redhat.

RED HAT GLUSTER STORAGE PRODUCT OVERVIEW

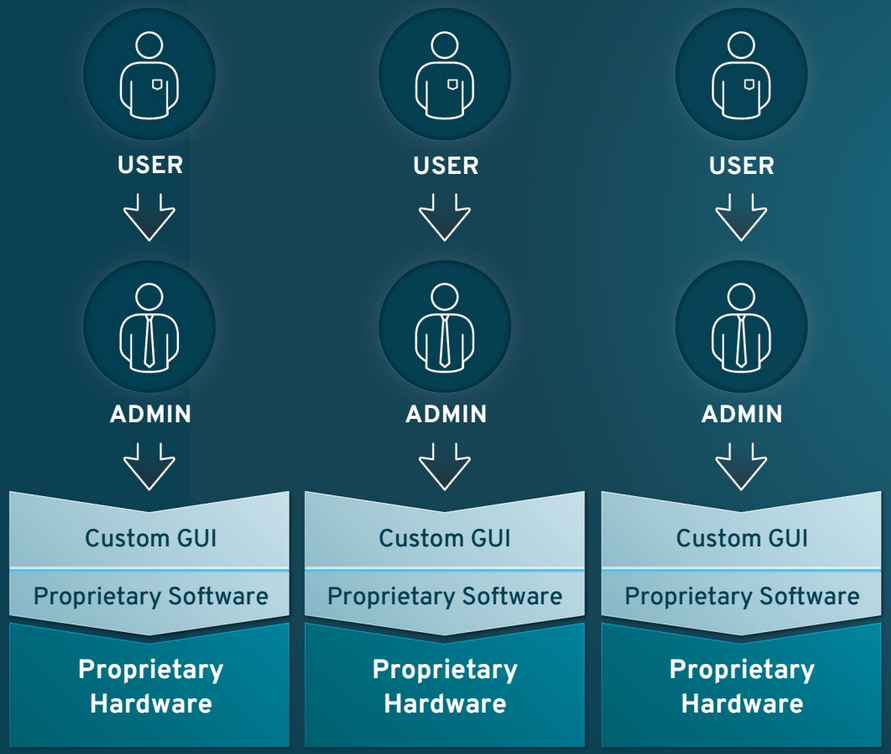
Michael Lessard
Senior Solutions Architect
October 2015

THE RED HAT STORAGE MISSION

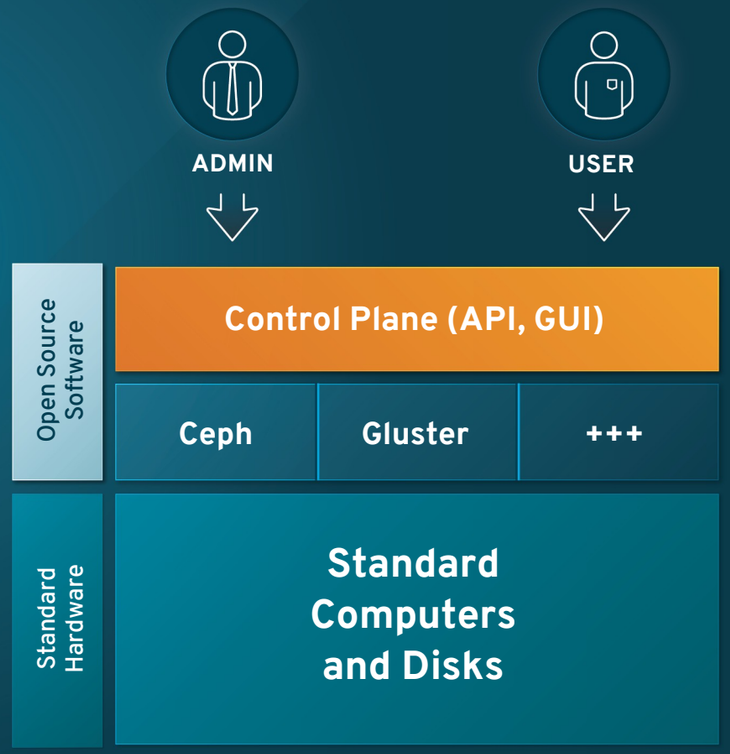
To offer a unified, open software-defined storage portfolio that delivers a range of data services for next generation workloads thereby accelerating the transition to modern IT infrastructures.

THE FUTURE OF STORAGE

Traditional Storage
Complex proprietary silos



Open, Software-Defined Storage
Standardized, unified, open platforms



WHY BOTHER?

PROPRIETARY HARDWARE	**Common, off-the-shelf hardware** Lower cost, standardized supply chain
SCALE-UP ARCHITECTURE	**Scale-out architecture** Increased operational flexibility
HARDWARE-BASED INTELLIGENCE	**Software-based intelligence** More programmability, agility, and control
CLOSED DEVELOPMENT PROCESS	**Open development process** More flexible, well-integrated technology

A RISING TIDE

Software-Defined Storage is leading a shift in the global storage industry, with far-reaching effects.

"By 2016, server-based storage solutions will lower storage hardware costs by 50% or more."

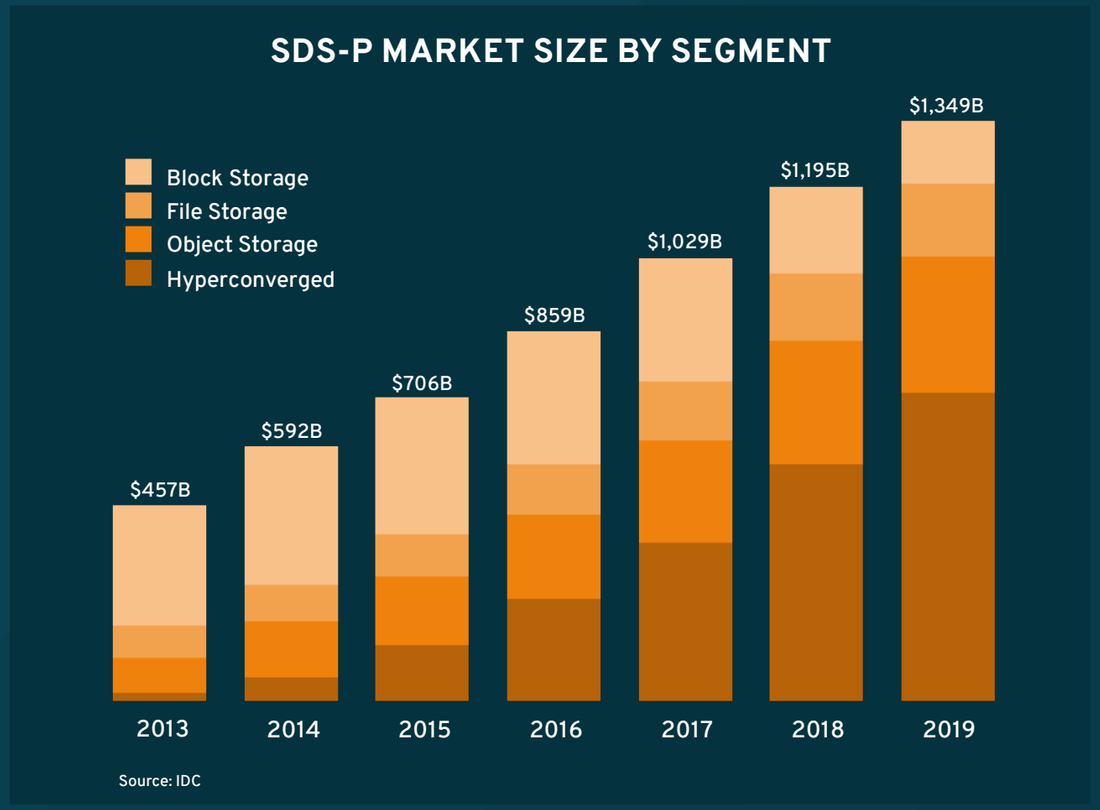
Gartner: "IT Leaders Can Benefit From Disruptive Innovation in the Storage Industry"

"By 2020, between 70-80% of unstructured data will be held on lower-cost storage managed by SDS environments."

Innovation Insight: Separating Hype From Hope for Software-Defined Storage

"By 2019, 70% of existing storage array products will also be available as software only versions"

Innovation Insight: Separating Hype From Hope for Software-Defined Storage

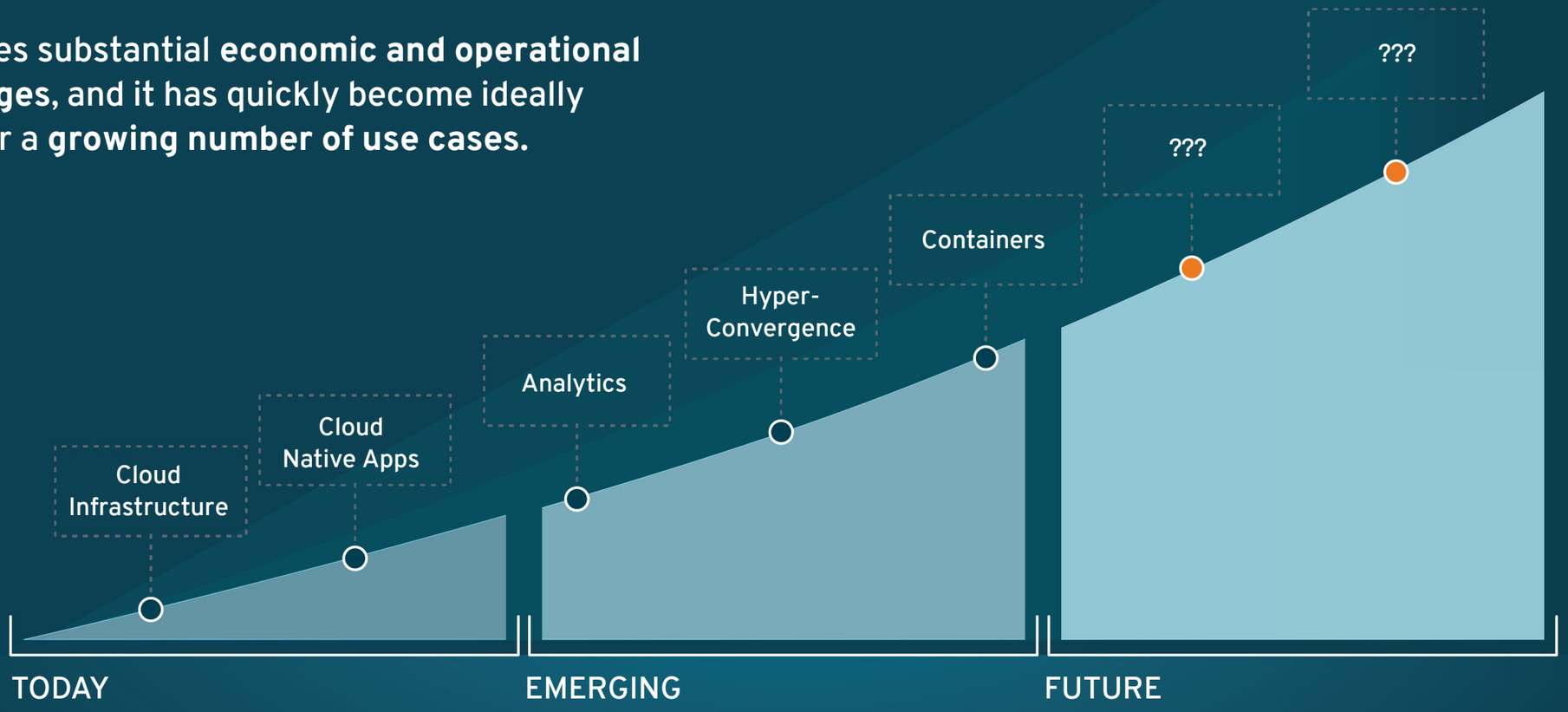


Market size is projected to **increase approximately 20% year-over-year** between 2015 and 2019.

THE JOURNEY

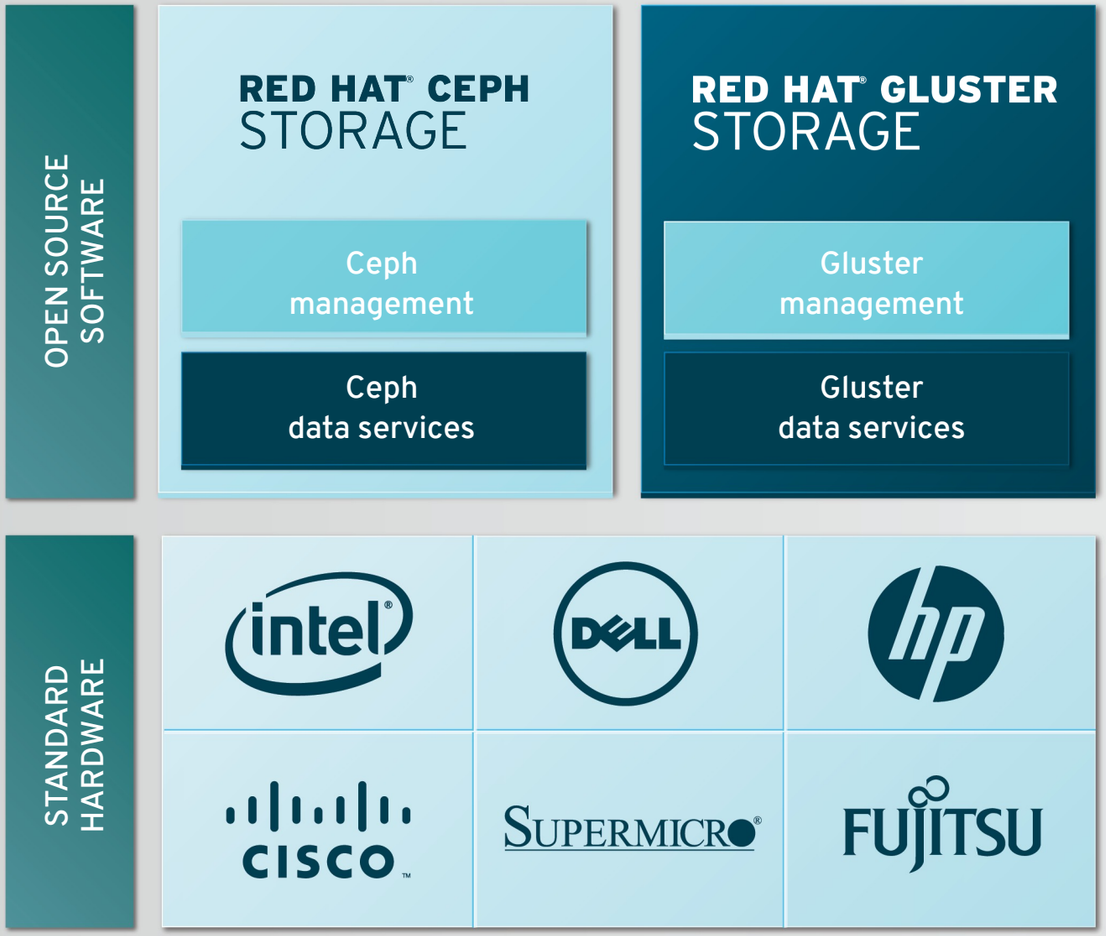
Open Software-Defined Storage is a fundamental reimagining of **how storage infrastructure works.**

It provides substantial **economic and operational advantages**, and it has quickly become ideally suited for a **growing number of use cases.**




THE RED HAT STORAGE
PORTFOLIO

THE RED HAT STORAGE PORTFOLIO



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

GROWING INNOVATION COMMUNITIES

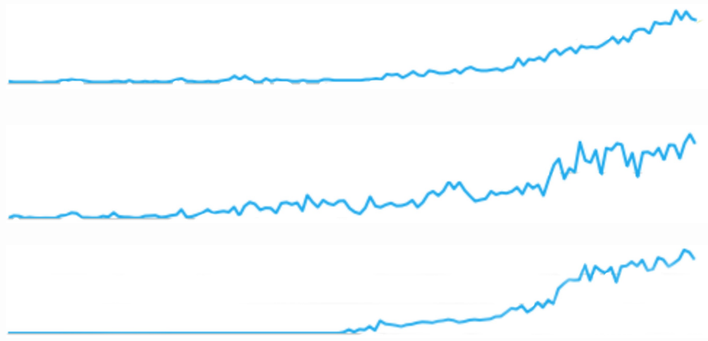
ceph

- Contributions from **Intel, SanDisk, SUSE,and DTAG**.
- Presenting Ceph Days in cities around the world and quarterly virtual Ceph Developer Summit events.

78 AUTHORS/mo

1500 COMMITS/mo

258 POSTERS/mo



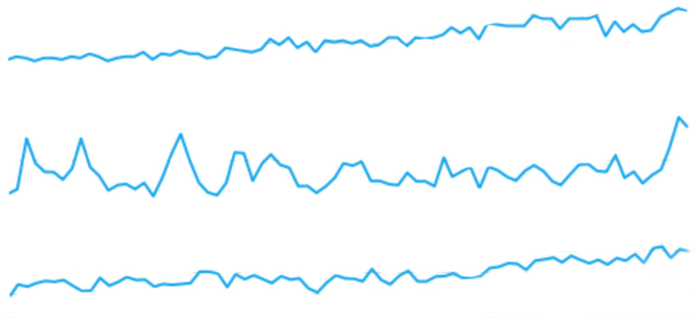
GLUSTER COMMUNITY

- Over **11M downloads** in the last 12 months
- Increased development velocity, authorship, and discussion has resulted in rapid feature expansion.

41 AUTHORS/mo

259 COMMITS/mo

166 POSTERS/mo



REFINEMENTS FOR PETABYTE-SCALE OPERATORS

RED HAT® CEPH STORAGE

Optimized for large-scale deployments

Version 1.3 of Red Hat Ceph Storage is the first major release since joining the Red Hat Storage product portfolio, and incorporates feedback from customers who have deployed in production at large scale.

Areas of improvement:

- Robustness at scale
- Performance tuning
- Operational efficiency

RED HAT® GLUSTER STORAGE

Enhanced for flexibility and performance

Version 3.1 of Red Hat Gluster Storage contains many new features and capabilities aimed to bolster data protection, performance, security, and client compatibility.

New capabilities include:

- Erasure coding
- Tiering
- Bit Rot detection
- NVSv4 client support


RED HAT GLUSTER STORAGE

OVERVIEW:
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

intuit

Customer Highlight: Intuit

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

RED HAT® GLUSTER STORAGE

TARGET USE CASES

Analytics

- Machine analytics with Splunk
- Big data analytics with Hadoop

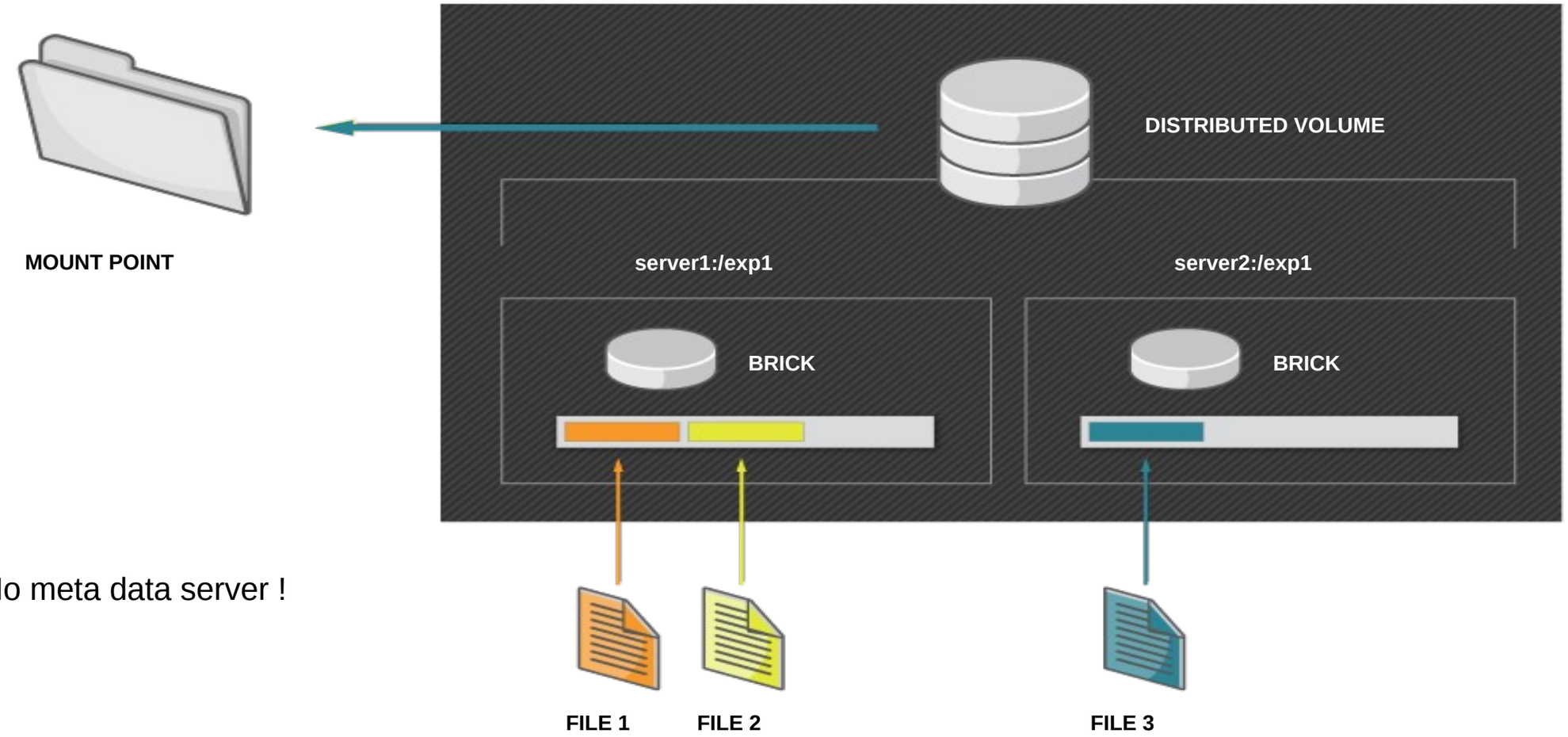
Enterprise File Sharing

- Media Streaming
- Active Archives

Enterprise Virtualization

Rich Media & Archival

RED HAT STORAGE USER PERSPECTIVE (DISTRIBUTED VOLUMES)



** No meta data server !

NODE PREPARATION

1 – Download RHGS iso from the customer portal then install RHGS on nodes

2 – Attach a disk (or use a separated partition) for gluster data to the servers

```
[root@rhgluster-node1 ~]#  systemctl status glusterd
glusterd.service - GlusterFS, a clustered file-system server
   Loaded: loaded (/usr/lib/systemd/system/glusterd.service; enabled)
   Active: active (running) since Mon 2015-09-28 12:12:28 MDT; 17min ago
  Process: 1309 ExecStart=/usr/sbin/glusterd -p /var/run/glusterd.pid (code=exited, status=0/SUCCESS)
 Main PID: 1329 (glusterd)
   CGroup: /system.slice/glusterd.service
           └─1329 /usr/sbin/glusterd -p /var/run/glusterd.pid
Sep 28 12:12:20 rhgluster-node1.mlc.dom systemd[1]: Starting GlusterFS, a clustered file-system server...
Sep 28 12:12:28 rhgluster-node1.mlc.dom systemd[1]: Started GlusterFS, a clustered file-system server.

[root@rhgluster-node1 ~]#  fdisk /dev/sdb
[root@rhgluster-node1 ~]#  mk.xfs -i size=512 /dev/sdb1
[root@rhgluster-node1 ~]#  mkdir /data
[root@rhgluster-node1 ~]#  vi /etc/fstab
 /dev/sdb1    /data     xfs    defaults  0  0
 [root@rhgluster-node1 ~]#  mount /data

# repeat these steps for each node
```

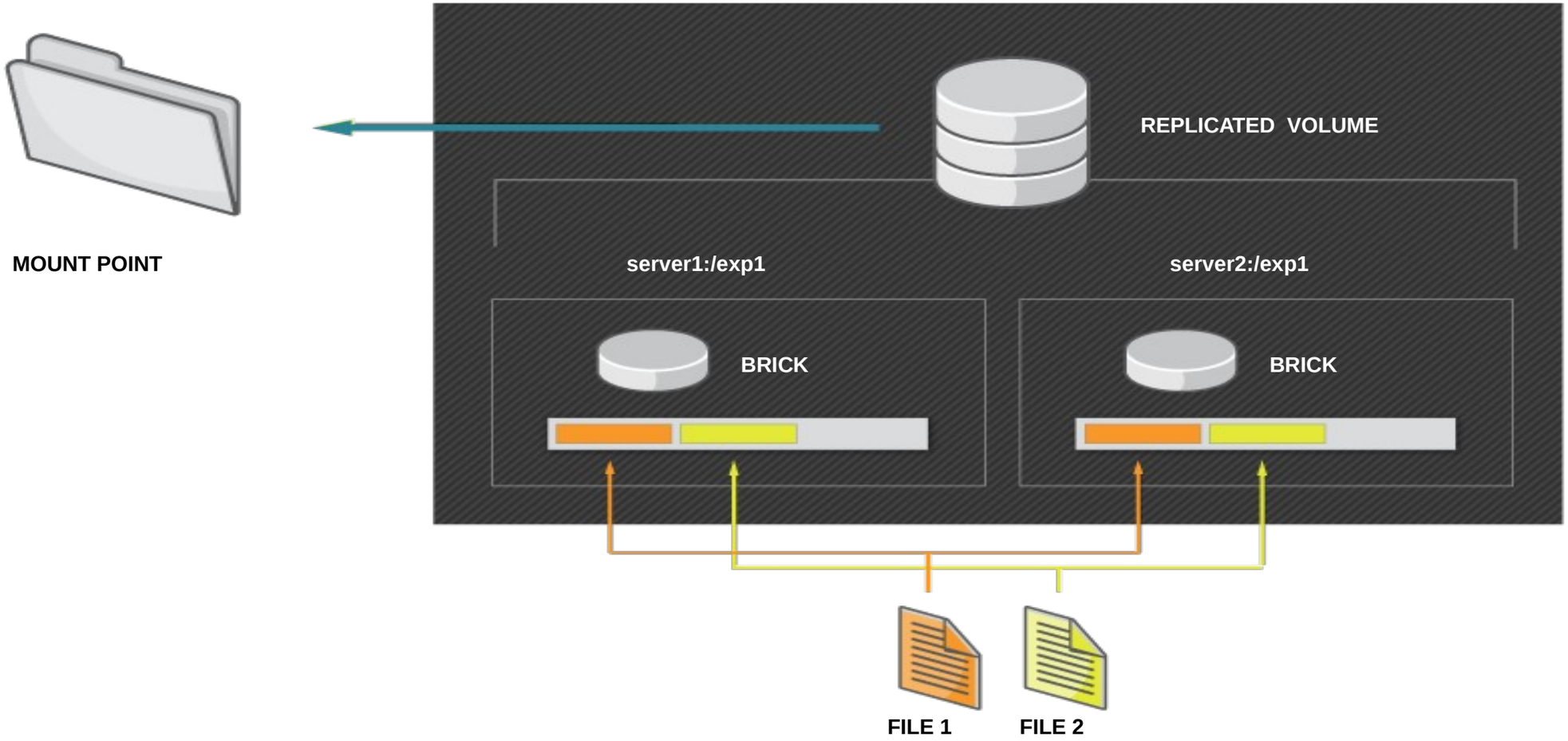
DISTRIBUTED CLUSTER

```
[root@rhgluster-node1 ~]#  gluster peer probe rhgluster-node2.mlc.dom
peer probe: success.

[root@rhgluster-node1 ~]# gluster peer status

Hostname: rhgluster-node2.mlc.dom
Uuid: f86a00c6-eb7f-430c-b0eb-26bf87a81094
State: Peer in Cluster (Connected)

[root@rhgluster-node1 ~]# gluster volume create distr1 rhgluster-node1.mlc.dom:/data/d1 rhgluster-
node2.mlc.dom:/data/d1
volume create: distr1: success: please start the volume to access data

 [root@rhgluster-node1 ~]# gluster volume info
Volume Name: dist1
Type: Distribute
Volume ID: 691e213e-c497-4fa2-9475-ecf70ce12213
Status: Created
Number of Bricks: 2
Transport-type: tcp
Bricks:
Brick1: gluster-node1.mlc.dom:/data/d1
Brick2: gluster-node2.mlc.dom:/data/d2
Options Reconfigured:
performance.readdir-ahead: on
```

DISTRIBUTED CLUSTER
Mount the filesystem from a remote computer (RHEL 7)

```
[root@laptop ~]# subscription-manager repos --enable=rhel-7-server-rh-common-rpms
[root@laptop ~]# yum install glusterfs glusterfs-fuse
[root@laptop ~]# mkdir /mnt/gluster
[root@laptop ~]#  mount -t glusterfs gluster-node1.mlc.dom:/distr1 /mnt/gluster
[root@laptop ~]#  touch 11 /mnt/gluster ; touch 22 /mnt/gluster


[root@rhgluster-node1 ~]#  ls /data/d1
11
[root@rhgluster-node2 ~]#  ls /data/d1
22
```

RED HAT STORAGE USER PERSPECTIVE (REPLICATED VOLUMES)



REPLICATED CLUSTER

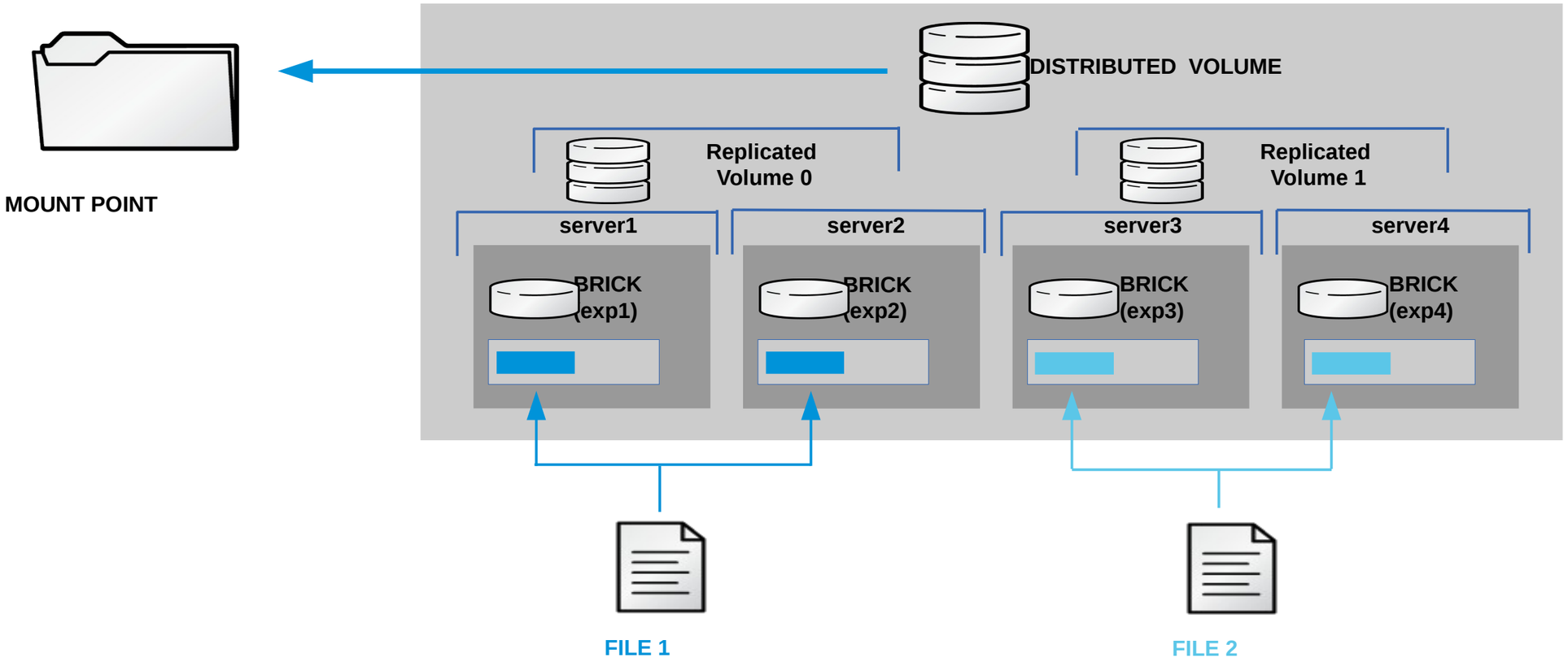
```
[root@rhgluster-node1 ~]#  mkdir /data/r1
[root@rhgluster-node2 ~]#  mkdir /data/r1

[root@rhgluster-node1 ~]#  gluster volume create rep1 replica 2 gluster-node1.mlc.dom:/data/r1 gluster-node2.mlc.dom:/data/r1

[root@rhgluster-node1 ~]#  gluster volume info

Volume Name: replicated
Type: Replicate
Volume ID: 841eccf1-abdc-4226-b7c3-81f449bd34d9
Status: Started
Number of Bricks: 1 x 2 = 2
Transport-type: tcp
Bricks:
Brick1: gluster-node1.mlc.dom:/data/gluster2
Brick2: gluster-node2.mlc.dom:/data/gluster2
Options Reconfigured:
changelog.changelog: on
geo-replication.ignore-pid-check: on
geo-replication.indexing: on
performance.readdir-ahead: on
```

REPLICATED CLUSTER
Mount the filesystem from a remote computer

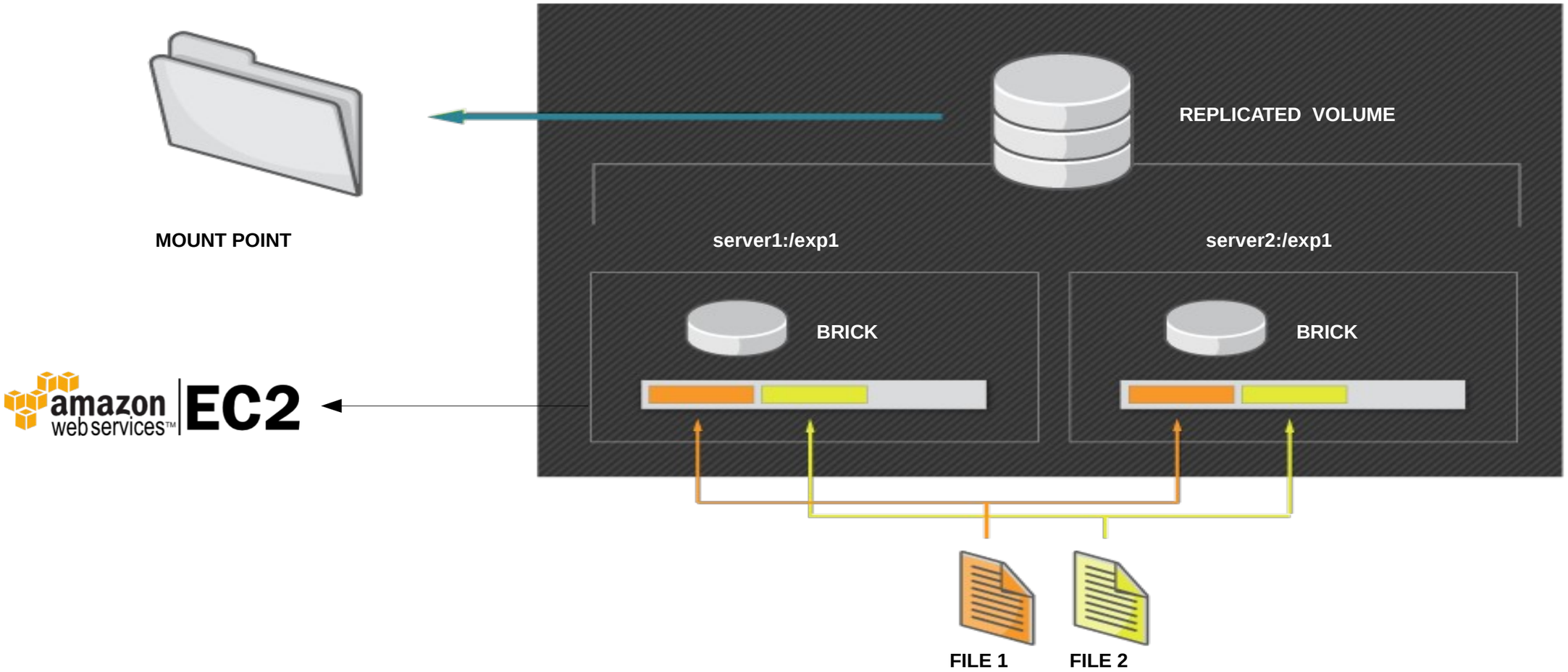
```
[root@laptop ~]# mkdir /mnt/gluster2
[root@laptop ~]#  mount -t glusterfs gluster-node1.mlc.dom:/rep1 /mnt/gluster2
[root@laptop ~]#  touch 11 /mnt/gluster2 ; touch 22 /mnt/gluster2


[root@rhgluster-node1 ~]#  ls /data/d1
11 22
[root@rhgluster-node1 ~]#  ls /data/d1
22 22
```

Fault-tolerant data placement (distributed replicated volume)

Creates a fault-tolerant distributed volume by mirroring the same file across 2 bricks



RED HAT STORAGE USER PERSPECTIVE
REPLICATION – DISASTER RECOVERY
GEO-REPLICATION



AMAZON NODE PREPARATION
1 – You need to transfer RH Gluster subscription in Amazon (http://www.redhat.com/en/technologies/cloud-computing/cloud-access)
2- Click on Images-Amis – Private Images and serch for RHGS
3 – Lunch an instance of RHGS 3.1 update 1

```
[root@rhgluster-node1 ~]#  vi /etc/hosts
52.23.169.60 rhgluster-amazon3.mlc.dom

[root@rhgluster-node2 ~]#  vi /etc/hosts
52.23.169.60 rhgluster-amazon3.mlc.dom

[root@laptop ~]# ssh -i laptop.pem ec2-user@52.23.169.60
[ec2-user@ip-172-30-0-148 ~]$ sudo su
[root@ip-172-30-0-148 ec2-user]# vi /etc/ssh/sshd_config
PasswordAuthentication yes
[root@ip-172-30-0-148 ec2-user]# systemctl restart sshd
[root@ip-172-30-0-148 ec2-user]# passwd

[root@ip-172-30-0-148 ec2-user]#  gluster volume create amazon rhgluster-amazon3.mlc.dom:/data force
[root@ip-172-30-0-148 ec2-user]#  gluster volume start amazon
Volume start: amazon: success
```

GLUSTER GEO-REPLICATION
DR IN AMAZON

```
[root@rhgluster-node1 ~]#  ssh-keygen
[root@rhgluster-node1 ~]#  ssh-copy-id -i .ssh/is_rsa.pub 52.23.169.60

[root@rhgluster-node1 ~]# systemctl start ntpd ; systemctl enable ntpd
[root@rhgluster-node2 ~]# systemctl start ntpd ; systemctl enable ntpd
[root@ip-172-30-0-148 ec2-user]# systemctl start ntpd ; systemctl enable ntpd

[root@rhgluster-node1 ~]# gluster system: execute gsec_create
[root@rhgluster-node1 ~]# gluster volume geo-replication replicated gluster-amazon1.mlc.dom::amazondata /
create push-pem
Creating geo-replication session between r1 & gluster-amazon1.mlc.dom::amazondata has been successful
[root@rhgluster-node1 ~]# gluster volume geo-replication replicated gluster-amazon1.mlc.dom::amazondata start
Starting geo-replication session between r1 & gluster-amazon1.mlc.dom::amazondata has been successful
[root@rhgluster-node1 ~]# gluster volume geo-replication replicated gluster-amazon1.mlc.dom::amazondata
status

MASTER NODE           MASTER VOL    MASTER BRICK    SLAVE USER    SLAVE                          SLAVE NODE              STATUS     CRAWL STATUS
LAST_SYNCED
-------------------------------------------------------------------------------------------------------------------------------------------------------------------
gluster-node1.mlc.dom    r1          /data/r1        root          gluster-amazon1.mlc.dom::amazondata    gluster-amazon2.mlc.dom    Active     Hybrid Crawl    N/A
gluster-node2.mlc.dom    r1          /data/r2        root          gluster-amazon1.mlc.dom::amazondata    gluster-amazon1.mlc.dom    Passive    N/A             N/A
```

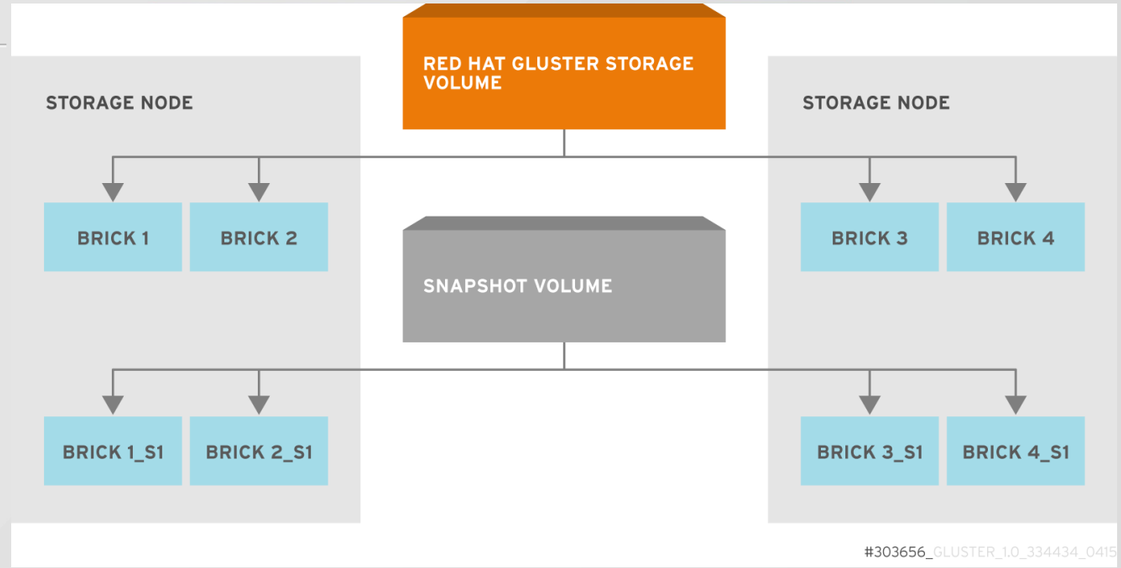
GLUSTER FEATURES

SNAPSHOT

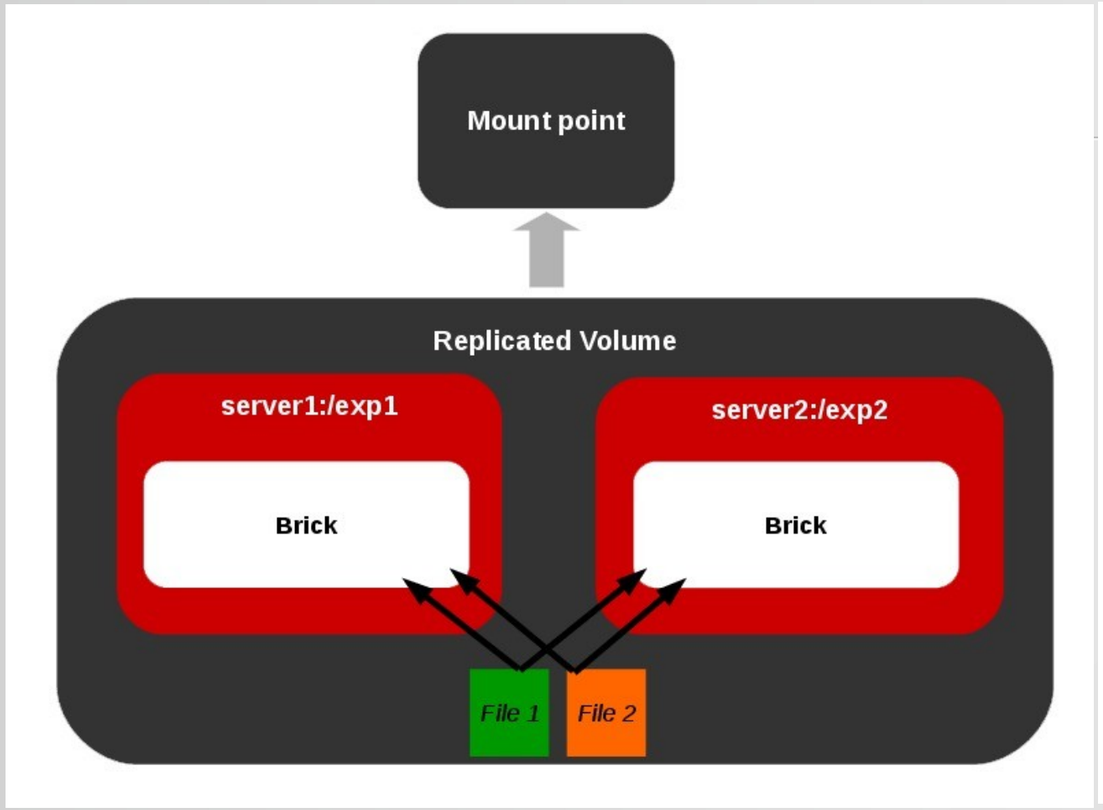
Gluster's snapshot feature

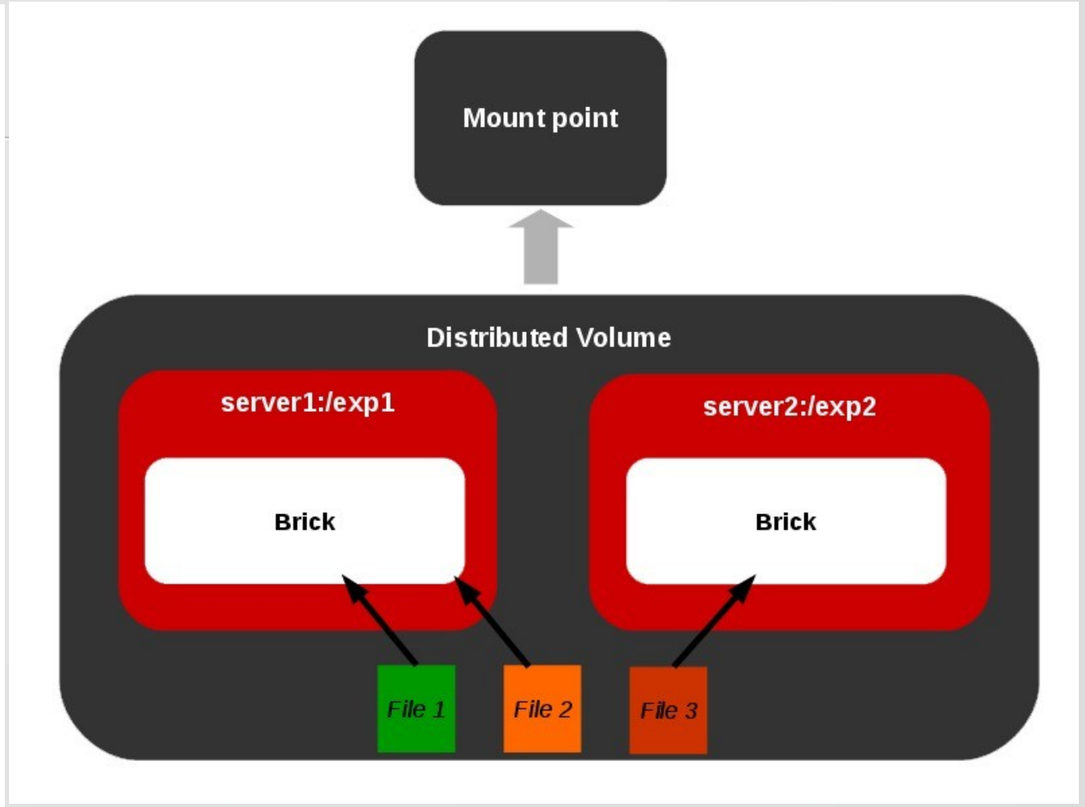
Gluster volume snapshot will provide point-in-time copy of a GlusterFS volume. This snapshot is an online-snapshot therefore file-system and its associated data continue to be available for the clients, while the snapshot is being taken.

Snapshot of a GlusterFS volume will create another read-only volume which will be a point-in-time copy of the original volume. Users can use this read-only volume to recover any file(s) they want. Snapshot will also provide restore feature which will help the user to recover an entire volume.



SELF HEALING + REBALANCE





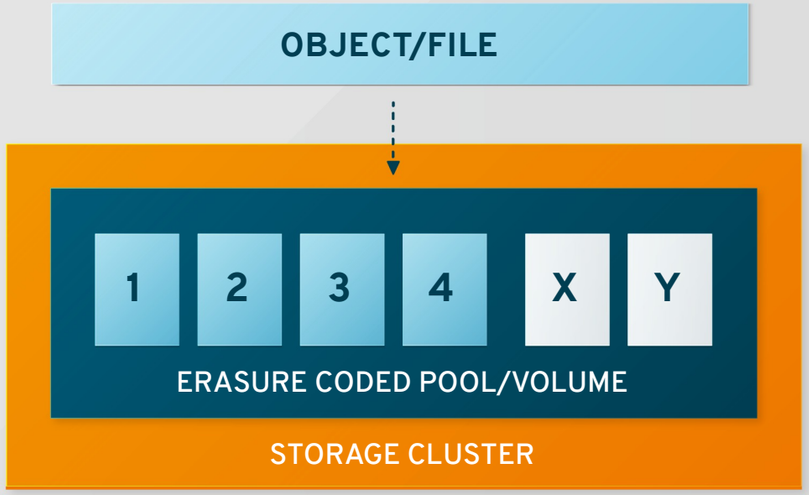
ERASURE CODING (dispersed volume)

Storing more data with less hardware

Standard replicated back-ends are very durable, and can recover very quickly, but they have an inherently large capacity overhead.

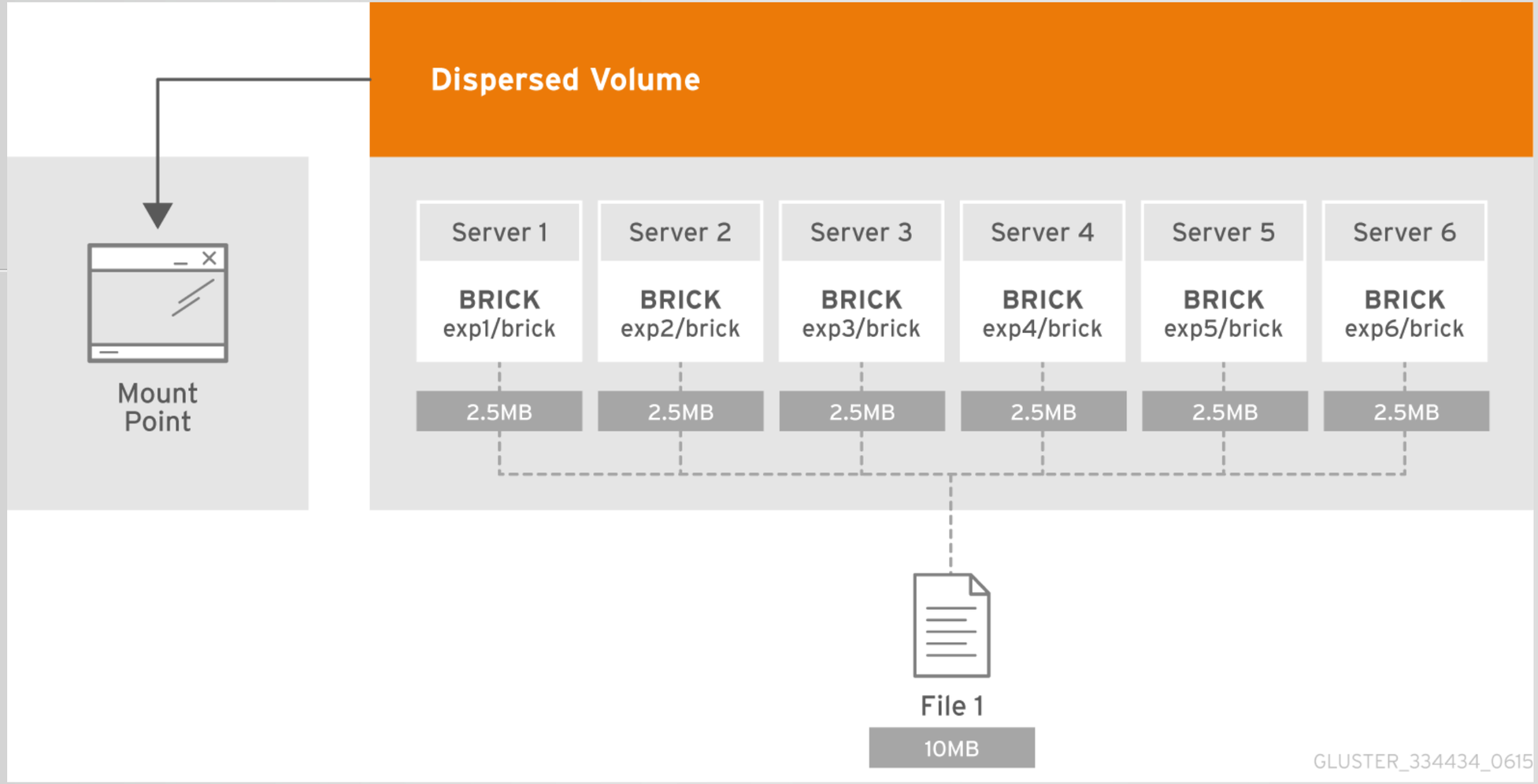
Erasure coding back-ends reconstruct corrupted or lost data by using information about the data stored elsewhere in the system.

Providing failure protection with erasure coding eliminates the need for RAID, consumes far less space than replication, and can be appropriate for capacity-optimized use cases.



Up to 75% reduction in TCO

RED HAT®
GLUSTER STORAGE


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
GLUSTER_334434_0615

ERASURE CODING

```
[root@rhgluster-node1 ~]#  gluster volume create dispers1 disperse-data 4 redundancy 2 transport tcp
node1:/data/dp1 node2:/data/dp1 node3:/data/dp1 node4:/data/dp1 node5:/data/dp1 node6:/data/dp1
Creation of dispers1 has been successful
Please start the volume to access data.

[root@rhgluster-node1 ~]#  gluster volume start dispers1
Starting test-volume has been successful
```

TIERING

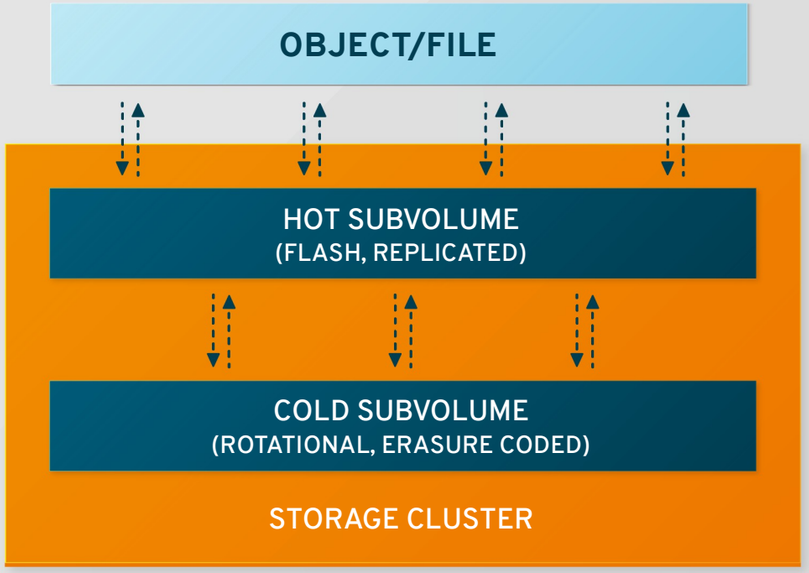
Cost-effective flash acceleration

Optimally, infrequently accessed data is served from less expensive storage systems while frequently accessed data can be served from faster, more expensive ones.

However, manually moving data between storage tiers can be time-consuming and expensive.

Red Hat Gluster Storage 3.1 now supports automated promotion and demotion of data between “hot” and “cold” sub volumes based on frequency of access.

***Tech preview



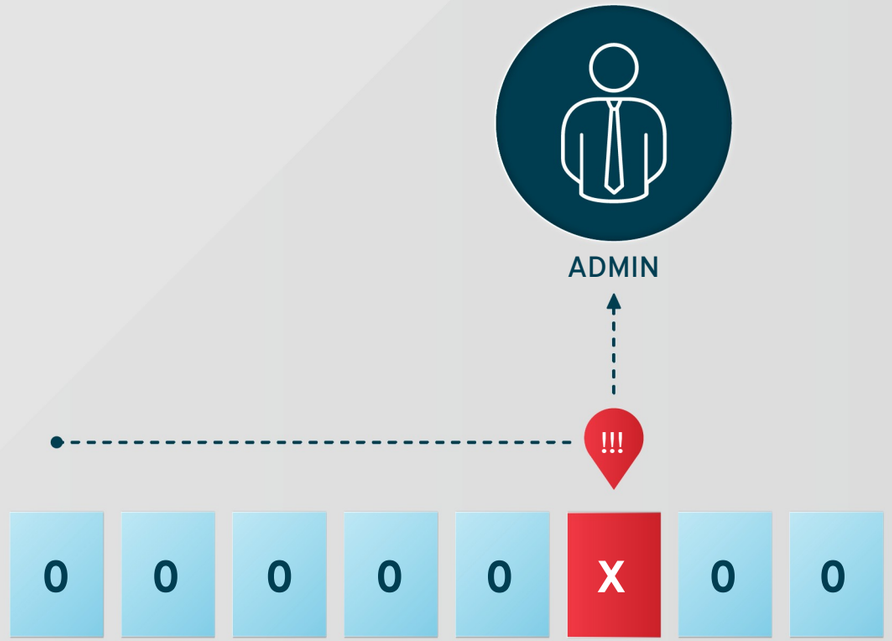
BIT ROT DETECTION

Detection of silent data corruption

Bit rot detection is a mechanism that detects data corruption resulting from silent hardware failures, leading to deterioration in performance and integrity.

Red Hat Gluster Storage 3.1 provides a mechanism to scan data periodically and detect bit-rot.

Using the SHA256 algorithm, checksums are computed when files are accessed and compared against previously stored values. If they do not match, an error is logged for the storage admin.



ENABLE BITROT

```
[root@rhgluster-node1 ~]#  l gluster volume bitrot r1 enable

[root@rhgluster-node1 ~]#  gluster volume bitrot r1 scrub-throttle lazy|normal|aggressive
[root@rhgluster-node1 ~]#  gluster volume bitrot r1 scrub-frequency daily|weekly|biweekly|monthly
```

ROADMAP DETAIL

ROADMAP: RED HAT GLUSTER STORAGE

TODAY (v3.1)

CORE
- Erasure Coding
- Tiering
- Bit Rot Detection
- Snap Schedule

FILE
- Active/Active NFSv4
- SMB 3 (basic subset)

SEC
- SELinux
- SSL encryption (in-flight)

MGMT
- Device Management
- Geo-Replication, Snapshots
- Dashboard

Gluster 3.7, RHEL 6, 7

V3.2 (H1-2016)

CORE
- Inode quotas
- Faster Self-heal
- Controlled Rebalance

FILE
- SMB 3 (advanced features)
- Multi-channel

SEC
- At-rest encryption

MGMT
- Dynamic provisioning of resources

Gluster 3.8, RHEL 6, 7

FUTURE (v4.0 and beyond)

CORE
- Compression
- Deduplication (Yadl)
- Highly scalable control plane
- Next-gen replication/distribution

FILE
- pNFS

PERF
- QoS
- Client side caching

MGMT
- New UI
- Gluster REST API

Gluster 4, RHEL 7

DETAIL:
RED HAT GLUSTER STORAGE 3.1

These features were introduced in the most recent release of Red Hat Gluster Storage, and are now supported by Red Hat.

MGMT	**Device management, dashboard**	Support in the console for discovery, format, and creation of bricks based on recommended best practices; an improved dashboard that shows vital statistics of pools.
MGMT	**Snapshots, Geo-replication**	New support in the console for snapshotting and geo-replication features.
CORE	**Tiering**	New features to allow creation of a tier of fast media (SSDs, Flash) that accompanies slower media, supporting policy-based movement of data between tiers and enhancing create/read/write performance for many small file workloads.
CORE	**Bit rot detection**	Ability to detect silent data corruption in files via signing and scrubbing, enabling long term retention and archival of data without fear of “bit rot”.
CORE	**Snapshot scheduling**	Ability to schedule periodic execution of snapshots easily, without the complexity of custom automation scripts.
CORE	**Backup hooks**	Features that enable incremental, efficient backup of volumes using standard commercial backup tools, providing time-savings over full-volume backups.
CORE	**Erasure coding**	Introduction of erasure coded volumes (dispersed volumes) that provide cost-effective durability and increase usable capacity when compared to standard RAID and replication.

DETAIL:
RED HAT GLUSTER STORAGE 3.1

These features were introduced in the most recent release of Red Hat Gluster Storage, and are now supported by Red Hat.

PERF	**Small file**	Optimizations to enhance small file performance, especially with small file create and write operations.
PERF	**Rebalance**	Optimizations that result in enhanced rebalance speed at large scale.
SECURITY	**SELinux enforcing mode**	Introduction of the ability to operate with SELinux in enforcing mode, increasing security across an entire deployment.
PROTOCOL	**NFSv4 (multi-headed)**	Support for data access via clustered, active-active NFSv4 endpoints, based on the NFS-Ganesha project.
PROTOCOL	**SMB 3 (subset of features)**	Enhancements to SMB 3 protocol negotiation, copy-data offload, and support for in-flight data encryption [Sayan: what is copy-data offload?]

EXTRA

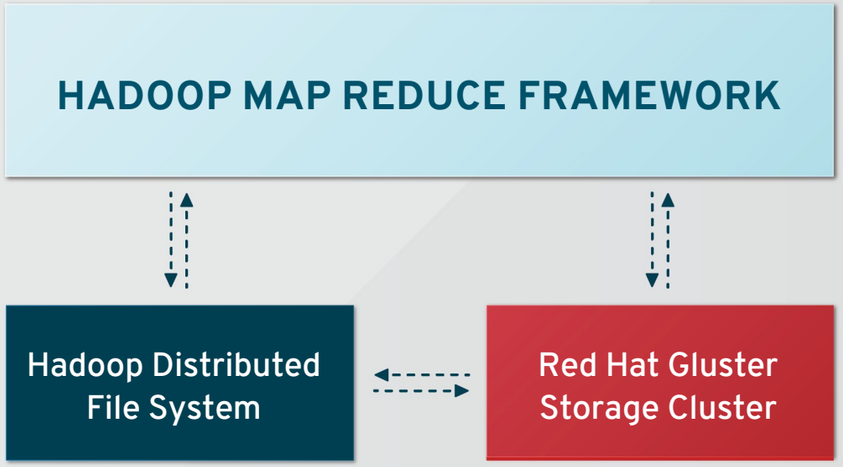

RED HAT GLUSTER STORAGE
TARGET WORKLOADS

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

BIG DATA ANALYTICS

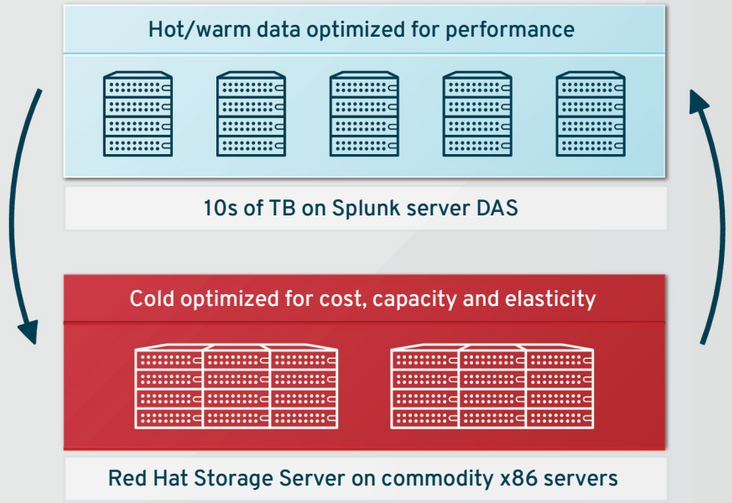
In-place Hadoop analytics in a POSIX compatible environment



FEATURES	BENEFITS
• Allows the Hortonworks Data Platform 2.1 to be deployed on Red Hat Gluster Storage • Hadoop tools can operate on data in-place • Access to the Hadoop ecosystem of tools • Access to non-Hadoop analytics tools • Consistent operating model: Hadoop can run directly on Red Hat Gluster Storage nodes	• Flexible, unified enterprise big data repository • Better analytics (Hadoop and non-Hadoop) • Familiar POSIX-compatible file system and tools • Start small, scale as big data needs grow • Multi-volume support (HDFS is single-volume) • Unified management (Hortonworks HDP Ambari and Red Hat Gluster Storage)

MACHINE DATA ANALYTICS

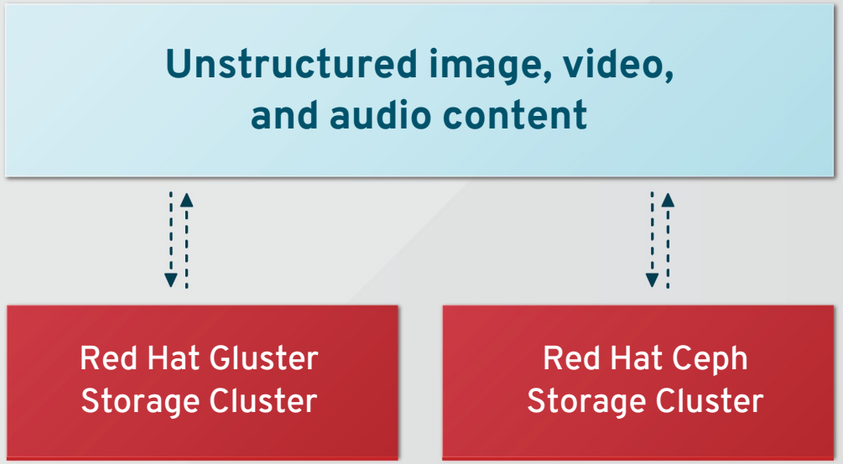
High-performance, scale-out, online cold storage for Splunk Enterprise



FEATURES	BENEFITS
• Multiple ingest options using NFS & FUSE • Expand storage pools without incurring downtime • Support for both clustered and non-clustered configurations	• Run high speed indexing and search on Splunk's cold data store • Pay as you grow economics for Splunk cold data • Reduce ingestion time for data with standard protocols • "Always online", fast, disk-based storage pools provide constant access to historical data

RICH MEDIA

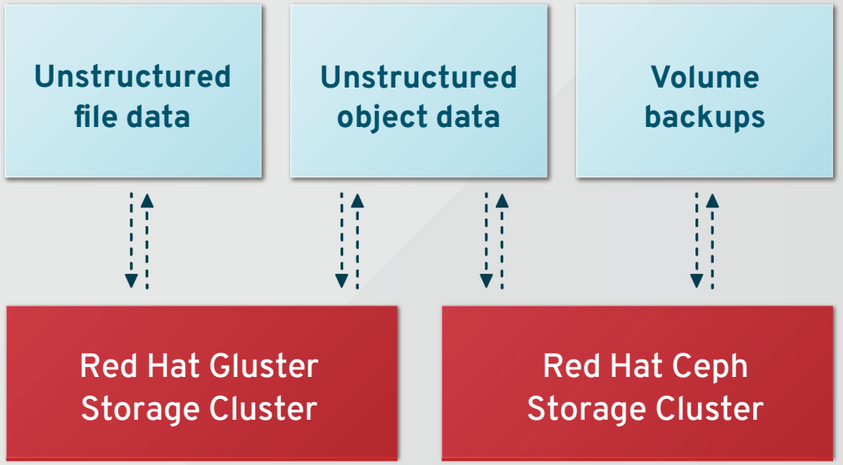
Massively-scalable, flexible, and cost-effective storage for image, video, and audio content



FEATURES	BENEFITS
• Support for multi-petabyte storage clusters on commodity hardware • Erasure coding and replication for capacity-optimized or performance-optimized pools • Support for standard file & object protocols • Snapshots and replication capabilities for high availability and disaster recovery	• Provides massive and linear scalability in on-premise or cloud environments • Offers robust data protection with an optimal blend of price & performance • Standard protocols allow access to broadcast content anywhere, on any device • Cost-effective, high performance storage for on-demand rich media content

ACTIVE ARCHIVES

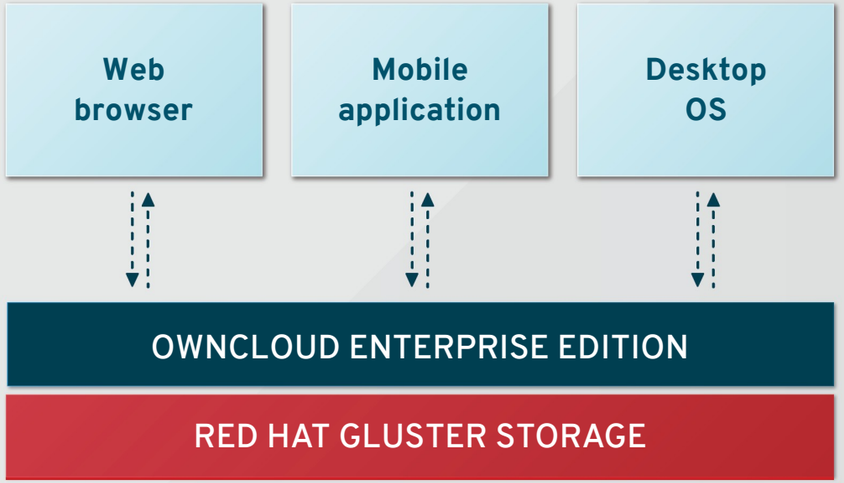
Open source, capacity-optimized archival storage on commodity hardware



FEATURES	BENEFITS
• Cache tiering to enable "temperature"-based storage • Erasure coding to support archive and cold storage use cases • Support for industry-standard file and object access protocols	• Store data based on its access frequency • Store data on premise or in a public or hybrid cloud • Achieve durability while reducing raw capacity requirements and limiting cost • Deploy on industry-standard hardware

FILE SYNC AND SHARE

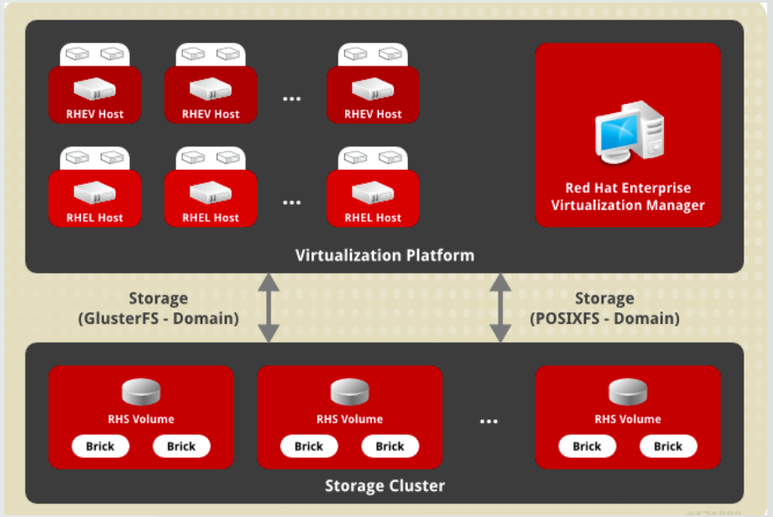
Powerful, software-defined, scale-out, on-premise storage for file sync and share with ownCloud



FEATURES	BENEFITS
• Secure file sync and share with enterprise-grade auditing and accounting • Combined solution of Red Hat Gluster Storage, ownCloud, HP ProLiant SL4550 Gen 8 servers • Deployed on-premise, managed by internal IT • Access sync and share data from mobile devices, desktop systems, web browsers	• Secure collaboration with consumer-grade ease of use • Lower risk by storing data on-premise • Conform to corporate data security and compliance polices • Lower total cost of ownership with standard, high-density servers and open source

ENTERPRISE VIRTUALIZATION

Scalable, reliable storage for Red Hat Enterprise Virtualization



FEATURES	BENEFITS
• Reliably store virtual machine images in a distributed Red Hat Gluster Storage volume • Manage storage through the RHEV-M console • Deploy on standard hardware of choice • Seamlessly grow and shrink storage infrastructure when demand changes	• Reduce operational complexities by eliminating dependency on complex and expensive SAN infrastructures • Deploy efficiently on less expensive, easier to provision, standard hardware • Achieve centralized visibility and control of server and storage infrastructure