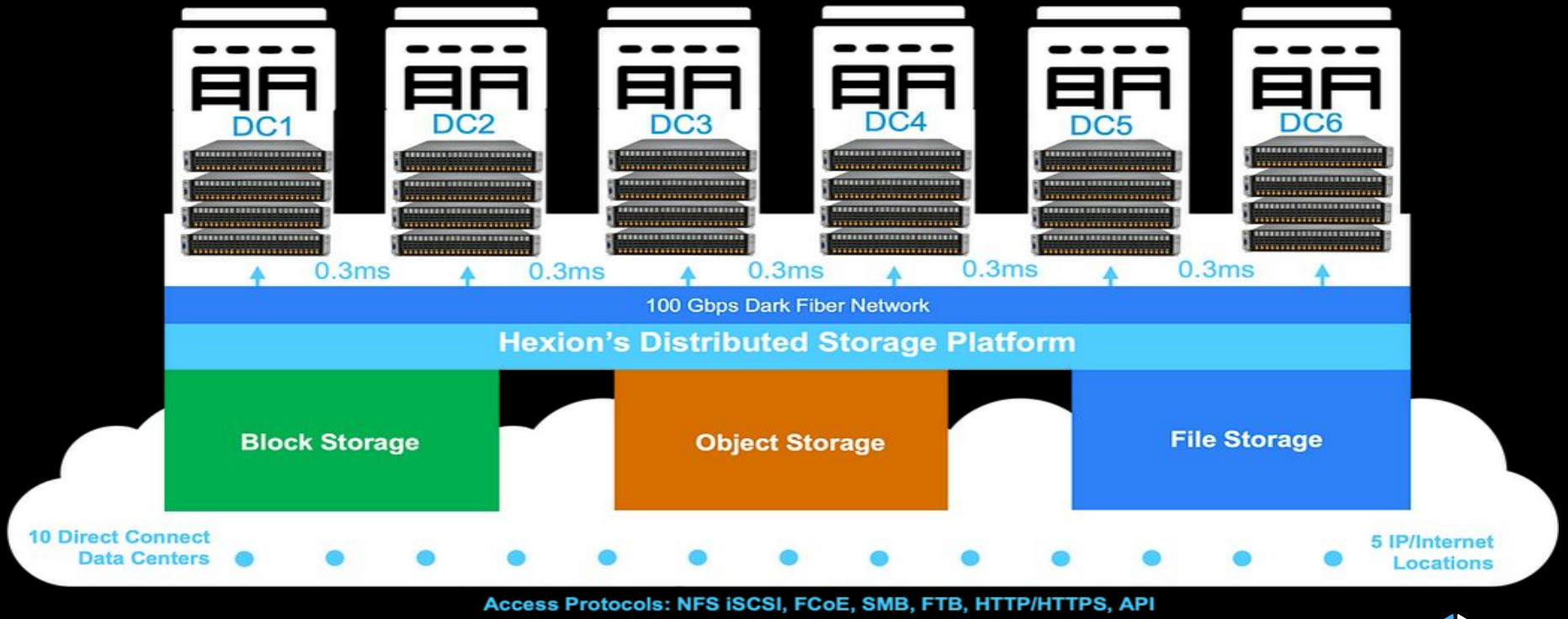
ceph days

Ceph Rados as an S3 Interface to Tape Storage

Stuart Hardy & Zaid Bester

HEXION

About Hexion Cloud



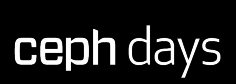
How would we use a Tape Library

ceph days

Core Idea	Vendor Independence	Hybrid Flexibility
Copies Ceph Rados Objects to and From the Tape Library	Ceph's open-source nature prevents lock-in common to proprietary systems. Abstract Tape Library functionality (shim).	A small, high-performance Ceph cluster handles S3 I/O. Tape provides cost-efficient, durable, long-term storage.

HEXION

The main driver: Cost / TB



If you have, or expect to have a certain level of scale:

Capex (only)	HDD Cluster	Tape Library Size
$100 000	2 PB	5 PB
$150 000	3 PB	13 PB
$250 000	5 PB	30 PB

Other costs / assumptions:
Reduce recovery time for a lower price.
Management Cost (estimated $0,70 / TB),
Power: **Tape library consumes far less power than an HDD cluster**

HEXION

HDD Cluster

ceph days

64 Dish Array in Karoo

Takes high-res images that has transformed our understanding of the Cosmos

Existing 50PB HDD Cluster
300 Hosts
3000 Disks

New Archive Layer

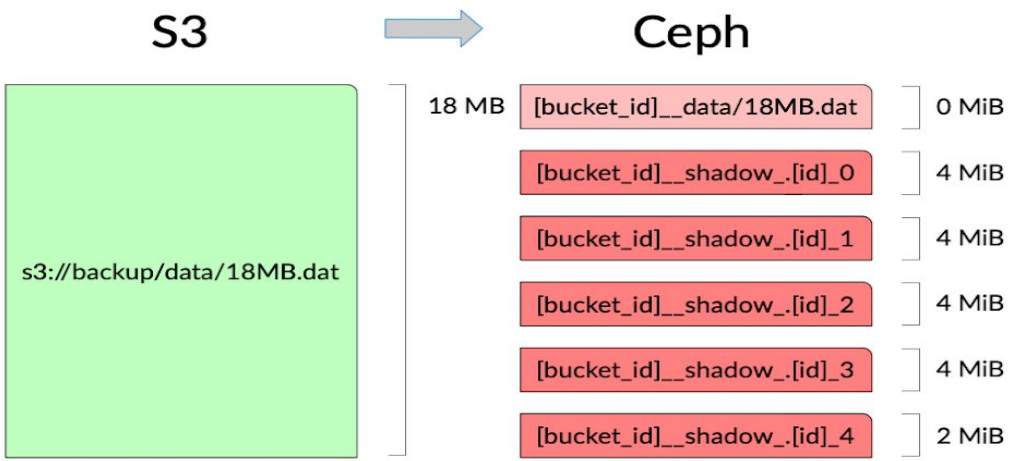
ceph days

New 80 PB Tape Archive

Base model TS4500
+ 4 x Expansion Frames - LTO 9

Production 80 PB Tape Library

The Anatomy of an S3 Object



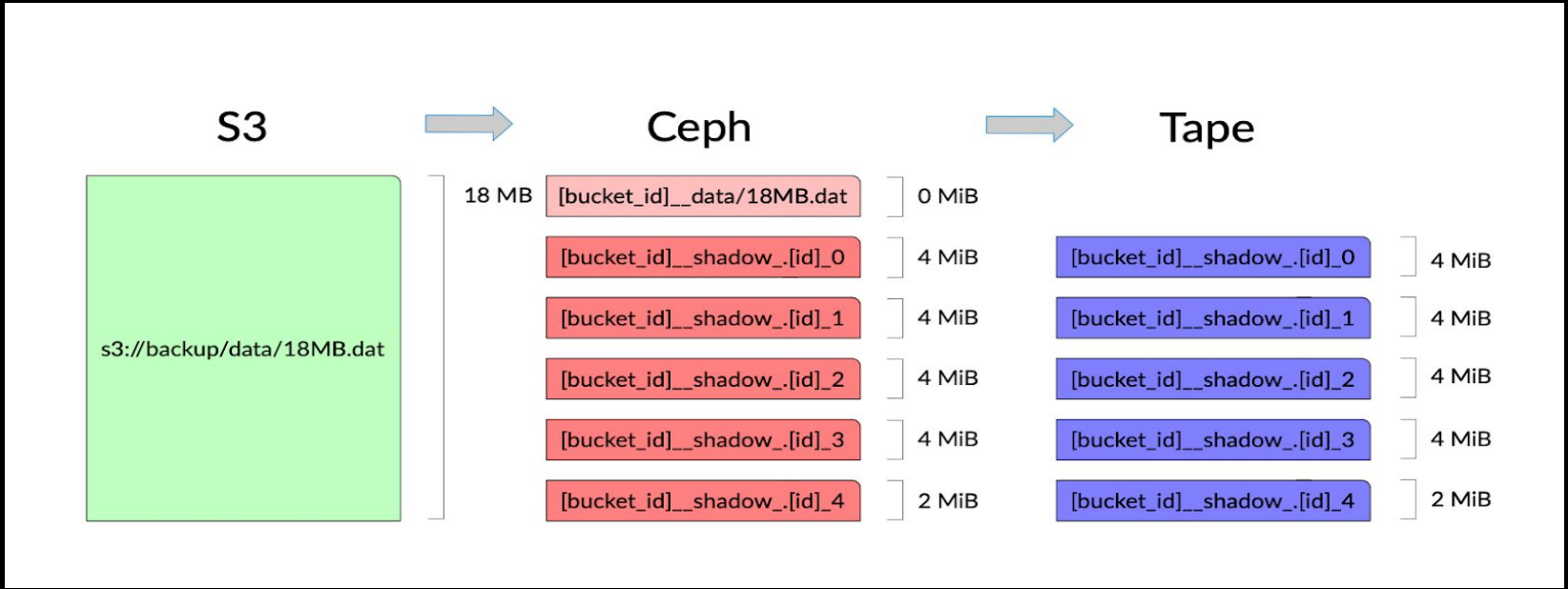
1. S3 object uploaded to Ceph ® split into multiple RADOS objects.

2. Ceph RGW manifest maps head and tail RADOS objects.

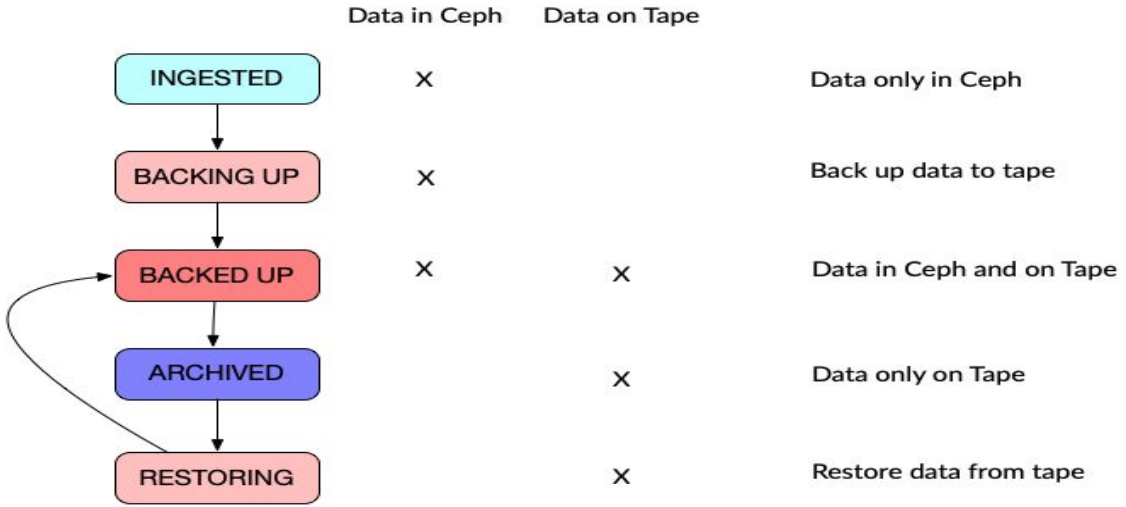
3. Shim Layer reads the manifest, locates tail objects, and streams them to tape, once completed the object can be truncated or deleted.

4. To restore, data is read from tape and reconstructed into RADOS objects.

Storing the RADOS Objects on tape



Simple Data State Representation



1. Data Lifecycle & Air-Gapped Mode

2. After archiving, RADOS tail objects can be truncated or deleted while metadata remains.

3. Data stored only on tape – providing physical air-gap protection.

4. S3 metadata remains visible; restores from tape rehydrate full access.

Management view

ceph days