

# Providing Backup Service via Restic to CTA

Fons Rademakers

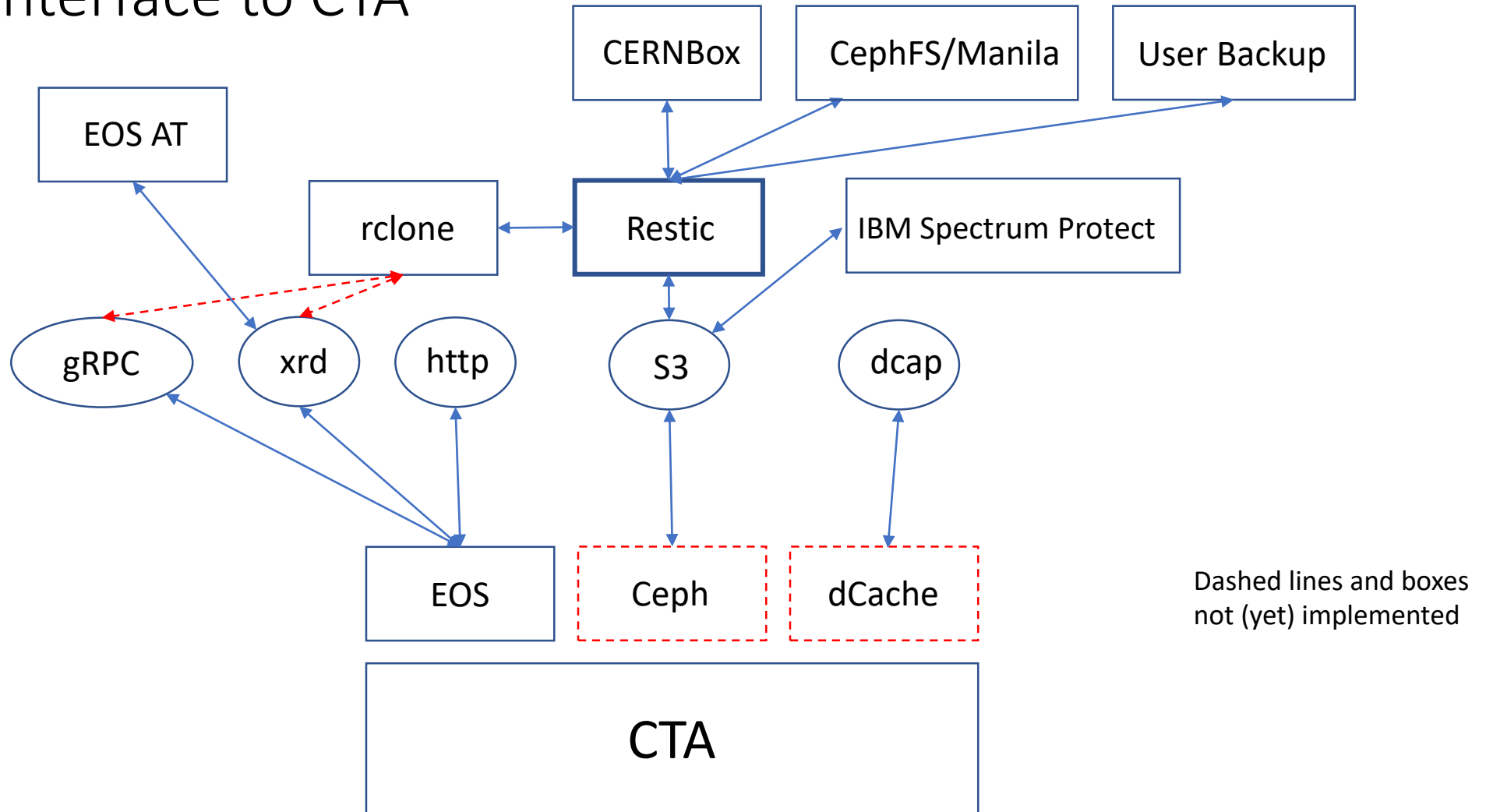
# Restic Support for Cold Storage

- A lot of recent interest in having restic supporting cold storage
  - AWS S3 Glacier and Glacier Deep Archive
  - Google Cloud Platform Coldline Storage and Archive Storage
  - Azure Blob Storage Cold and Archive
  - CERN CTA
- Restic meta data not separated from the packed data
  - Need hot/cold separation
  - Reduce read access to the cold storage as much as possible
  - Reduce all API calls to the cold storage as much as possible
  - Need functionality to get info which cold data will be needed (for warmup)

# Restic Support for Cold Storage

- First steps made in support by recent restic PR [#3235](#)
  - Adds support for option `--repo-hot` to specify a “hot repo” for all repository files except packed data files.
  - The `-repo` is still fully written (packed and meta data)
  - The hot repo acts as cache and takes small fraction of the space of the full repo

# Restic Interface to CTA



These file systems support other, like POSIX compliant, access methods less relevant for CTA

# Restic Interface to CTA

- From the above picture the most efficient, useful and elegant solution looks to be the one provided by implementing Ceph front-end support for CTA
  - Efficient as no extra rclone process overhead is needed
  - Useful as Ceph supports S3 and many potential outside CTA users like an S3 interface
  - Elegant as providing Ceph front-end support solves several issues in one go by making restic support transparent (except for the usage of the `-repo-hot` flag) and providing CTA S3 support which has been asked for by many (potential) users
  - The Ceph front-end storage serves as short lived cache, except for the partition storing restic hot data

# Discussion

- Main decision:
  - Implementation of a CTA Ceph front-end (like EOS CTA)
    - Understanding the effort, code-wise, time-wise?
  - Implementation of rclone EOS interface
    - Possible continuation of work of summer student K. Derrien
- Testing of modified restic with the `-repo-hot` option
- Policies to be discussed:
  - How many incremental restic snapshots should be maintained?
  - Size of "little" Ceph CTA fast buffer, size of "hot repo" partition?

# Summary

- We discuss a way to use CTA as cold storage backup service via restic
- An S3 interface to CTA is a frequently requested option, especially by (potential) outside CTA users who already have Ceph S3 deployed
  - AARNet mentions they will work on an S3 interface... collaboration?
- A CTA Ceph S3 front-end would solve this issue and make restic usage transparent