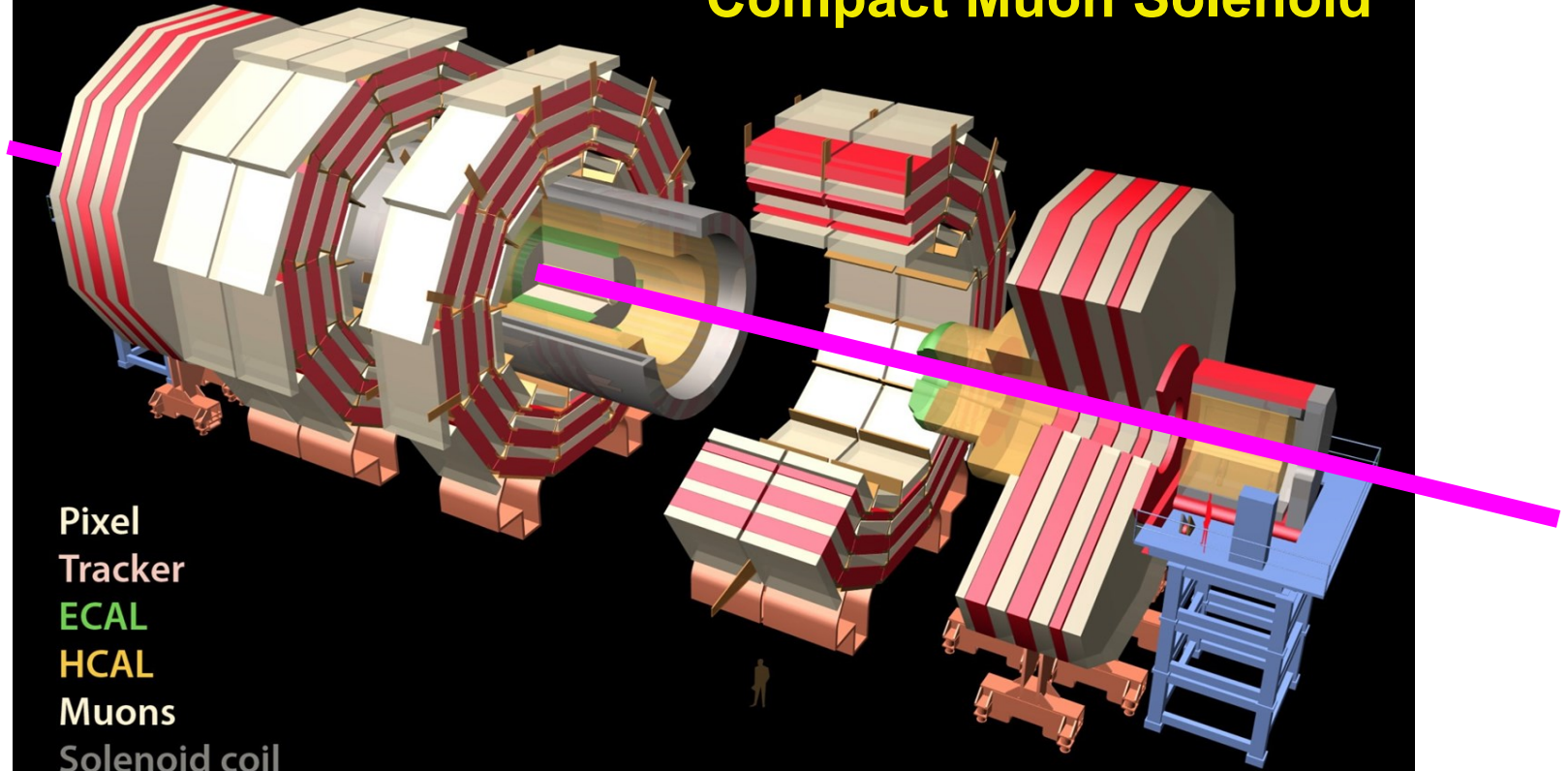
Alice

Atlas



CMS @ LHC

Compact Muon Solenoid



Pixel
Tracker
ECAL
HCAL
Muons
Solenoid coil

Total weight 12500 t, Overall diameter 15 m, Overall length 21.6 m, Magnetic field 4 Tesla



CMS Experiment at the LHC, CERN

Data recorded: 2016-Oct-14 09:56:16.733952 GMT

Run / Event / LS: 283171 / 142530805 / 254



A deluge of data

- Proton bunches collide every **25 ns**
- **150 million** sensors delivering data
- **30+ Petabytes** of new data per year
- All data must be **distributed worldwide**

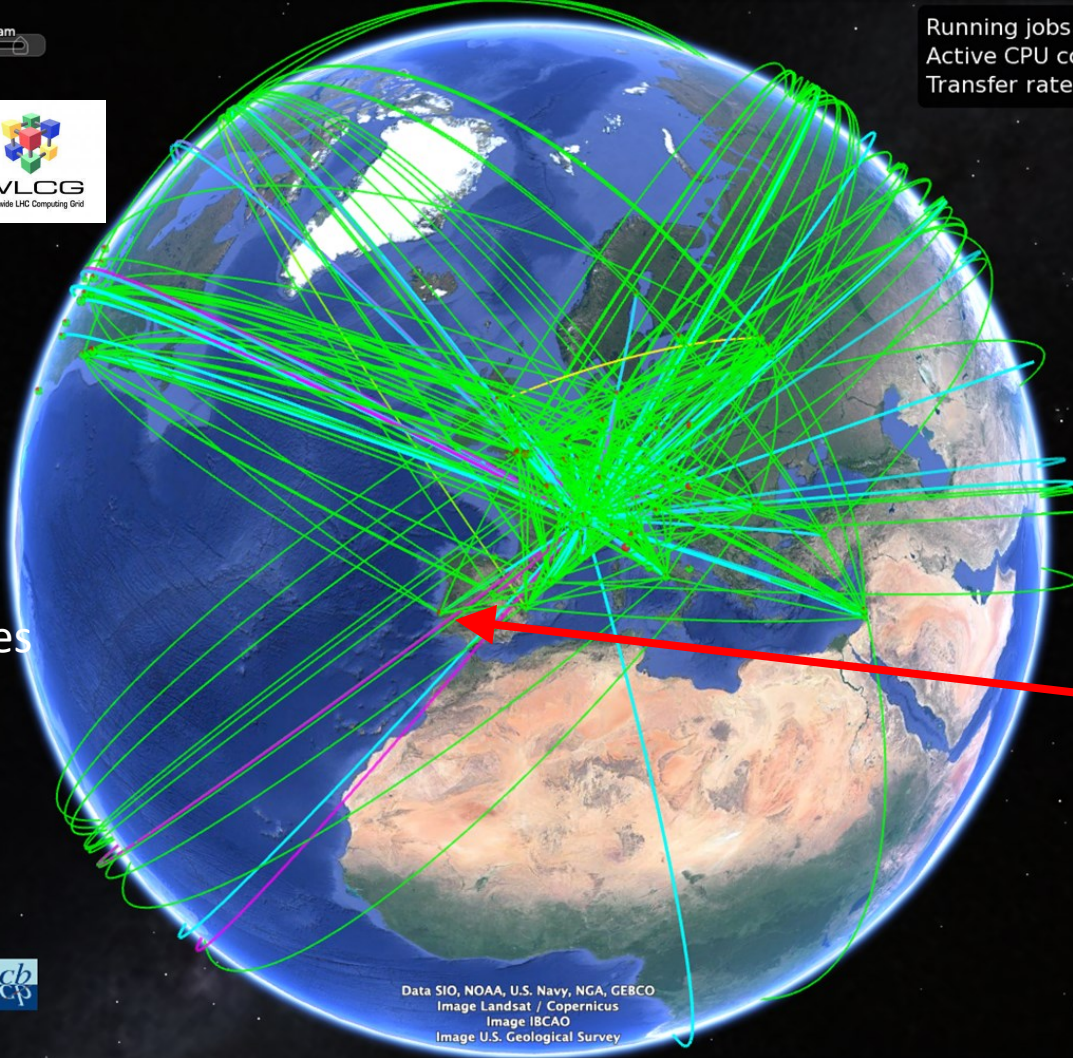


Worldwide LHC Computing Grid



Computing and data infrastructure for the LHC experiments

- 170+ sites
- 42 countries
- 2 million tasks/day
- 1.4 million CPU cores
- 1.5 Exabytes
- 12 000 users

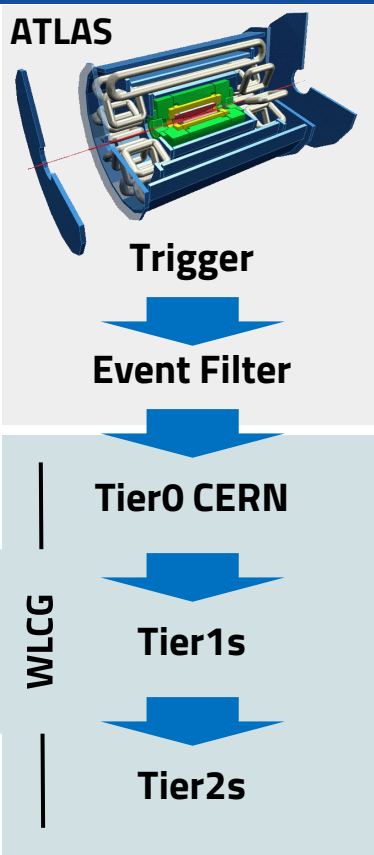


Portugal has a Tier-2 site operated by **LIP + INCD**

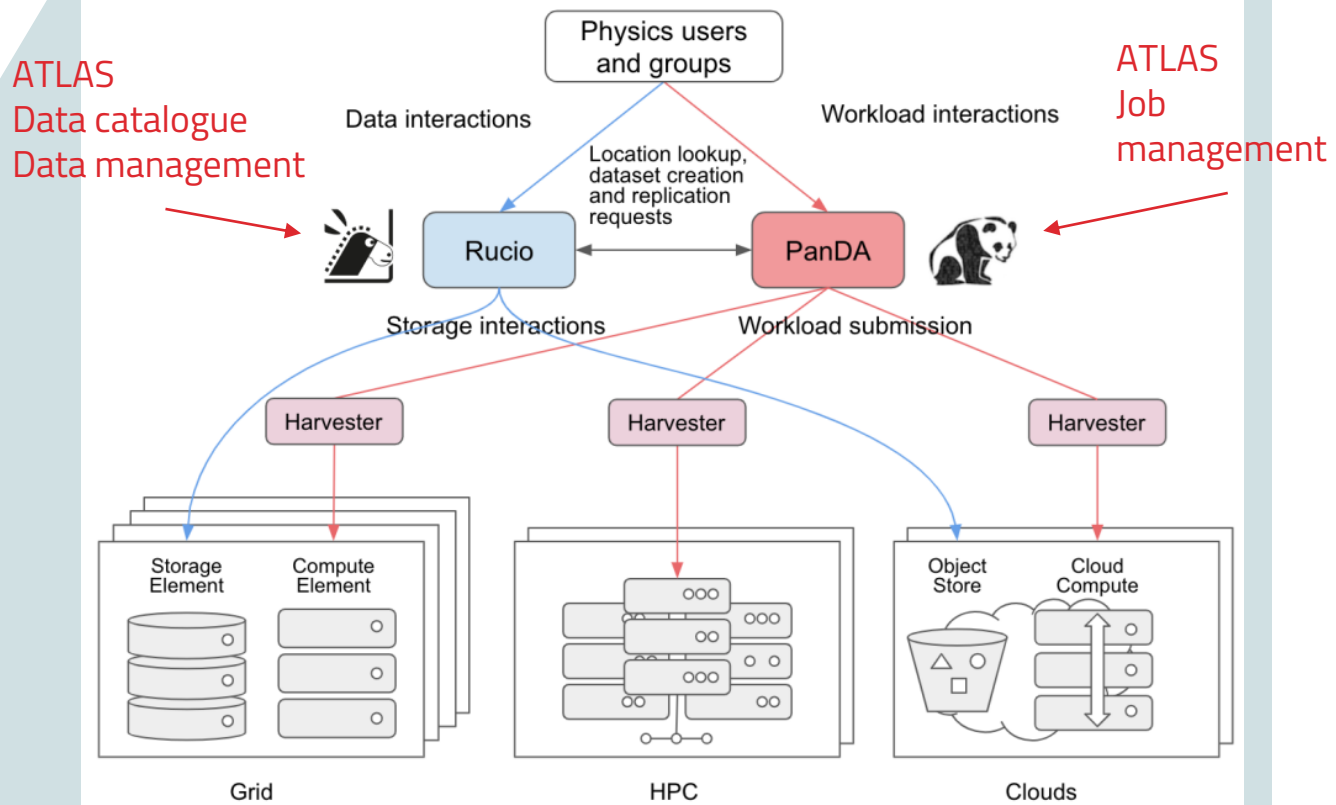
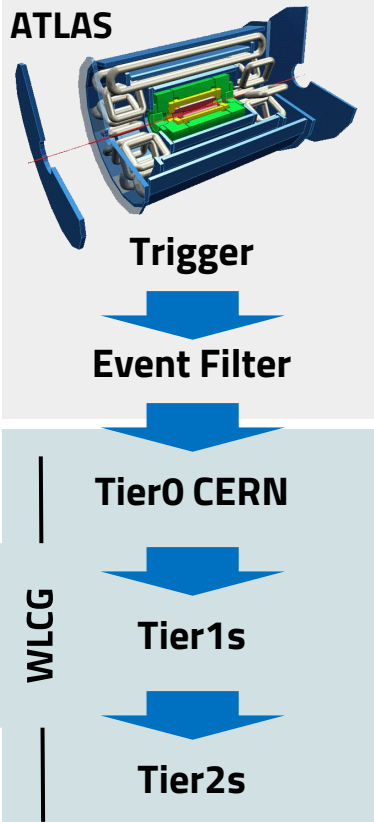


Data SIO, NOAA, U.S. Navy, NGA, GEBCO
Image Landsat / Copernicus
Image IBCAO
Image U.S. Geological Survey

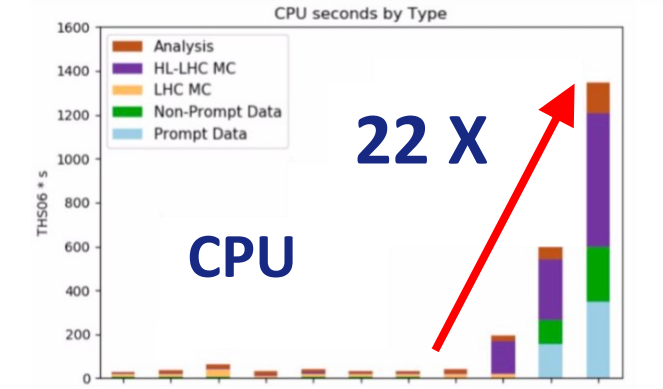
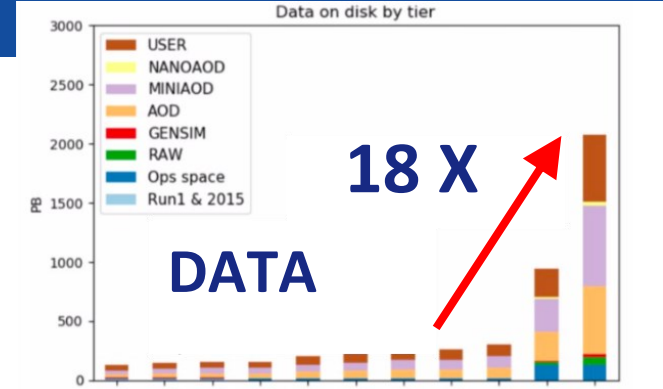
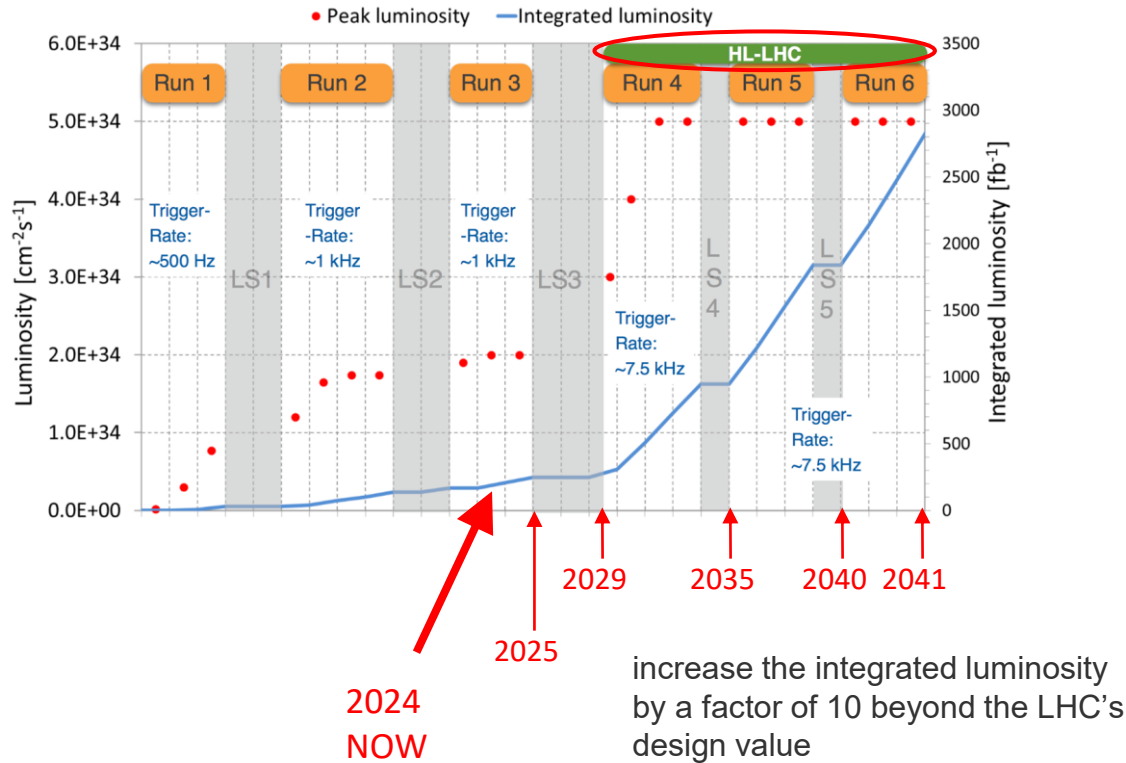
Overview of the data processing in the ATLAS experiment



Overview of the data processing in the ATLAS experiment



A growing challenge \Rightarrow High Luminosity LHC (HL-LHC)



INCD and the Portuguese Tier-2

Infraestrutura Nacional de Computação Distribuída
Portuguese Worldwide LHC Computing Grid Tier-2





INCD centers in 2023



Cloud Computing
cloud computing



HTC Computing
high throughput
computing (GRID)



HPC Computing
high performance
computing



RNCA



INCD-A @ LNEC in Lisbon

HPC / HTC / Cloud / Federation
6000 CPU cores
5 Petabytes online raw
100 Gbps + 100 Gbps LHCONE
Includes the WLCG Tier-2



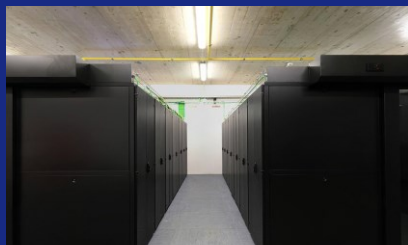
INCD-B @ REN in Riba-de-Ave (partially moved to Lisbon)

HPC / HTC
2600 CPU cores (640 cores)
384 Terabytes raw
1 Gbps



INCD-L @ LIP in Lisbon

Tape storage
1 Petabyte backups
10 Gbps



INCD-D @ UTAD in Vila Real

HPC / HTC / Cloud / Federation
5000 CPU cores + IB HDR200
4 Petabytes online raw
10 Gbps



INCD-C @ UC in Coimbra (BEING IMPROVED)

Tape storage expansion
20 Petabytes
10 Gbps

LIP in IBERGRID and EGI



36
EU funded
projects

1790⁺
Enabled
articles

7.1B
HTC CPU hours
consumed

29
EGI Council
participants

407M
Computational
jobs

84.000
Users

70M
Cloud CPU hours
consumed

Federating compute and storage from hundreds of data centres for research.

WLCG is one of the supported communities
in EGI and IBERGRID

LIP responsibilities and activities:

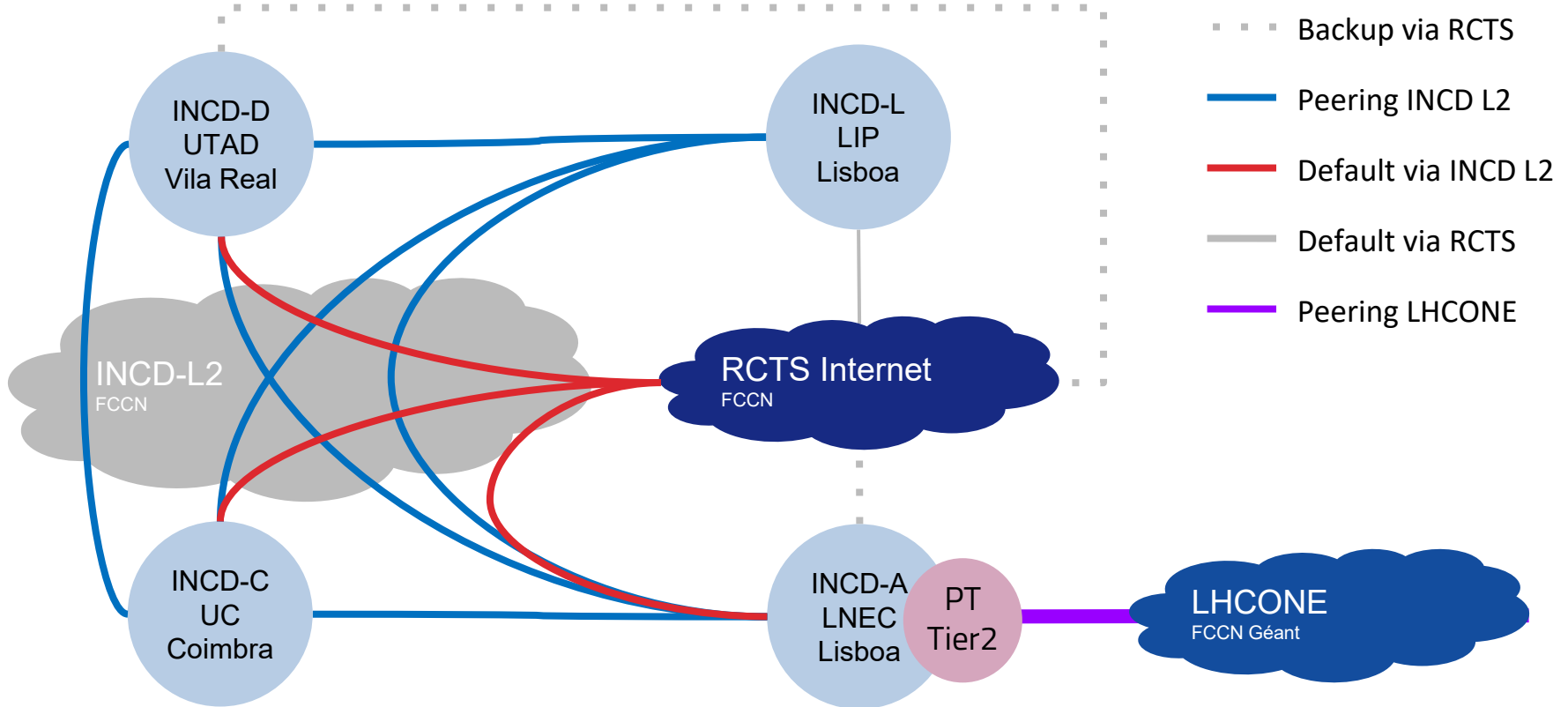
- IBERGRID and EGI provide the backbone for WLCG
- Infrastructure operations coordination at Iberian level and interface with EGI operations
- Software management for the EGI and IBERGRID federations
- National technical contact point
- Security contact for Portugal
- Support to user communities
- Developing and operating core services e.g. software repositories for the EGI federation
- Integration of thematic and/or user services

cloud+grid+data





Peerings



Portuguese WLCG Tier-2 ATLAS and CMS

The **Tier-2 / Tier-3** uses the INCD infrastructure and is **operated by the LIP** computing team.

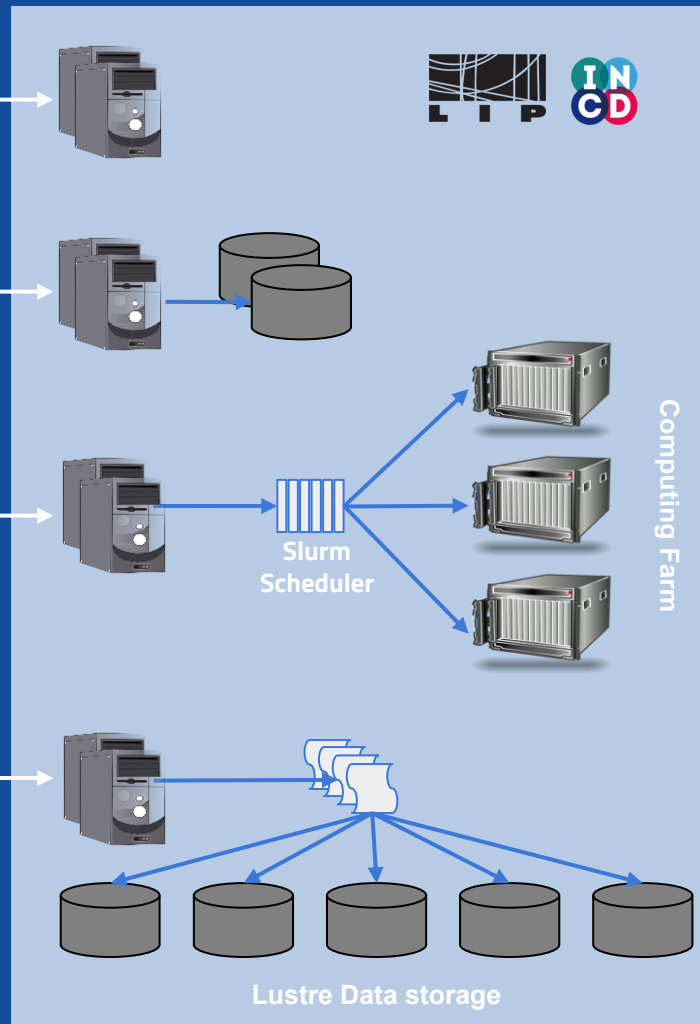
- High Throughput Computing facility
- Based at INCD using the Lisbon site
- Shares the INCD Slurm batch system
- Shares the INCD Lustre storage system
- Intensive data processing and analysis
- Montecarlo simulations
- Continuous data WAN data transfers

TOPBDII
SiteBDII
APEL
ARGUS
PerfSonar
...

CVMFS
Worldwide FS
Experiments
Software

ARC-CE
with
Slurm as
scheduler

XRootd
Webdav
StoRM SRM
with
Lustre as
Underlying
Storage



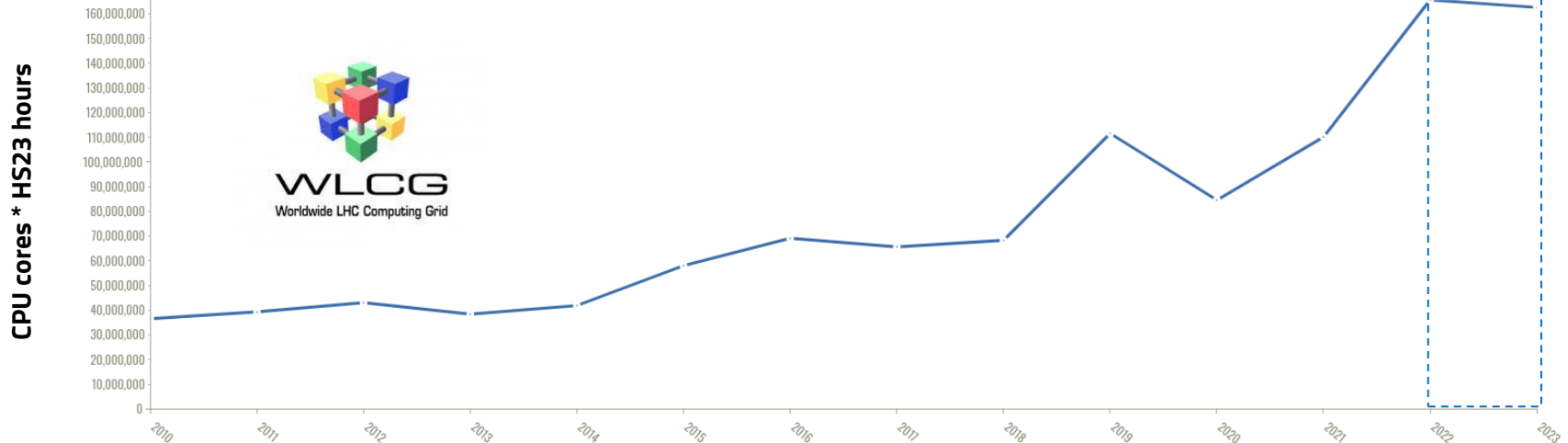
Portuguese WLCG Tier-2 ATLAS and CMS

Operating 24x365 since 2010:

- 22,471,371 jobs
- 1,131,173,940 HS23 hours

Supported experiments:

- ATLAS (50%)
- CMS (50%)



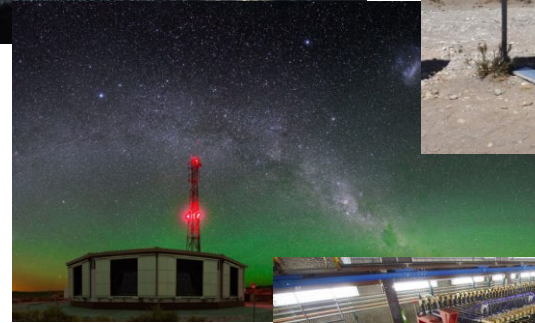
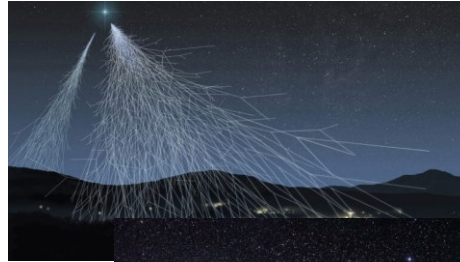
BELLA

High Energy Physics data across the atlantic



LIP common projects with Latin America

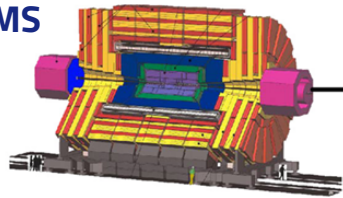
- Common CERN LHC experiments
CMS and **ATLAS**
- **AUGER @ Pierre Auger Laboratory** cosmic rays observatory in Argentina
- **MARTA** muon detector Argentina
- **DUNE** at LBNF - Fermilab US
- **SWG0** Southern Wide-Field Gamma Ray Observatory



BELLA-S



CMS



300 Tbps

Online System

CERN Switzerland

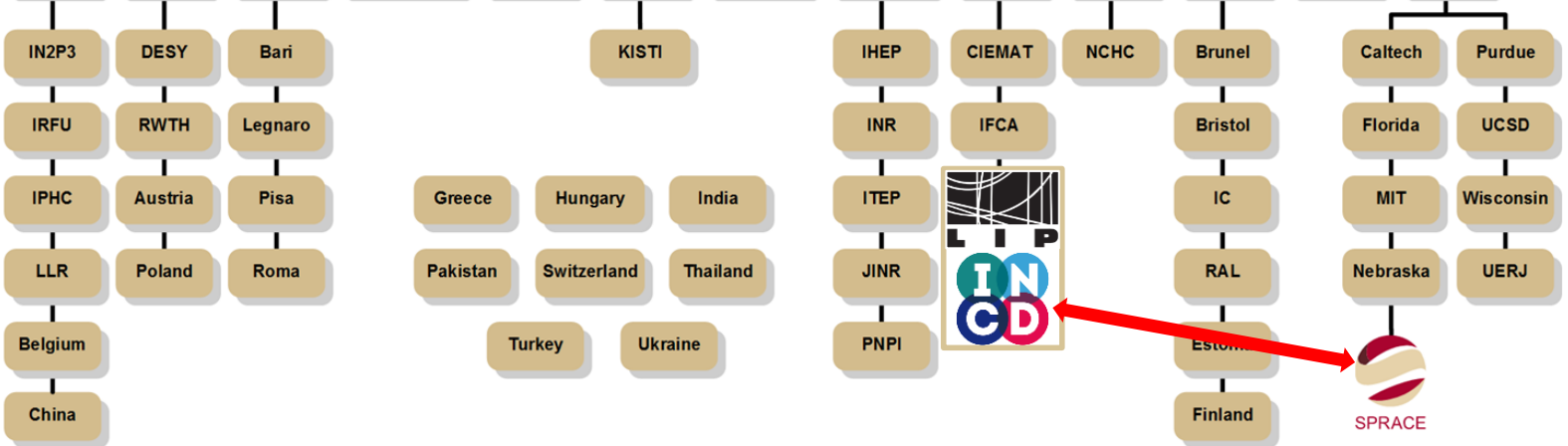
Tier-0
CERN

Structure for CMS in 2021
(for ATLAS the sites are different)

Tier-1s



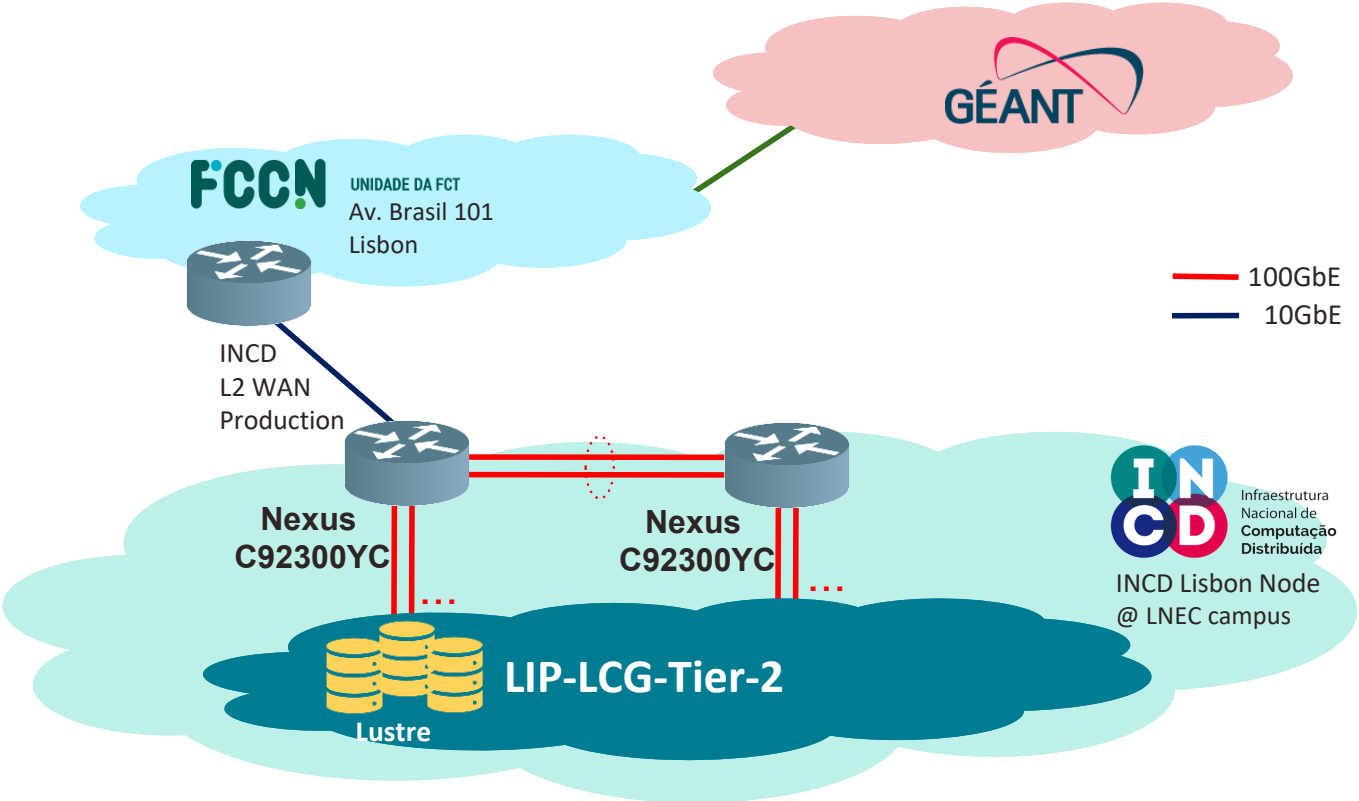
Tier-2s



LIP-Tier2 connectivity for Bella 2021

LIP-Tier-2

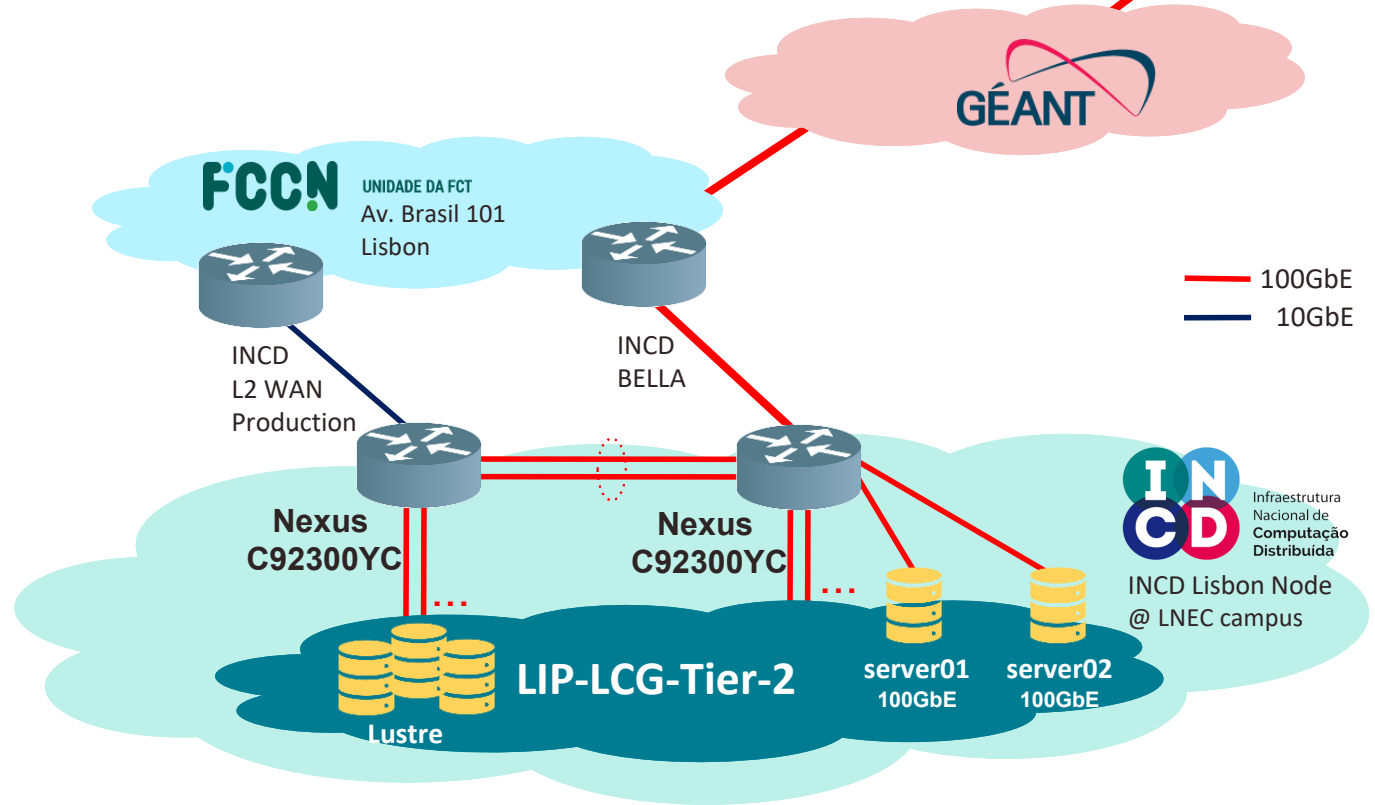
- Supports CMS and ATLAS experiments
- The Tier-2 is part of the INCD infrastructure



LIP-Tier2 connectivity for Bella 2021

LIP-Tier-2

- Supports CMS and ATLAS experiments
- The Tier-2 is part of the INCD infrastructure



server01 Xeon Silver 4110 2.1GHz 16C 160GB RAM + Mellanox NIC + NVME
 server02 Epyc 7501 2.0GHz 64C 512GB RAM + Mellanox NIC + NVME



Latency between the PT CMS Tier-2 and BR CMS tier-2

Before Bella

- Latency to SPRACE ~ 254 ms

```
1 172.16.203.254 (172.16.203.254) 0.437 ms
2 172.16.100.1 (172.16.100.1) 0.346 ms
3 Router63.Lisboa.fccn.pt (193.137.1.233) 0.698 ms
4 Router30.Lisboa.fccn.pt (194.210.6.112) 0.617 ms
5 Router1.Lisboa.fccn.pt (194.210.6.103) 0.752 ms
6 fccn.mx2.lis.pt.geant.net (62.40.124.97) 0.407 ms
7 ae4.mx1.mad.es.geant.net (62.40.98.97) 9.513 ms
8 ae7.mx1.gen.ch.geant.net (62.40.98.67) 44.189 ms
9 ae6.mx1.par.fr.geant.net (62.40.98.183) 36.771 ms
10 ae5.mx1.lon2.uk.geant.net (62.40.98.178) 43.299 ms
11 ae6.mx1.lon.uk.geant.net (62.40.98.36) 44.102 ms
12 internet2-gw.mx1.lon.uk.geant.net (62.40.124.45) 118.094 ms
13 ae-1.4079.rtsw.atla.net.internet2.edu (198.71.45.6) 131.068 ms
14 et-3-0-0.4079.rtsw.jack.net.internet2.edu (162.252.70.43) 136.614 ms
15 198.71.45.189 (198.71.45.189) 148.902 ms
16 ae0-2005.rt04.ce.ampath.net (190.103.185.11) 257.684 ms
17 143-108-254-242.ansp.br (143.108.254.242) 253.750 ms
18 200.136.80.225 (200.136.80.225) 253.616 ms !X
```

With Bella

- Latency to SPRACE ~ 106 ms

```
1 172.16.203.254 (172.16.203.254) 0.382 ms
2 194.210.4.169 (194.210.4.169) 1.162 ms
3 Router30.Lisboa.fccn.pt (194.210.6.108) 0.562 ms
4 Router1.Lisboa.fccn.pt (194.210.6.103) 0.646 ms
5 fccn.mx2.lis.pt.geant.net (62.40.124.97) 0.495 ms
6 redclara-gw.lis.pt.geant.net (62.40.127.151) 62.728 ms
7 for-sao.redclara.net (200.0.204.7) 106.989 ms
8 sprace01.redclara.net (200.0.207.116) 106.452 ms !X
```

Traceroutes from the
LIP Tier-2 to SPRACE

Latency between the PT CMS Tier-2 and BR CMS tier-2

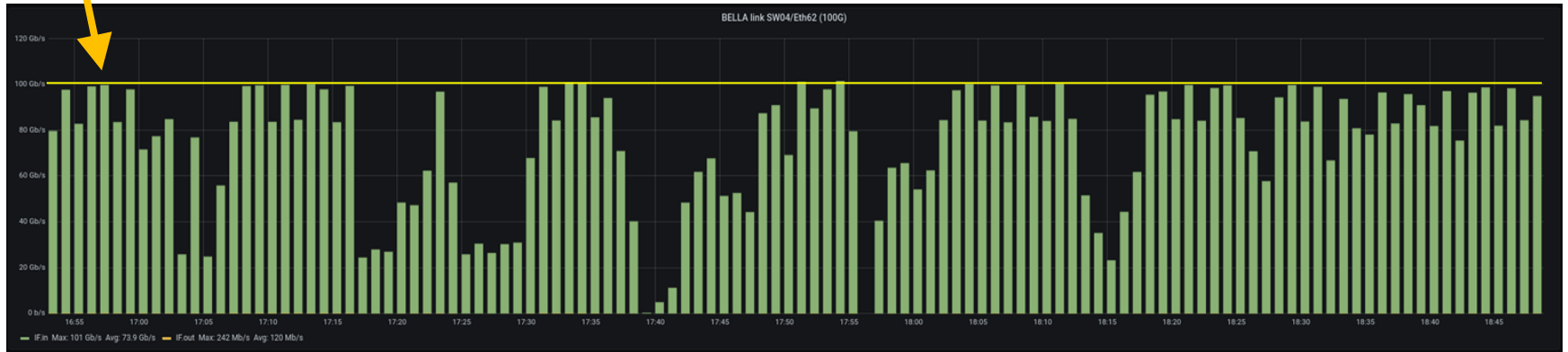
Production 10GbE

- Peak to SPRACE ~ 6 Gbps
- Using ~ 800 iperf streams

Using Bella 100GbE

- Peak to SPRACE ~ 100 Gbps
- Using ~ 480 iperf streams

100Gb/s



Using 2x INCD servers and 2x SPRACE servers (with 100GbE interfaces)

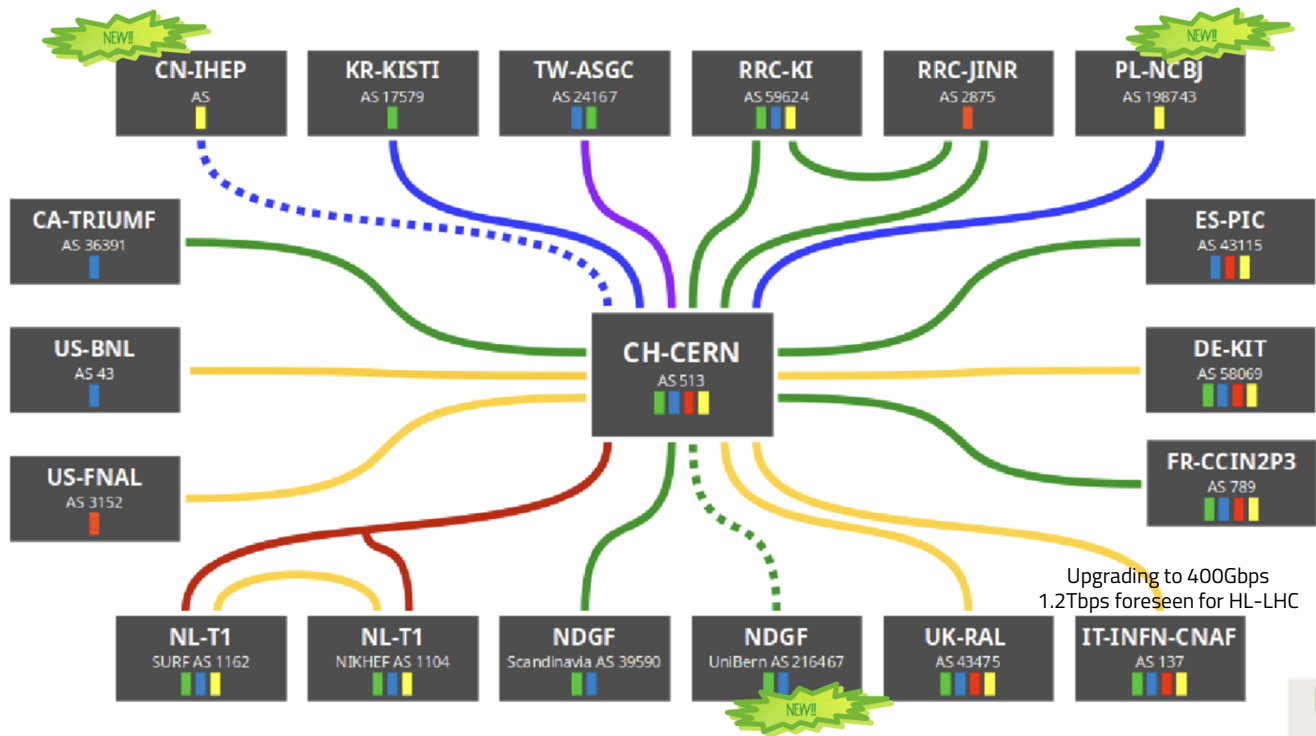
LHCOPN

LHC Optical private Network

LHCOPN

LHCOPN

LHC Optical Private Network (LHCONE)
Interconnects the WLCG Tier-1 centres worldwide



Star topology

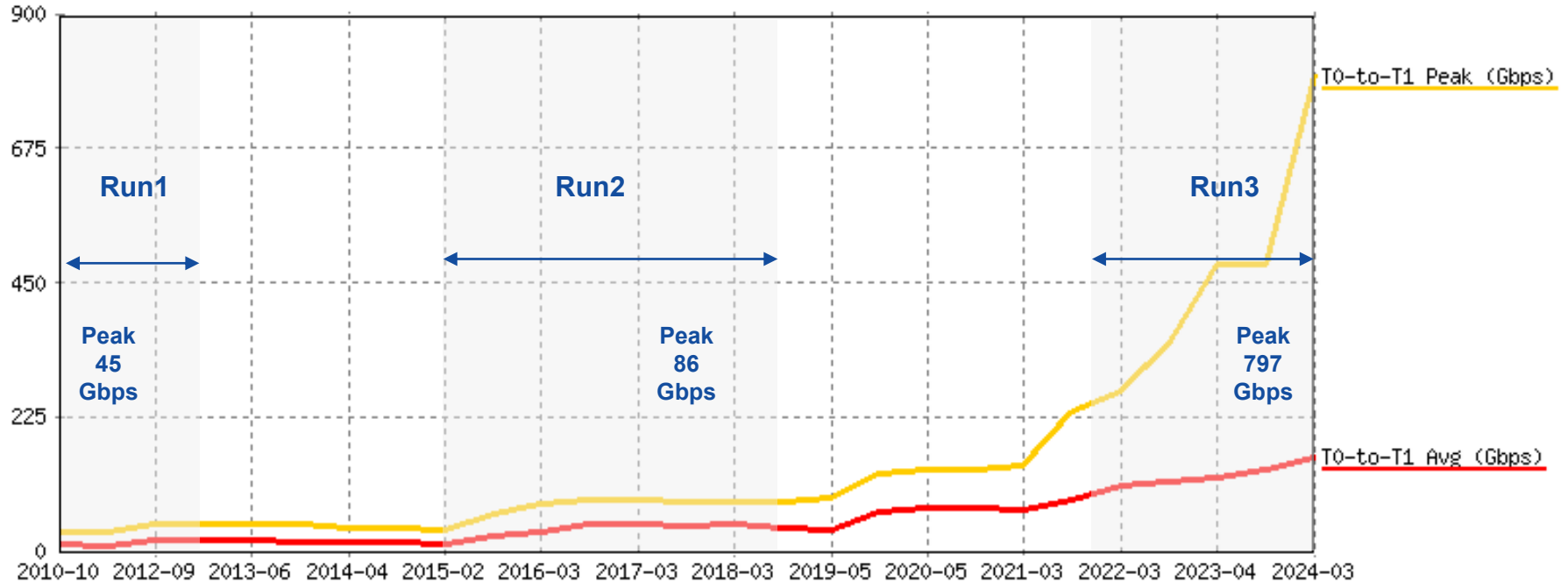
- 1x Tier-0
- 15x Tier-1s
- 18x Sites
- 15x countries
- 3x continents

Requirement for 2029: 1 Tbps per site

■ = Alice ■ = Atlas ■ = CMS ■ = LHCb
■ 10Gbps
■ 20Gbps
■ 100Gbps
■ 200Gbps
■ 400Gbps

edoardo.martelli@cern.ch 20231003

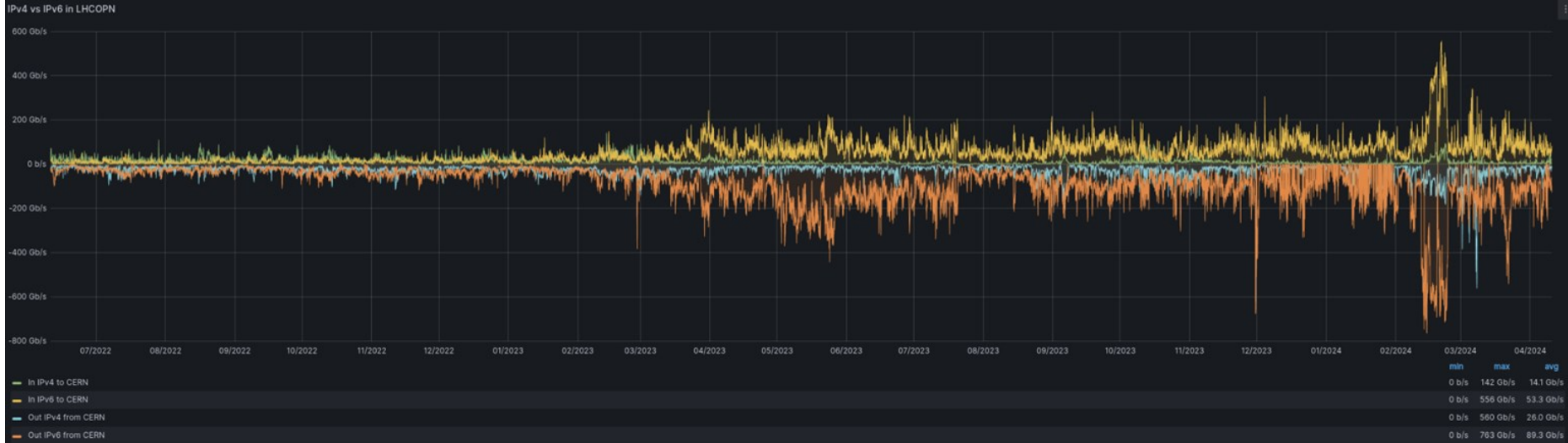
LHCOPN network traffic



LHCOPN network traffic from CERN Tier0 to all the aggregated Tier1s

CERN Tier0 outgoing → WLCG Tier1 centres (flow of data from the experiments at CERN to T1s)

LHCOPN traffic IPv4 / IPv6



CERN Light - LHCOPN IPv6 and IPv4 (IN+OUT)
Last 22 months

LHCONE

LHC Open Network Environment



Private Network connecting WLCG sites (Tier-1s + Tier-2s)

- Dedicated network for LHC
 - **ATLAS, CMS**, ALICE LHCb
- A collaborative effort among Research & Education Network Providers
- Multi domain **L3 VPN**
- Routed Virtual Private Network connecting Science-DMZs
- BGP communities for traffic engineering
- Only for declared IP prefixes

The LHCONE is open to other HEP scientific collaborations

- Belle II
- Xenon
- Juno
- Nova
- **Pierre Auger Observatory**
- **DUNE**

The

- > 33 countries
- > 144 sites



LHCONE recent upgrades

- CERN upgrade to 400G with ESnet
- CERN upgrade to 2x 400G with GEANT

New sites in LHCONE:

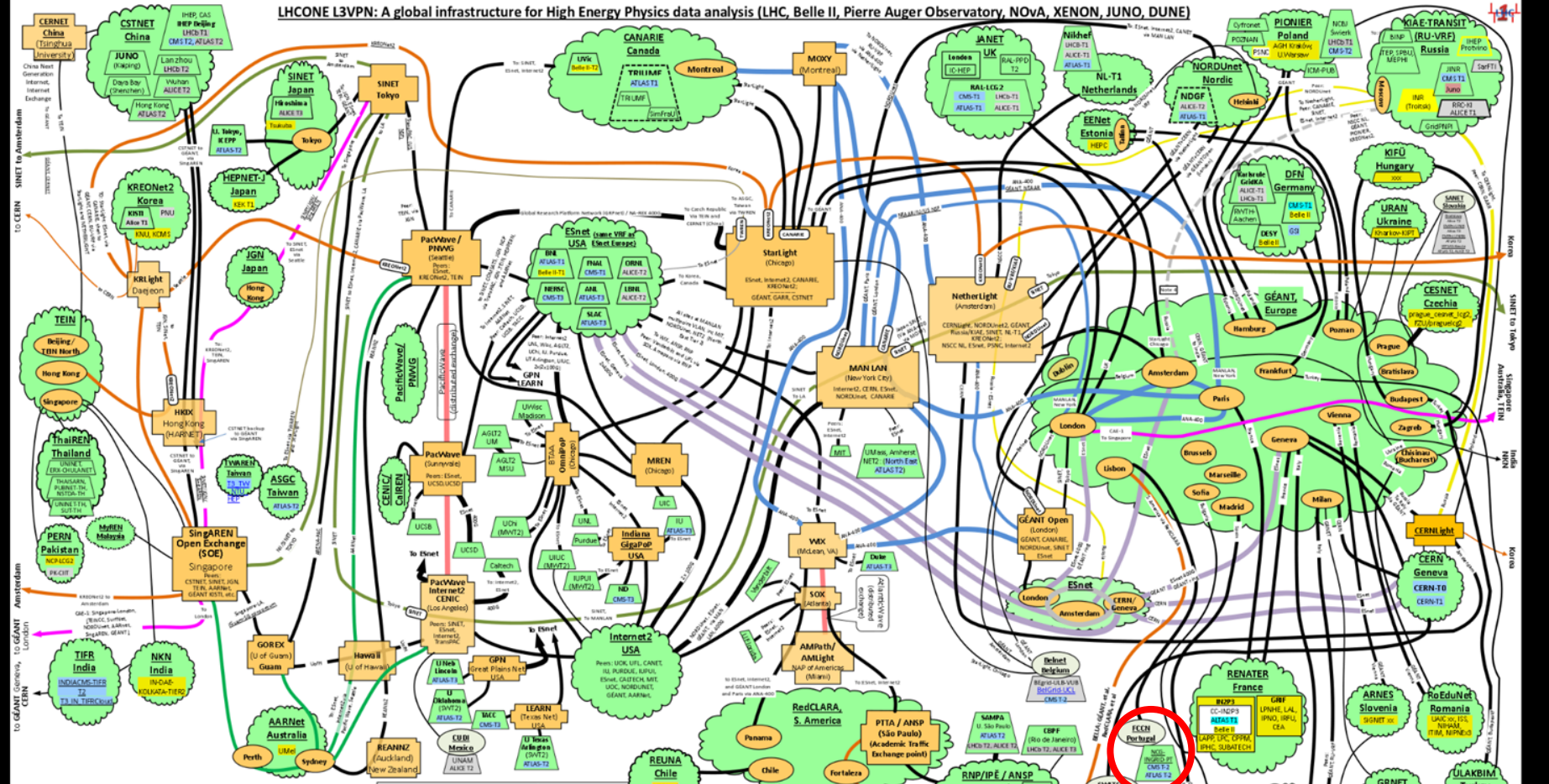
- Lawrence Berkeley National Laboratory (US)
- University of Massachusetts (US)
- **LIP (Portugal) ←**

Traffic:

- Seen first LHCONE peak above 1 Tbps in GEANT



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



LHCONE Map Ver. 9.0, 2024-03-24 - WEJohnston, ESnet, wej@es.net

LHCONE VRF domain/agggregator
 - A provider network.
 - Connector network or institution - provides, e.g. an L2 path between VRFs.
 - Underlined link information indicates link information, not use.
 - W/O sites that are not connected to LHCONE.
 - Exchange point.
 - Future site.

International Infrastructure by provider/collaboration

Various	Various	SINET
ADDnet	NORDUnet	NAE, Russia
GEANT	NAE, Russia	KREONet, Korea
SINET, Japan, global ring	BELLA: GEANT, et al	RedCLARA, et al
TWAREN, Taiwan	ESnet transatlantic, USA	
ESnet transatlantic, USA		
SINET/ISN/SingAren		
ANA-300/400 - Various links provided by CANARIE, ESnet, GEANT, Internet2, NORDUnet, SURFnet, SNET IU/NSF		

SHRTE: LHC ALICE or LHCb site
CHAF-T1: LHC Tier 1 ATLAS and CMS
JOH: LHC Tier 2/3 ATLAS and CMS
BEI: Belle II Tier 1/2
JUNO: JUNO

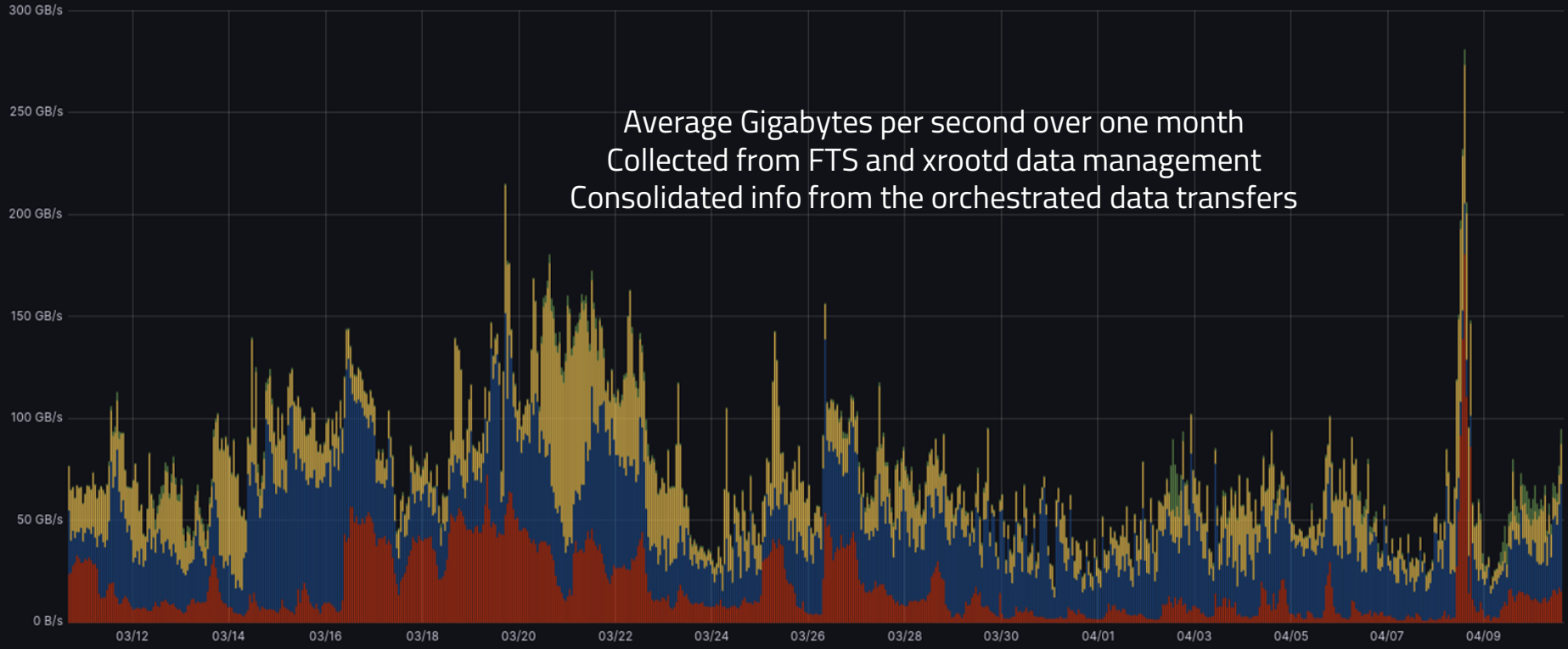
Sites that are standalone VRFs

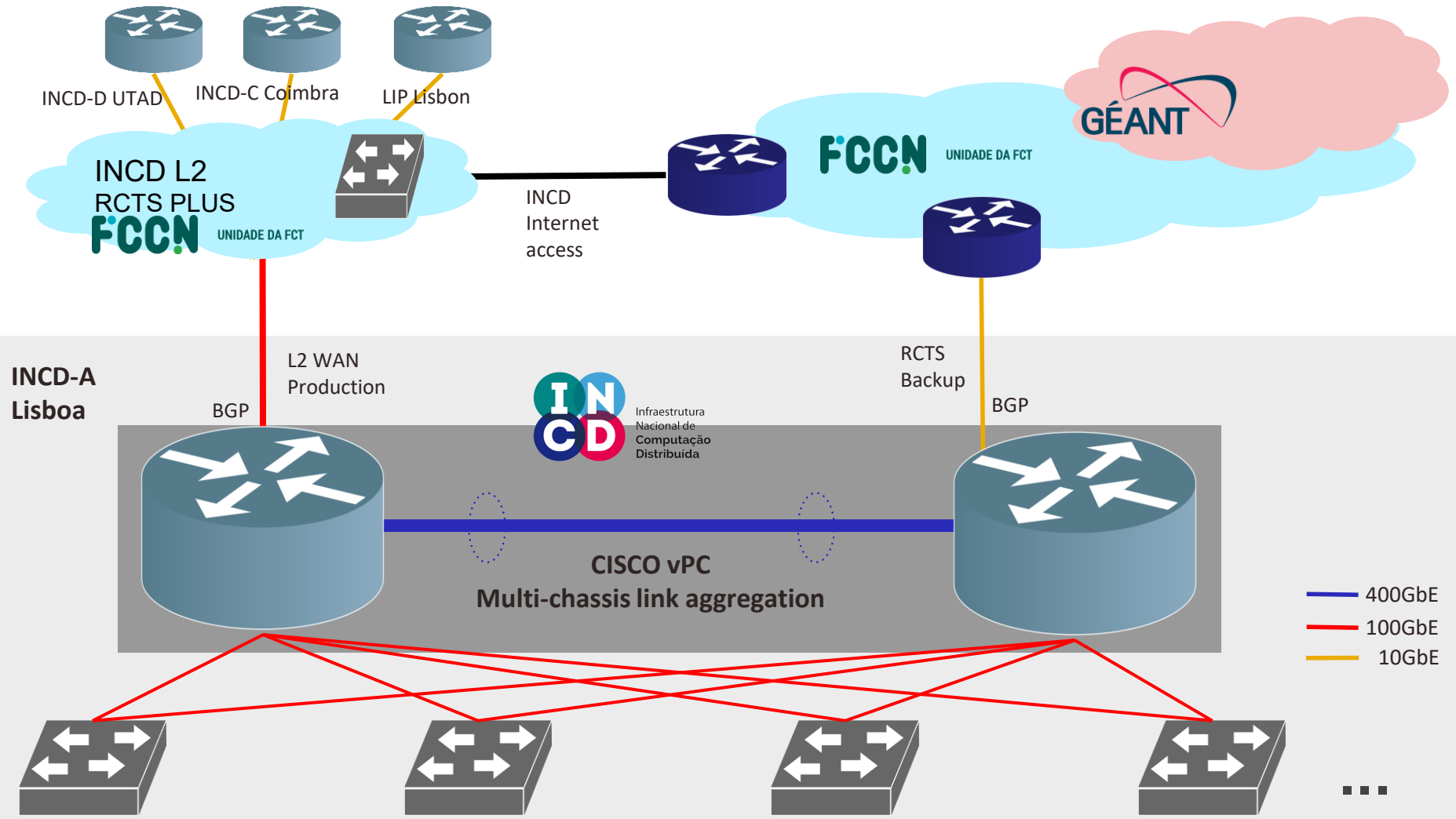
NOTES
 1) Only links involved in LHCONE are shown
 2) LHCOPN links are not shown on this diagram
 3) For more exploration see "Interpreting the LHCONE Map" at <https://www.cern.ch/en/press/2024/03/24/interpreting-the-lhc-one-map>
 4) GEANT and CANARIE have shutdown the peering between their VRF and NAE, as a result of the Ukraine war.

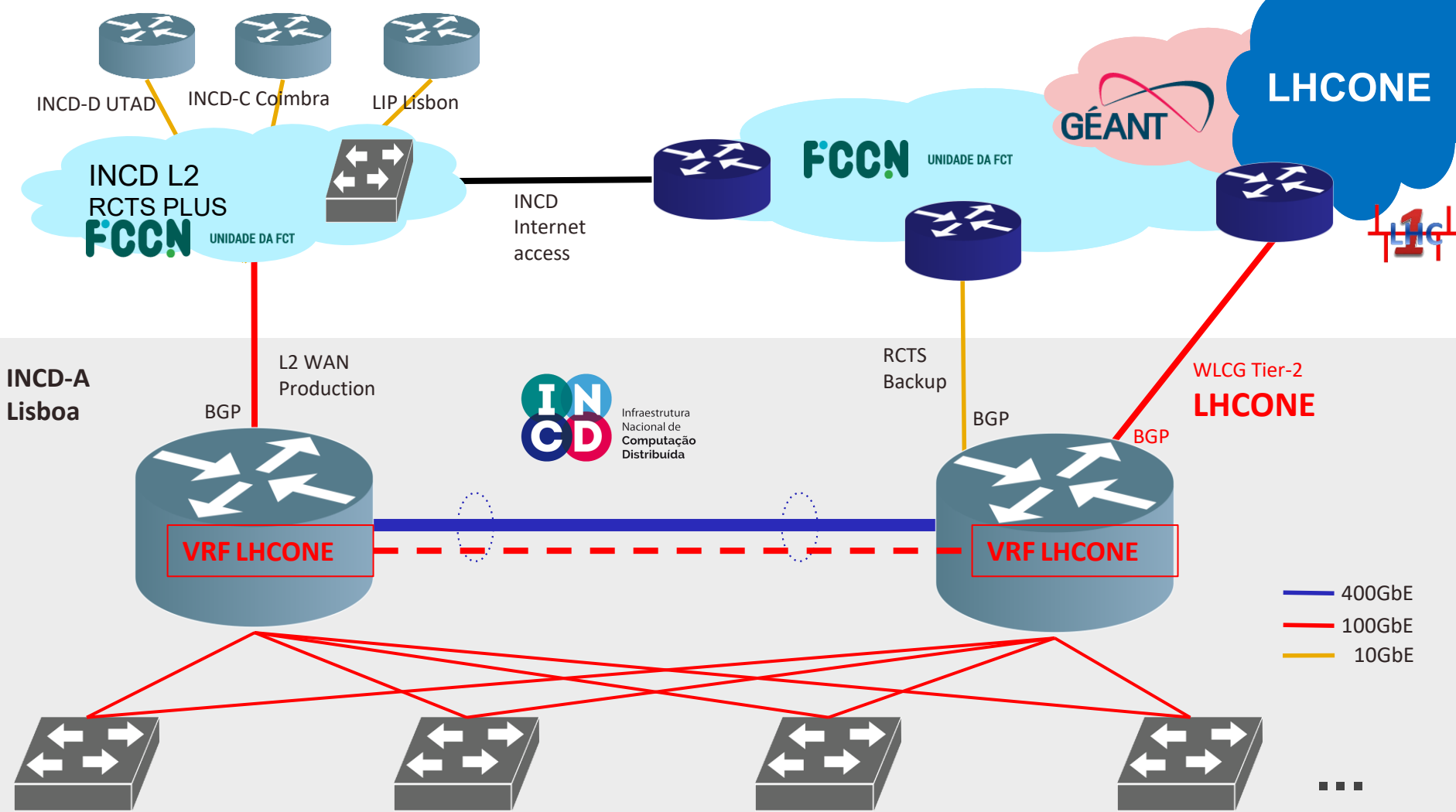
Legend:
 - Green circle: LHCONE VRF domain/agggregator
 - Black line: Provider network
 - Orange line: Connector network or institution
 - Underlined link: Underlined link information
 - W/O sites: W/O sites that are not connected to LHCONE
 - Exchange point: Exchange point
 - Future site: Future site

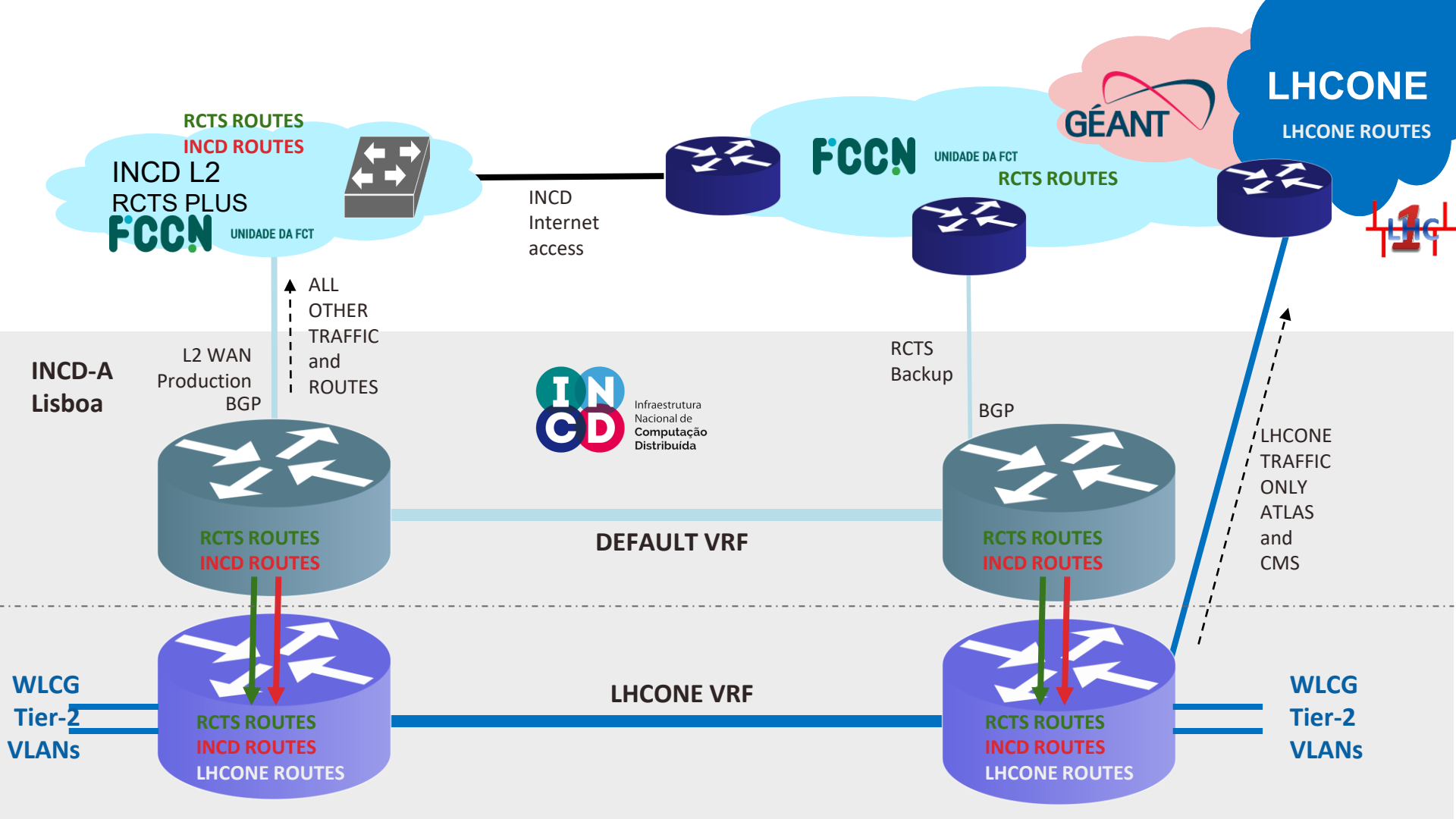
WLCG data transfers

Transfer Throughput [🔗](#)









LHCONE vs Internet

- Conventional Internet connectivity
- Personar latency node @Brookhaven (US)
- `mtr -6 lhccperfmon.bnl.gov`

- LHCONE connectivity
- Personar latency node @Brookhaven (US)
- `mtr -6 lhccperfmon.bnl.gov`

My traceroute [v0.85]
wn302.a.incd.pt (::)
2024
Wed Apr 10 19:00:23

Host StDev	Packets		Pings			
	Loss%	Snt	Last	Avg	Best	Wrst
1. 2001:690:2150:aa:28:ffff:ffff:3 0.0	0.0%	61	0.7	0.8	0.6	1.1
2. Router61.Lisboa.fccn.pt 0.0	15.0%	61	0.9	1.0	0.8	1.9
3. Router60.Lisboa.fccn.pt 0.0	0.0%	61	6.5	6.5	6.4	7.4
4. Router40.Porto.fccn.pt 0.0	0.0%	61	6.7	6.6	6.4	6.8
5. Router2.Porto.fccn.pt 0.8	56.7%	61	6.7	7.0	6.6	10.9
6. fccn-bckp-gw.rt1.por.pt.geant.net 3.8	0.0%	61	6.3	7.3	6.2	26.5
7. ae2.rt1.bil.es.geant.net 0.2	0.0%	61	36.2	35.6	35.1	36.9
8. ae4.mx1.par.fr.geant.net 2.0	0.0%	61	35.3	35.7	35.1	50.9
9. ae8.mx1.lon2.uk.geant.net 0.0	0.0%	61	31.8	32.3	31.8	32.5
10. ae2.mx1.lon.uk.geant.net 0.0	0.0%	61	32.5	32.7	32.2	32.9
11. esnet-eex.lon.uk.geant.net 0.0	0.0%	61	102.6	103.0	102.5	103.5

My traceroute [v0.85]
gftp01.ncg.ingrid.pt (::)
Wed Apr 10 18:58:45 2024

Host	Packets		Pings				
	Loss%	Snt	Last	Avg	Best	Wrst	StDev
1. 2001:690:2150:aa:35:ffff:ffff:4	0.0%	61	1.0	0.7	0.6	1.2	0.0
2. 2001:690:810:154::1	37.7%	61	1.2	1.2	0.7	6.9	0.9
3. Router30.Lisboa.fccn.pt	0.0%	61	0.9	0.7	0.5	0.9	0.0
4. ???							
5. fccn-ap1-lhccone.mx2.lis.pt.geant.net	0.0%	61	0.4	1.4	0.3	27.4	3.9
6. ae3.rt1.por.pt.geant.net	0.0%	60	26.3	26.5	26.1	36.0	1.3
7. ae2.rt1.bil.es.geant.net	0.0%	60	28.6	28.5	28.2	36.9	1.1
8. 2001:798:111:1::75	0.0%	60	28.7	28.9	28.1	41.0	2.1
9. newy32a0a-cr6--geant--se366-v6.lhccone.es	0.0%	60	101.2	101.2	101.1	101.7	0.0
10. bn1515b-cr6--bn1-se-101.lhccone.es.net	0.0%	60	100.6	100.5	100.4	101.0	0.0
11. bn1--bn1-se-101.lhccone.es.net	0.0%	60	102.6	102.6	102.5	103.1	0.0
12. ???							
13. lhccperfmon.bnl.gov	0.0%	60	102.5	102.5	102.5	102.8	0.0

ATLAS DDM (data transfers) monitoring

ATLAS ctrl+k

Home > Dashboards > DDM Monitoring > DDM Transfers > View panel ★ 🔗 🕒 Last 18 days UTC 🔍

Binning auto **Group by** dst_cloud **Activity** Analysis Input + Analysis Output + Data Consolidation + Data Reba... **Protocol** All **Src tier** All

Src country All **Src cloud** All **Src site** All **Src endpoint** All **Src token** All **Dst tier** All **Dst country** Portugal **Dst cloud** All

Before LHCONE

Efficiency

	CA	CERN	DE	ES	FR	IT	ND	NL	RU	UK	US
NCG-INGRID-PT_DATADISK	66%	43%	76%	11%	67%	88%	11%	77%	94%	85%	87%

With LHCONE

Efficiency

	CA	CERN	DE	ES	FR	IT	ND	NL	RU	UK	US
NCG-INGRID-PT_DATADISK	98%	100%	100%	91%	82%	90%	96%	100%	100%	95%	99%

Evolution and final remarks

Ongoing

IPv6 in WLCG

- WLCG now supports IPv6-only clients
- Tier-1s: complete; Tier-2s: 97% storage is IPv6 capable
- Most data transfers use IPv6
- We have concentrated on ensuring use of IPv6
 - LHCOPN/LHCONE can be 90-95% IPv6 - **but not always!**
- We continue to address more obstacles to IPv6 in WLCG
 - All WLCG CPU services and CPU clients to dual-stack is priority now
 - Approved by WLCG MB on 17 Oct 2023
- **End point is still IPv6-only services (IPv4 is “legacy” networking)**
- ***Message to new research communities - build on IPv6 from start***

multiONE

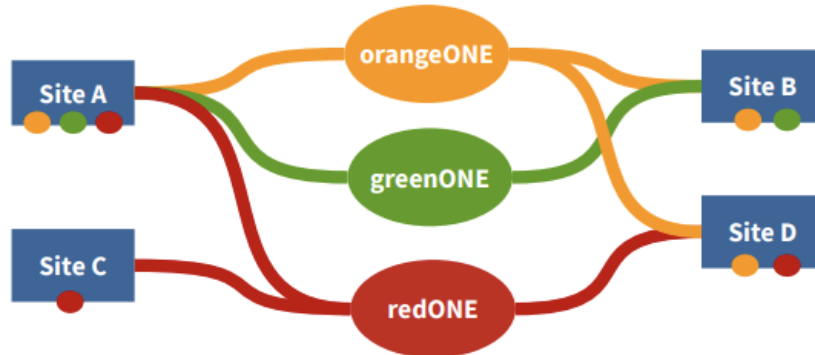
LHCONE already very large, it could become risky to include other large science projects

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



Networking for the LIP Tier-2

Developments ongoing

- Monitoring
 - perfSONAR 5 instances (bandwidth, latency)
 - Improve monitoring
- Jumbo frames
 - For data transfers within LHCONE
 - First for storage gateways then for compute nodes
- Storage
 - Increase the number of storage gateways significantly
- Compute nodes
 - Separate traffic and route to LHCONE destinations
 - IPv6 as preferred protocol and bypass firewall for file transfers
- Enlarge usage to other projects
 - DUNE and AUGER



LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS
partículas e tecnologia

Thanks!

**Discovery
through
science**

**Innovation
through
technology**

**Sharing
with People**

Use of the IPv6 Flow Label for WLCG Packet Marking

draft-cc-v6ops-wlwg-flow-label-marking-02

Workgroup: Internet Engineering Task Force
Internet-Draft: draft-cc-v6ops-wlwg-flow-label-marking-02
Published: 10 July 2023
Intended Status: Informational

**Network Flow and Packet
Marking for Global Scientific
Computing (scitags.org)**

- **IPv6 flowlabel tagging**

or another option

- **Firefly flow marking**

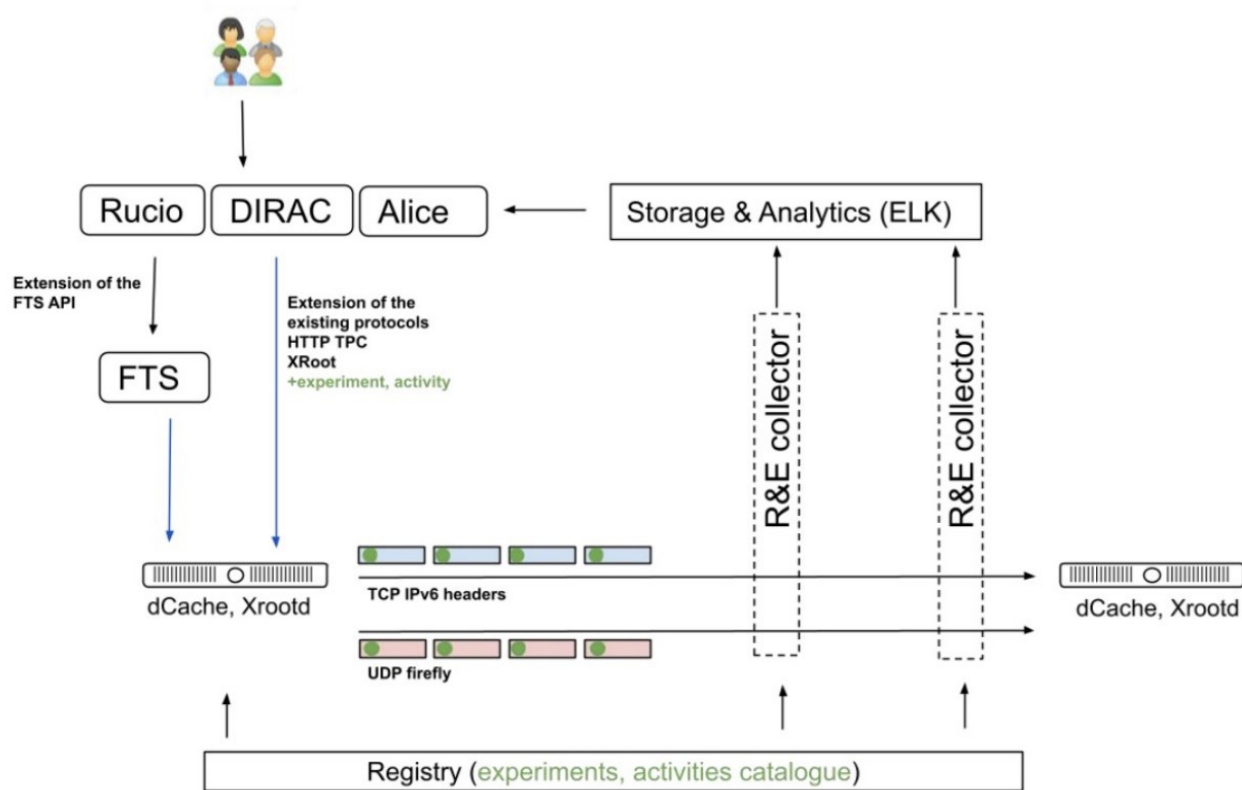
Use of the IPv6 Flow Label for WLCG Packet Marking

Abstract

This document describes an experimentally deployed approach currently used within the Worldwide Large Hadron Collider Computing Grid (WLCG) to mark packets with their project (experiment) and application. The marking uses the 20-bit IPv6 Flow Label in each packet, with 15 bits used for semantics (community and activity) and 5 bits for entropy. Alternatives, in particular use of IPv6 Extension Headers (EH), were considered but found to not be practical. The WLCG is one of the largest worldwide research communities and has adopted IPv6 heavily for movement of many hundreds of PB of data annually, with the ultimate goal of running

Use of the IPv6 Flow Label for WLCG Packet Marking

scitags.org



SENSE to move data

Project led by UCSD and Caltech

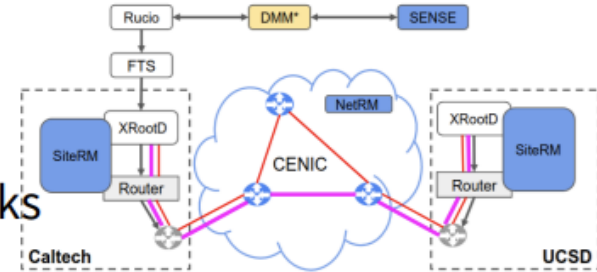
The increased requirements of the HL-LHC requires to use any resource in the most efficient way, including networks

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



CERN

connectivity

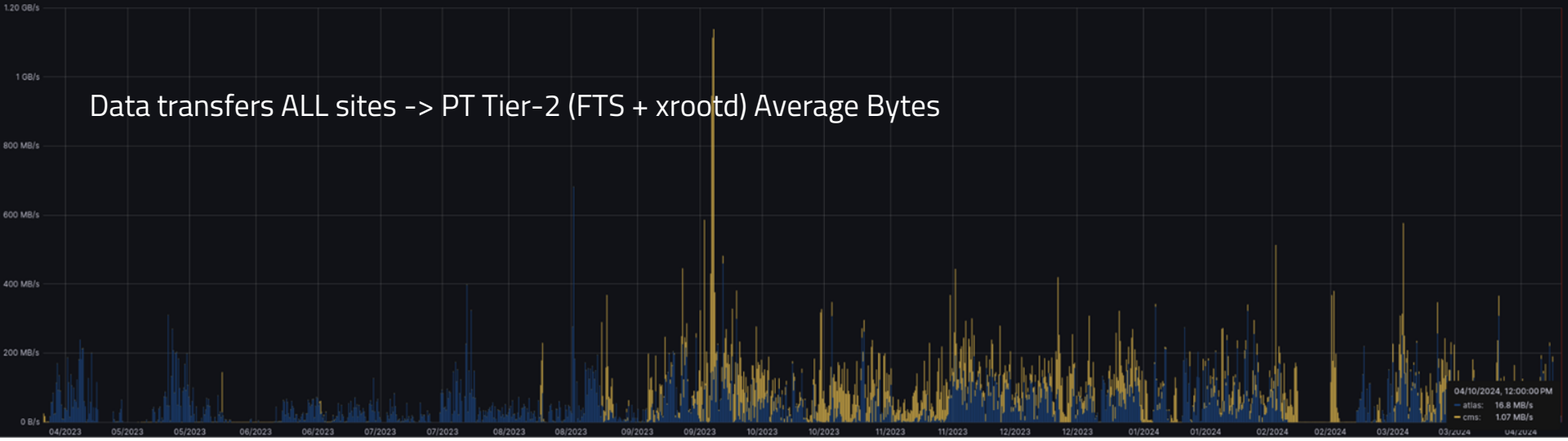
Experiments' DAQ lines to CERN IT data-centre, Capacity in place for LHC Run3:

- ALICE: 3.2 Tbps
- LHCb: 400 Gbps
- CMS: 400 Gbps
- ATLAS: 200 Gbps

CERN External Network: some numbers:

- LHCOPN capacity: 2.1 Tbps
- LHCONE capacity: 1.2 Tbps
- Internet capacity: 1 Tbps
- Stateful Firewall capacity: 0.6 Tbps
- HTAR (firewall bypass) capacity: 0.6 Tbps

Data transfers ALL sites -> PT Tier-2 (FTS + xrootd) Average Bytes



Data Transfers PT Tier-2 -> ALL sites (FTS + xrootd) Average Bytes

