



Empowering CERNBox users with self-service restore functionality

Gianmaria Del Monte (IT-SD, CERN)

27/04/2023 - EOS Workshop 2023

cback backup system

Introduction

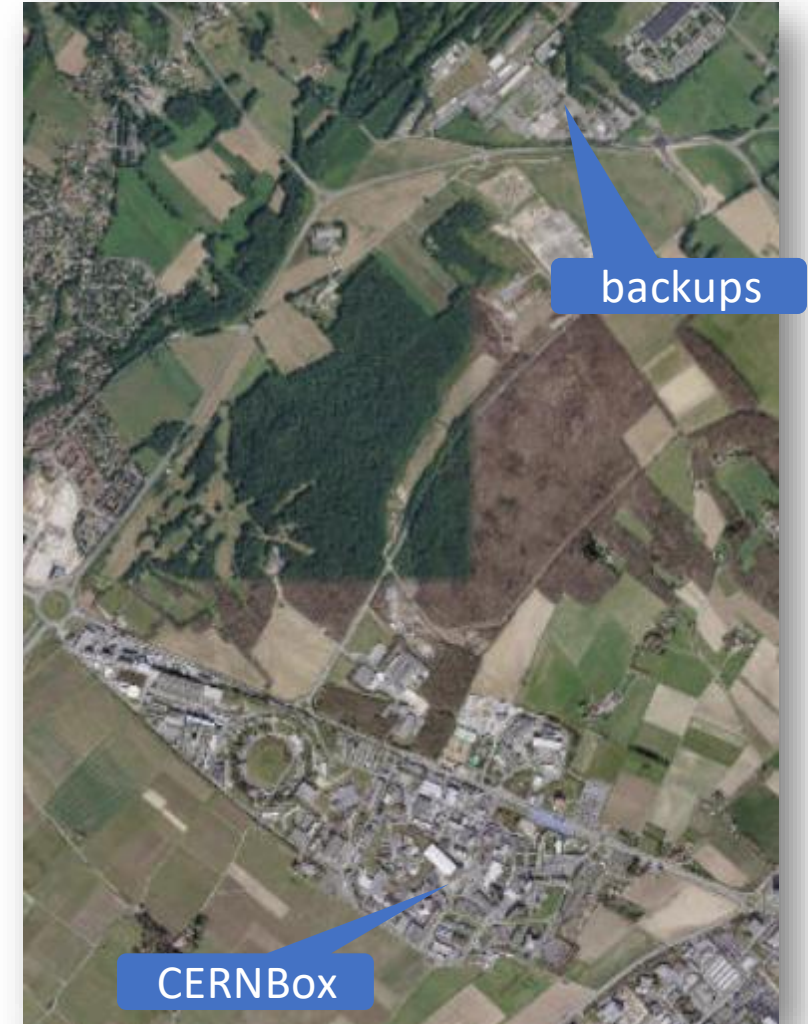
- **Backup orchestration tool** developed at CERN
- Creates **daily snapshots** for all *CERNBox users and projects* spaces and some services *cephfs* bases
- Based on **restic**:
 - Fast
 - Easy
 - Repository/snapshots based
 - Deduplication
 - Encryption
 - Compression
 - Prune old snapshots
 - Support for multiple backends



cback backup system

Overview

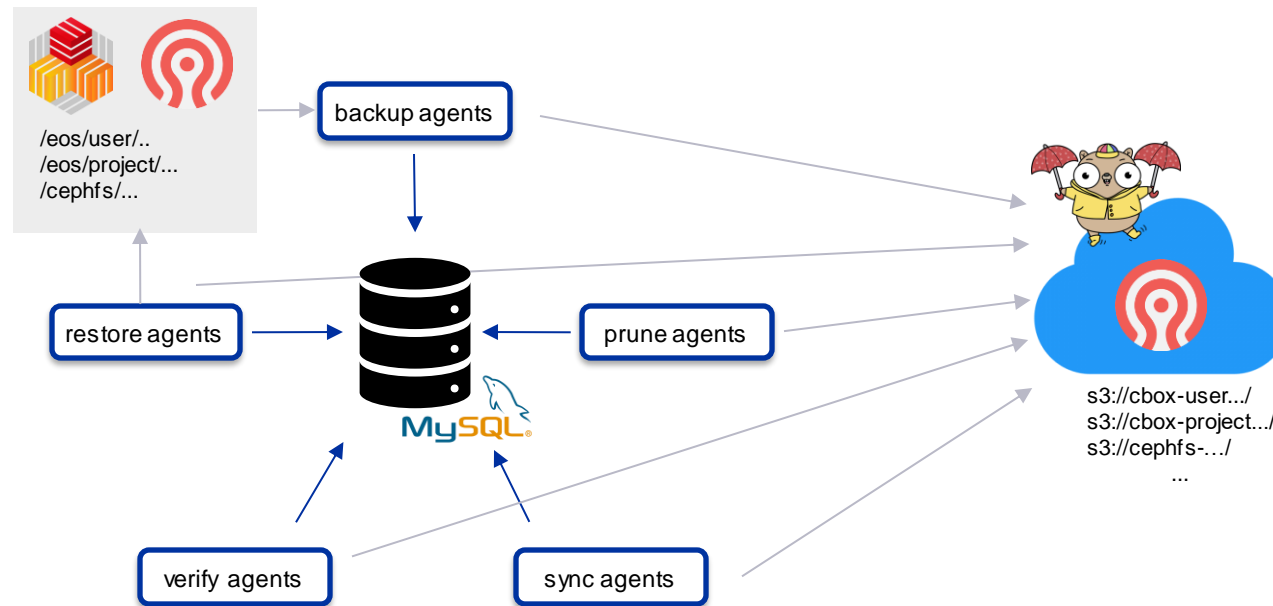
- Data stored in an **other geographical location** in the CERN campus
- **+850M** of files processed every day
- **+5.3PB** of files backed up
- **~40k** backup jobs
- Support for different sources
 - EOS
 - CephFS
 - ...



cback backup system

Architecture

- Composed by **stateless** agents (backup, restore, prune, verify and sync)
- They share the state (i.e. the jobs to run) in a **central database**



cback backup system

Architecture

- Agents pick a random job from a **pending queue** and run it (**asynchronous**)
- The operator can run a job **synchronously**
- After successfully run a job, the agent enqueues it in a **completed queue**
- A **switch** agent move a job from the completed to the pending queue periodically
 - For example every 24 after the job finished to run

Integrations

- **Expose** to CERNBox users their daily backups
- **Self-service restore**
- Expose other backups to **service admins**
 - without sharing internal restic secrets or s3 keys

Portal

default

GET /login Login

GET /whoami Whoami

backups

GET /backups/ Get Backups List

POST /backups/ New Backup

GET /backups/{id} Get Backup Id

PUT /backups/{id} Update Backup Id

GET /backups/{id}/snapshots Get Snapshots From Backup Id

GET /backups/{id}/snapshots/{snapshot_id}/{path} Get File From Snapshot

OPTIONS /backups/{id}/snapshots/{snapshot_id}/{path} Stat Resource From Snapshot

GET /backups/{id}/mount Mount

DELETE /backups/{id}/mount Umount

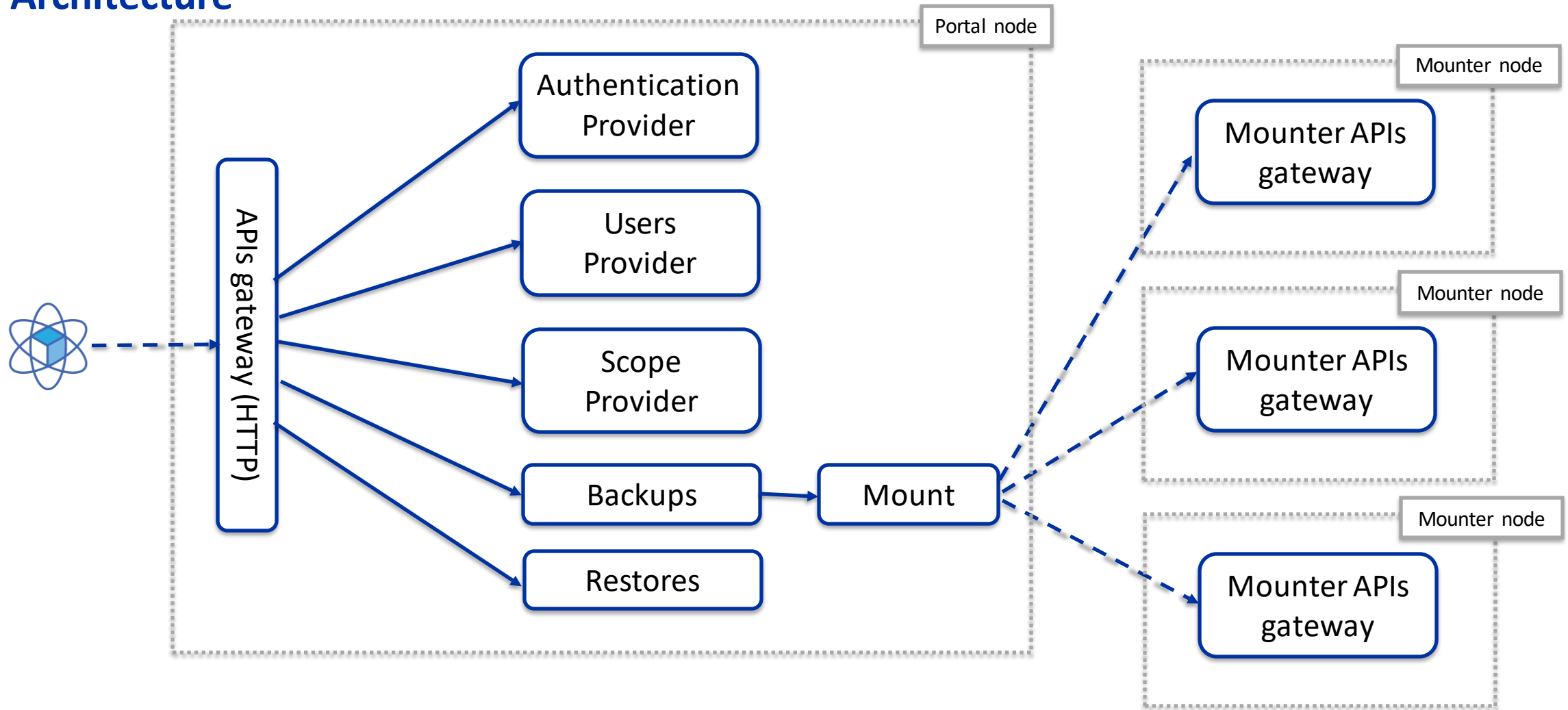
restores

GET /restores/ Get Restores List

POST /restores/ New Restore

GET /restores/{id} Get Restore Id

Portal Architecture



Portal

Authentication provider

- Integration with **ldap**
 - Users can use the cback APIs using CERN credentials
- **Impersonator** driver
 - Impersonate a user using a token previously generated
 - Useful with integration with external applications (CERNBox)



Portal

Scope provider

- Provide the list of backups the user has access to (+ permissions)
- Provided by **drivers**
 - **cbox**: *"a user has access to his home folder + projects in which is admin"*
 - **manual**: operator can specify *manually* the permissions of a user

Portal

Mounters

- Have the responsibility to **mount** a restic repository
- Manage the **life cycle** of a mount
 - 5 min of inactivity the mount expires
- Expose the local mounted restic repository by some **HTTP APIs** to the clients
 - Protected by bearer auth using a token valid only during the life cycle of the mount



> eos > user > g > gdelmont ⋮



Files ✓



Backups



oad



Favorites



Shares



Projects



Deleted files



HPC Data



Name ↓

Shares

Si



asd

5.



backup

29



Documenti



4.



Documenti riconoscimento

6.



Downloads

9.



Immagini

18



Lavoro

2



Nozze

1.



PRIVATE

39.



Programmazione

1.



ssd

1.

<

☰

Restore jobs

↶

My backups

↶

Project backups

My backups

backups-me

<input type="checkbox"/>	Name	Modified
<input type="checkbox"/>	 2022-10-31T19:00:14+01:00	4 months ago
<input type="checkbox"/>	 2022-11-30T19:00:10+01:00	3 months ago
<input type="checkbox"/>	 2022-12-31T19:00:10+01:00	2 months ago
<input type="checkbox"/>	 2023-01-31T19:00:05+01:00	1 month ago
<input type="checkbox"/>	 2023-02-12T19:00:05+01:00	23 days ago
<input type="checkbox"/>	 2023-02-19T19:00:07+01:00	16 days ago
<input type="checkbox"/>	 2023-02-26T19:00:05+01:00	9 days ago
<input type="checkbox"/>	 2023-02-28T19:00:09+01:00	7 days ago
<input type="checkbox"/>	 2023-03-01T19:00:04+01:00	6 days ago
<input type="checkbox"/>	 2023-03-02T19:00:07+01:00	5 days ago
<input type="checkbox"/>	 2023-03-03T19:02:13+01:00	4 days ago
<input type="checkbox"/>	 2023-03-04T19:00:11+01:00	3 days ago
<input type="checkbox"/>	 2023-03-05T19:00:07+01:00	2 days ago
<input type="checkbox"/>	 2023-03-06T07:00:05+01:00	2 days ago
<input type="checkbox"/>	 2023-03-06T19:00:07+01:00	1 day ago
<input type="checkbox"/>	 2023-03-07T07:00:05+01:00	1 day ago
<input type="checkbox"/>	 2023-03-07T19:00:05+01:00	14 hours ago
<input type="checkbox"/>	 2023-03-07T19:00:05+01:00	14 hours ago
<input type="checkbox"/>	 2023-03-08T07:00:11+01:00	2 hours ago

































 Restore jobs

 My backups

 Project backups

My backups

backups-me > cback > eos > user > g > gdelmont > 2022-10-31T19:00:14+01:00 ⋮

<input type="checkbox"/> Name	Modified	Status	Act
<input type="checkbox"/>  PRIVATE	5 months ago	 Restore	
<input type="checkbox"/>  Downloads	1 year ago	 Restore	
<input type="checkbox"/>  Immagini	5 months ago	 Restore	
<input type="checkbox"/>  gatto.jpg	6 months ago	 Restore	
<input type="checkbox"/>  New file.pptx	4 months ago	 Restore	
<input type="checkbox"/>  Programmazione	2 years ago	 Restore	
<input type="checkbox"/>  Università	11 months ago	 Restore	
<input type="checkbox"/>  backup	5 months ago	 Restore	
<input type="checkbox"/>  ssd	1 year ago	 Restore	
<input type="checkbox"/>  test	5 months ago	 Restore	
<input type="checkbox"/>  Fronalpstock_big.jpg	7 months ago	 Restore	
<input type="checkbox"/>  Lavoro	1 year ago	 Restore	
<input type="checkbox"/>  Nozze	6 months ago	 Restore	
<input type="checkbox"/>  Documenti	1 year ago	 Restore	
<input type="checkbox"/>  Documenti riconoscimento	2 years ago	 Restore	

































My backups

Restore jobs

My backups

Project backups

backups-me > cback > eos > user > g > gdelmont > 2022-10-31T19:00:14+01:00

<input type="checkbox"/> Name	Modified	Status	Act
<input type="checkbox"/>  PRIVATE	5 months ago	 Restore	
<input type="checkbox"/>  Downloads	1 year ago	 Restore	
<input type="checkbox"/>  Immagini	5 months ago	 Restore	
<input type="checkbox"/>  gatto.jpg	6 months ago	 Restore	
<input type="checkbox"/>  New file.pptx	4 months ago	 Restore	
<input type="checkbox"/>  Programmazione	2 years ago	 Restore	
<input type="checkbox"/>  Università	11 months ago	 Restore	
<input type="checkbox"/>  backup	5 months ago	 Restore	
<input type="checkbox"/>  ssd	1 year ago	 Restore	
<input type="checkbox"/>  test	5 months ago	 Restore	
<input type="checkbox"/>  Fronalpstock_big.jpg	7 months ago	 Restore	
<input type="checkbox"/>  Lavoro	1 year ago	 Restore	
<input type="checkbox"/>  Nozze	6 months ago	 Restore	
<input type="checkbox"/>  Documenti	1 year ago	 Restore	
<input type="checkbox"/>  Documenti riconoscimento	2 years ago	 Restore	






- 
-  Restore jobs


 My backups

 Project backups

Restore jobs

<input type="checkbox"/>	Name	Storage	Restore
<input type="checkbox"/>	 gatto.jpg		 Restor
<input type="checkbox"/>	 ocis web commits.md		<div>Restore available</div>



 **Success**
Backup restore was initiated



CERNBox



CERNBox > eos > user > g > gdelmont ⋮



All files



Favorites



Shares



Projects



Deleted files



HPC Data



New



Upload



Name ↓



__restored_03_08_2023_09_04_11_



asd



backup



Documenti



Documenti riconoscimento

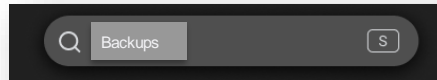


Download

Integration with CERNBox

Vision for UI

- Notification, when restore job is done
- Search in backups
- Compare versions of a file from different backups





Empowering CERNBox users with self-service restore functionality

Gianmaria Del Monte (IT-SD, CERN)
gianmaria.del.monte@cern.ch