

Snapshot Cauterize

Sandeep Joshi

Narain C Ramadass

EMC, Isilon Storage Division

Agenda

- ❑ OneFS overview
- ❑ Snapshot – basics
- ❑ Cauterize
 - ❑ Why
 - ❑ Technical overview
- ❑ Challenges
- ❑ Future work

OneFS Overview

- ❑ NAS file server
- ❑ Scalable
 - ❑ Add more storage in 5 mins
- ❑ Reliable
 - ❑ 8x mirror / +4 parity
 - ❑ Striped across nodes
- ❑ Single volume file system (20.25 PB)
- ❑ 3 to 144 nodes
- ❑ Fully symmetric peers
 - ❑ No metadata servers
- ❑ Commodity hardware
 - ❑ CPU, Mem, Disks (12 to 36)



OneFS Overview contd .



- ❑ Concurrent access to all files with the following protocols
 - ❑ CIFS/SMB
 - ❑ NFS
 - ❑ HTTP/FTP
 - ❑ HDFS

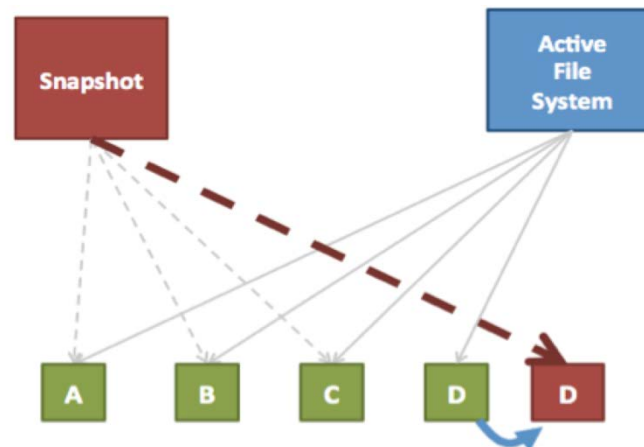
OneFS Overview contd ..

- What is it?

File system, volume manager and RAID combined into a single software layer across all nodes in a cluster

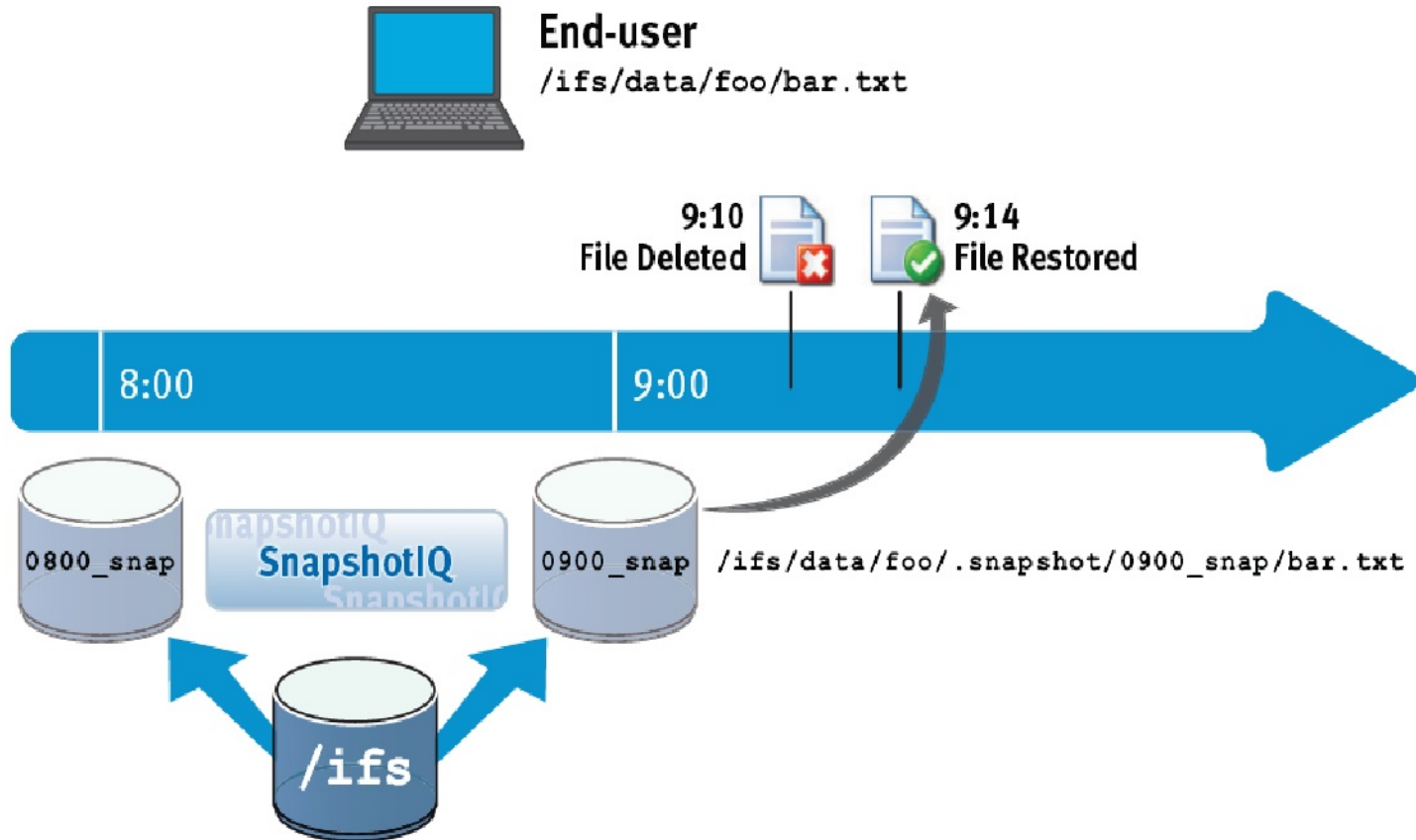
Snapshots

- ❑ A file system snapshot provides a consistent image of a file system at a particular point in time
- ❑ OneFS
 - ❑ Stores differences in the snapshot
 - ❑ Directory level



- ❑ Benefits
 - ❑ Data Protection
 - ❑ Tape backup
 - ❑ Replication/mirroring
 - ❑ Fast Recovery (directly from snapshot)
 - ❑ File system consistency
 - ❑ Analytics
 - ❑ Testing

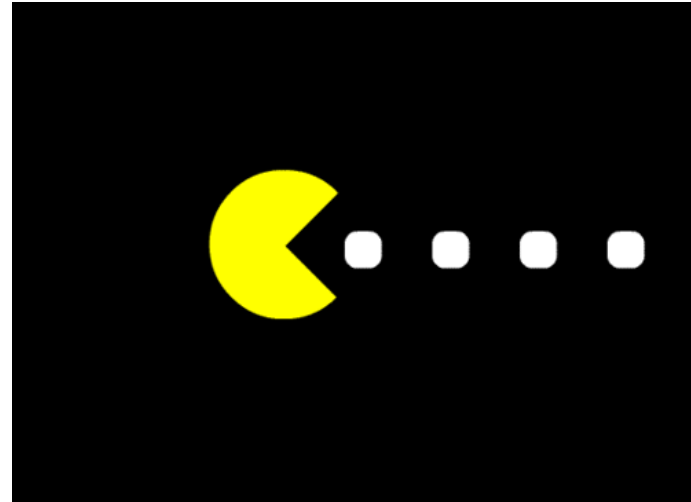
Snapshots



- ❑ How to reclaim space?
 - ❑ Current approaches
 - ❑ Limit number of snapshots
 - ❑ Delete
 - ❑ Stubbing

Another approach - cauterize

- ❑ Per file operation/property
- ❑ Keep meta-data
 - ❑ Size attribute not changed
 - ❑ Free only data blocks
- ❑ Limitations
 - ❑ Oldest snapshot first
 - ❑ No intermediate
- ❑ Side effect
 - ❑ EIO on read
 - ❑ No useful data in Cauterized files
 - ❑ Handle user-level quotas



Application impact

- ❑ Backup window – no cauterization
- ❑ Remote replication window
- ❑ fsck and other FS tools
- ❑ 3rd Party application

Snapshot Cauterization – user impact

- ❑ It's all about perception - snapshots are "untouched"
- ❑ Administrative and compliance requirements
- ❑ In a snapshot,
File Size != space reclaimable on file cauterization

Challenges

- ❑ Long running
- ❑ Normal IO impact
- ❑ Hard-links
- ❑ Clones
- ❑ ADS

Future work

- ❑ Directory cauterization
- ❑ Meta-data only snapshot
 - ❑ Pre-emptive cauterization of all files in the snapshot
 - ❑ Useful for FS analytics
- ❑ Performance measurement
- ❑ Guidance
 - ❑ Which files to cauterize
 - ❑ Estimation of disk space that can be reclaimed
 - ❑ File cauterize in "simulation" or "dry-run" mode

Thank You

sandeep.joshi@isilon.com