



ALICE



Calibration and Conditions Database (CCDB)

costin.grigoras@cern.ch
For the ALICE Collaboration

ALICE CCDB at a glance

- Central store of calibration and condition data of in Run3+
Metadata stored separately from the serialized calibration data
Data distribution using a set of reliable Grid SEs
- Millisecond resolution for object Interval of Validity (IoV)
- X.509 certificate authenticated writes, open reads
- HTTP(s) for restful metadata queries
HTTP(s) and/or XrootD for data access
- Multicast feedback loop in the online reconstruction pass for data compression and calibration
Consumes and produces new calibration objects in real time during experiment data taking

Path format

`/Detector/Category/Param/tStart[/tEnd][/UUID][/key=value]`

Folder structure, 3 levels deep by convention

For most requests a **reference time** is mandatory

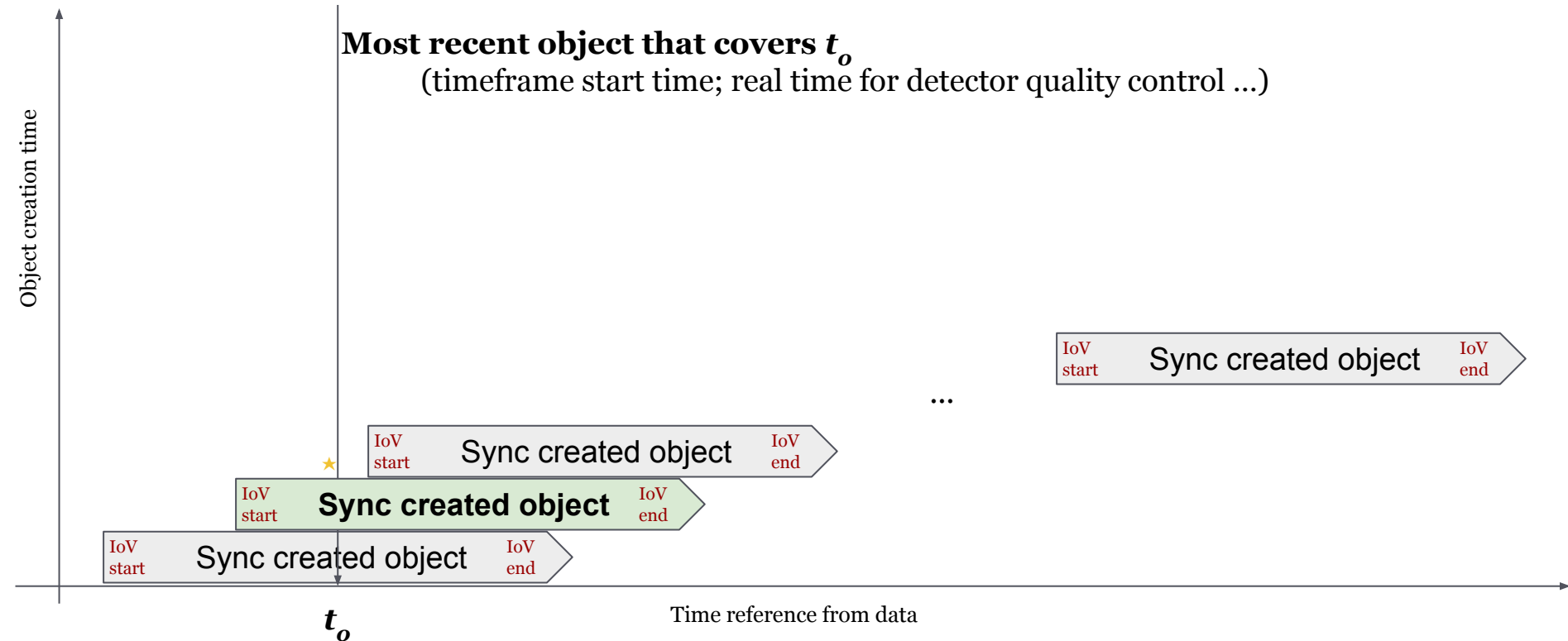
User-defined **metadata** associated to each object, can also filter by it

Additional HTTP headers:

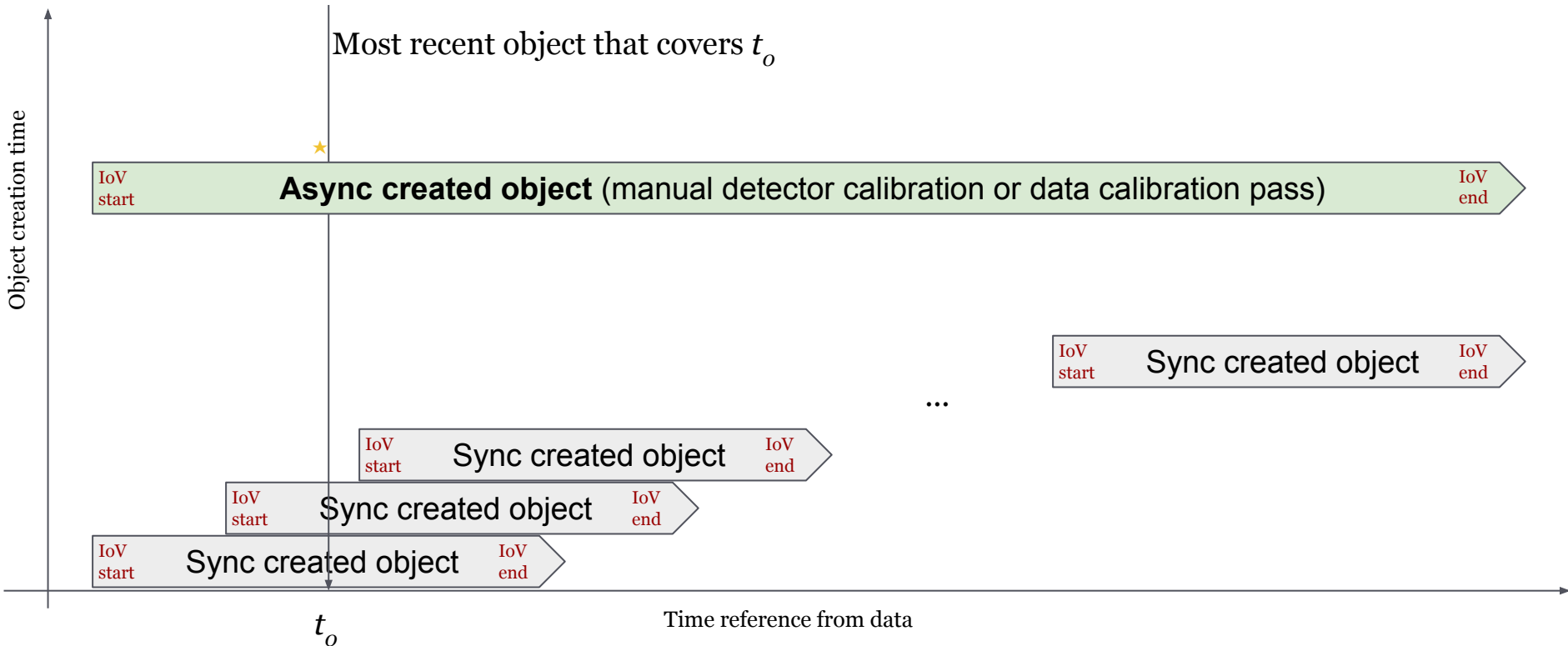
If-None-Match : client cached object(s) to validate

If-Not-After : snapshot / time machine functionality

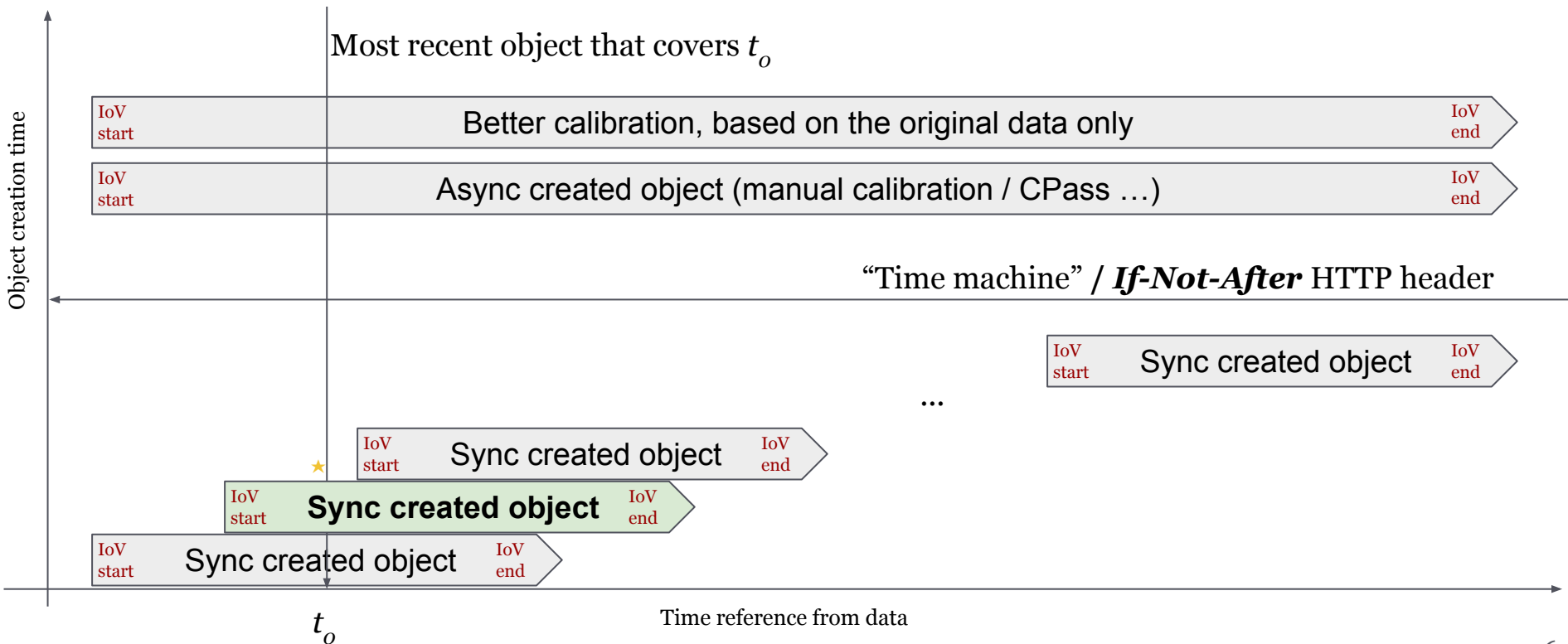
IoV queries during data taking



IoV queries, manual calib.



IoV queries, snapshots



cURL-based REST examples

#upload an object to the repository

```
curl -F blob=@/tmp/file http://alice-ccdb.cern.ch/Detector/Calib/Align/1/100000/quality=2
HTTP/1.1 201
Location: http://alice-ccdb.cern.ch/download/a329fcc6-9818-4d2e-a5af-16ca73686cf2
```

#query to find the object valid at given moment in time

```
curl http://alice-ccdb.cern.ch/Detector/Calib/Align/50000
HTTP/1.1 303
Location: alien:///alice/data/CCDB/.../a329fcc6-9818-4d2e-a5af-16ca73686cf
ETag: "a329fcc6-9818-4d2e-a5af-16ca73686cf2"
Valid-From: 1
Valid-Until: 100000 } IoV endpoints (in epoch milliseconds)
quality: 2
Content-Location: alien:///alice/data/CCDB/.../a329fcc6-9818-4d2e-a5af-16ca73686cf
Content-Location: http://alice-ccdb.cern.ch/download/a329fcc6-9818-4d2e-a5af-16ca73686cf2
Content-Disposition: inline;filename="o2-tpc-IDCZero_1681052400217.root"
ObjectType: o2::tpc::IDCZero
runNumber: 534275 } Metadata from production
```

#with non-matching metadata constraints

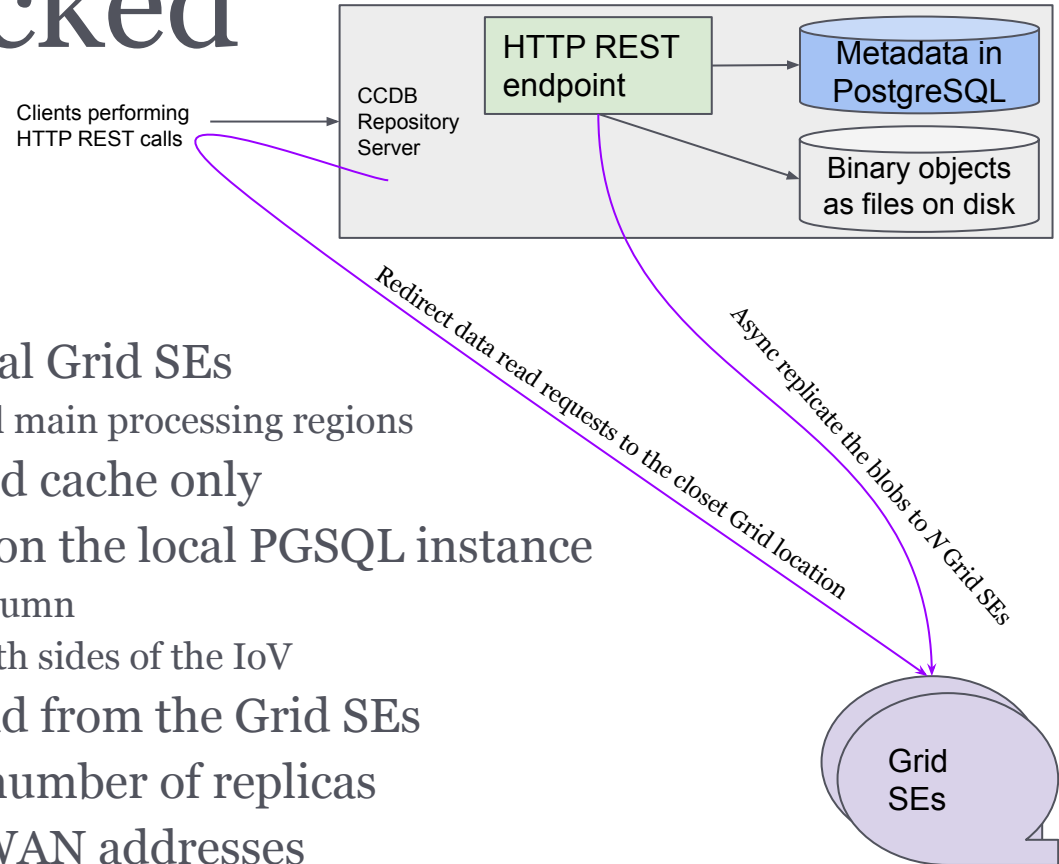
```
curl http://alice-ccdb.cern.ch/Detector/Calib/Align/50000/quality=1
HTTP/1.1 404
```

#check if the object is still valid at a later moment in time, i.e. processing the subsequent data block

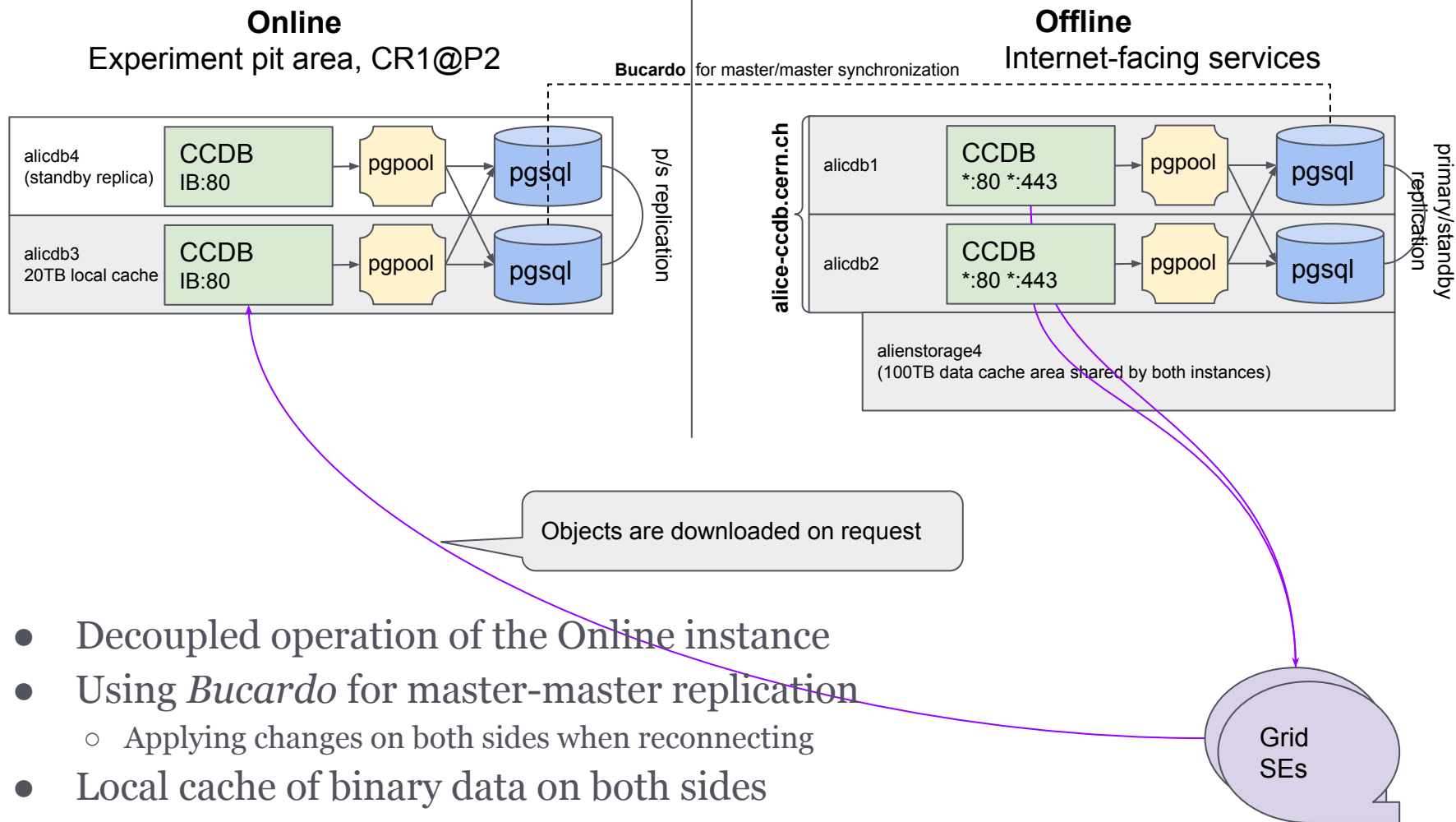
```
curl -H 'If-None-Match: a329fcc6-9818-4d2e-a5af-16ca73686cf2' http://alice-ccdb.cern.ch/Detector/Calib/Align/76543
HTTP/1.1 304
```

Not modified

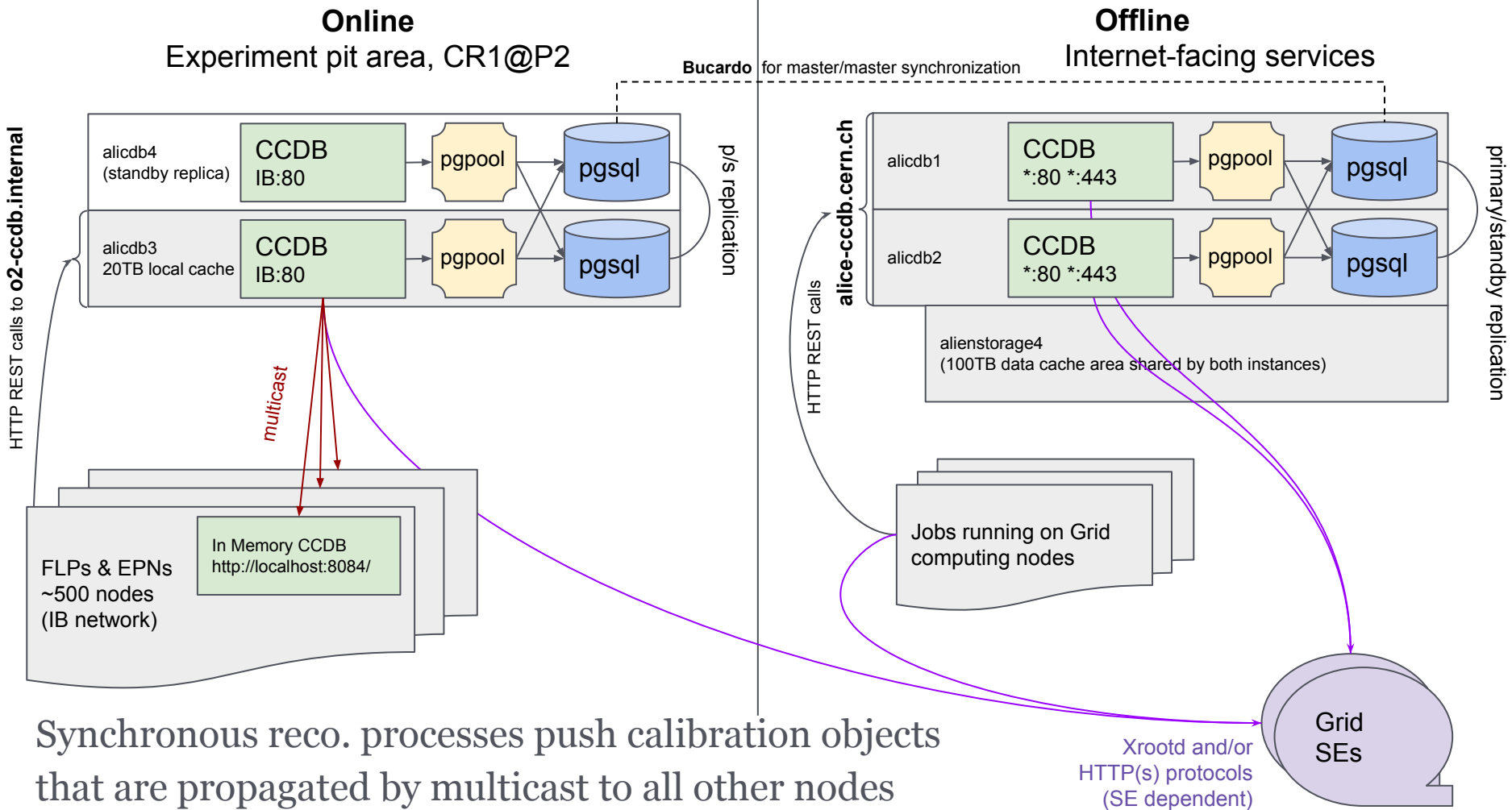
Grid SE-backed



- Blobs are uploaded to several Grid SEs
 - Geographically distributed in all main processing regions
- Local disk used as buffer and cache only
- Metadata queries executed on the local PGSQL instance
 - GiST index on a *tsrange* IoV column
 - Efficient insert and match of both sides of the IoV
- Clients are redirected to read from the Grid SEs
- Bandwidth scales with the number of replicas
- Location-aware sorting of WAN addresses



- Decoupled operation of the Online instance
- Using *Bucardo* for master-master replication
 - Applying changes on both sides when reconnecting
- Local cache of binary data on both sides



Synchronous reco. processes push calibration objects that are propagated by multicast to all other nodes

Some figures

1.2TB of data in 5M calibration objects

Append-only policy

8 Grid SE replicas on HTTP-enabled endpoints

450Hz of requests to Offline instances (1w avg)

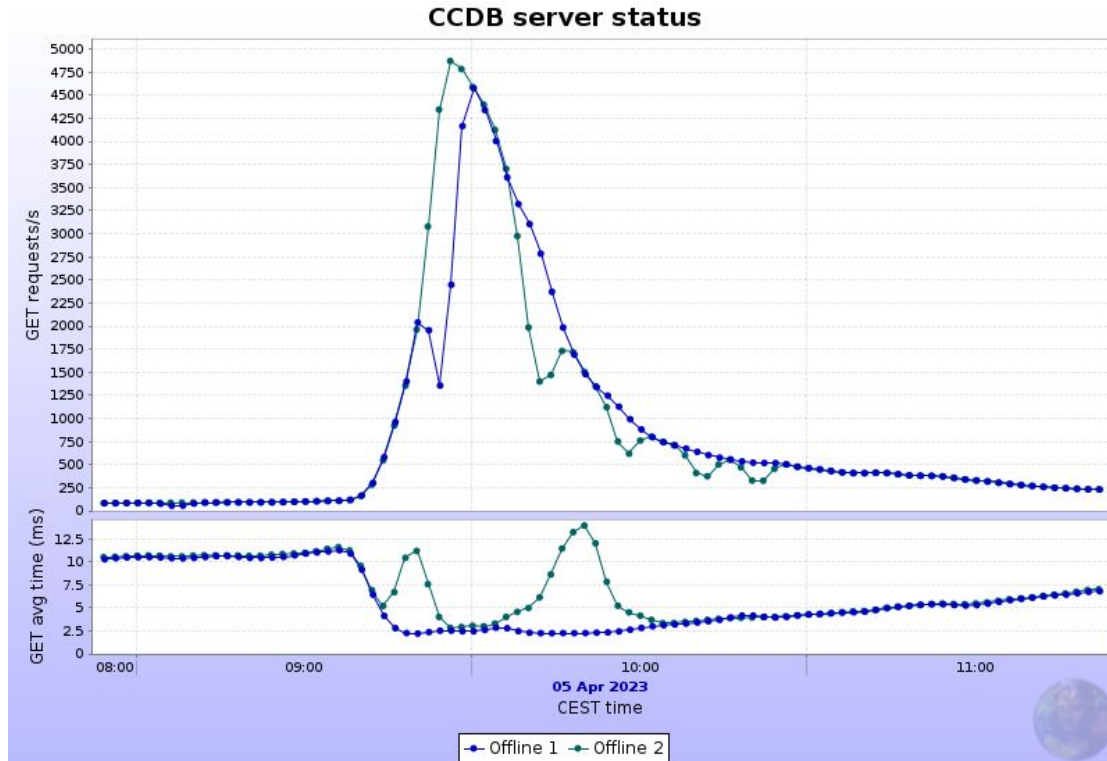
12ms average response time to Grid jobs

2.3Hz of new objects while data taking

Most of them TPC integrated digital current data

83MB in 195 paths used by Online workflows

Scale test of Offline services



20KHz / server in synthetic benchmarks

Real-world Grid test

1K concurrent jobs

10KHz of cache

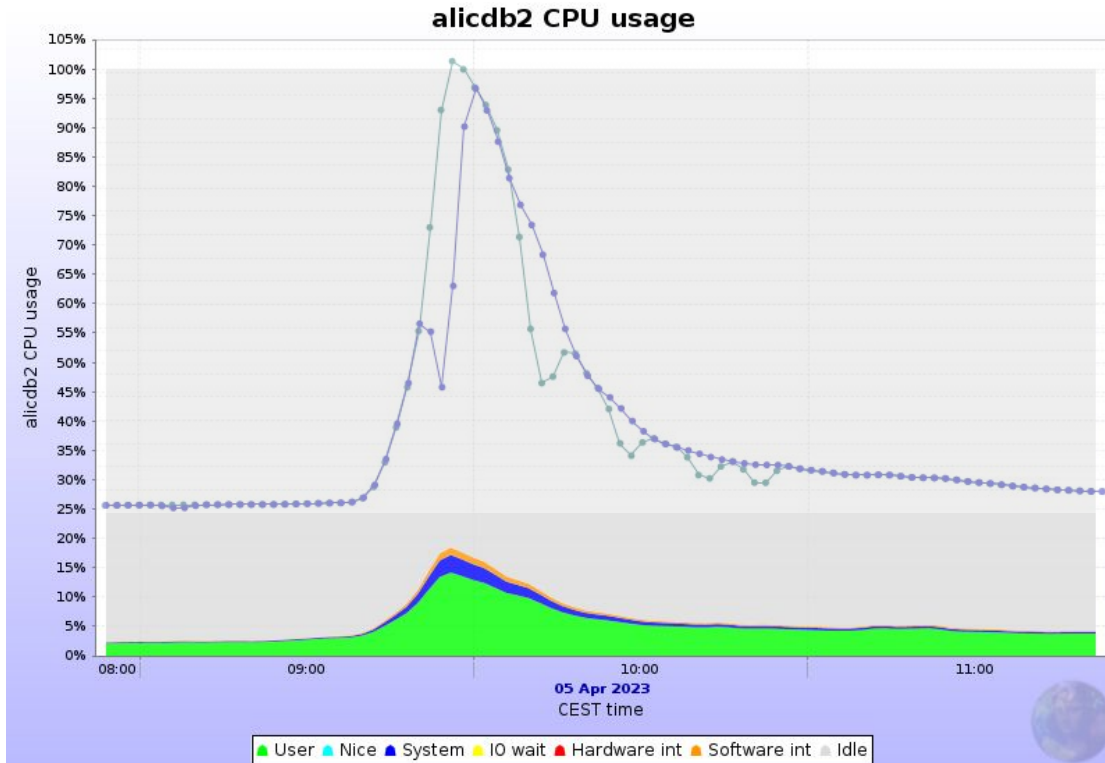
validating requests

Full O2 framework

No HTTP keep-alives yet

request rate is $f(RTT)$

Scale test of Offline services



20KHz / server in synthetic benchmarks

Real-world Grid test

1K concurrent jobs
10KHz of cache
validating requests

20% server CPU usage
during that time

Summary

Java open source project embedding a Tomcat server

REST service for storing calibration/condition/QC data

- ROOT serialization & streaming support
 - TGrid plugin and CCDB helper functions to query and load objects in memory

CCDB serves both real-time and offline data processing

Offloading data management to the Grid middleware

Three server flavors for

- Local machine / development endpoint
- In-memory cache with multicast receiver (real time data compression)
- PostgreSQL, Grid SE-backed & multicast sender