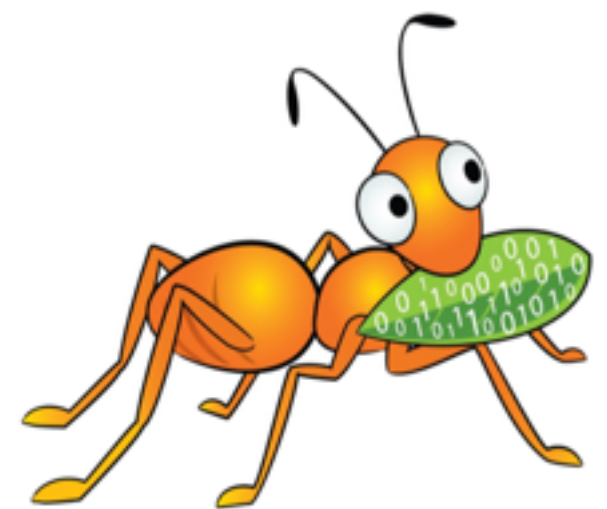


GLUSTERFS INTRODUCTION



BIN LIU
2016/9/5



INTRODUCTION

- Scalable network filesystem
- Free and open source software、 community
- Gluster inc. , 2005, acquired by RedHat, 2011
- Gluster 3.8 is the latest major release as of June 2016
- Gluster 4.0 - December 2016

FEATURES

- Linux(CentOS/Ubuntu)
- No master/Stateless
- UserLand
- NFS/CIFS/Swift API
- Distributed/Mirroring/Striped
- Economies of Scale

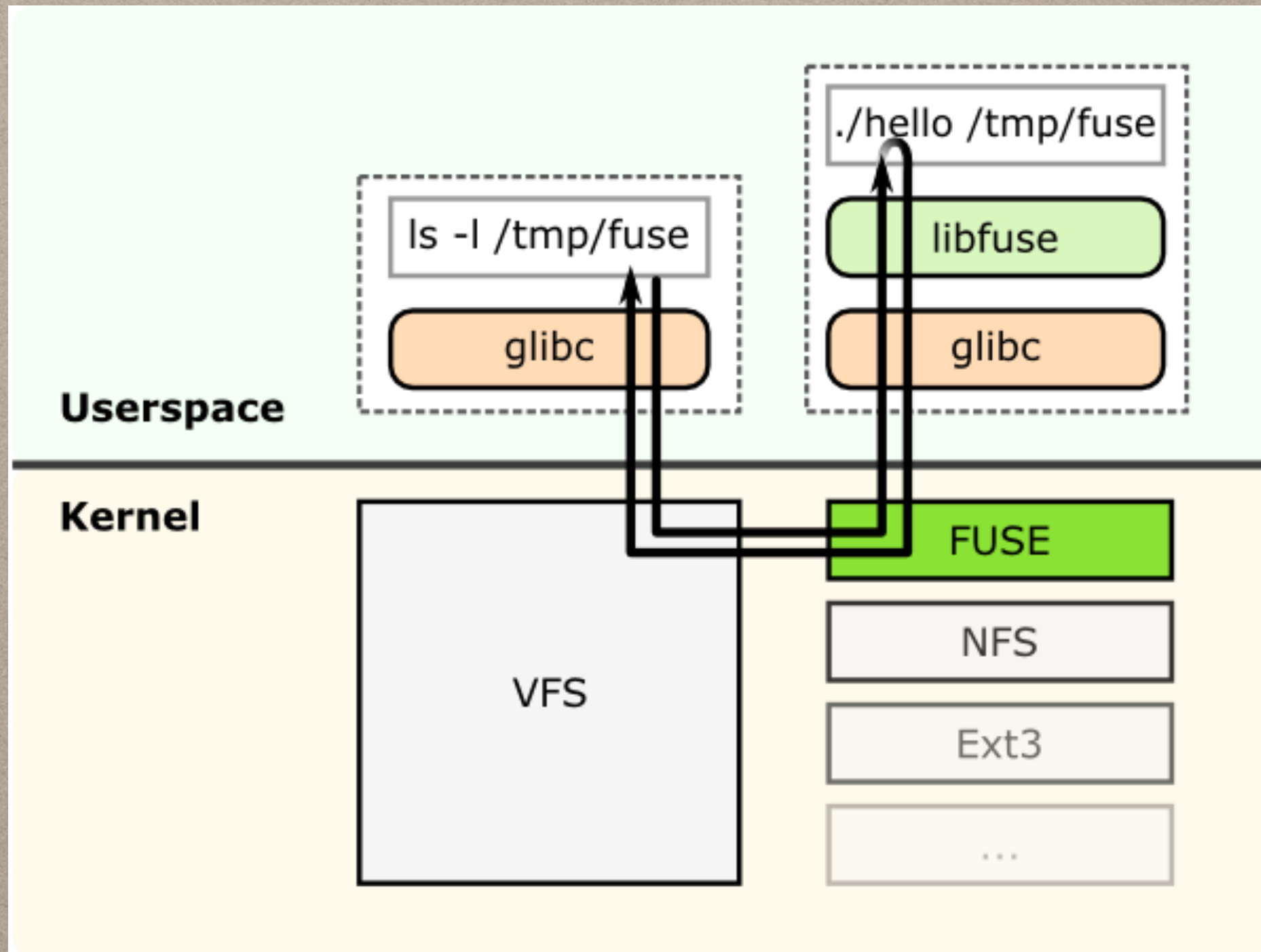
DISADVANTAGES

- Must on xfs/ext4 ...
- Volume & bricks are combined manually
- Volume & bricks are combined manually
- Volume & bricks are combined manually

KEY TECHS

- FUSE
- DHT

FUSE



DHT(DISTRIBUTED HASH TABLE)

- Consistent hashing for bricks
- Hashing by filename
- Different by directory

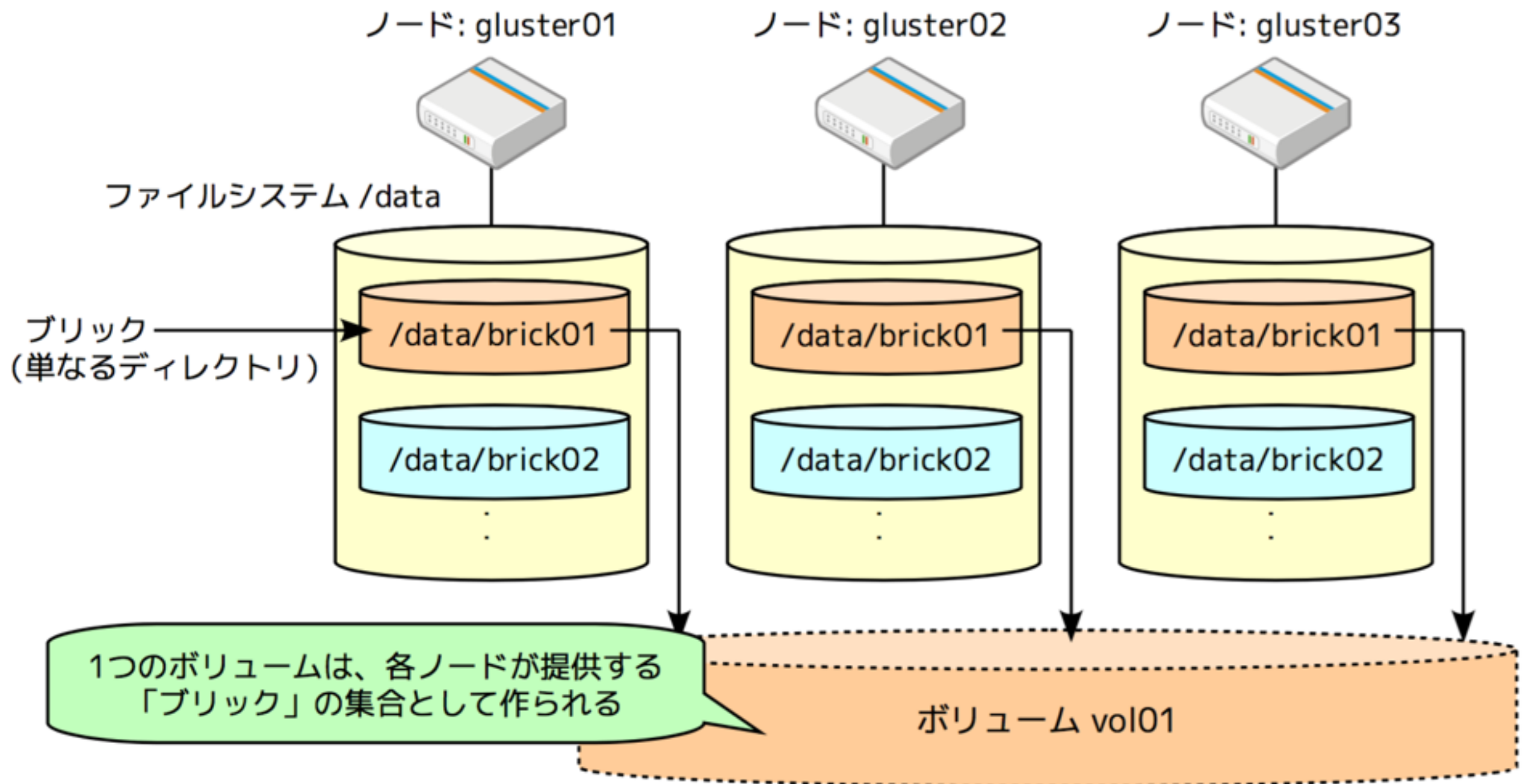
DHT



BASIC CONCEPT

- Trusted Storage Pool - Collection of Storage Servers
- Brick - An export directory on a server
- Volume - Logical collection of bricks
- Sub-volume - part of a volume

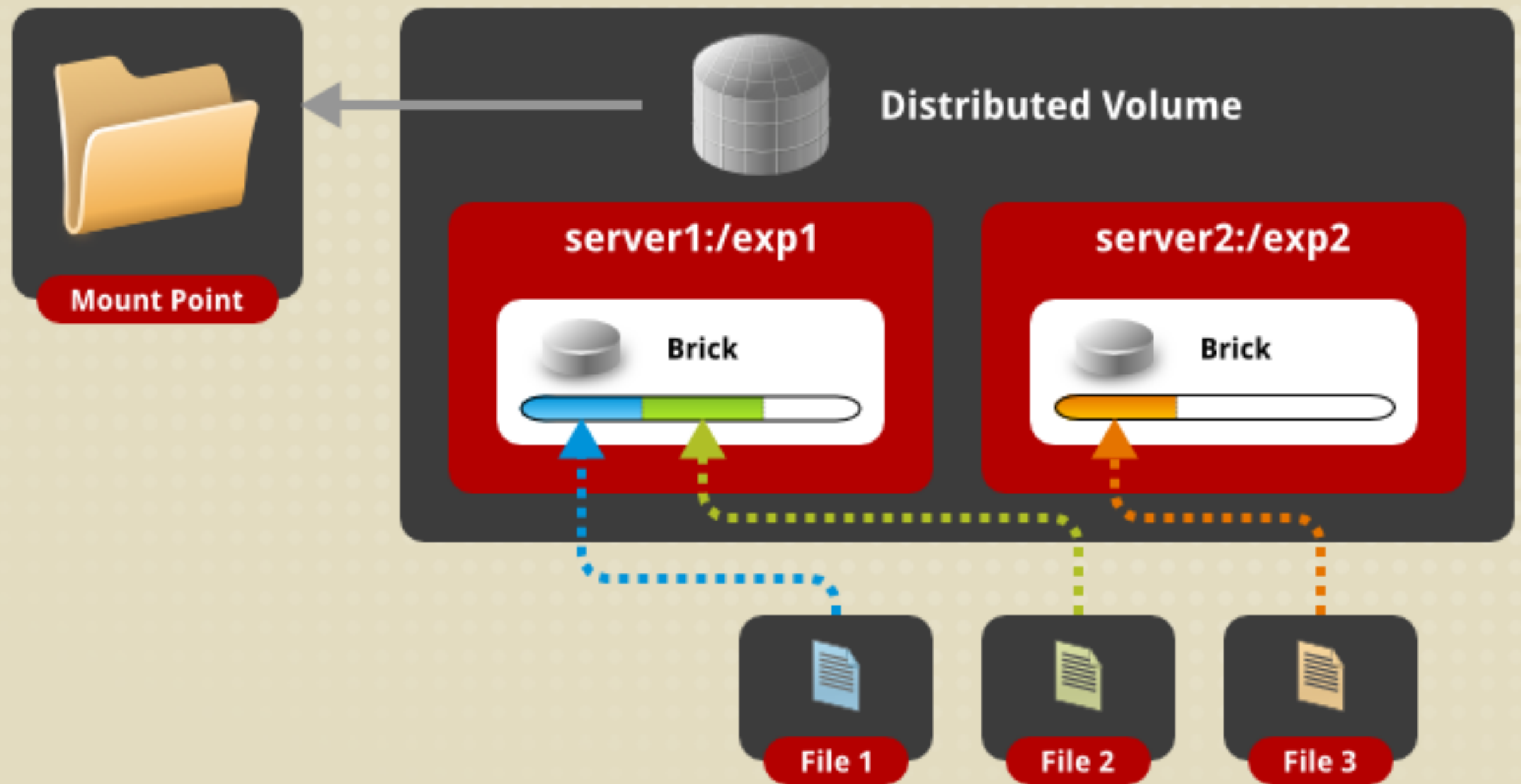
BASIC ARCH



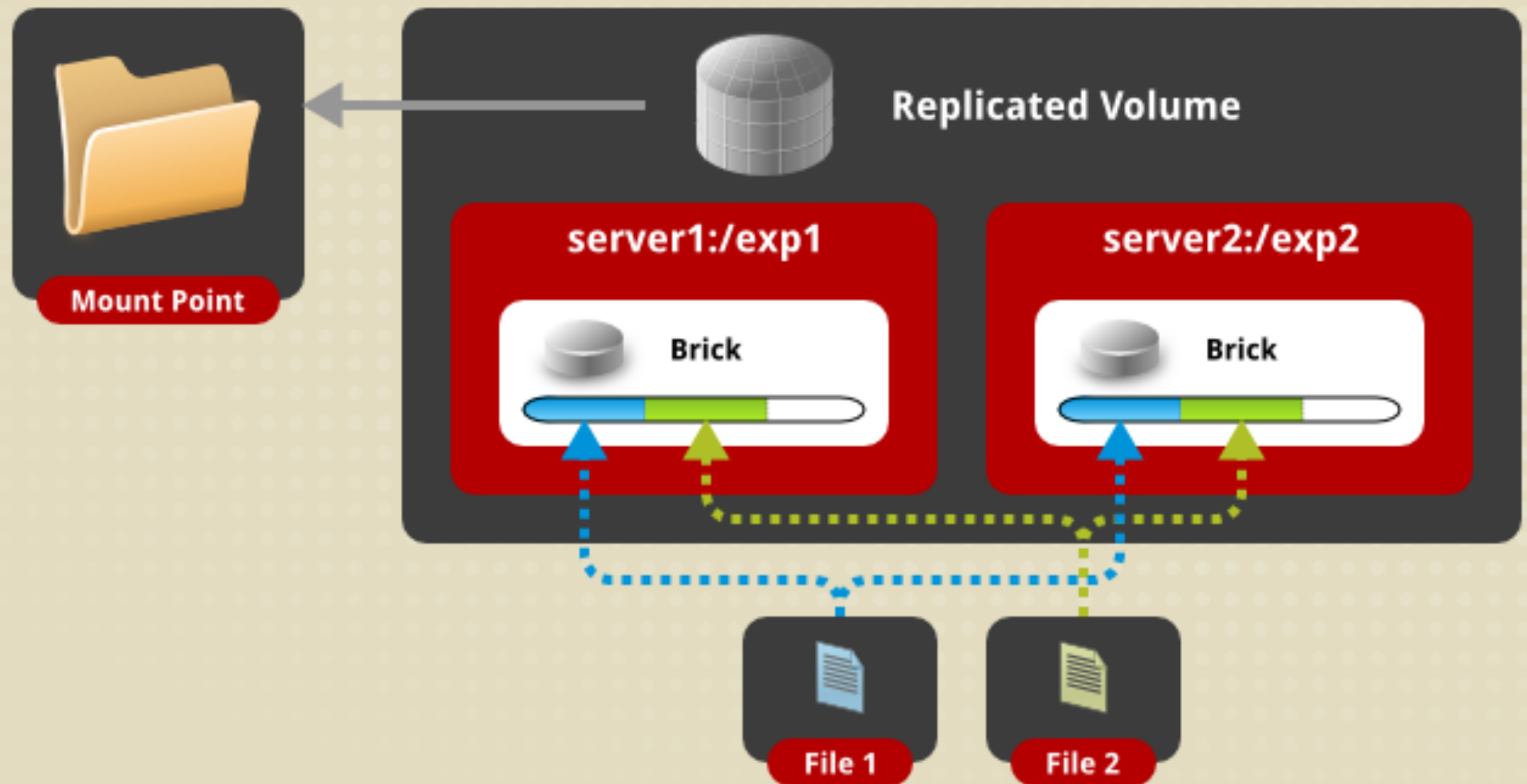
GLISTERS VOLUME TYPES

- Distributed Volume
- Replicated Volume
- Distributed Replicated Volume
- Striped Volume
- Distributed Striped Volume

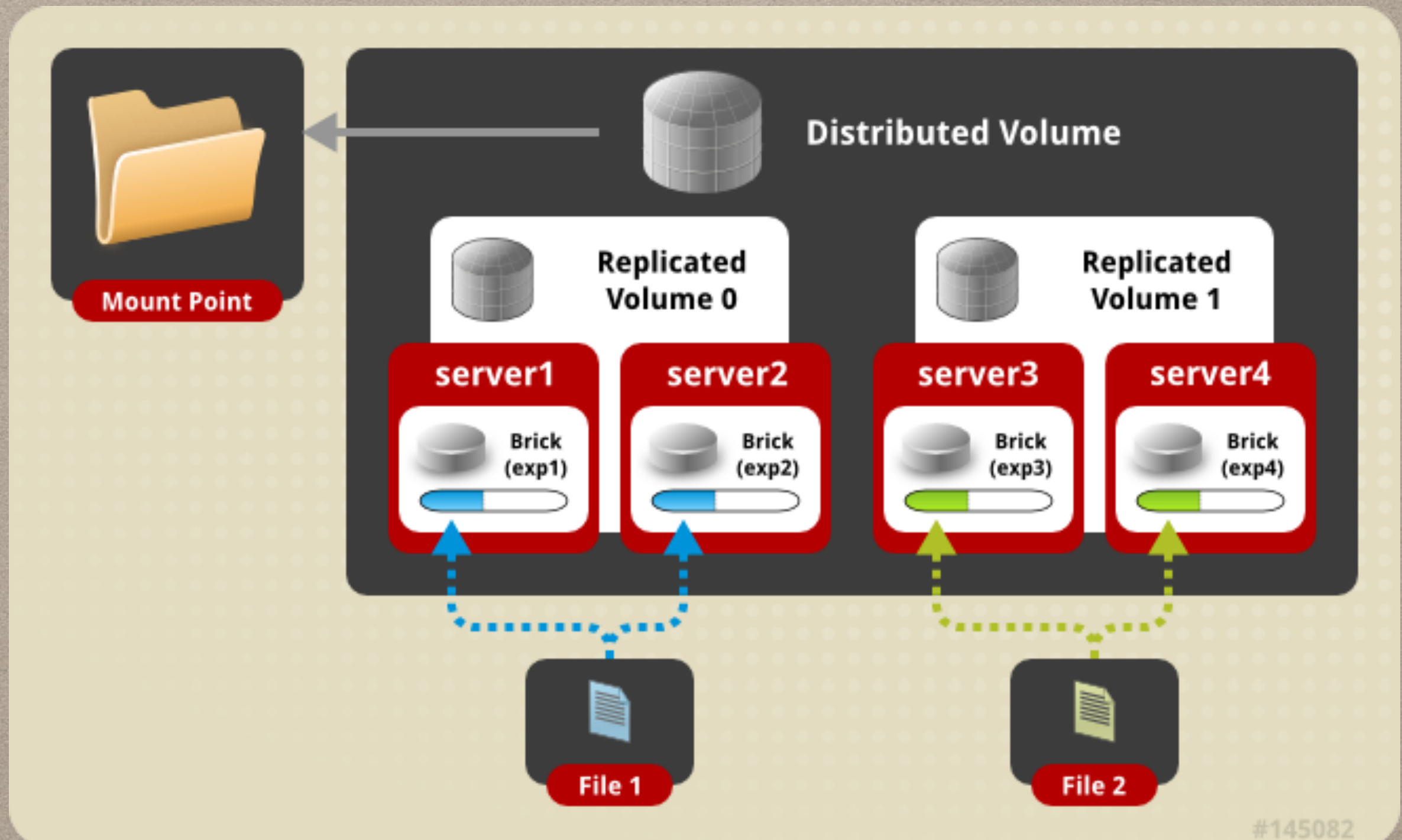
DISTRIBUTED VOLUME



REPLICATED VOLUME



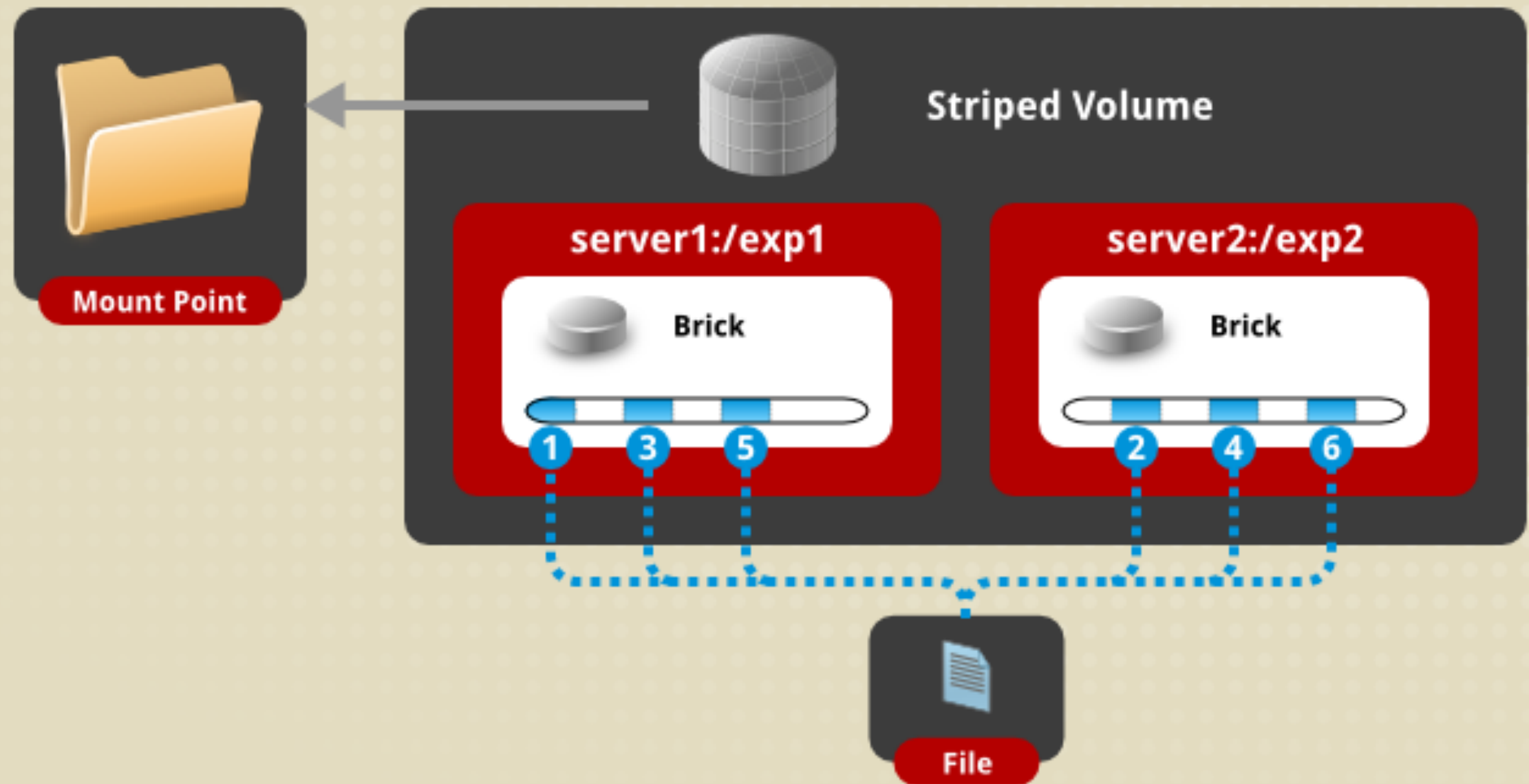
DISTRIBUTED REPLICATED VOLUME



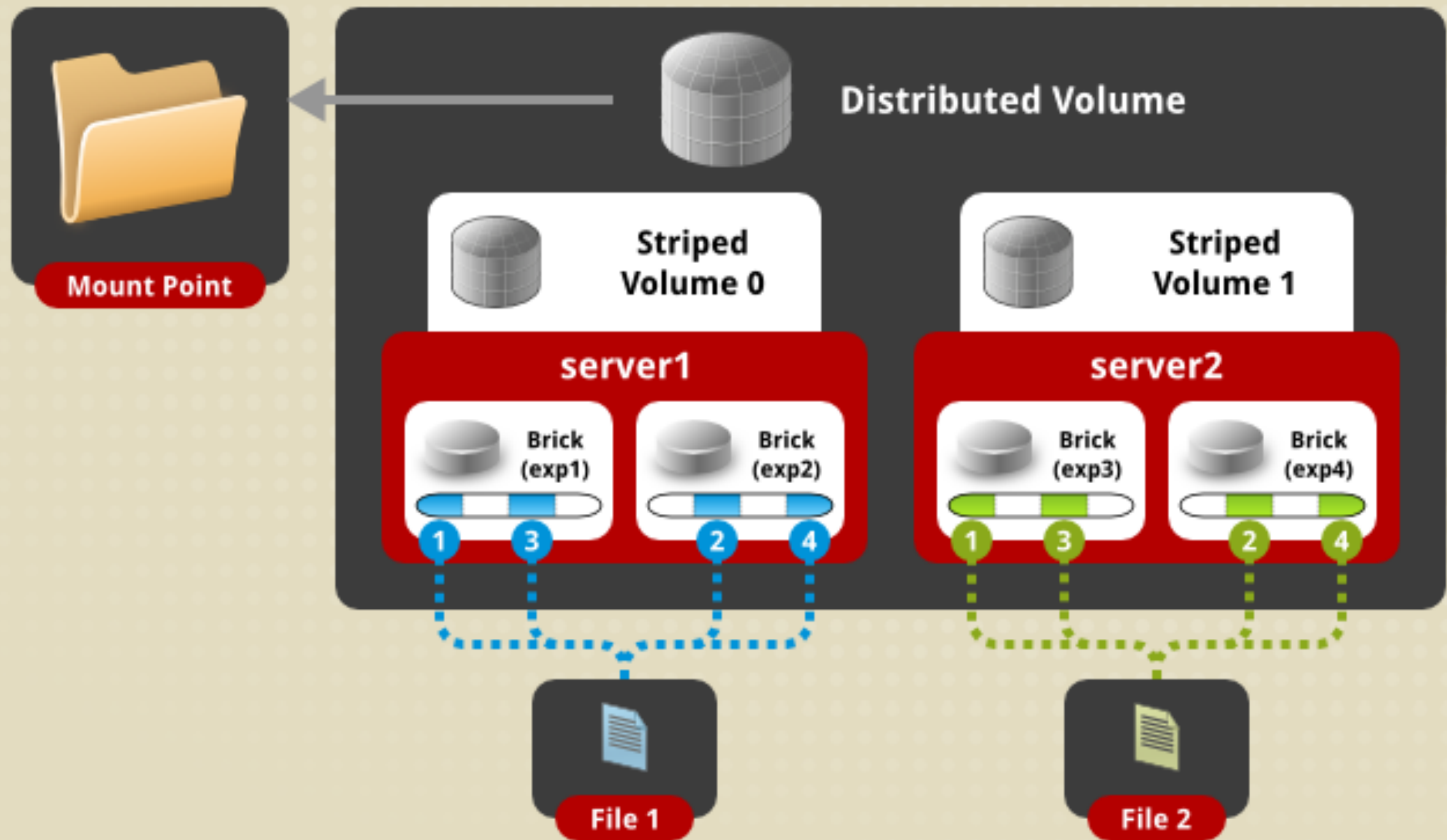
DISTRIBUTED REPLICATED GLUSTERFS VOLUME

```
# gluster volume create test-volume \  
replica 2 \  
server1:/exp1 server2:/exp2 \  
server3:/exp3 server4:/exp4
```


STRIPED VOLUME



DISTRIBUTED STRIPED VOLUME



DISTRIBUTED STRIPED VOLUME

```
# gluster volume create test-volume stripe 4 \  
server1:/exp1 server2:/exp2 \  
server3:/exp3 server4:/exp4 \  
server5:/exp5 server6:/exp6 \  
server7:/exp7 server8:/exp8
```

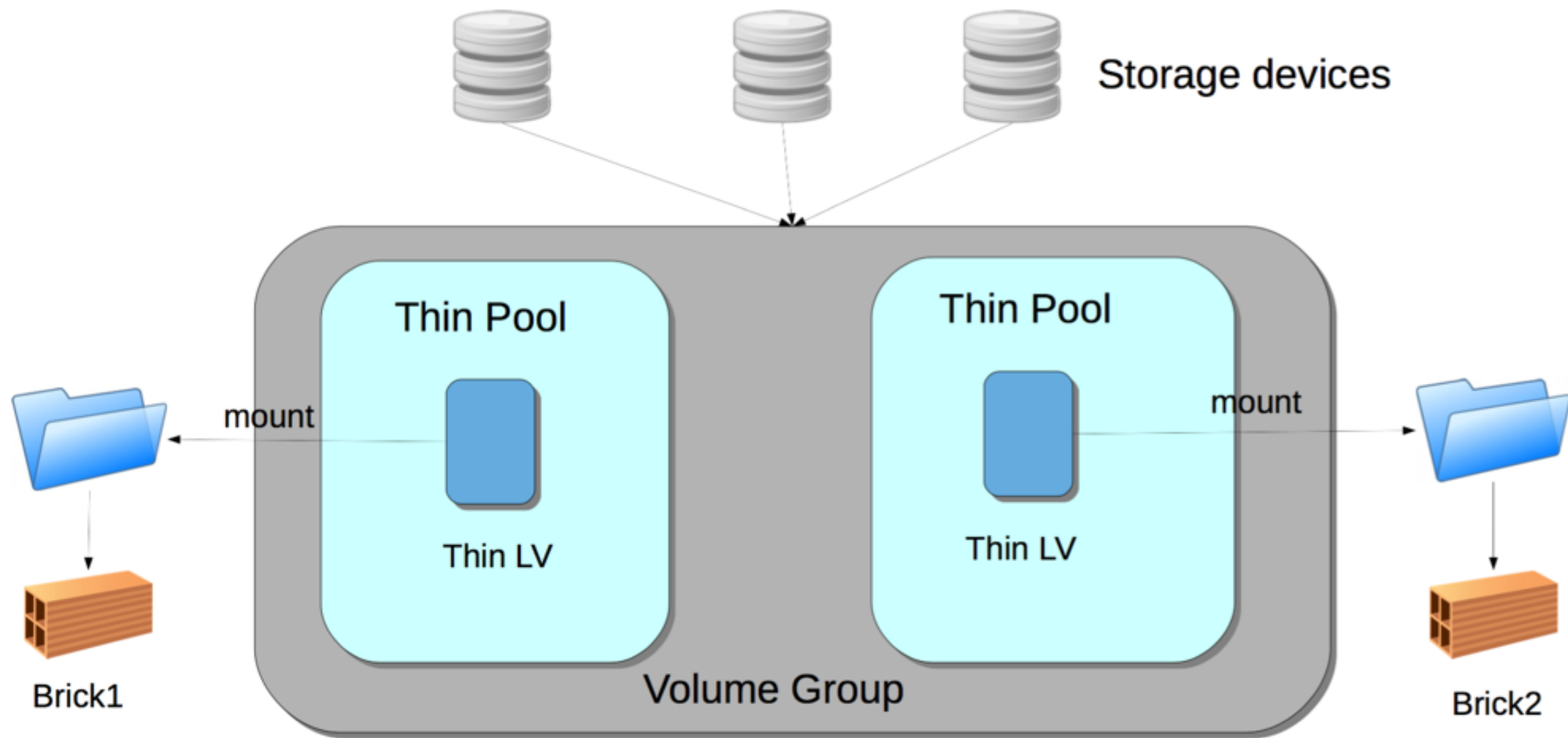

VOLUME QUOTA

- `gluster volume quota vol-name enable`
- `gluster volume quota vol-name limit-usage / 1GB`
- `gluster volume quota vol-name limit-usage /foo
30GB`
- `gluster volume quota vol-name list`

SNAPSHOT

- Each brick should be on an independent thin provisioned LVM.
- Brick LVM should not contain any other data other than brick.
- None of the brick should be on a thick LVM.
- gluster version should be 3.6 and above.

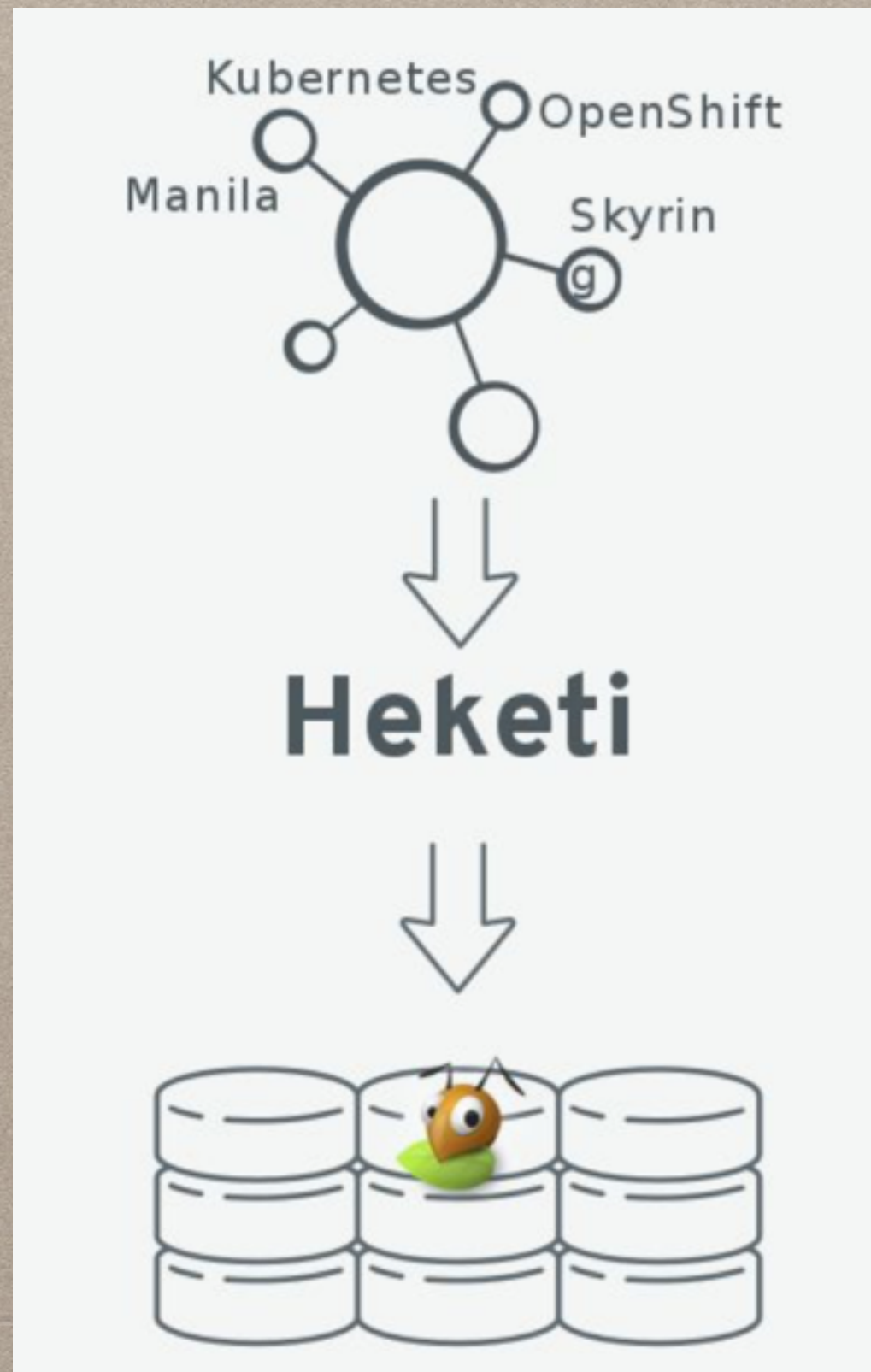
SNAPSHOT



GEO-REPLICATION

- From 3.2
- LAN/WAN/Internet/Cascading
- Master-slave model
- How to replication
 - * Change detection
 - * Replication(Rsync)

STAAS - STORAGE AS A SERVICE



THANKS