



Red en los centros de computación del LHC (CERN): desafíos futuros y su repercusión en España

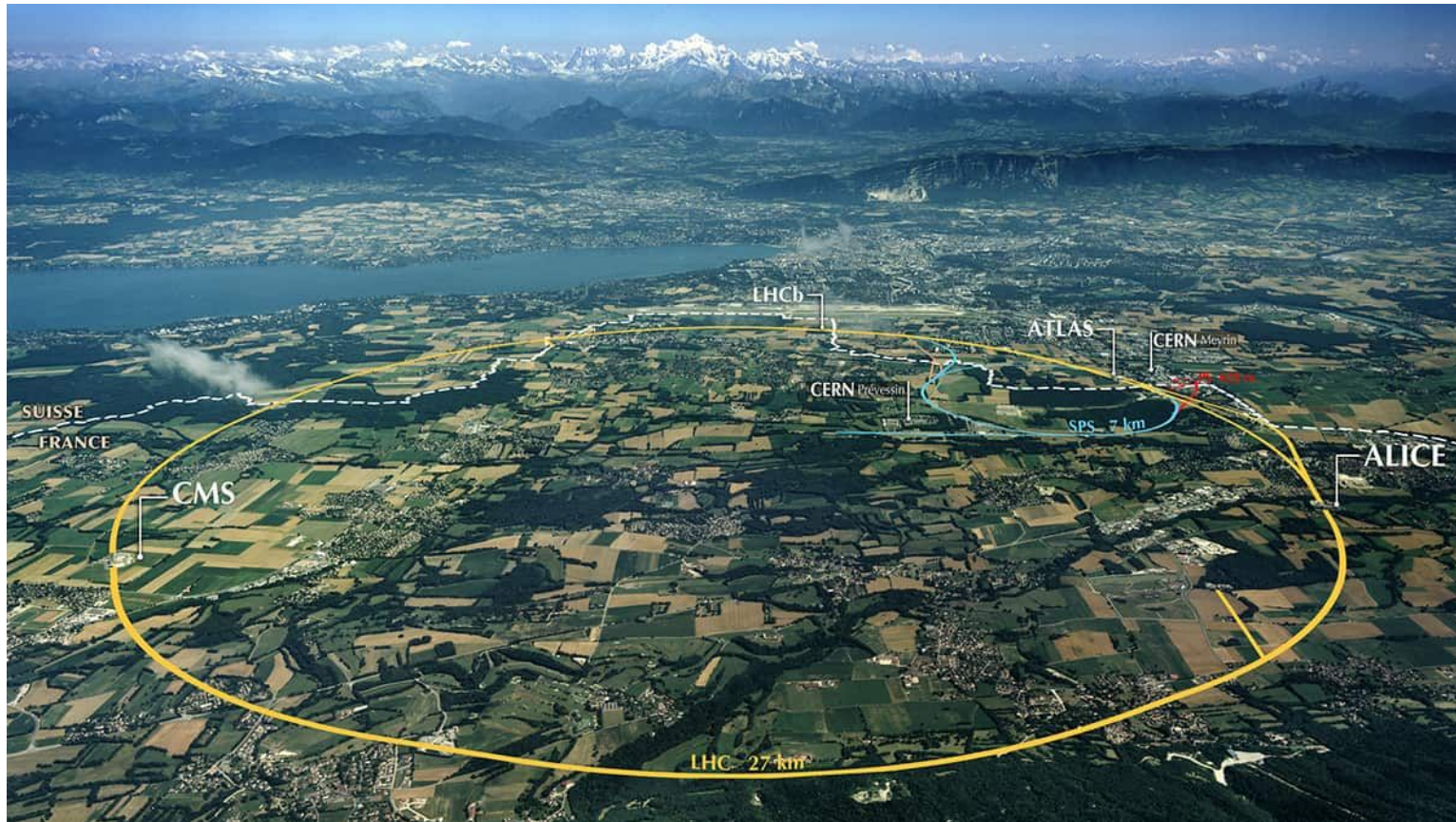
J. Flix

Jornadas Técnicas RedIRIS 2024

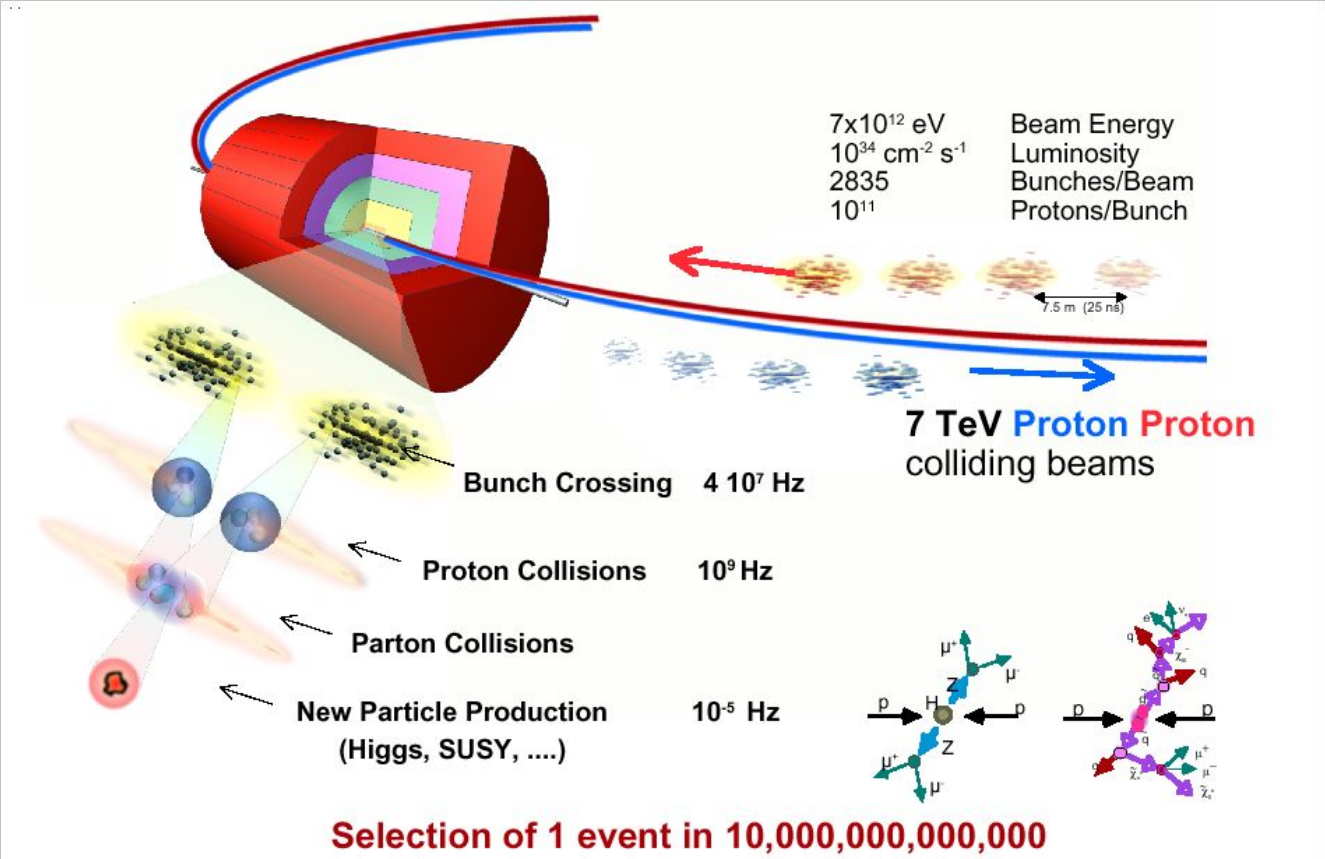
Mallorca
28-30 May 2024



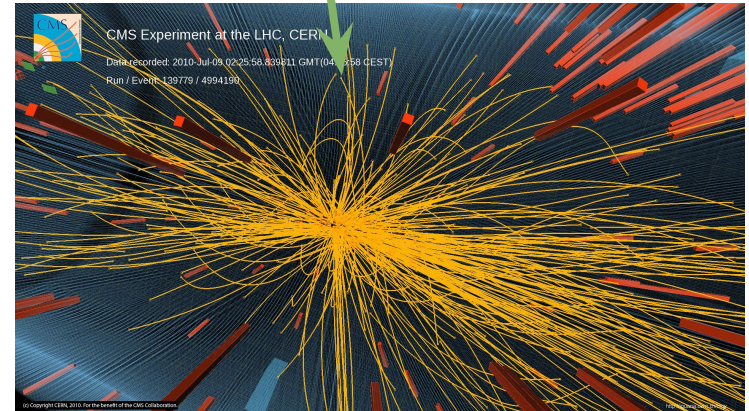
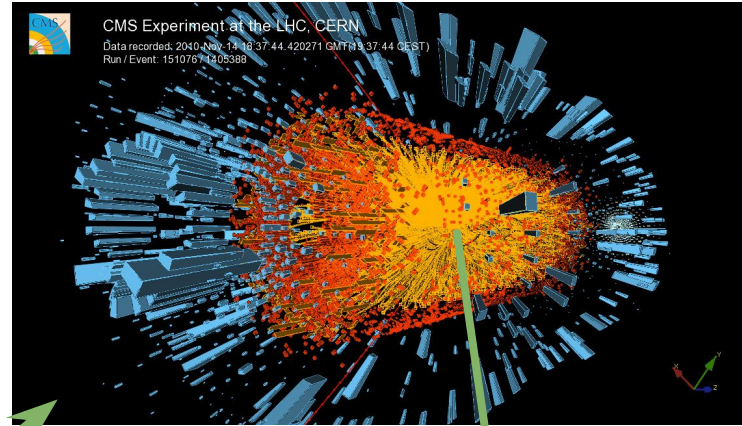
Large Hadron Collider (CERN)



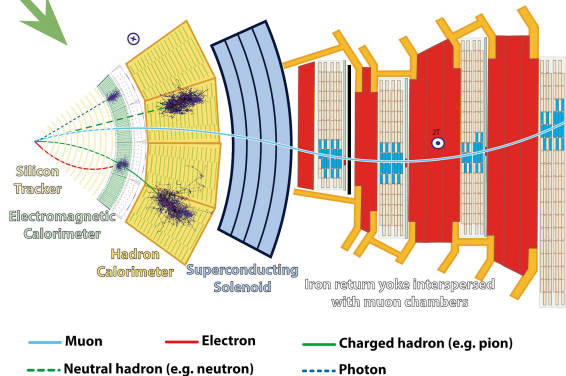
Large Hadron Collider (CERN)



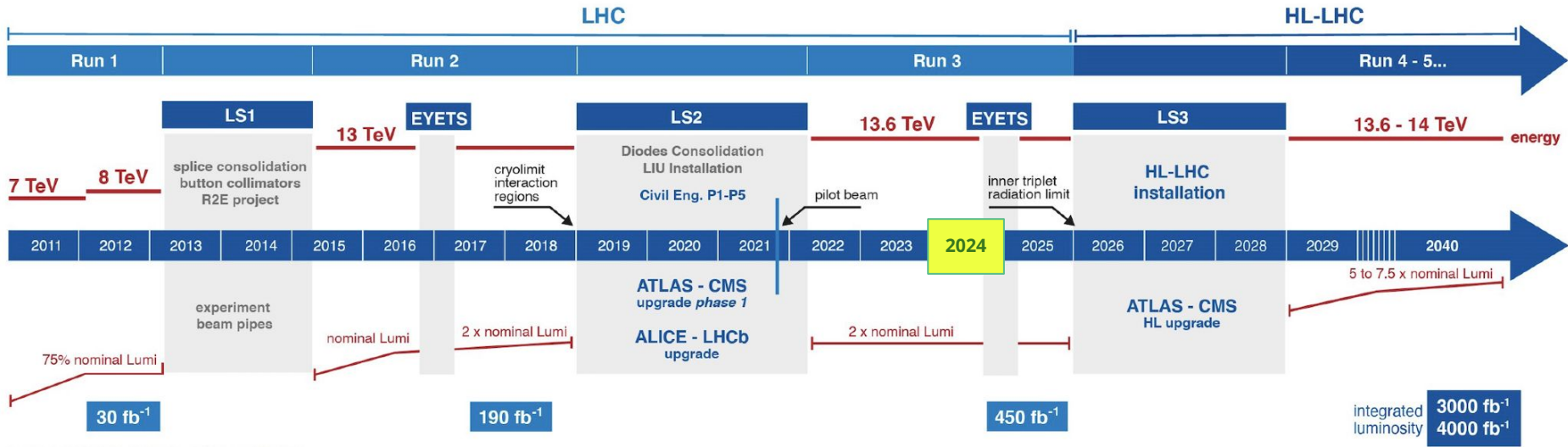
Large Hadron Collider (CERN)



~3 kHz por detector
1 MB/evento



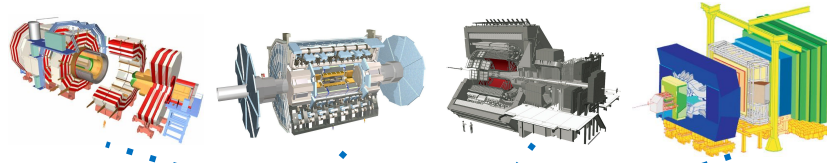
Large Hadron Collider (CERN)



~15 años operando → **75 PB/año de datos** (raw + reducidos + simulaciones)

Ha sido el primer experimento científico en generar **1 Exabyte de datos**

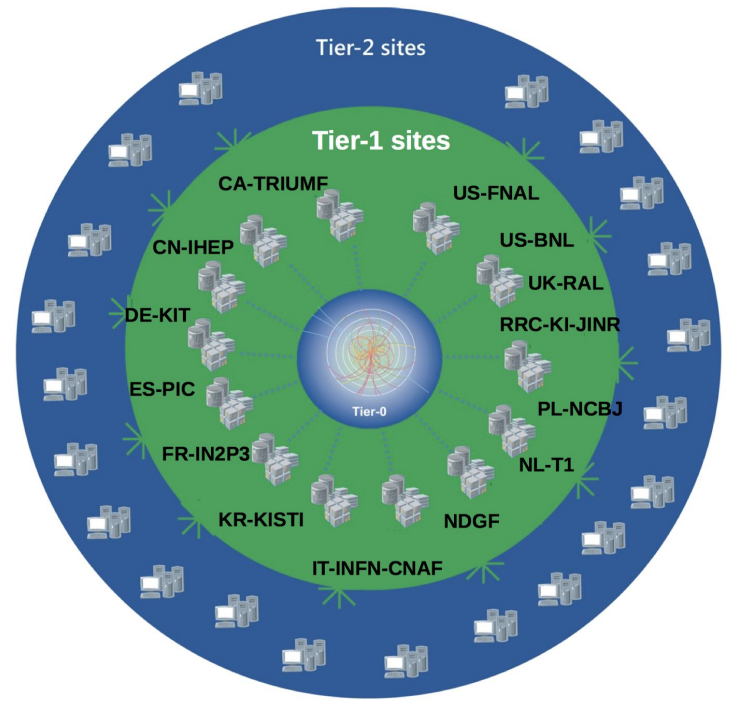
Modelo computación distribuida



Tier-0
Data source
Full data on Tape
Data reconstruction

Tier-1s
Distributed 2nd copy on Tape
Data Reconstruction
Simulations, Data analyses

Tier-2s
Simulations, Data analyses



Worldwide LHC Computing Grid (WLCG)



WLCG

Worldwide LHC Computing Grid



Global collaboration

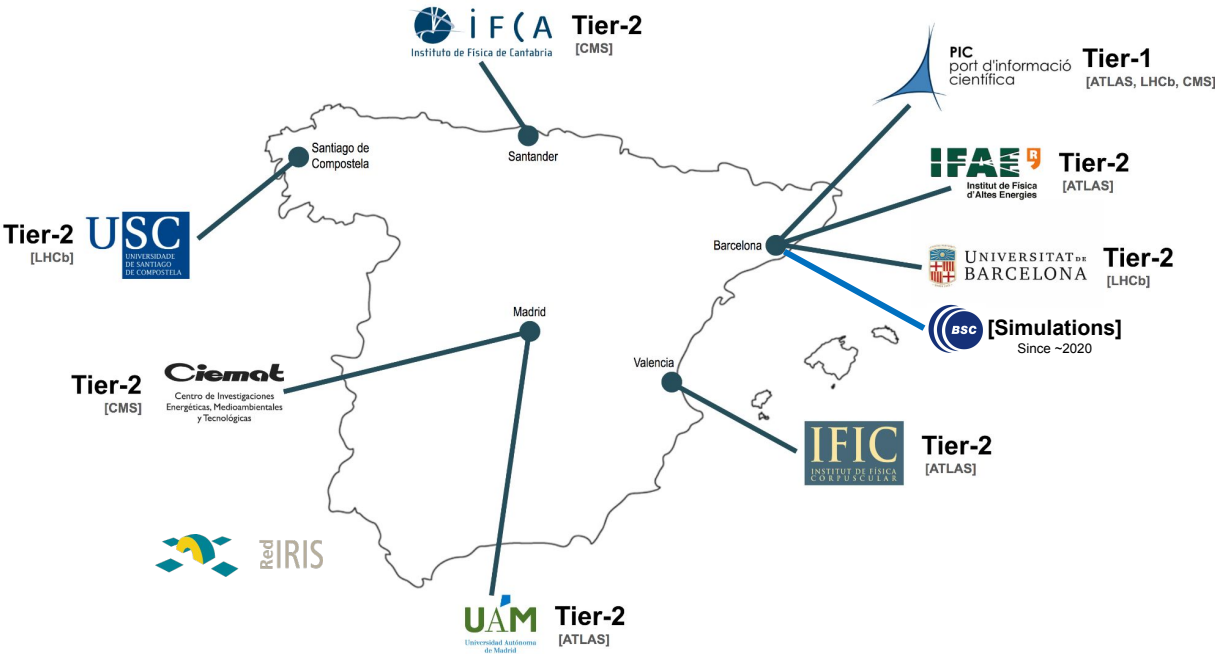
42 countries

170 computing centres

Over 1 million computer cores

2 exabytes of storage

Recursos WLCG en España



	CPU-cores	Disk (PB)	Tape (PB)
PIC-T1	12000	14	45
ATLAS-T2 IFIC+IFAE+UAM	5500	8	
CMS-T2 CIEMAT+IFCA	5000	6	
LHCb-T2 UB+USC	500		
BSC	~8000		
Total	31000	28	45

**Sistema distribuido de computación
desplegado desde ~2003**

~200M CPU-hours + ~70M CPU-hours (BSC)

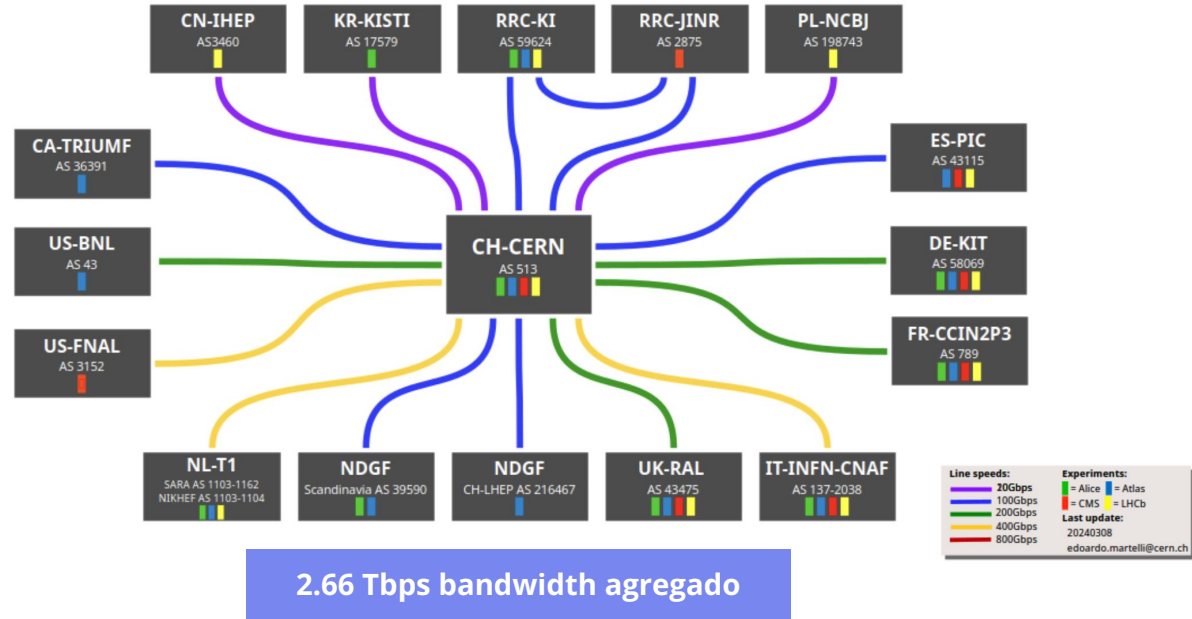
~4% del total de recursos de WLCG tanto a nivel Tier-1 como Tier-2
(con el objetivo de volver a una contribución del 5%)

Redes privadas en WLCG

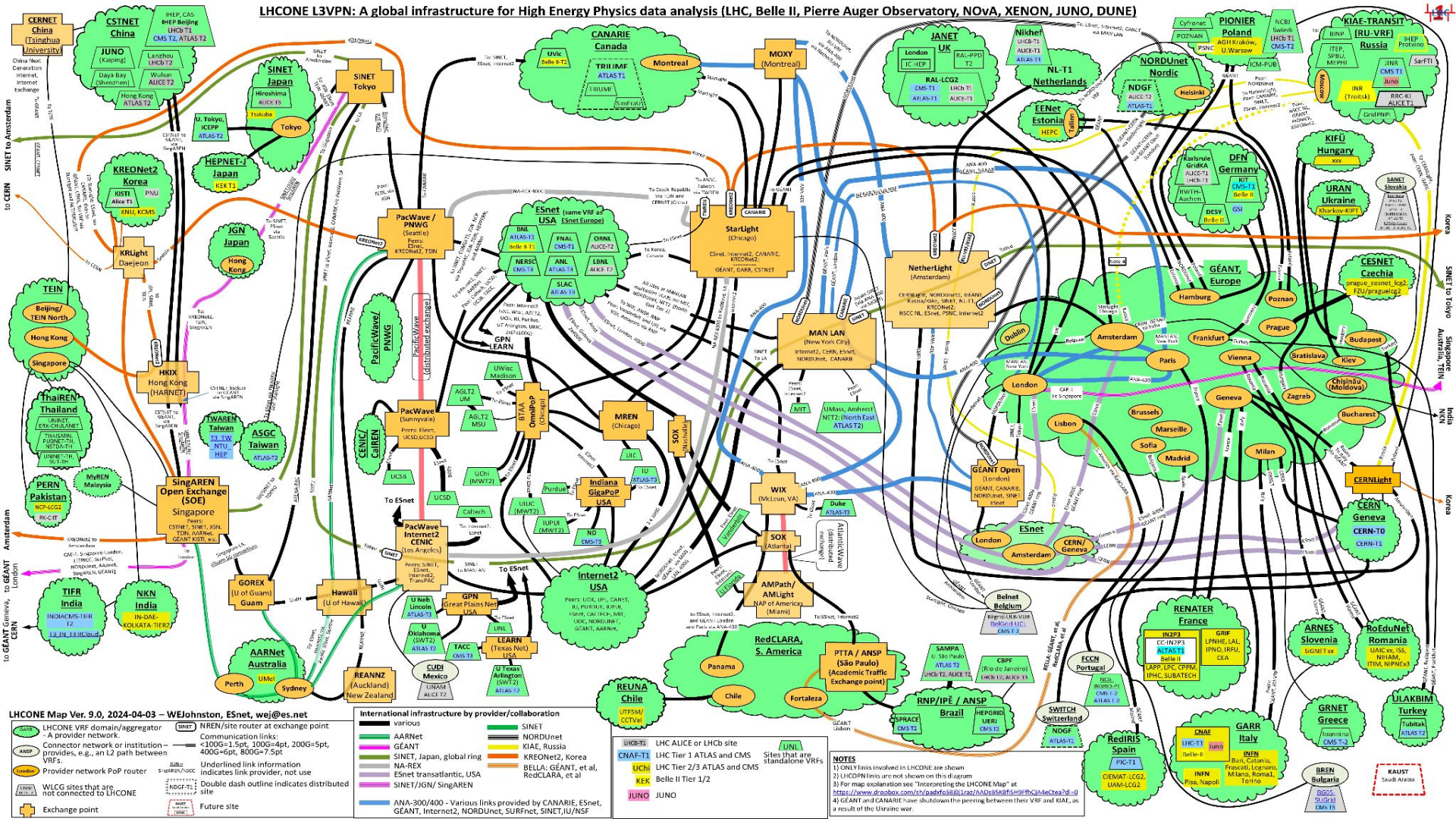
LHCOPN (Tier-0 ↔ Tier-1s): enrutament avançat BGP. Tràfic exclusiu LHC

LHCONE (Tier-1s ↔ Tier-2s: Layer3 VPN multi-dominio. LHC + altres ciències

LHCOPN



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, 2024-04-03 – WEJohnston, ESnet, wej@es.net

LHCONE VRF domain/aggregator
A provider network
Connector network or institution—provides, e.g., an L2 path between VRFs

Provider network PoP router
Wildcag sites that are not connected to LHCONE

Exchange point

NREN/Site router at exchange point
A provider network
<100G=1.5pt, 200G=3pt, 400G=6pt, 800G=7.5pt

Underlined link information indicates link provider, not use

Double dash outline indicates distributed

Future site

International infrastructure by provider/collaboration

various	SINET
ARNet	NORDUnet
GEANT	KREONet2, Korea
SINET, Japan, global ring	BELLA, GEANT, et al
NA-REX	RedCLARA, et al
SINET/US/Canada, USA	
SINET/Japan/SingAren	
ANA-300/400+ Various links provided by CANARIE, ESnet, GEANT, Internet2, NORDUnet, SURFnet, SINET, US/NSF	

US/UK LHC ALICE or LHCb site

CNAF-T1 LHC Tier 1 ATLAS and CMS

UCHI LHC Tier 2/3 ATLAS and CMS

KEK Belle II Tier 1/2

JUNO JUNO

UNL Sites that are standalone VRFs

NOTES

- ONLY links involved in LHCONE are shown
- LHCOPN links are not shown on this diagram
- For more explanation see "Introducing the LHCONE Map" at <https://www.dfn.de/2024/04/03/Introducing-the-LHCONE-Map/>
- GEANT and CANARIE have shutdown the peering between their VRF and KIAI, as a result of the Ukraine war.

France
RENATER

Spain
RedIRIS

Italy
GARR

Germany
DFN

UK
JANET

USA
ESnet, MAN LAN, STARLIGHT, NETHERLIGHT, PACWAVE/PNWG, GARR, etc.

Other Regions
Korea (KREONet2), Japan (SINET), India (NKN), etc.

Conexiones LHCOPN/LHCONE en España



PIC Tier-1 + IFAE Tier-2: 100 Gbps LHCOPN + 100 Gbps LHCONE (CSUC → RedIRIS)

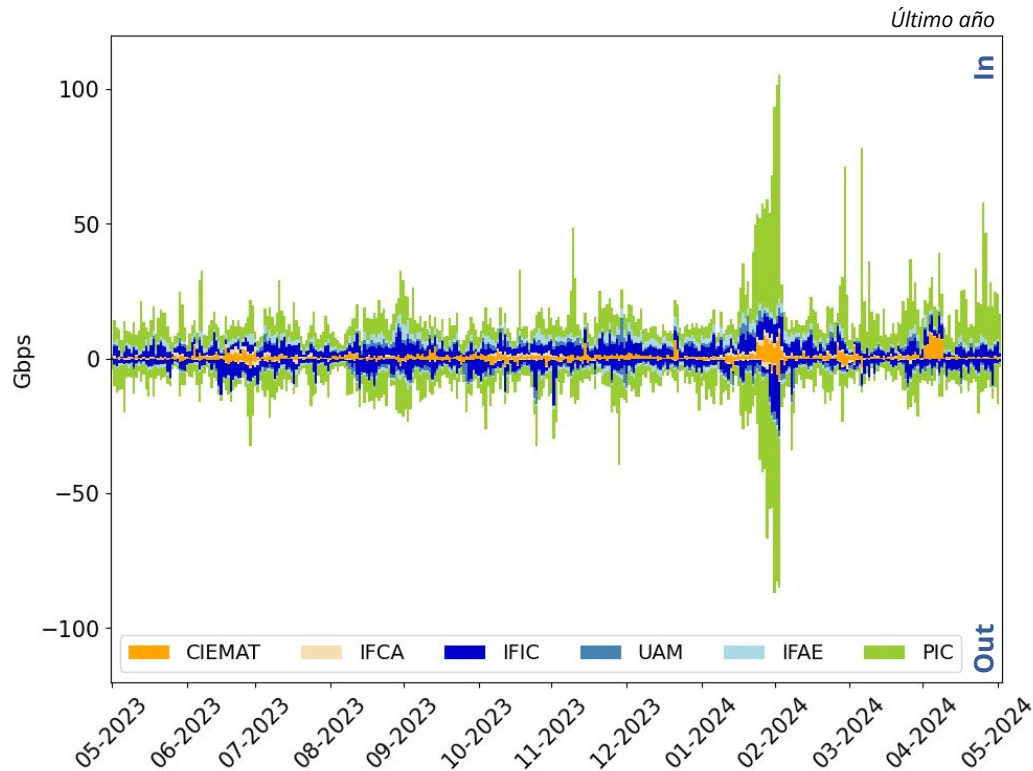
IFIC Tier-2: LHCONE a 25 Gbps → planes para actualizar y conectarse al punto de RedIRIS de la UV a 100 Gbps

UAM Tier-2: LHCONE a 25 Gbps → planes para actualizar y conectarse al punto de RedIRIS a 50 Gbps

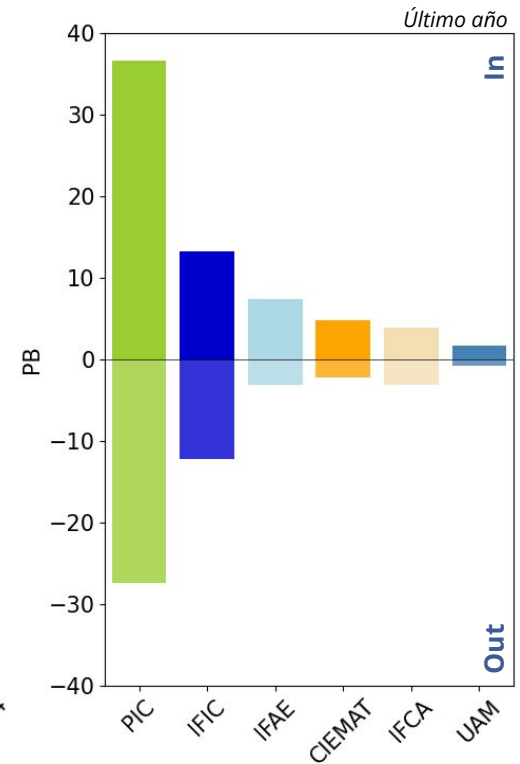
CIEMAT Tier-2: LHCONE a 10 Gbps → planes para actualizar y conectarse al punto de RedIRIS del CIEMAT a 100 Gbps

IFCA Tier-2: LHCONE 100+100 Gbps directo a RedIRIS (infraestructura preparada para 200+200 Gbps)

Tráfico de red LHC en España

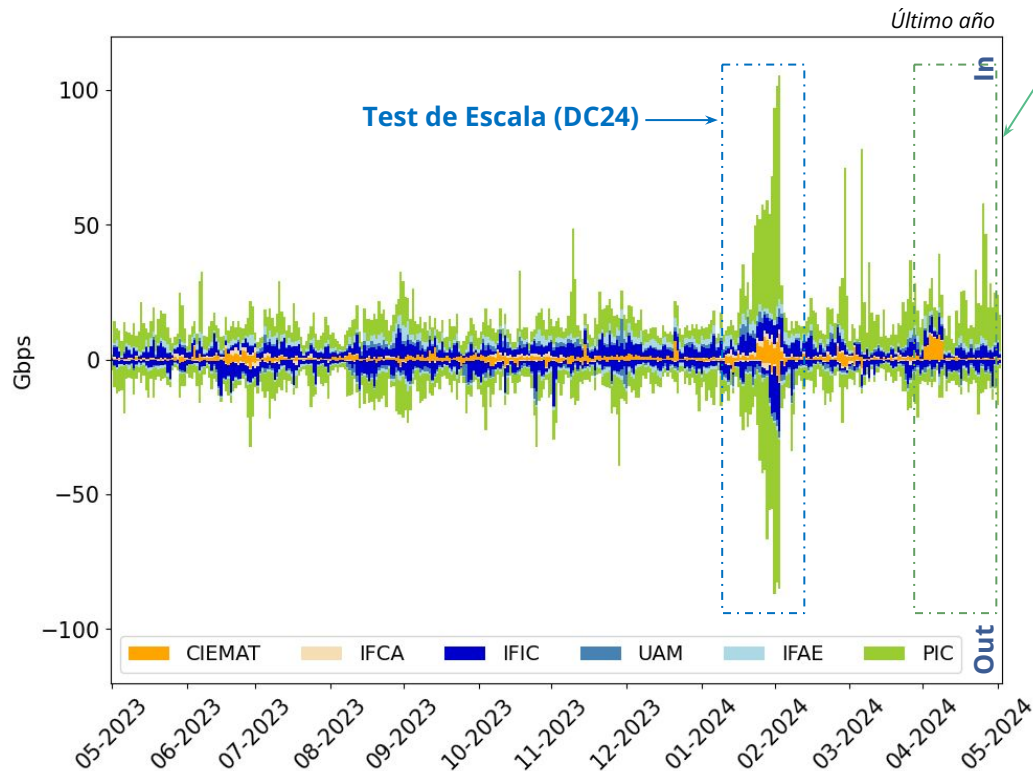


93M ficheros transferidos
Avg in+out: ~30 Gbps

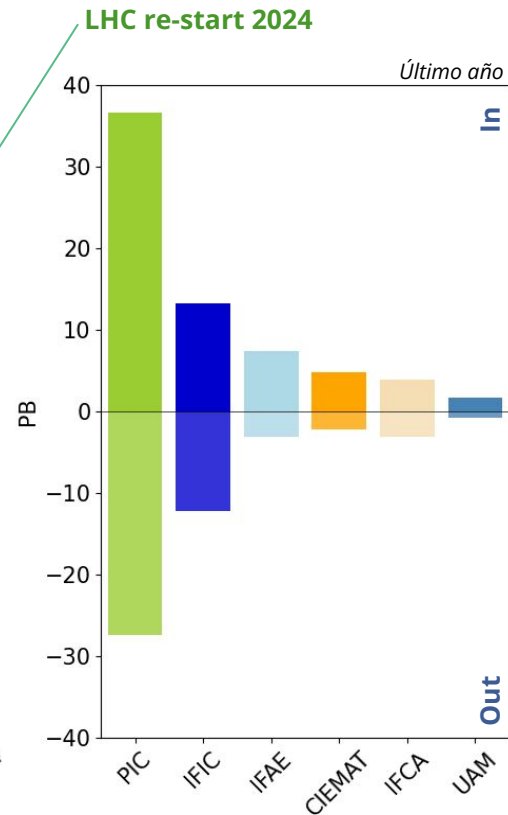


120 PB transferidos (España)
2.5 Exabytes transferidos (WLCG)

Tráfico de red LHC en España



93M ficheros transferidos
Avg in+out: ~30 Gbps



120 PB transferidos (España)
2.5 Exabytes transferidos (WLCG)

Tests de escala hacia HL-LHC

El LHC de Alta Luminosidad (**HL-LHC**) es una versión mejorada del LHC, que operará a una **luminosidad más alta** o, en otras palabras, producirá más datos científicos

El HL-LHC entrará en servicio en **2029**, aumentando el volumen de datos generados por los experimentos del LHC en un **factor 10**

Requerimientos de red hacia HL-LHC:

- **Minimal Model:** $\text{Sum(ATLAS,ALICE,CMS,LHCb)} * 2(\text{for bursts})$ [1]
4.8 Tbps HL-LHC bandwidth
- **Flexible Model:** Supone la lectura de datos mencionados anteriormente para reprocesamiento/reconstrucción en 3 meses
9.6 Tbps HL-LHC bandwidth

[1] Para dimensionar la red en los centros, se debe aplicar otro factor 2 (overprovisioning)

Tests de escala hacia HL-LHC

Se han planificado una serie de **tests de escala (Data Challenges)** para demostrar que estamos preparados a nivel de red, infraestructura y servicios de caras a HL-LHC

2021: 10% de los requerimientos de HL-LHC [DC21] ✓

2024: 25% de los requerimientos de HL-LHC [DC24] ✓

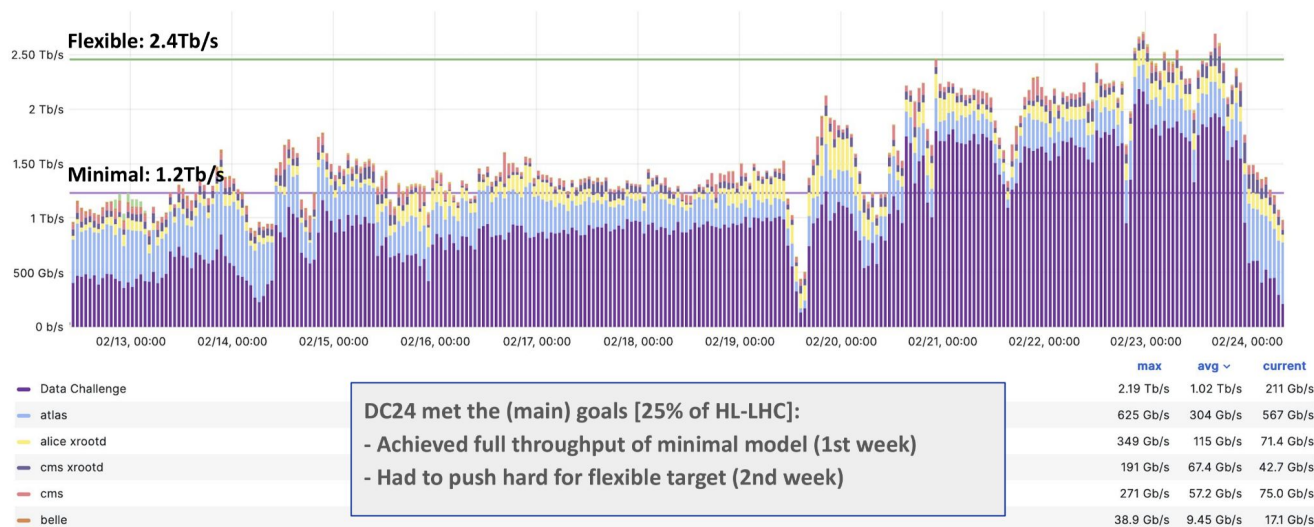
2026: 50% de los requerimientos de HL-LHC

2028: 100% de los requerimientos de HL-LHC

2029: Arranque del HL-LHC (Run4)

DC24: Resultados [Feb. 2024]

WLCG Throughput ⓘ



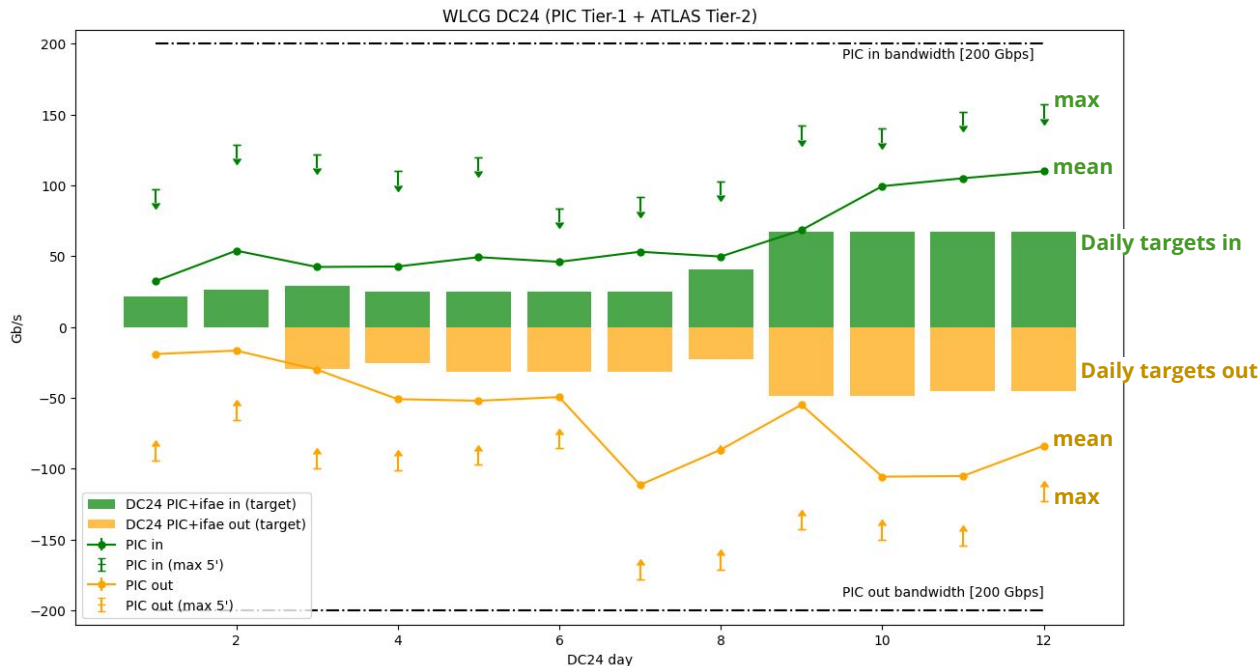
DC24 met the (main) goals [25% of HL-LHC]:

- Achieved full throughput of minimal model (1st week)
- Had to push hard for flexible target (2nd week)

En relación a la **red** se llevaron a cabo una serie de proyectos, por ejemplo:

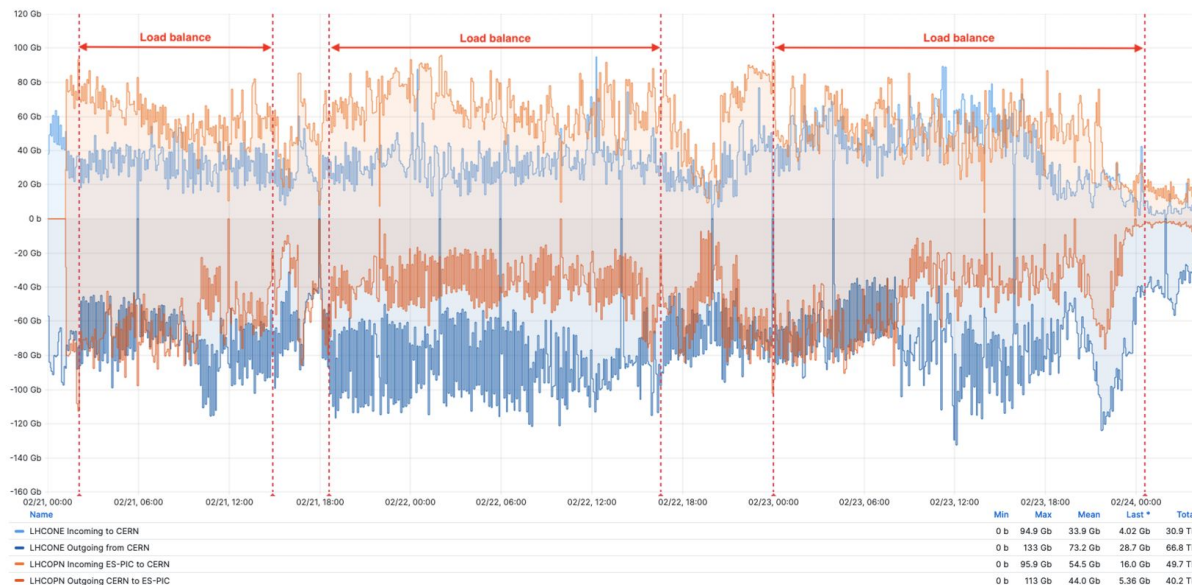
- **Packet marking** (tags en paquetes IPv6); mejoras en la **monitorización global** del uso de red en WLCG; **NOTED** (FTS-driven SDN, balanceo carga LHCOPN ↔ LHCONE)

DC24: Resultados PIC



A lo largo de los días de prueba DC24, PIC mantuvo un tráfico promedio entrante (saliente) de **63 (64) Gbps**, con valores máximos alcanzando **157 (178) Gbps**, respectivamente (medidos en intervalos de cinco minutos)

DC24: Resultados PIC



- ES-PIC load balancing between LHCOPN and LHCONE (from 21st to 23rd of February 2024)

NOTED se activó con éxito en el PIC, a modo de demostración en la implementación de **balanceo de carga entre LHCOFN y LHCONE** (se activa cuando hay alta carga en la LHCOFN para evitar saturar los 100 Gbps disponibles)

Red en WLCG: mirando a futuro

En los próximos 10 años, la Red de WLCG enfrentará **dos desafíos principales**:

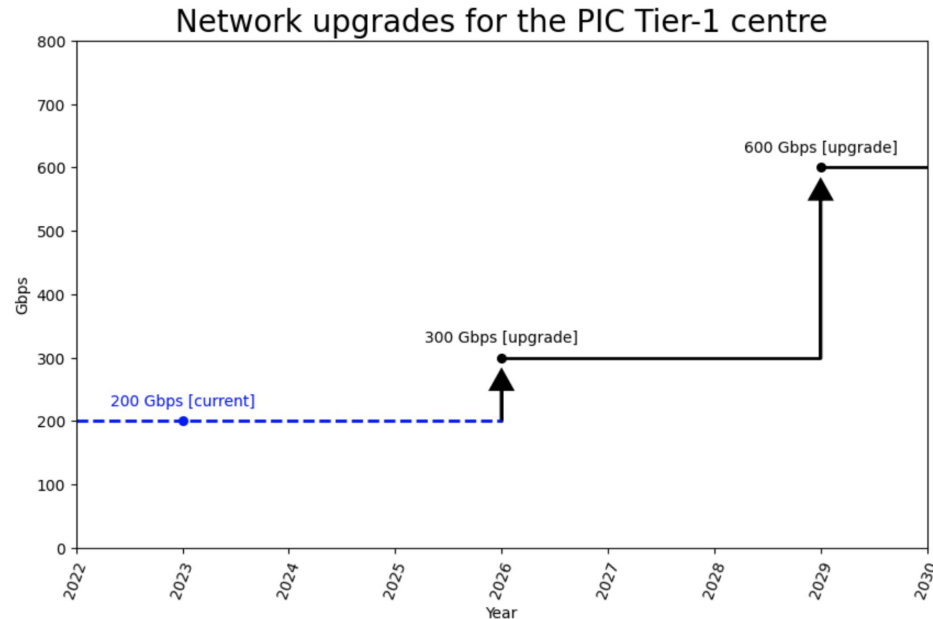
- Manejar los **volúmenes** de datos y la **complejidad** del HL-LHC
- La **convivencia con otros experimentos y ciencias** en la misma infraestructura

La comunidad **WLCG** y **RedIRIS** desempeñan un **papel principal en España**:

- **Modernizar e incrementar los servicios de red**, que permitan avanzar con las actividades de I+D en curso
- **Cooperar con otros experimentos y ciencias** para impulsar la evolución de las redes de investigación y educación compartidas

Red: mirando a futuro (caso PIC)

El **PIC** actualmente da soporte a unos **~60 proyectos científicos** y es una **ICTS** (Red Española de Supercomputación - nodo de datos), siendo el LHC el *top-user*. Se han evaluado las **necesidades de capacidad de red**, a tomar como base en la planificación interna de red del PIC y de las infraestructuras de RedIRIS y CSUC.



Conclusiones

El LHC se enfrenta a un **gran desafío hacia HL-LHC (2029)**

El **volumen de datos** a transferir en WLCG se **incrementará de forma significativa**

Las **redes** que interconectan los centros de computación deberán ser **actualizadas**

Por primera vez, tendremos WAN **~Tbits en centros grandes de WLCG**

En España, nuestras estimaciones son una conexión en el **PIC Tier-1 de ~600 Gbps**, y conexiones **multi-100 Gbps para los Tier-2 españoles**, hacia **2029**

RedIRIS ha desempeñado un papel crucial al proporcionar redes redundantes de alta capacidad en España, lo que ha contribuido al excelente desempeño de nuestros centros en WLCG desde que comenzamos nuestra andadura en 2003

¡Agradecemos enormemente todo el apoyo brindado a nuestra comunidad a lo largo de estos años!

Mos deim coses!

Acknowledgements: This project is partially financed by the Spanish Ministry of Science, Innovation and Universities through grants PID2022-142604OB-C21/C22, PID2022-136323OB-C21/C22 and PID2022-139519OB-C21

Contact: jflix@pic.es

Thanks for feedback provided by my colleagues: A. Pacheco, C. Acosta, A. Pérez-Calero, G. Merino, M. Delfino, V. Acín, S. González de la Hoz, J. Salt, J. M. Hernández, A. Delgado, J. del Peso, I. Cabrillo, F. Matorras, J. Saborido, A. Fernández, X. Vilasís, F. Martínez