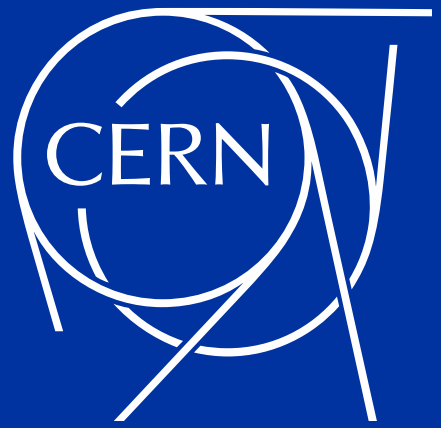


MySQL Community Edition at CERN

Abel Cabezas Alonso

1st May 2024

Abel Cabezas Alonso

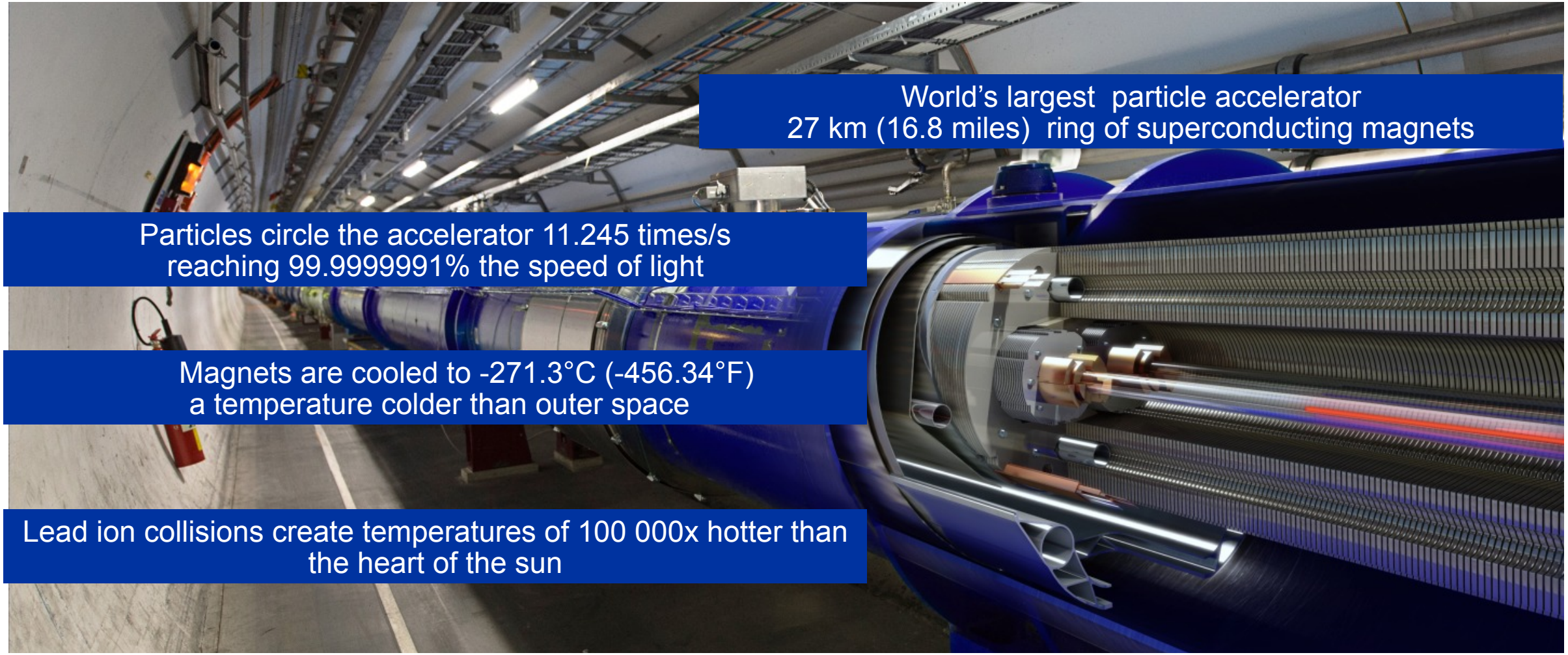
- Database Engineer at CERN since 2019
- Transition as DevOps engineer
- Early career as Software Developer

Abel Cabezas Alonso

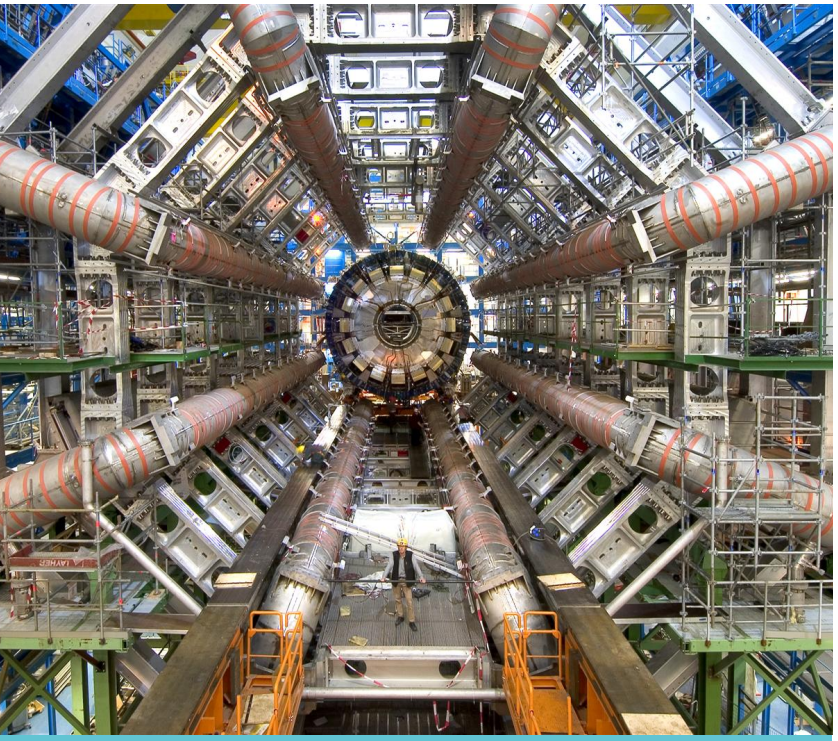
abel.cabezas.alonso@cern.ch

- Established in 1954
- 23 Member states
- Our mission:
  - Unveil how the universe works and what it is made of
  - Provide a unique range of particle accelerator facilities to enable research at the forefront of the human knowledge
  - Unite people from all over the world to push the frontiers of science and technology

The Large Hadron Collider

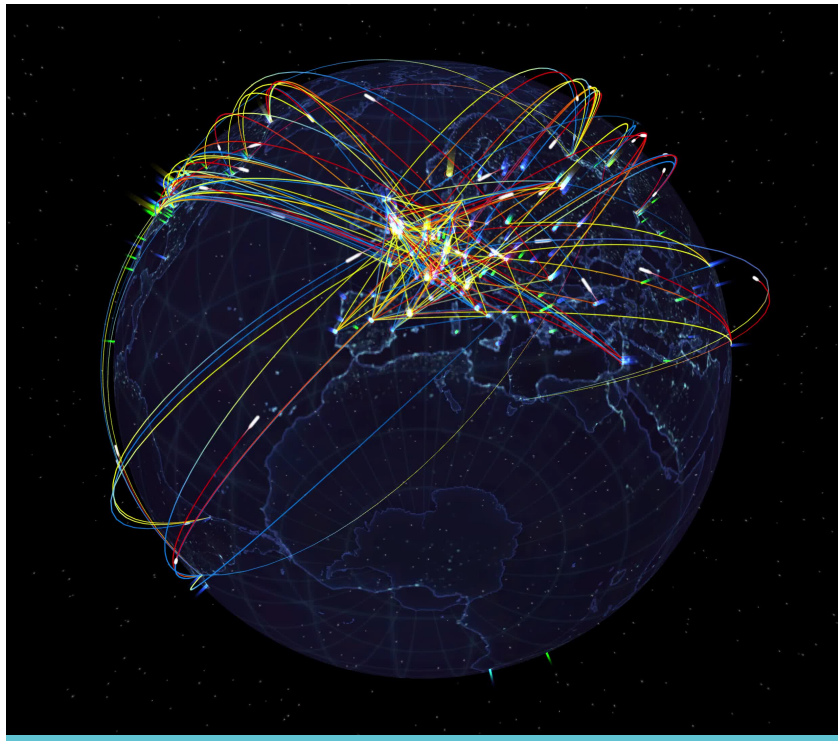


The Worldwide LHC Computing Grid (WLCG)



1 PB of data per second
Only 1% is kept (events with specific characteristics)

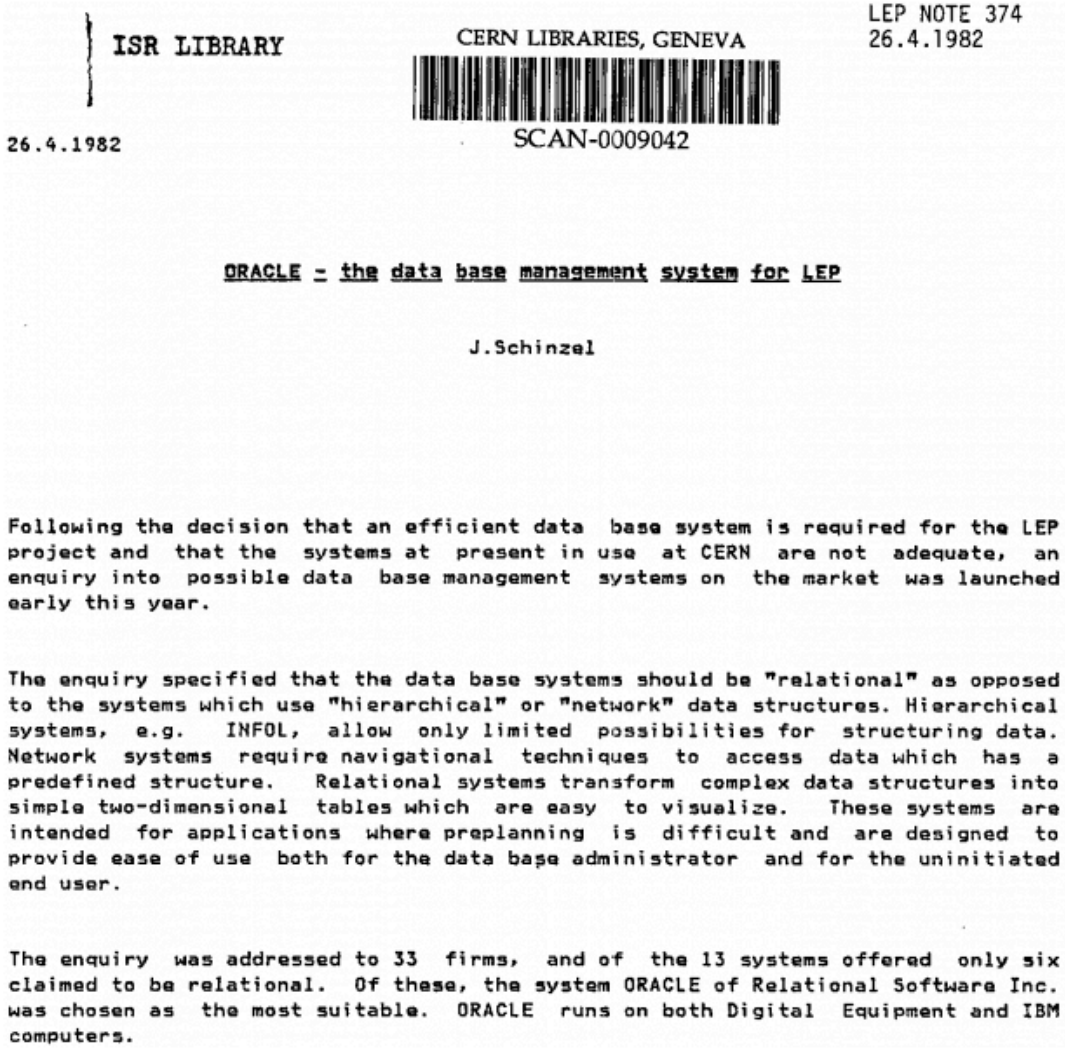
Tier0:
Data reconstruction + Tape archival
+ data distribution to other tiers
~ 200 PB of data per year



WLCG:
- 170 collaborating centers
- 36 countries
- Data analysis

Databases at CERN: Oracle

- **Oracle databases since 1982**
  - 105 Oracle databases
  - More than 11.800 Oracle accounts
  - RAC, Active DataGuard, OEM, RMAN…
  - Complex environment
  - Used by
    - Administrative Information Services
    - Engineering teams
    - Accelerator and experiments
    - etc.
  - Full DBA support
  - ≈ 5PB of data



ISR LIBRARY

CERN LIBRARIES, GENEVA

LEP NOTE 374
26.4.1982

26.4.1982

SCAN-0009042

ORACLE - the data base management system for LEP

J.Schinzel

Following the decision that an efficient data base system is required for the LEP project and that the systems at present in use at CERN are not adequate, an enquiry into possible data base management systems on the market was launched early this year.

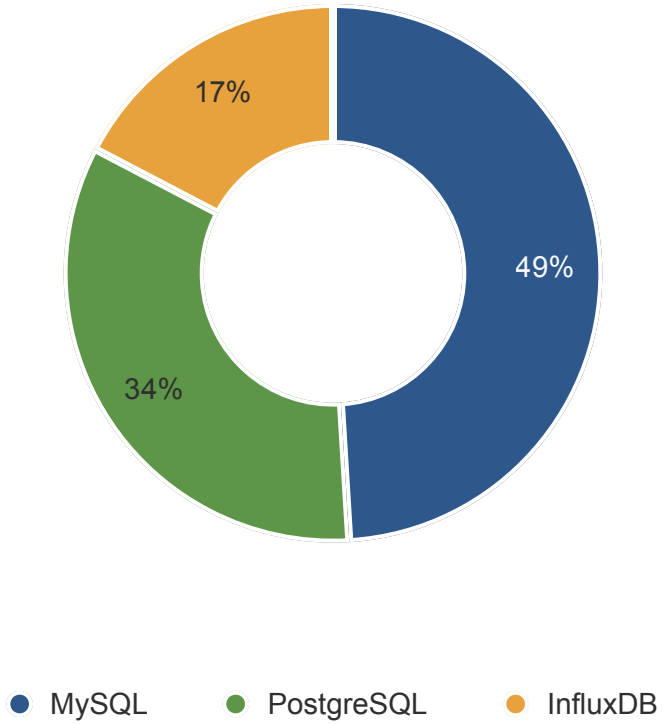
The enquiry specified that the data base systems should be "relational" as opposed to the systems which use "hierarchical" or "network" data structures. Hierarchical systems, e.g. INFOL, allow only limited possibilities for structuring data. Network systems require navigational techniques to access data which has a predefined structure. Relational systems transform complex data structures into simple two-dimensional tables which are easy to visualize. These systems are intended for applications where preplanning is difficult and are designed to provide ease of use both for the data base administrator and for the uninitiated end user.

The enquiry was addressed to 33 firms, and of the 13 systems offered only six claimed to be relational. Of these, the system ORACLE of Relational Software Inc. was chosen as the most suitable. ORACLE runs on both Digital Equipment and IBM computers.

Databases at CERN: DBOD

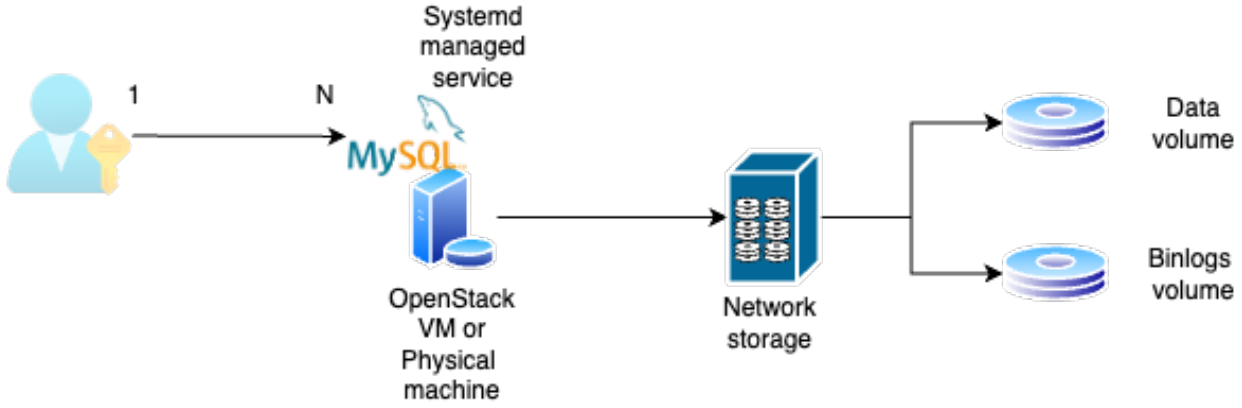
- **Database On Demand (DBOD)**

  - DBaaS conceived in 2011
  - A number of key database applications were running on
  - user-managed MySQL database instances
  - MySQL was the chosen/only supported technology for some applications
  - Empowers users to be their own DBA
  - Flexible architecture allowing to easily integrate other DBMS
  - More than 1200 database server instances
    - ≈600 MySQL, ≈400 PostgreSQL, ≈200 InfluxDB
  - Used by
    - CERN' Authorization and authentication service
    - Experiments (ATLAS, LHCb, etc.)
    - WLCG file transfer service
  - ≈150 TB of data



MySQL deployment

- Several MySQL binaries per host
- Several database instances per host
- Two different NetApp NFS volumes per DB instance:
  - data directory + binary log directory
- Types of deployment
  - Single instance
  - Replication for disaster recovery
  - Replication to scale out reads (app has to tolerate eventual consistency)
  - ProxySQL + primary-replica
  - MySQL InnoDB cluster



Towards high availability: ProxySQL + async replication

Objective

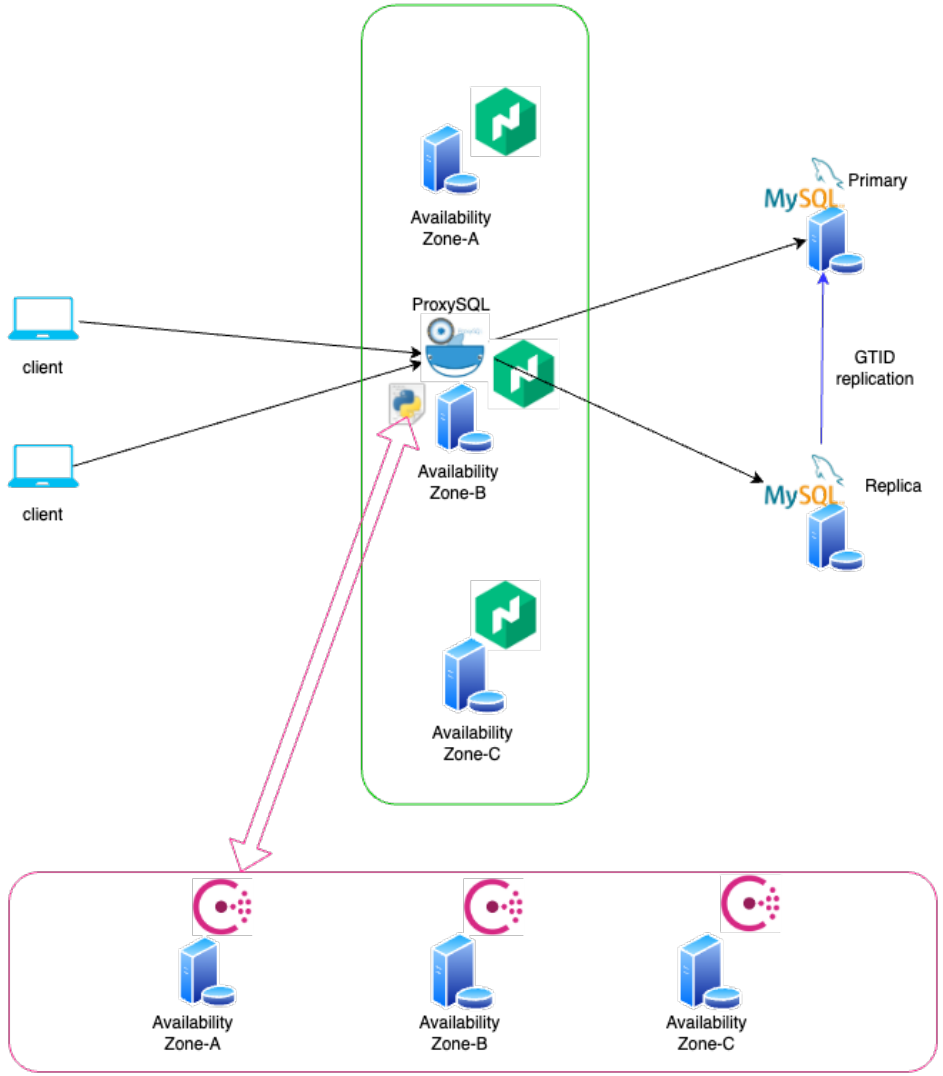
- Automate failover

About ProxySQL

- Built-in monitoring module
- Not designed for reconfiguring the topology
  - No built-in failover/failback
- Scheduler module to extend logic:
  - Failover logic
  - Resolution of conflicts (split brain)
  - Monitoring of replication channel

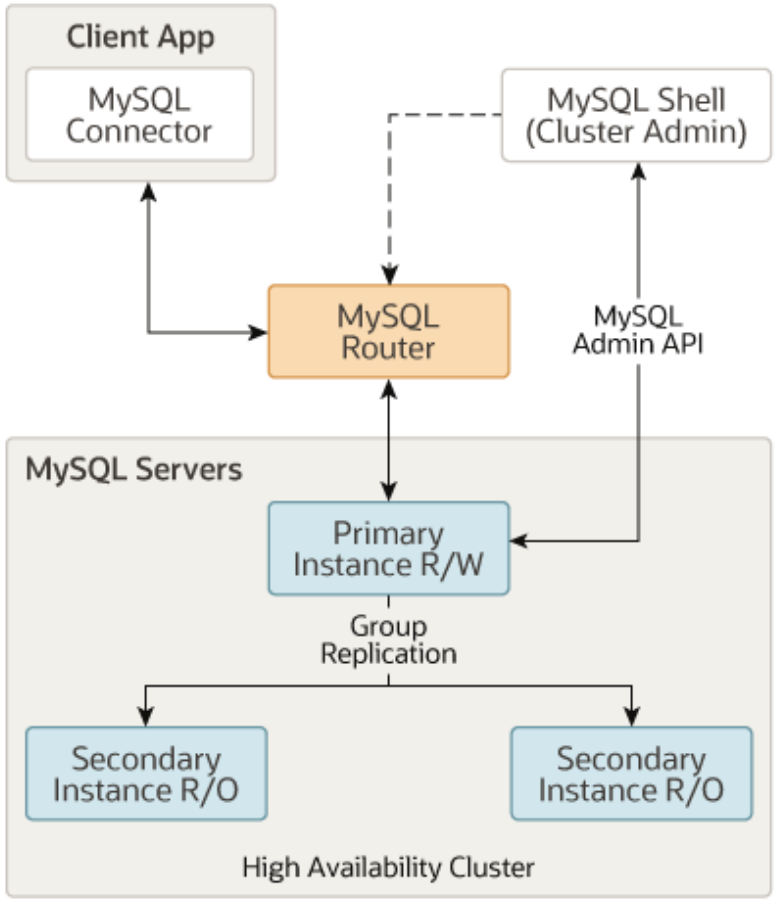
Limitations

- Not a pure HA solution
- SPOF
- Not possible to deploy several proxies for our use case
- Big maintenance effort



High Availability with MySQL InnoDB cluster

- No SPOF by deploying several MySQL Routers
- Fully fledged HA solution
- Easy to integrate*
  - Storage watchdog
- Easy to deploy with MySQL Shell
- Easy to maintain
  - Simplified management with MySQL Shell
  - Very good documentation
    - Quorum loss
    - Cluster reboot
- Seamlessly scale out reads through MySQL Router



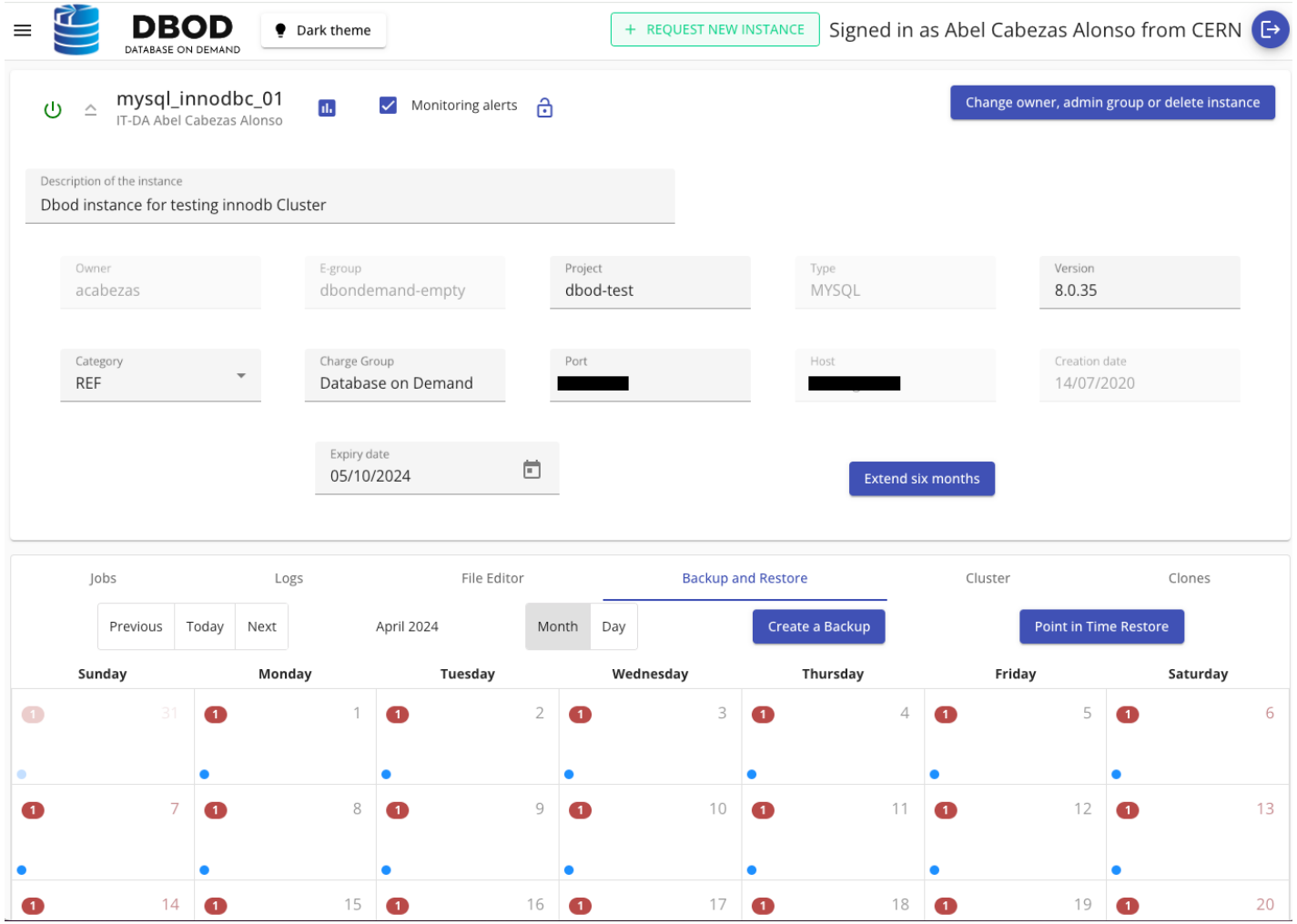
Automation

Web automation

- Automated backup and recovery services
- MySQL Shell upgrade checker
- Management of configuration files
- Cloning
- Integrated monitoring
- Integrated upgrades
  - Primary-replica upgrade logic

Ops automation

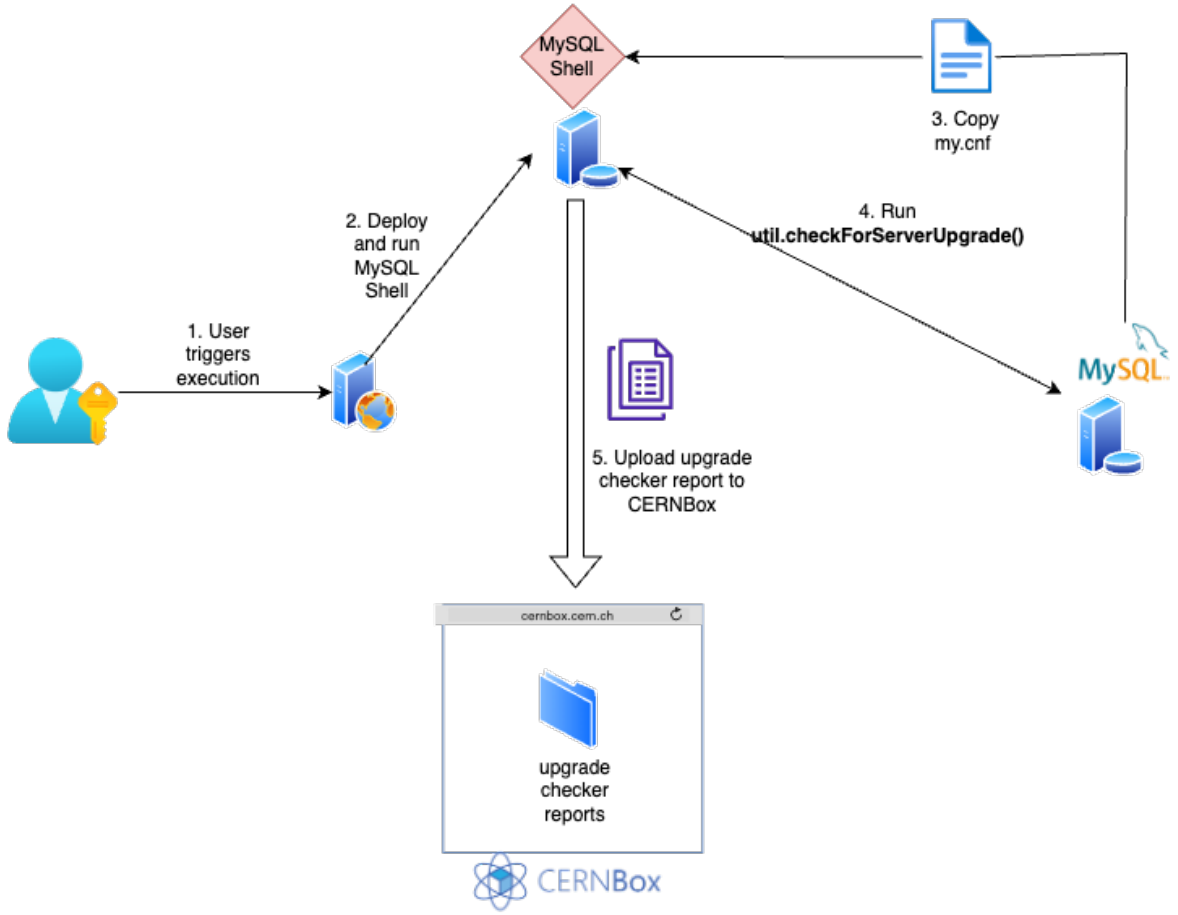
- Continuous validation of backups
- Instance and storage migration
- Automated replica provisioning
- Automated replication switchover
- Detection of idle instances
- Integrated password hash cracker



Automating instance upgrades with MySQL Shell

Upgrade checker utility

- MySQL Shell integration with extended logic
- Can be run on demand
- Upgrades disabled by default
- Only enabled once the upgrade checker report is “clear”
- Report shared via CERNBox (cloud storage)
- Users can correct errors and warnings before upgrading autonomously
- Extended logic for replication setups



Service evolution

- MySQL is a key element of many services at CERN
- Ensure that MySQL evolution fits in the service
  - Integration of MySQL InnoDB cluster to discontinue ProxySQL
  - Promote the usage of utf8mb4
  - Integrate InnoDB replica set to ease the management of existing replication setups
  - Evaluation of MySQL clone plugin for clone provisioning
- Fully automate MySQL InnoDB cluster lifecycle
- Explore possibilities to further automate MySQL deployments at CERN
- Finalise the integration of our Business Continuity/Disaster Recovery plan

home.cern

Automating utf8mb3 character set conversion

utf8mb3 is deprecated

- Instances coming from 5.6.x / 5.7.x
- Run once a day an automated check looking for utf8mb3 usage
- Enable the automated charset conversion on the web interface for the affected instances
- Allow dry-run:
  - Generates only DDL to be applied
- Run conversion
  - Generate DDL dump from before and after + conversion log
- Recommended to first test in a cloned instance to avoid surprises like:

```
ERROR 1074
Column length too big for column 'foo' (max = 16383); use BLOB or TEXT instead
```

  - A VARCHAR column can only accommodate up to 16383 characters for the utf8mb4 character set