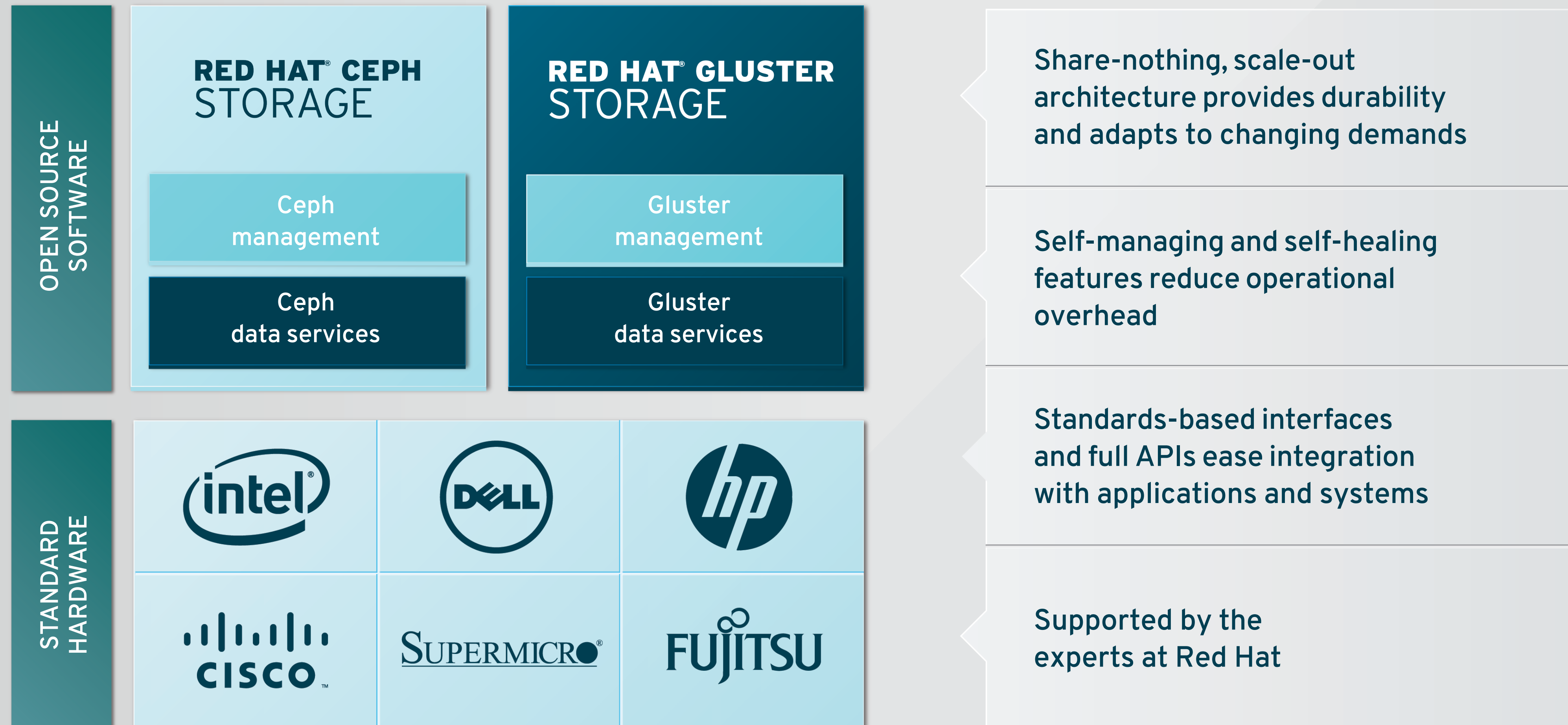# AGENDA

The Portfolio

A Quick Look Back

The Present

What's Next: The Future
Gluster Upstream Roadmap
Red Hat Gluster Storage Integration Roadmap

redhat.

THE RED HAT STORAGE PORTFOLIO

OPEN SOURCE SOFTWARE

RED HAT® CEPH STORAGE

Ceph management

Ceph data services

RED HAT® GLUSTER STORAGE

Gluster management

Gluster data services

STANDARD HARDWARE

intel®

D&LL

hp

CISCO™

SUPERMICRO®

FUJITSU

Share-nothing, scale-out architecture provides durability and adapts to changing demands

Self-managing and self-healing features reduce operational overhead

Standards-based interfaces and full APIs ease integration with applications and systems

Supported by the experts at Red Hat

redhat

redhat

RED HAT GLUSTER STORAGE

Nimble file storage for petabyte-scale workloads

Purpose-built as a scale-out file store with a straightforward architecture suitable for public, private, and hybrid cloud

Simple to install and configure, with a minimal hardware footprint

Offers mature NFS, SMB and HDFS interfaces for enterprise use

Customer Highlight: Intuit
Intuit uses Red Hat Gluster Storage to provide flexible, cost-effective storage for their industry-leading financial offerings.

intuit®

RED HAT® GLUSTER STORAGE

TARGET USE CASES

Analytics
• Machine analytics with Splunk
• Big data analytics with Hadoop

Enterprise File Sharing
• Media Streaming
• Active Archives

Enterprise Virtualization

redhat.

FOCUSED SET OF USE CASES

ANALYTICS
- Big Data analytics with Hadoop
- Machine data analytics with Splunk

CLOUD INFRASTRUCTURE
- Virtual machine storage with OpenStack
- Object storage for tenant applications

RICH MEDIA AND ARCHIVAL
- Cost-effective storage for rich media streaming
- Active archives

SYNC AND SHARE
- File sync and share with ownCloud

ENTERPRISE VIRTUALIZATION
- Storage for conventional virtualization with RHEV

redhat

Delivering on-demand TV with Red Hat Storage

"We now have access to a much faster system. Compared to the EMC solution, our costs are 8 times lower."

Niels Jonkman
Business Manager IPTV, Glashart Media

BUSINESS CHALLENGE

- Use of video-on-demand services at Glashart Media, subsidiary company of KPN, tripled in 2 years and exponential growth is expected
- Needed a large scale-out, highly available and cost-effective storage solution to meet growing demand of video-on demand TV services
- Emphasis on scale-out and low cost storage since margins in business are small

SOLUTION

- Red Hat Gluster Storage
- Supermicro X9 servers and Cisco 10Gbe switch

BENEFITS

- Reduced storage costs by 8x
- Unified management of file and object data
- Increased flexibility, scalability and redundancy
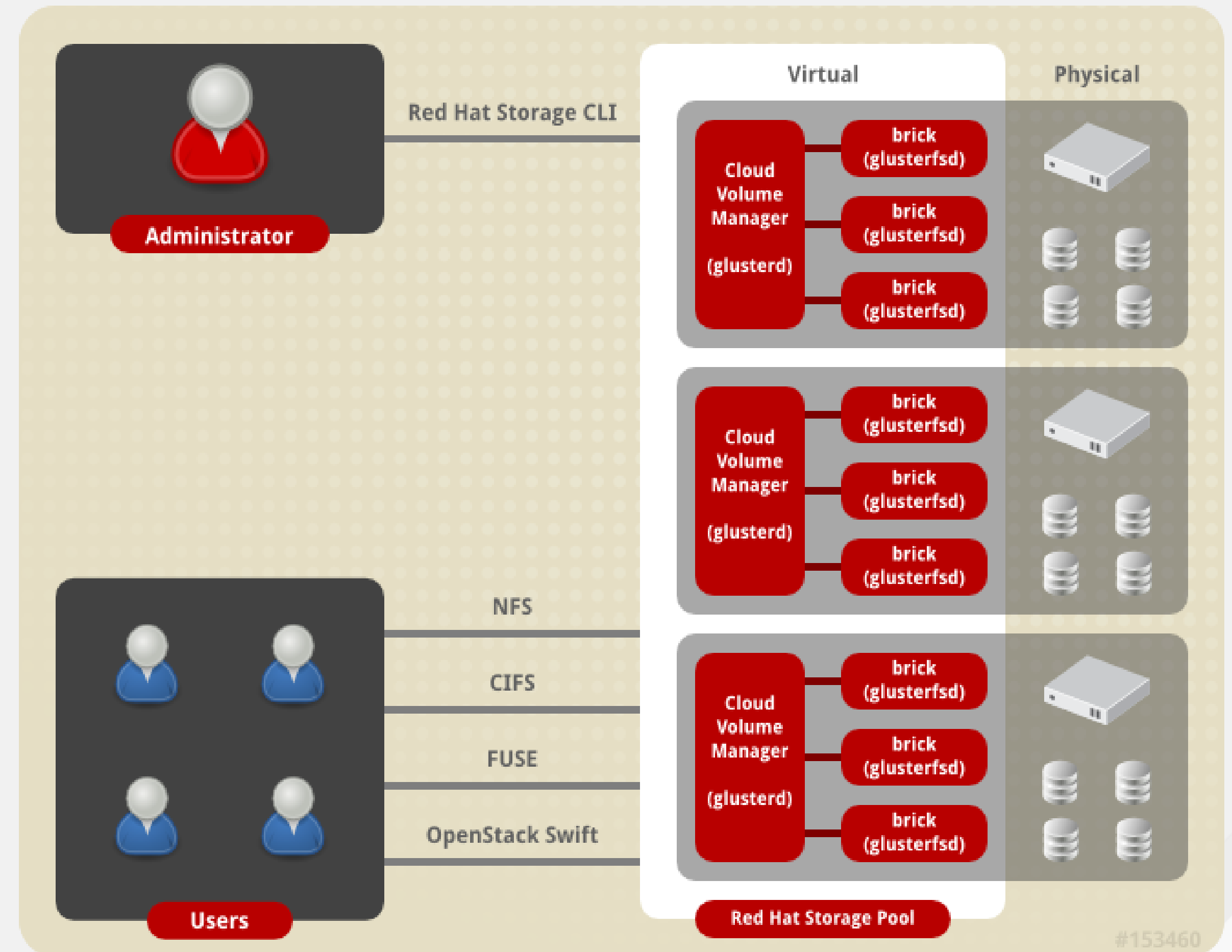- Gained ability to expand service portfolio

Gluster Architecture

Scale-out with Global namespace. Built for massive scale.

No <u>centralized meta-data server</u> = no Single Point Of Failure

<u>Modular architecture</u> for ultimate flexibility (e.g. stripe or not!)

Deployment Agnostic – run on bare metal, virtual machine and cloud (and containers!)

A QUICK LOOK BACK

LOOKING BACK

Red Hat Storage Server 2.0 (GA June 2012)
- 6 updates released
- Features: VM image store, performance & stability
- EOL-ed on June 2014

Red Hat Storage Server 2.1 (GA Sep 2013)
- 6 updates released. Planned EOL October 2015.
- Features: Quota, Geo-Rep, management console, SMB 2.0

redhat.

RHGS 3.0 (DENALI)
PREVIOUS MAJOR RELEASE

Launched September 2014

- GlusterFS 3.6, RHEL 6

Key Features
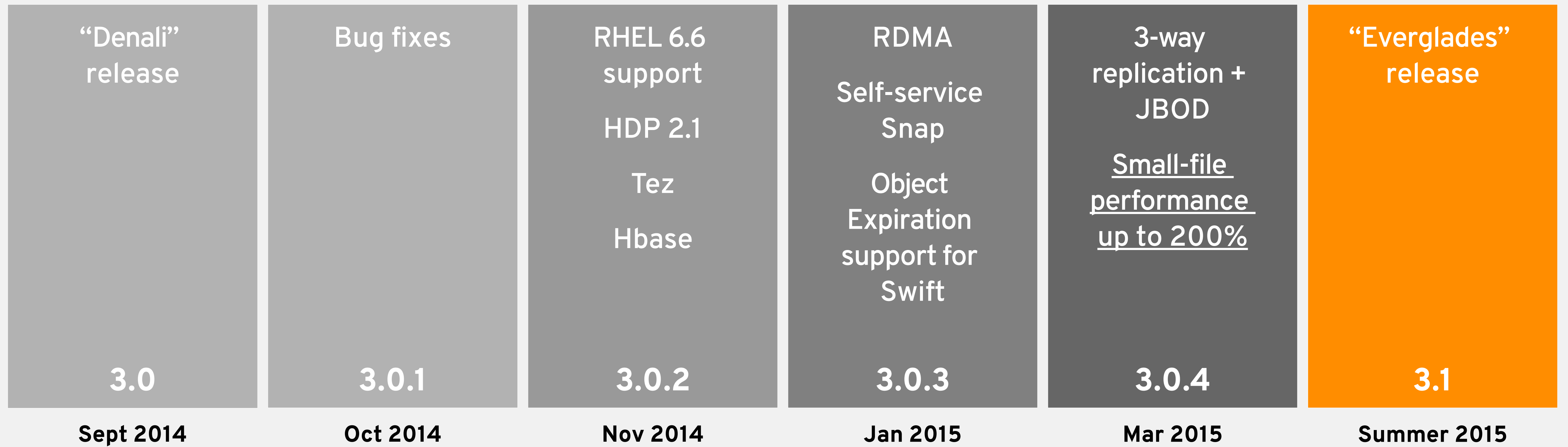- Volume snapshots for disk based backup

Management
- Monitoring using Nagios
- SNMP Support
- Rolling upgrade support, Red Hat CDN based delivery

Hadoop Plug-in for HortonWorks Data Platform 2.0.6

Scale
- 60 drives per server, 128 nodes per cluster

redhat.

BETWEEN 3.0 AND TODAY

"Denali" release	Bug fixes	RHEL 6.6 support HDP 2.1 Tez Hbase	RDMA Self-service Snap Object Expiration support for Swift	3-way replication + JBOD <u>Small-file performance up to 200%</u>	"Everglades" release
3.0	**3.0.1**	**3.0.2**	**3.0.3**	**3.0.4**	**3.1**

Sept 2014 **Oct 2014** **Nov 2014** **Jan 2015** **Mar 2015** **Summer 2015**

redhat.

THE PRESENT

RHGS 3.1 (EVERGLADES)

Baseline
- GlusterFS 3.7, RHEL 6, RHEL 7

Key Features
- Erasure Coding, Tiering, Bit-Rot Detection

Protocols
- Active/Active NFSv4
- SMB 3 (protocol negotiation, in-flight encryption, server-side copy)

Red Hat Gluster Storage Console
- **Device Management**, Geo-Rep, Snapshot, Dashboard, Snapshot Scheduling

Security
- SSL based network encryption
- **SELinux Enforcing Mode**

Performance
- Rebalance performance enhancement (100% improvement)

redhat.

ERASURE CODING
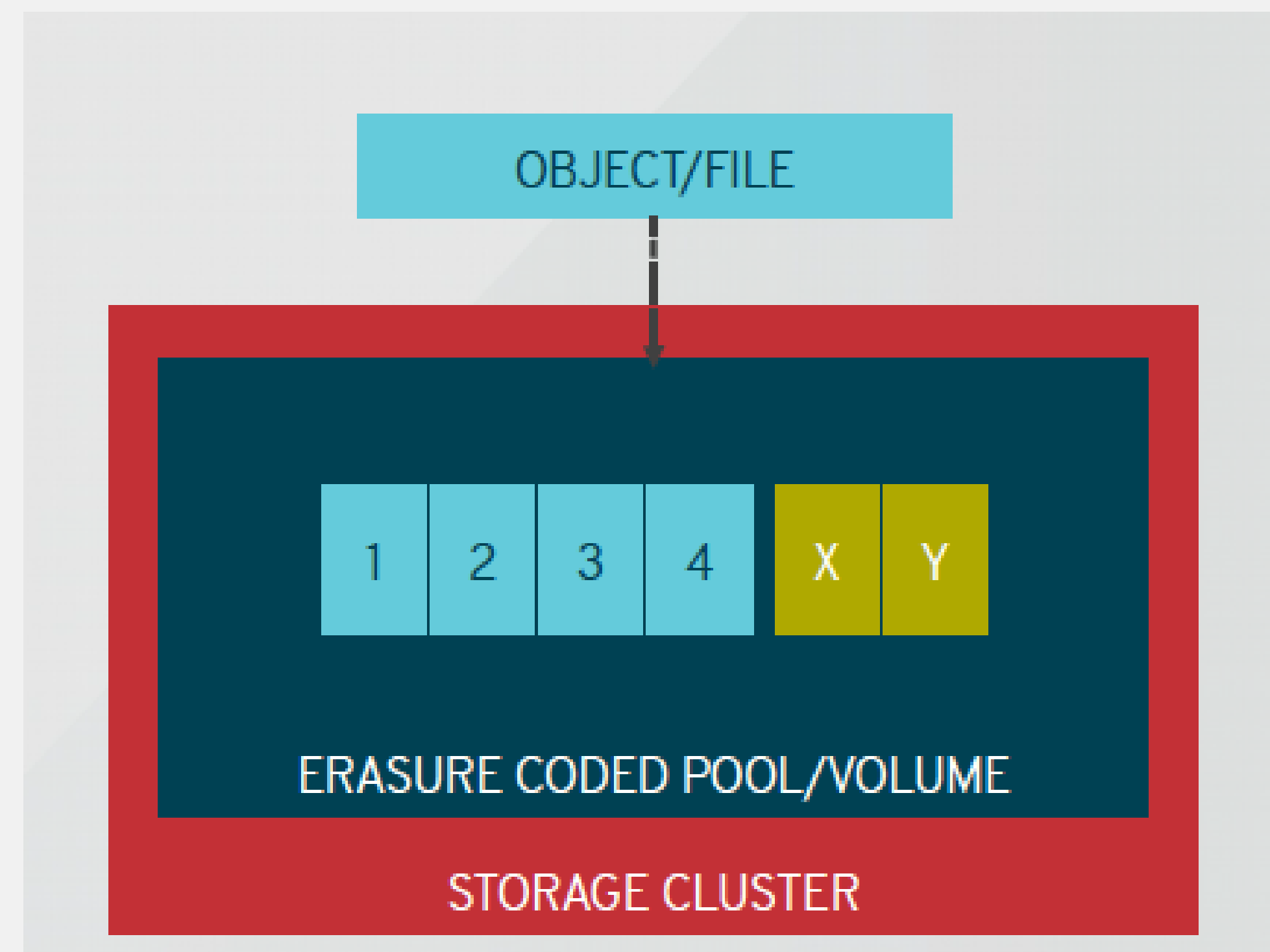
Data protection without using RAID & replication

Break data into smaller fragments, store and recover from a smaller number of fragments
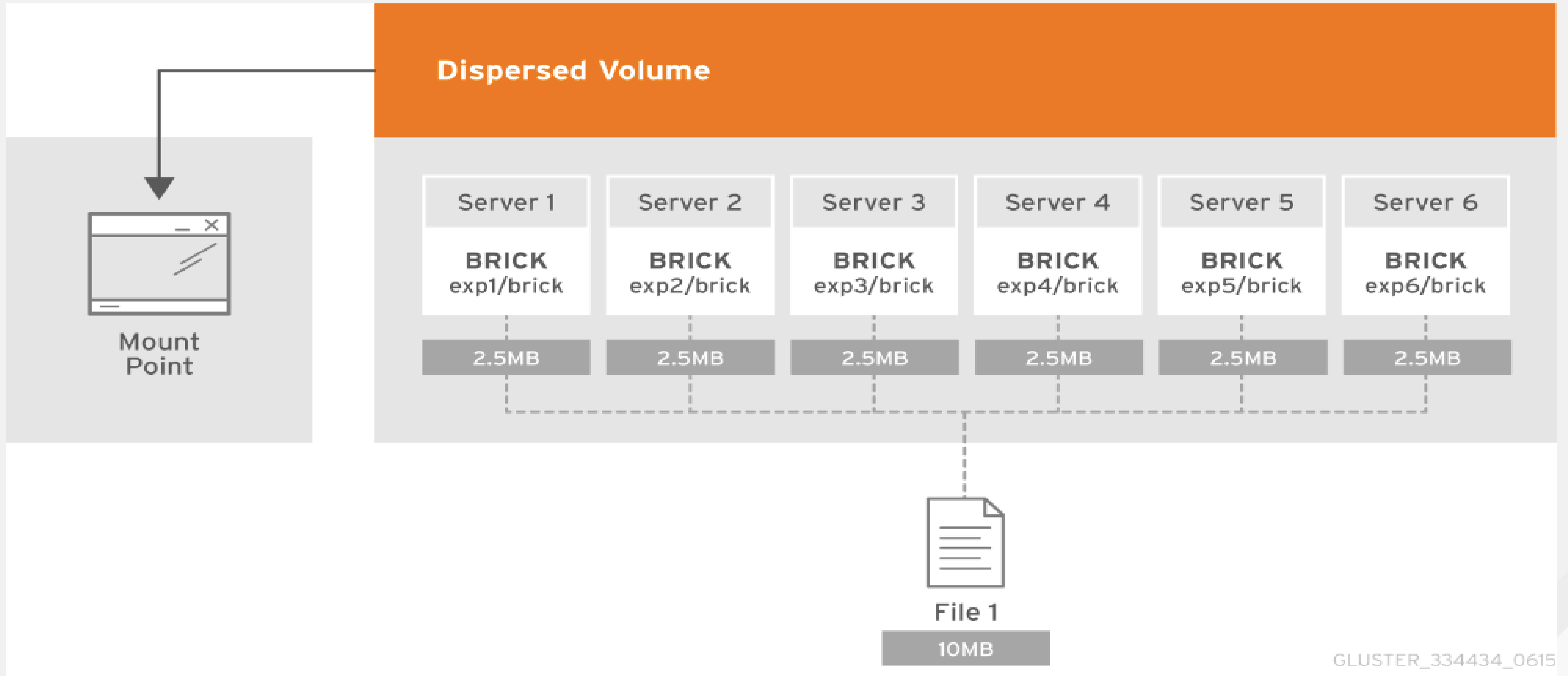
Algorithm used is REED-Solomon

New type of volumes: Dispersed, dist-dispersed

Initial supported configurations: 8+3, 8+4 & 4+2 configuration
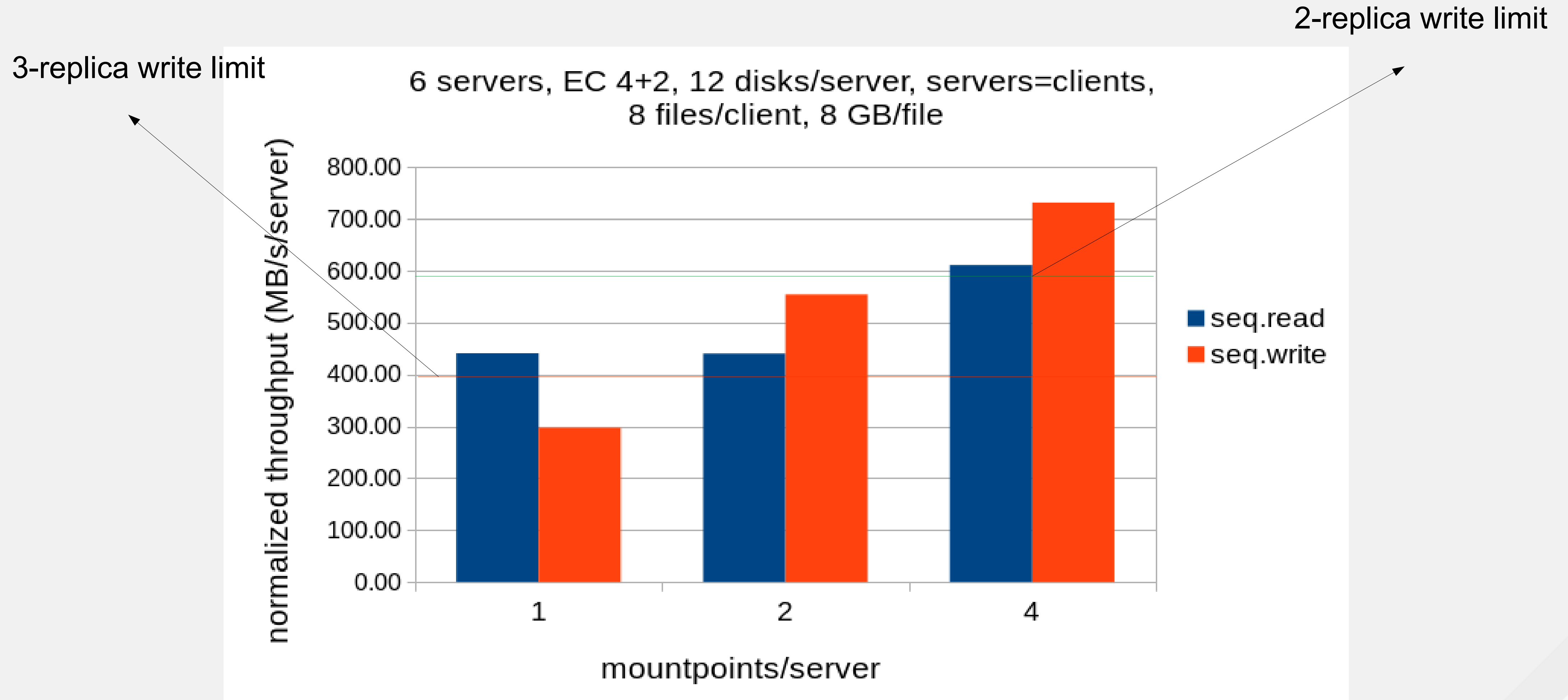
2.4x to 1.37x. ~75% savings

Dispersed Volume

Dispersed Volume

Mount Point

Server 1	Server 2	Server 3	Server 4	Server 5	Server 6
BRICK exp1/brick	**BRICK** exp2/brick	**BRICK** exp3/brick	**BRICK** exp4/brick	**BRICK** exp5/brick	**BRICK** exp6/brick
2.5MB	2.5MB	2.5MB	2.5MB	2.5MB	2.5MB

File 1

10MB

redhat.

SEQUENTIAL IO PERFORMANCE WITH ERASURE CODING

6 servers, EC 4+2, 12 disks/server, servers=clients, 8 files/client, 8 GB/file

2-replica write limit

3-replica write limit

seq.read
seq.write

mountpoints/server

normalized throughput (MB/s/server)

TIERING
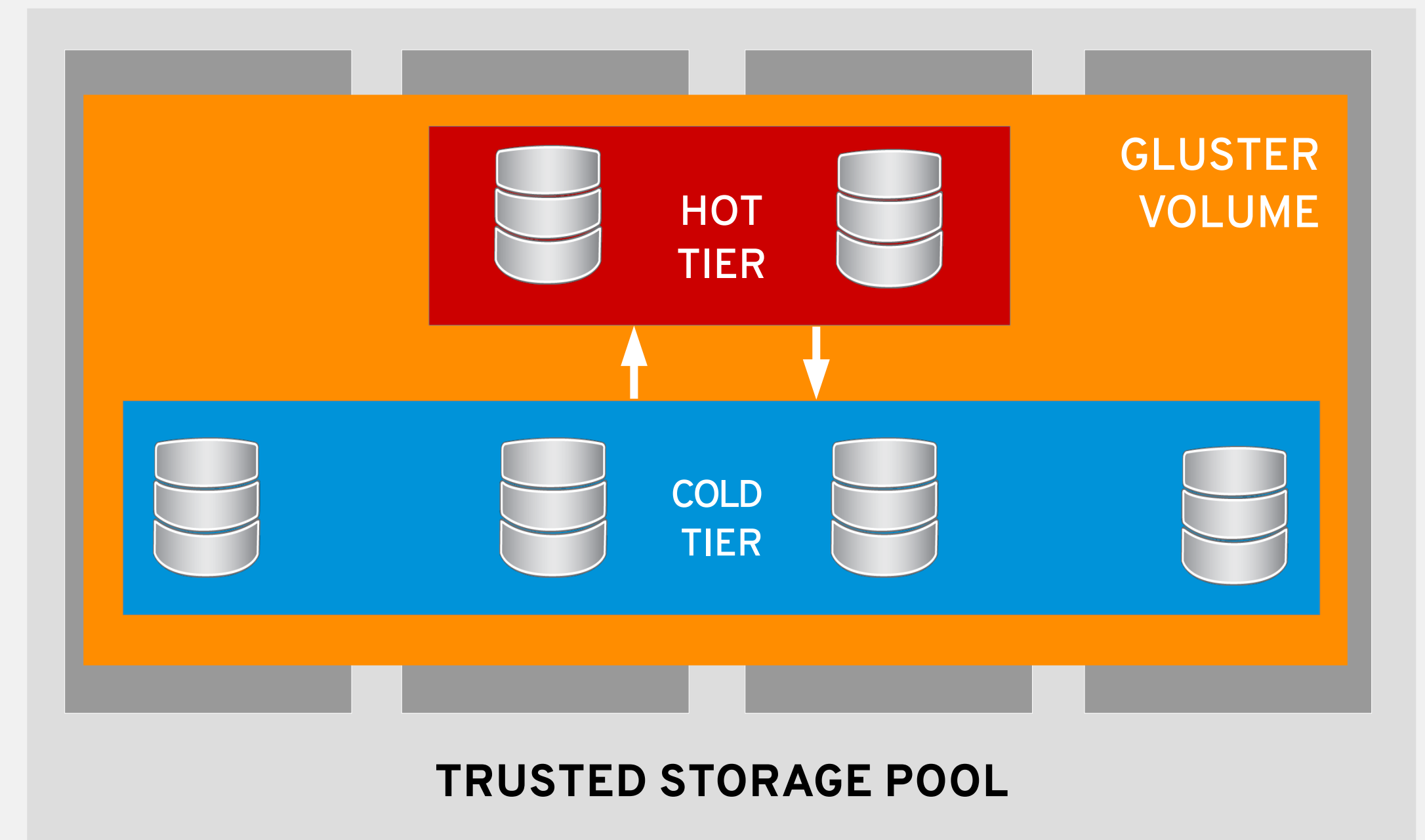
Automated data movement between hot & cold tiers

Movement based on access frequency
- Hot tiers could be SSDs, cold tiers are normal disks (replicated or EC)

Attach & detach a tier to and from an existing Gluster volume

All I/Os forwarded to hot tier

I/O misses promote data to hot tier

GLUSTER VOLUME

HOT TIER

COLD TIER

TRUSTED STORAGE POOL

BIT ROT DETECTION

Protection against "silent data corruption"

Two fundamental procedures
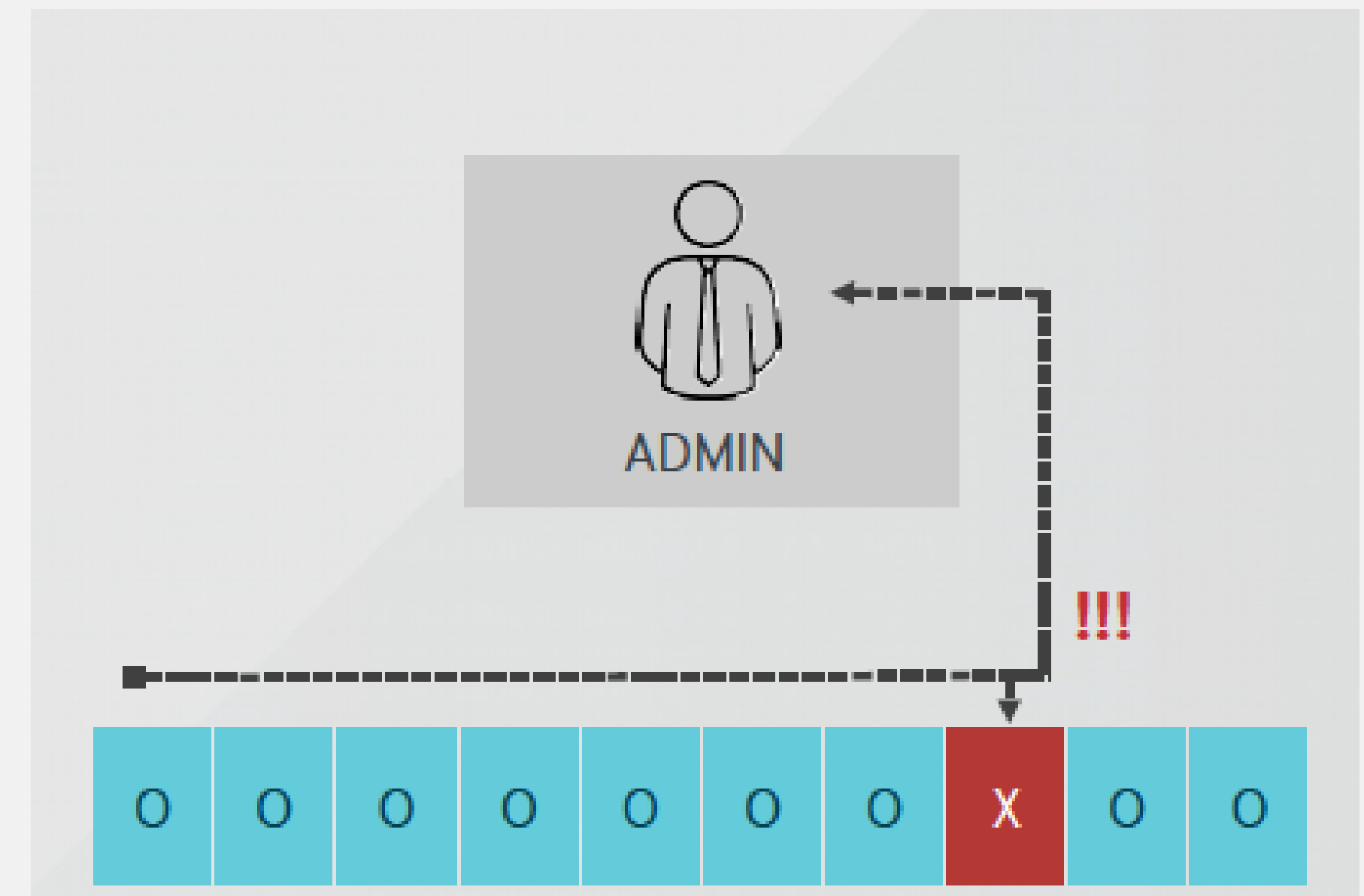- Signing using SHA256
- Scanning/scrubbing for rot

Lazy checksum maintenance
- (not inline to data path)

Checksum calculation when a file is "stable"

Alert/log on mismatched checksums

Scanning mode is admin selectable to control impact

ACTIVE/ACTIVE NFSv4

NFSv4 ACLs – Granular and richer

PseudoFS support – See all exports served out as a single file-system

Dynamic Exports – Add/Delete exports without restarting NFS-ganesha
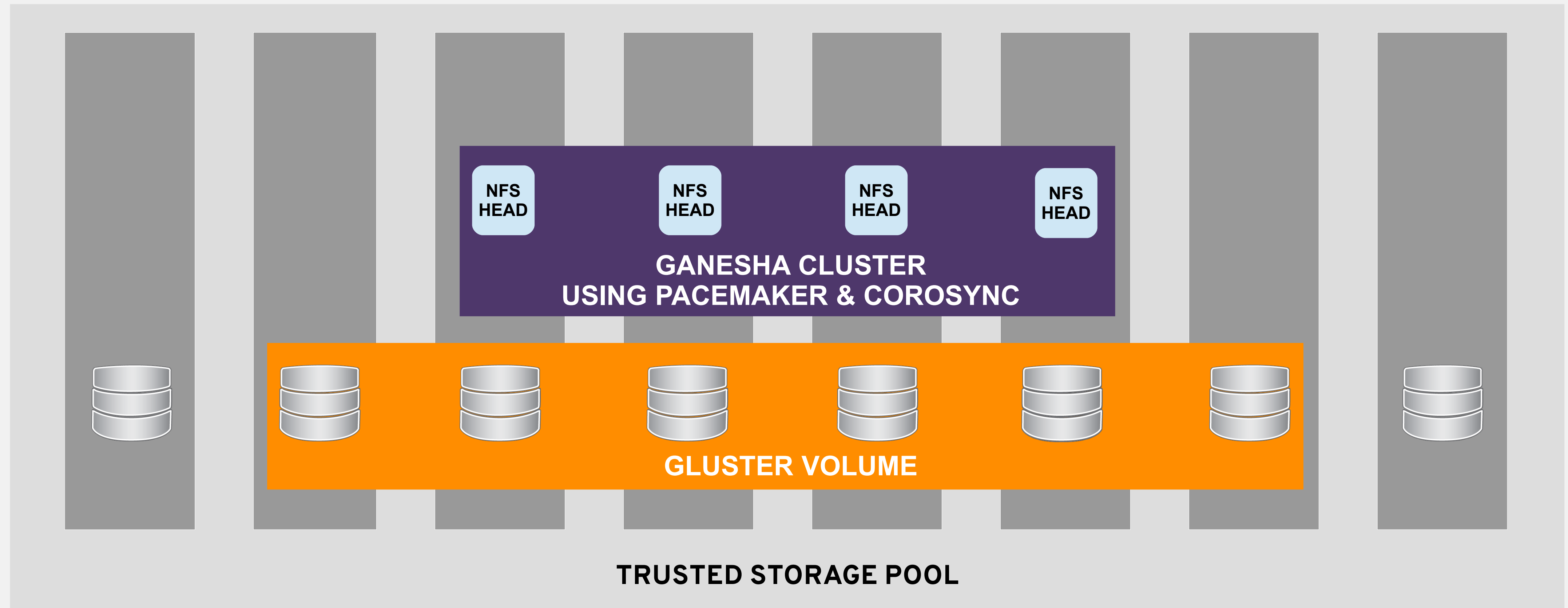
Security
- Kerberos authentication using RPCSEC_GSS, krb5/i/p

Active/Active cluster-on-cluster
- With up to 16 active-active NFS heads
- Gluster storage pool scales out as usual

Delegations to be supported in an update release

redhat.

ACTIVE/ACTIVE NFSv4

NFS HEAD

NFS HEAD

NFS HEAD

NFS HEAD

GANESHA CLUSTER
USING PACEMAKER & COROSYNC

GLUSTER VOLUME

TRUSTED STORAGE POOL

Summarizing RHGS 3.1 (EVERGLADES)

Most feature-rich RHGS release ever!

- Delivering on TCO reduction, data protection & security

Use-Cases

- Active Archives

- Generic File Shares

- Back-end for messaging solutions that need NFSv4

- OpenStack Manila

redhat.

WHAT'S NEXT – THE FUTURE

STORAGE TRENDS:
MODERN IT INFRASTRUCTURES

TRADITIONAL STORAGE	NEXT GENERATION STORAGE
Manual provisioning of LUNs and volumes with some degree of automation	Self-service provisioning by lines of businesses and application developers
Static selection of storage platforms based on application needs	Catalog based storage service offerings with metering & charge-back
Scale-up with some scale-out	Expand, Shrink and scale on demand
Little to no flexibility in selecting optimum storage back-end for workloads	Policy based storage back-end selection

redhat

APPLIED TO GLUSTER

Key elements for modern storage infrastructure (Manila, containers, hyper-converged)

- Consumption Model ("File As A Service" or "NAS on Demand")
  - Dynamic provisioning, healing, tuning & balancing
  - Security & multi-tenancy
  - Cloud scale & stability at scale

- Performance: performant storage back-end for a wide variety of workloads

- Advanced data services: tiering, compression, de-duplication

redhat.

GLUSTER UPSTREAM ROADMAP

GLUSTERFS 4.0 CONTEXT

Gluster 4.0 will be our technology base for the next five years or so
- Based on 3.x experience

Design must be based on estimates of where we'll be in 2021

Higher node counts and more complex networks

Heterogeneous storage
- e.g. NVMe for performance, SMR for capacity

New workloads and usage models
- Hyper-convergence
- Containers
- "XYX as a service" and multi-tenancy

FLEXIBLE STORAGE MANAGEMENT

Declarative and constraint-based
- Not "this brick and this brick and this brick"
- More like "this big, replicated this many times, these features"
- We figure out which combinations match user requirements

Overlapping replica/stripe/erasure sets
- Ease requirement to add bricks in multiples
- Better load distribution during and after failures

Multiple replication levels (and types) within one volume

More sophisticated tiering, rack- or security-aware placement, etc.

redhat.

GLUSTERD CHANGES

More scalable membership protocol

Stronger consistency for configuration data

Improved modularity
- Most-changed code in 3.x
- Increasing complexity and merge conflicts slow down the entire project
- Plugin approach allows independent development of new features

Prerequisite for other 4.0 features

redhat.

PERFORMANCE ENHANCEMENTS

Client-side caching
- Fully consistent via "upcall" mechanism

Third-party copy
- Already part of NFS and SMB protocols

Multiple networks and Quality of Service
- Leverage faster private networks e.g. for replication
- Isolate internal traffic
- Protect tenants from each other

FUTURE FOCUS AREAS

Theme: Storage/File as a Service

Use-Cases: Storage for containers, OpenStack Manila

Technology Enablers:
- Dynamic Provisioning
- At-rest Encryption
- Inode-quotas
- Cloud Scale & Stability at Scale
- Performant back-end for diverse workloads
- Autonomous operations
- Multi-tenancy

redhat.

RED HAT GLUSTER STORAGE ROADMAP

RHGS 3.2 (Fundy)
H1-CY2016

RHGS 4 (Gir)
(In Planning)

Baseline
- GlusterFS 3.8, RHEL 6, RHEL 7

Management
- Dynamic provisioning of volumes

Key Features
- Inode quotas

Protocols
- SMB 3.0 (advanced features)
  - Multi-channel support

Performance
- Rebalance
- Self-heal

Security
- At-rest encryption

Baseline
- GlusterFS 4, RHEL 7

Key Features
- Compression, Deduplication

Core Infrastructure
- Next gen replication
- Highly scalable control plane
- DHTv2

Protocols
- pNFS

Performance
- QoS
- Client side caching

Management
- New UI, Gluster ReST API

RED HAT GLUSTER STROAGE INTEGRATION ROADMAP

OPENSTACK MANILA

Tech preview level support for RHELOSP 7 (Kilo)
- Create/delete/rename/list share
- Create/delete snapshots
- Allow/deny access to shares
- OSP Director integration planned for September release (ver 1.1)

Full support expected in RHELOSP 8 (Liberty)
- Create/delete share dynamically
- Create share from snapshot
- Exploring integration with Barbican for managing certificates

CONTAINERS

2-key storage use-cases
- Persistent data store for containers
- Container image registries

Focused on "Persistent data store for containers" use-case
- Containerized applications mount Gluster as their data store
  - Separate compute & storage pools
  - Hyper-Converged

- NFS or GlusterFS native client integration in kubernetes

Key attributes that makes Gluster interesting
- Not impacted my mount-storm
- Built in HA
- Lends itself well to hyper-converged environments

redhat.

RED HAT SUMMIT

LEARN. NETWORK.
EXPERIENCE OPEN SOURCE.

#redhat #rhsummit

redhat.