



**CS³
MESH⁴
EOSC**

Connecting European Data



C(ERN) BACK(UP)

Consolidated multi-petabyte backup solution for heterogenous storages

Gianmaria Del Monte

Software engineer - CERN



CS3MESH4EOSC has received funding from the European Union's Horizon 2020 Research and Innovation programme under **Grant Agreement No. 863353**.

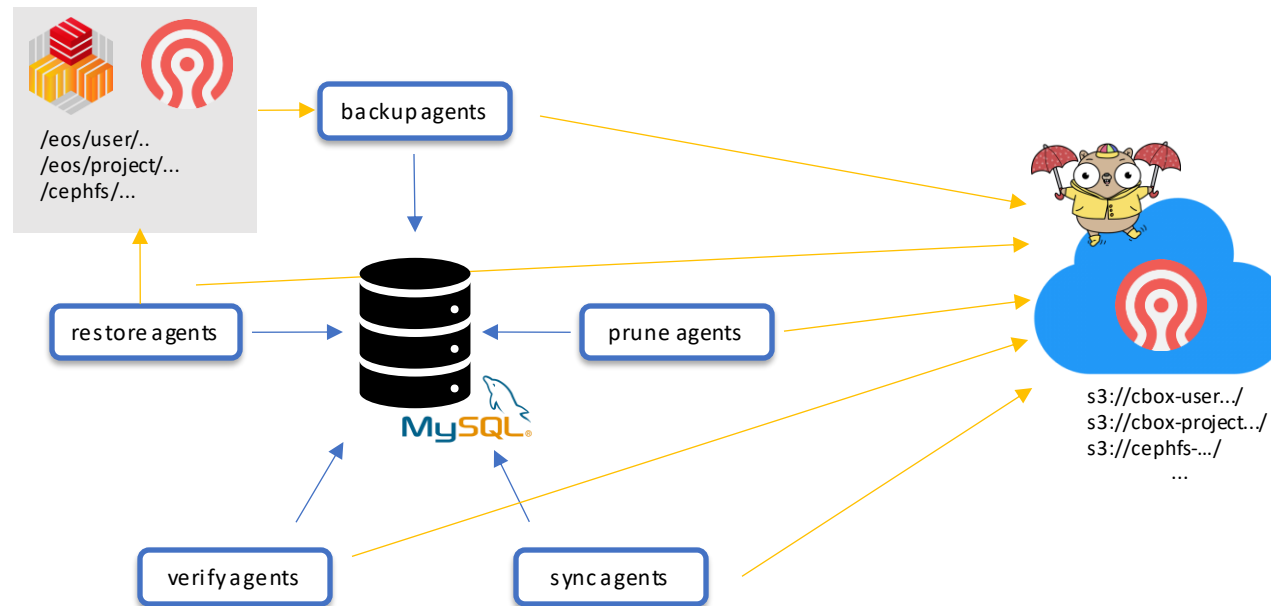
- # **Backup orchestration tool** developed at CERN
- # Creates **daily snapshots** for all *CERNBox users and projects* and some services *cephfs* based
- # Based on **restic**
 - # fast
 - # easy
 - # repository/snapshot based
 - # deduplication
 - # encryption
 - # compression
 - # prune old snapshots



- # 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 job
- # Support for different sources
 - # EOS
 - # CephFS
 - # ...



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



- # 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 24h after the job finished to run

Used to backup **several sources**:

- # CERNBox (EOS)
- # Manila shares (CephFS)
- # Gitlab (CephFS)

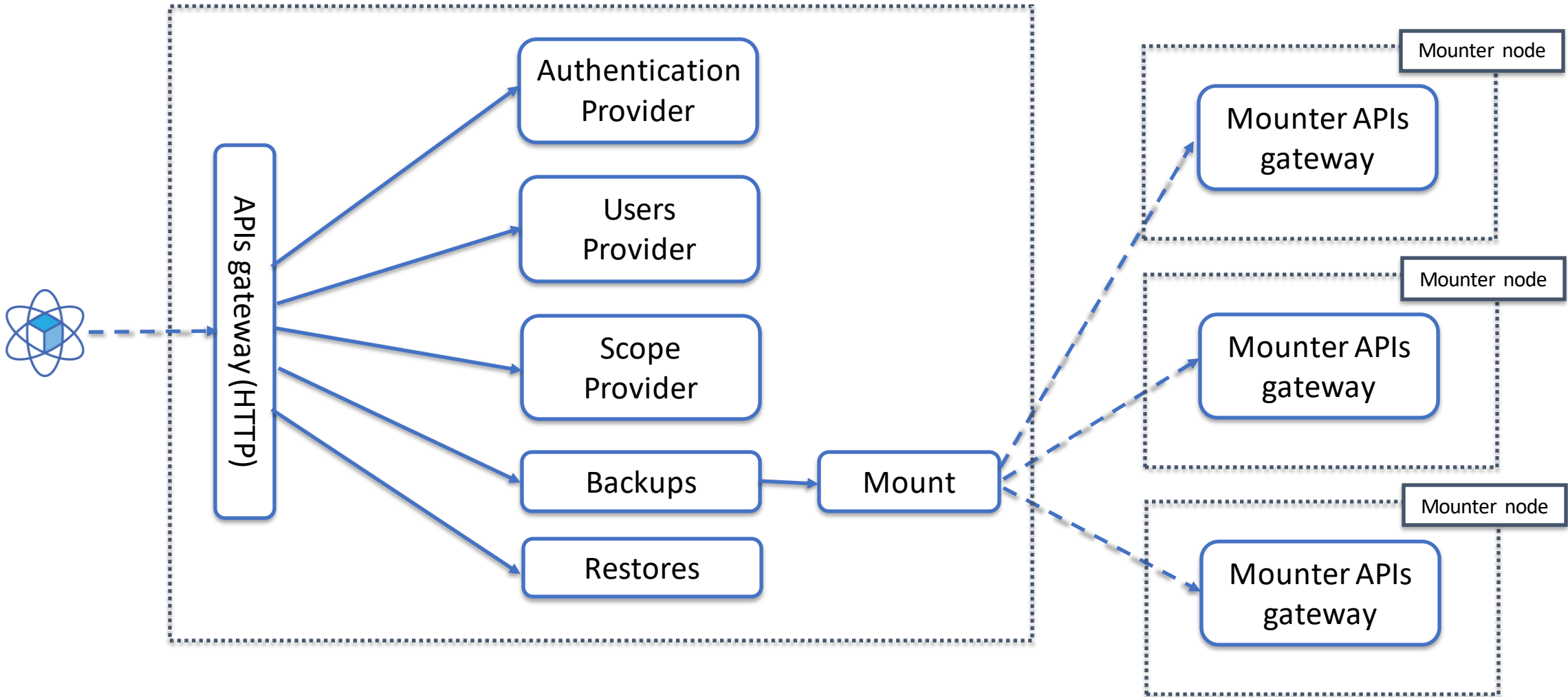


Having different instances of cback is *expensive*:

- # **Resources**: different machines per service means replication of resources, that in some cases could not be fully exploited.
- # **Managing**: an update of cback requires an update in all the instances. Different configurations are spread into different machines.

- # cback introduces the concept of **groups**
 - # Every group has its own configuration
 - # Different source filesystems (EOS, CephFS, ...)
 - # Different sets of s3 credentials/endpoints
 - # Different restic secrets
 - # Different schedules
- # *CERNBox, manila shares* and *gitlab* are backed up by the same cback cluster

- # **Expose** to CERNBox users their daily snapshots
- # **Self-service restore**
- # Expose other backups to **service admins**
 - # Without sharing internal restic secrets





Files ✓

> eos > user > g > gdelmont ⋮



Backups



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.




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

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	



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

- Restore jobs
- My backups
- Project backups

- Name
-  gatto.jpg
-  ocis web commits.md



Sto
Resto
Restore available

 **Success**
Backup restore was initiated








CERNBox > eos > user > g > gdelmont ⋮

 All files

 New

 Upload

-  Favorites
-  Shares
-  Projects
-  Deleted files
-  HPC Data

Name ↓

 **__restored_03_08_2023_09_04_11_** 

 asd

 backup

 Documenti

 Documenti riconoscimento

 Downloade

- # CERNBox integration
 - # notification, when restore is done*
 - # search in backups
 - # compare versions of a file from different backups
- # Tape integration
 - # restic metadata are included in the packed data
 - # reduce read access to tapes as much as possible
 - # need a copy of metadata in hot storage to create a restore plan

*Modern notification system for Sync and Share using NATS – Javier @ 12.15



CS³
MESH⁴
EOSC

Connecting European Data

Thank you!
Discover more on...

 [cs3mesh4eosc.eu](https://www.cs3mesh4eosc.eu)

 [company/cs3mesh4eosc](https://www.linkedin.com/company/cs3mesh4eosc)

 [CS3org](https://twitter.com/CS3org)

 [CS3MESH4EOSC Project](https://www.youtube.com/channel/UCHKcZEKmqXjCvc3MLFjFxbw)
<https://www.youtube.com/channel/UCHKcZEKmqXjCvc3MLFjFxbw>

Gianmaria Del Monte

 gianmaria.del.monte@cern.ch

 <https://www.linkedin.com/in/gianmaria-del-monte-81575b162/>



CS3MESH4EOSC has received funding from the European Union's Horizon 2020 Research and Innovation programme under **Grant Agreement No. 863353**.



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	✓	🔒