



Welcome to the OpenInfrastructure Meetup at CERN!

Enrica Porcari

Department Head, CERN IT



cephalocon

4-5 DECEMBER 2024

CERN
GENEVA, SWITZERLAND

#Cephalocon

SUBMIT A PROPOSAL

SPONSOR

REGISTER NOW

VENUE INFORMATION

CERN SCIENCE GATEWAY

Espl. des Particules 1
1217 Meyrin, Switzerland



Running 4-5 December, Cephalocon brings together a global community of operators, developers, and researchers to celebrate Ceph, the open source distributed storage system designed to deliver excellent performance, reliability, and scalability.

<https://events.linuxfoundation.org/cephalocon/>

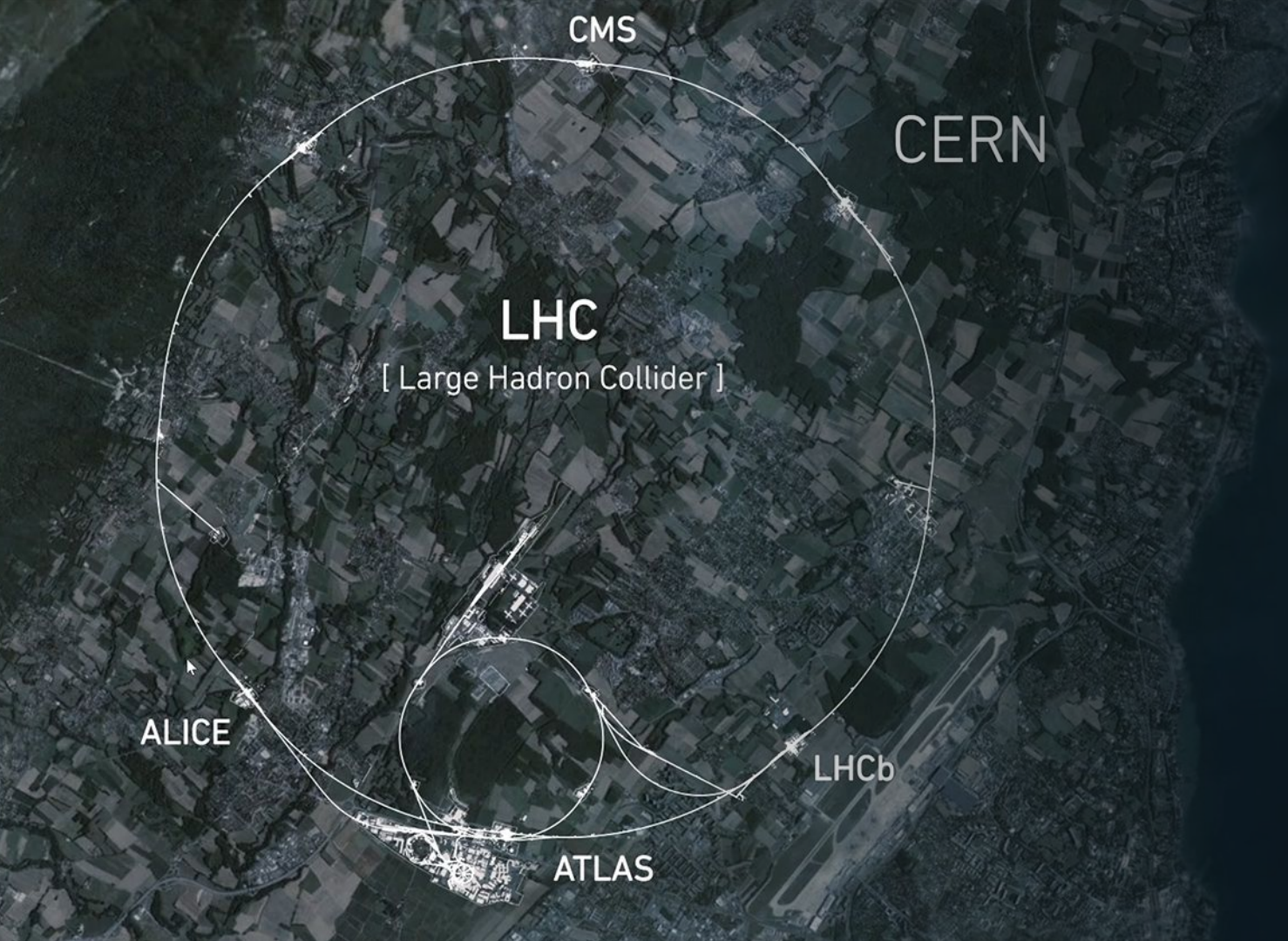


The CERN Private Cloud Journey

Arne Wiebalck

Head of Compute and Devices, CERN IT

OpenInfrastructure Meetup CERN, 6 June 2024



CMS

CERN

LHC

[Large Hadron Collider]

ALICE

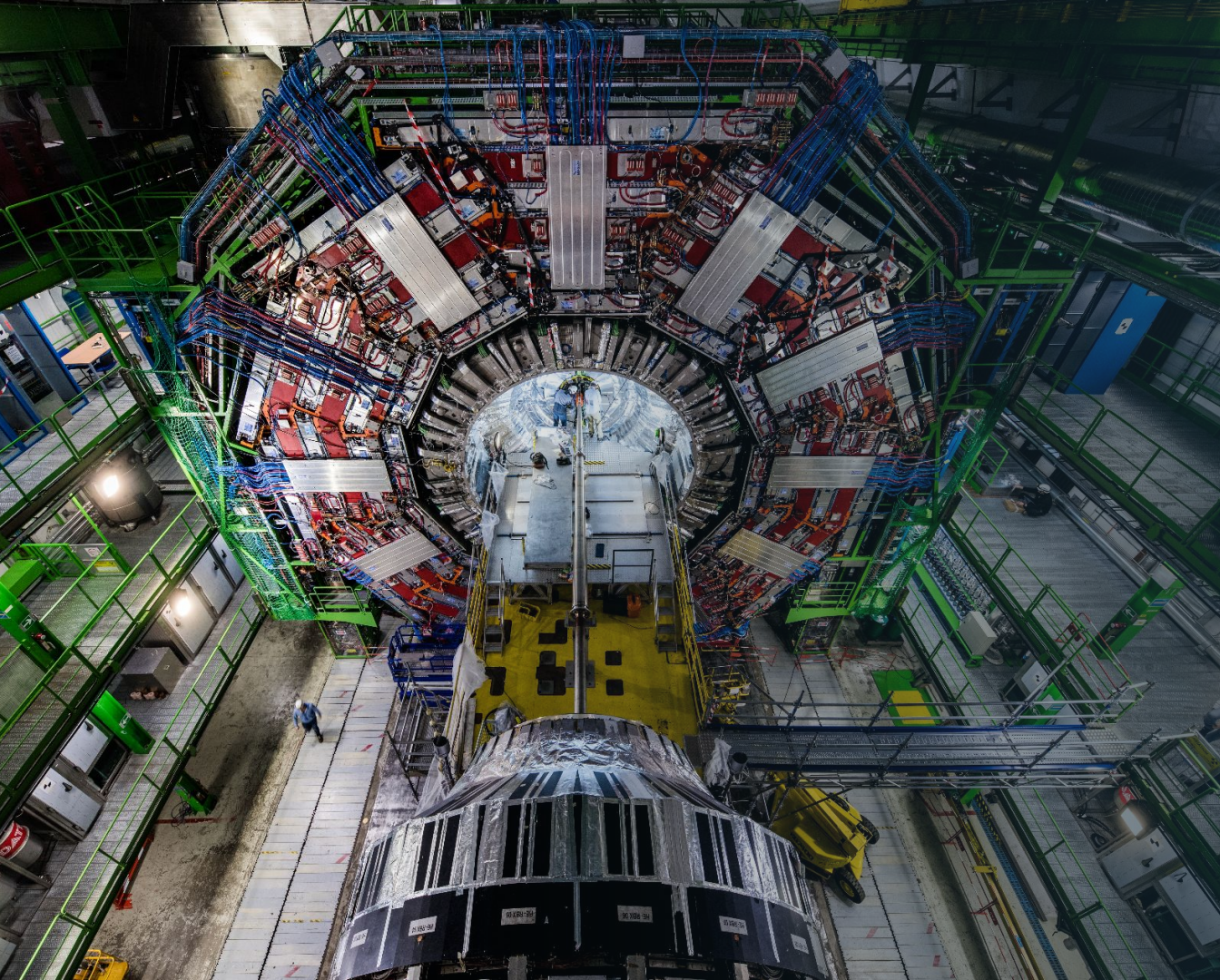
LHCb

ATLAS

**Centre
Européen pour la
Recherche
Nuclear**

Founded 1954
Science for Peace
23 Member States
Fundamental Research
Largest HEP Lab world-wide
LHC flag-ship tool

- 27km circumference
- 100m underground
- 11k turns/sec
- 1B collisions/sec



4 LHC Detectors

- Cathedral-sized, 10'000 tons
- 'Pictures' at 40MHz / 25ns
- 100M channels, ~1PB/sec
- 5'000+ collaborators

5 additional LHC experiments

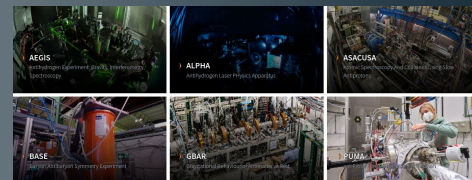
Fixed target experiments

Anti-matter experiments

Experimental Facilities

Non-accelerator experiments

...

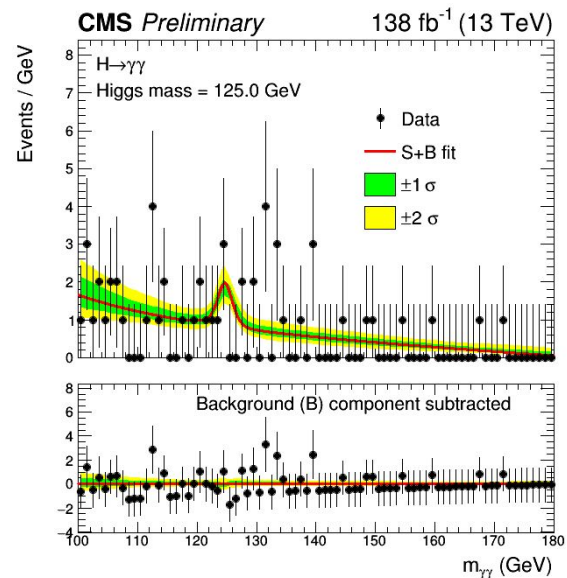
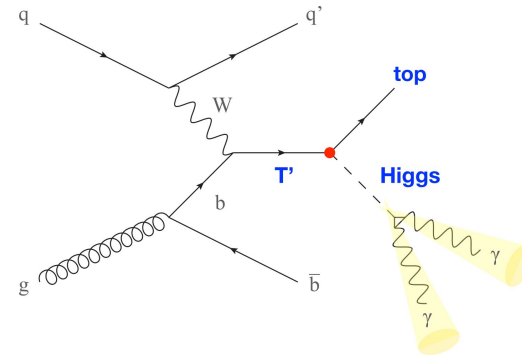
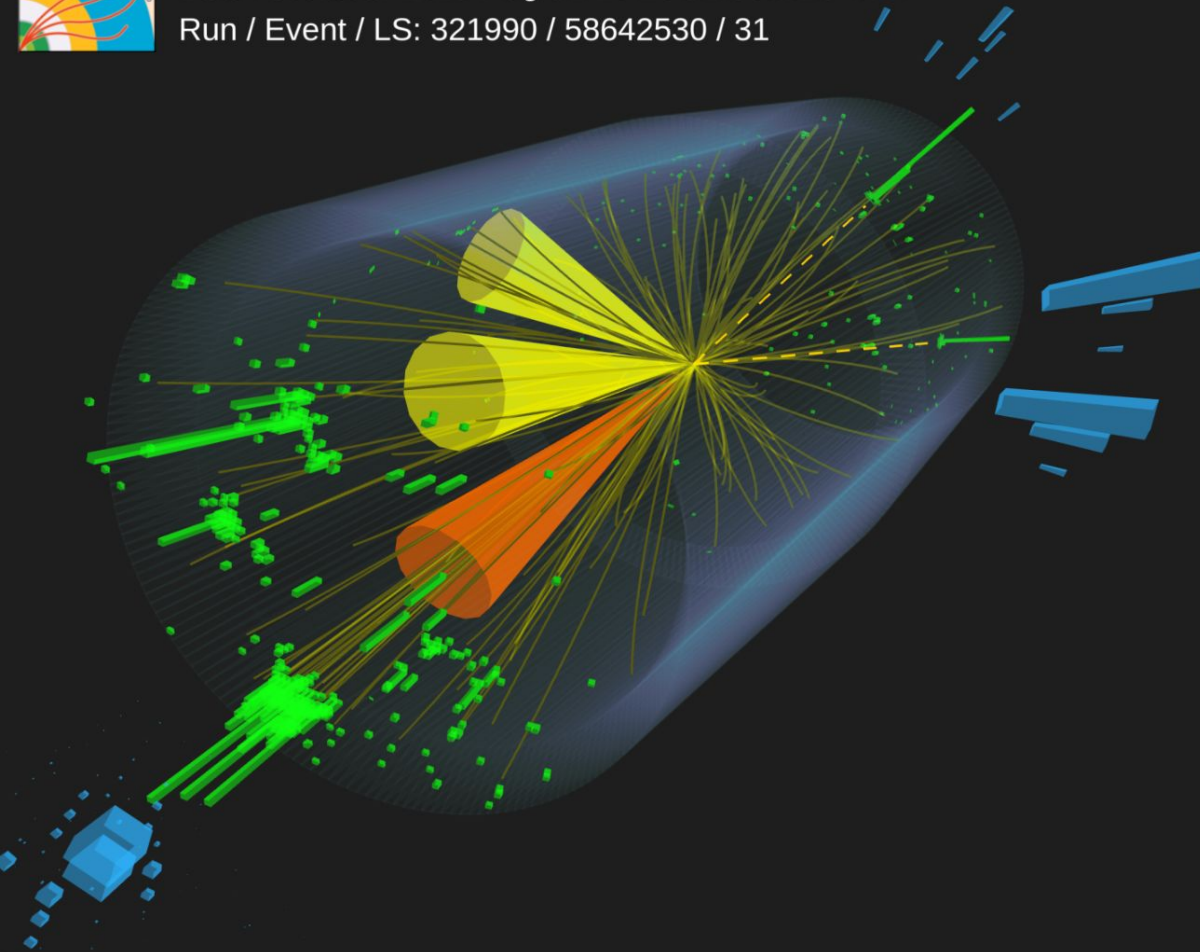


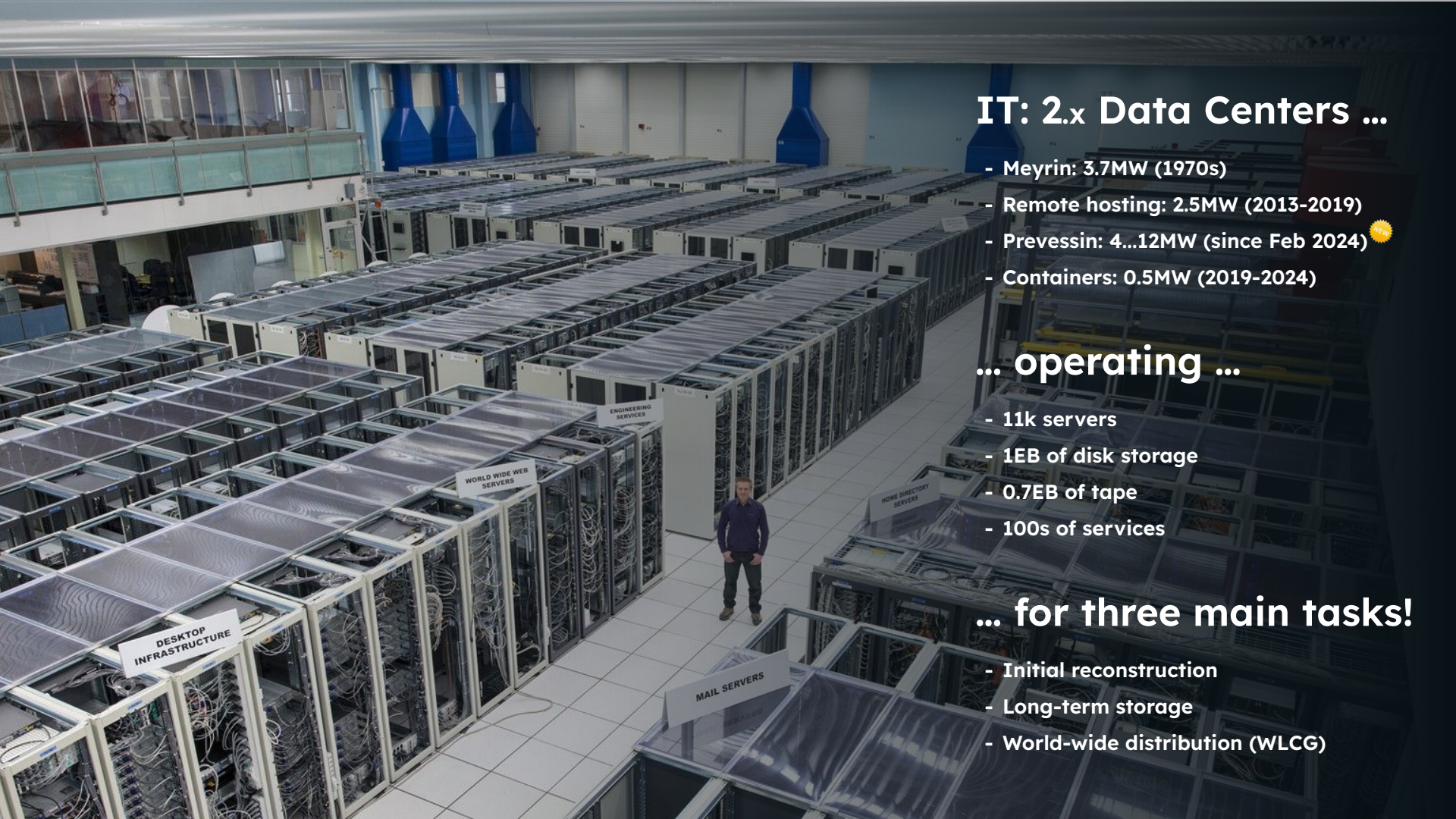


CMS Experiment at the LHC, CERN


Data recorded: 2018-Aug-31 13:28:32.983615 GMT

Run / Event / LS: 321990 / 58642530 / 31





IT: 2.x Data Centers ...

- Meyrin: 3.7MW (1970s)
- Remote hosting: 2.5MW (2013-2019)
- Preessin: 4...12MW (since Feb 2024) 
- Containers: 0.5MW (2019-2024)

... operating ...

- 11k servers
- 1EB of disk storage
- 0.7EB of tape
- 100s of services

... for three main tasks!

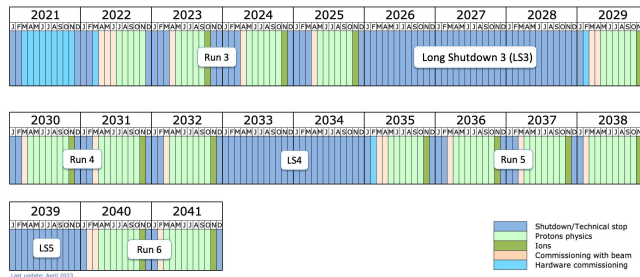
- Initial reconstruction
- Long-term storage
- World-wide distribution (WLCG)

Why did CERN set up a private cloud in 2012?



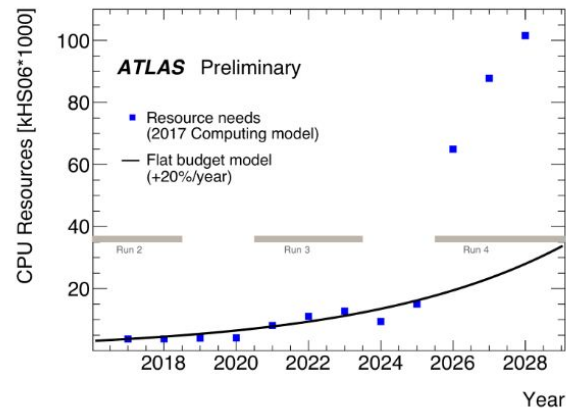
→ CERN's working cycle is determined by the LHC

- Multi-year RUNs (currently in "RUN3")
- Multi-year Long Shutdowns ("LS3" in 2026)



→ Situation in 2012

- LS1 about to start ← **opportunity for change!**
- EU projects finished ← **person power & dev effort dropped!**
- LHC Computing CPU & Storage needs increasing massively
- Other deployments have surpassed CERN's scale
- "Agile Infrastructure" Project: Config' Mgmt, Monitoring, **laaS**



Why did CERN build its private cloud on *FOSS*?



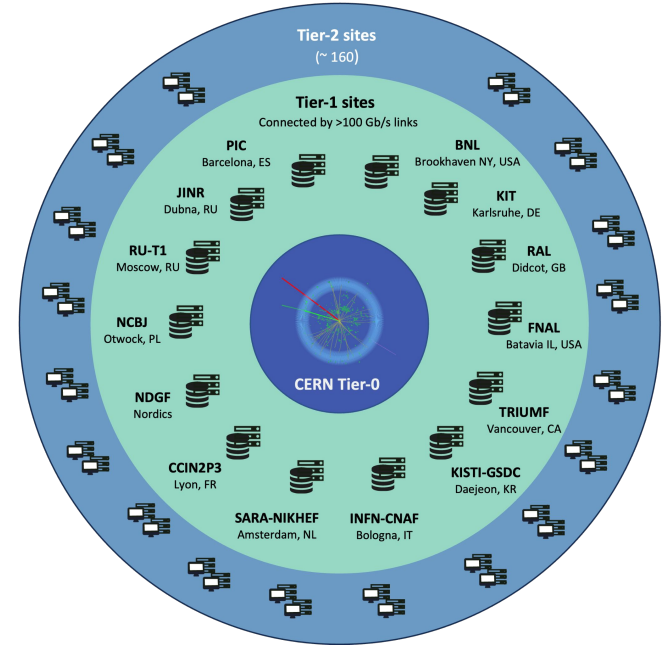
→ CERN is the Tier-0 in WLCG

- Buying for compute on special conditions ⇐ **cost!**
- Limit server configurations ⇐ **efficiency!**
- FOSS: Adaptation feasible ⇐ **flexibility!**

→ Aligns with CERN's goals

- “Giving back” ⇐ **societal impact!**
- Talent management ⇐ **member states!**

→ (Helps with some data governance questions)



What were the goals and KPIs?



→ New Data Centre support

- Overcome limits of Meyrin DC & BC/DR
- **Integrated Wigner Data Center!**

→ Sustainable tool support

- Scale to our size & not CERN-only
- **Managed increased capacity w/ work force!**

→ Improve user response time

- Procurement of resources & self-service kiosk
- **Provisioning down from months to minutes!**

→ Precise monitoring & accounting

- Consolidate across batch, grid, virtual, physical
- **Single accounting chain!**

→ Improve resource efficiency

- Tailored & streamlined (burn-in, benchmarks)
- **All resources handled with one tool!**

→ Enable cloud interfaces

- Experiments started to use EC2 already
- **Cloud interfaces available & expanded!**



How did CERN set up a private cloud?



TOOLS

→ Resource Provisioning (IaaS)

- Based on OpenStack



→ Configuration Management

- Based on Puppet



→ Centralised Monitoring

- Based on Lemon (sensor)
- ELK stack



POLICIES

→ Virtualized services!

- Within reason
- Exceptions are expensive!

Faster provisioning
Increase resource efficiency
Simplify infra management

→ Puppet managed!

Profit from others
Auto-document
Ensure reproducibility

→ Centrally monitored!

Efficient debugging
Integrated alarming
Facilitate accounting



How does the interaction with upstream work?



→ CERN contributed from the start!

- Code contribution, reviews, operator experiences
- Took on roles: Core contributor, Project Leads, TC & Board members
- Presented at summits and meetups
- Hosted OpenInfra Days and Meetups
- Numerous blog posts
- CERN is member of {Linux, OpenInfra Foundation, Cloud Native Computing, AlmaLinux} foundation...



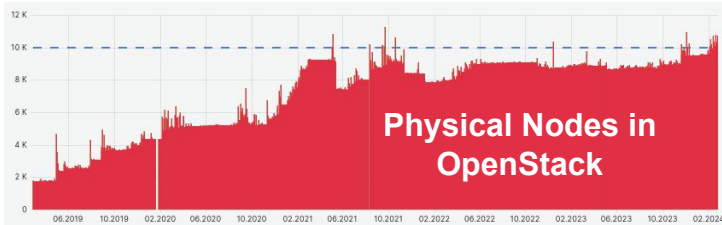
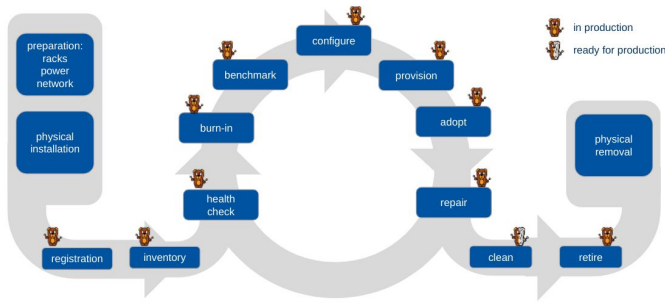
→ Contributing to build up credibility is the only sustainable way!

- We have seen how it works, and how it does not when you do not engage.



What about adapting to new paradigms?

➔ Bare Metal Fleet Management: 10,000 nodes from hire to retire



BACK TO THE FUTURE : Physical Batch

- ➔ Conversion of virtual to physical batch
 - > with the availability of a bare metal API, we revisited the virtualisation tax
- ➔ ~3'800 hypervisors recreated as physical batch instances
 - > done in multiple chunks over several months
- ➔ Terraform as the 'Infrastructure-as-Code' tool to interface with OpenStack/Ironic



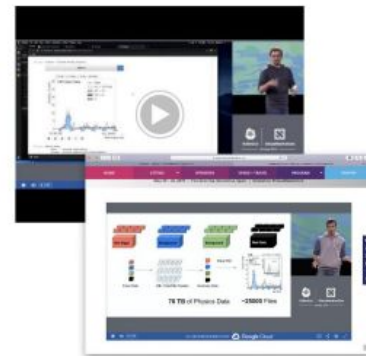
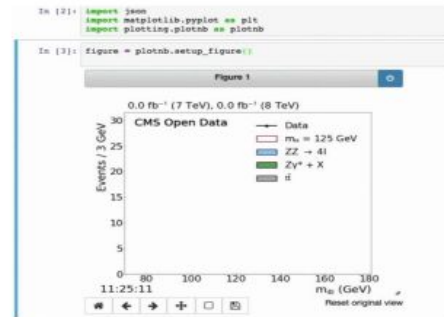
Bonus: 16'000 VMs less than one year ago ... 10k+ IPv4 addresses free'd up.



What about adapting to new paradigms?



→ Containers: Re-running a Higgs analysis



<https://www.youtube.com/watch?v=CTfp2woVEkA>



What about Public Clouds at CERN IT?



→ Public Clouds to complement on-premises offering

- Non-available services or infrastructure, e.g. Quantum Computers
- Technologic preview and faster turn-around, e.g. latest processors
- Bridge-gap solution or peak handling (bursting), e.g. for batch compute
- Business continuity and disaster recovery, e.g. for site failure

→ CERN IT is looking into integrating such offerings.



What about security in an (open) private cloud?



→ **Security is *your* responsibility, more than ever**

- CERN cloud runs ~10k physical nodes & ~15k virtual machines

→ **Cloud team had to run multiple cloud-wide campaigns**

- Spectre and Meltdown
- Reboot the cloud
- In control, but massive impact



<https://www.openstack.org/summit/vancouver-2018/summit-schedule/events/21844/defending-the-cloud-castle-the-openstack-weapons-and-warriors-that-are-keeping-security-threats-at-bay>



So, all shiny then?

➔ Introduction not easy at all!








- Cloud orchestrators are complex frameworks
- Technical integration can be expensive
- User adoption was challenging

➔ Deployment 'surprises'

- Finger trouble
- Extremely intricate technical problems

➔ Learning curve!

Mysteries?

-  **Innocent instances being killed ...?**
- init script sending SIGTERM to wrong PID
-  **Host shut down shortly after boot ...?**
- IPMI bypass upset nova's power state synchronization
-  **Deletions grinding Cinder to a halt ...?**
- "In-place upgrades are evil"
-  **Bare metal database losing entries ...?**
- everyone has mixed up dev and prod once
-  **Upgrades before the upgrade ...?**
- some have even mixed up dev and prod twice ☹️
-  **No monitoring when doing Manila tests...?**
- reduce logging when launching 10k pods!
-  **Volume data loss on reboot ...?**



Arne Wiebalck: Operational War Stories

<https://youtu.be/3HiQmWYp1Sk>

Summary & Current Status



- CERN IT built a private cloud w/ FOSS
- Underpinning services since 10+ years
- Achieved our goals and KPIs
- Steep learning curve and investment
- Community/upstream interaction crucial



OpenInfra DAYS EUROPE SPECIAL EDITION

OpenInfra Day Sweden 📅 May 7, 2024 📍 Volvohallen , Gothenburg, Sweden → Register Now CFP	OpenInfra Day Germany x Sovereign Cloud Stack Summit 📅 May 14 & 15, 2024 📍 Villa Elisabeth , Berlin, Germany → Register Now CFP Sponsor
OpenInfra Day Turkiye 📅 May 20, 2024 📍 Albert Long Hall Cultural Center , Boğaziçi University, Istanbul, Turkiye → Register Now CFP Sponsor	OpenInfra Day France 📅 May 2024 🕒 More info coming soon
OpenInfra Meetup: Switzerland 📅 June 6, 2024 📍 CERN , Geneva, Switzerland → Register Now	



Additional resources



10 years of OpenStack at CERN:

<https://www.openstack.org/videos/summits/virtual/10-years-of-OpenStack-at-CERN-From-0-to-300k-cores>

OpenStack in Production: Mysteries, Challenges, and Achievements:

<https://superuser.openinfra.dev/articles/openstack-production-cern-lightning-talk/>

CERN Tech Blog:

<https://techblog.web.cern.ch/techblog/>

CERN Summit videos:

<https://www.openstack.org/videos/search?search=cern>





www.cern.ch