GlusterFS introduction. Volumes architecture.

Done by: Anar Bazarhanova bazarhanova@gmail.com
Course: Computing clusters, Grids and Clouds
Course supervisor: Andrey Y. Shevel

ITMO, Laboratory of the Network Technologies in Distributed Computing Systems
02.06.2015

What is it?

- Gluster is a distributed scalable network-attached storage filesystem that allows rapid provisioning of additional storage based on your storage consumption needs.
- by Gluster, Inc., then by Red Hat, Inc., after their purchase of Gluster in 2011 (GNU License v3).
  - Cluster management and configuration
  - Data distribution
  - Common control and data distribution

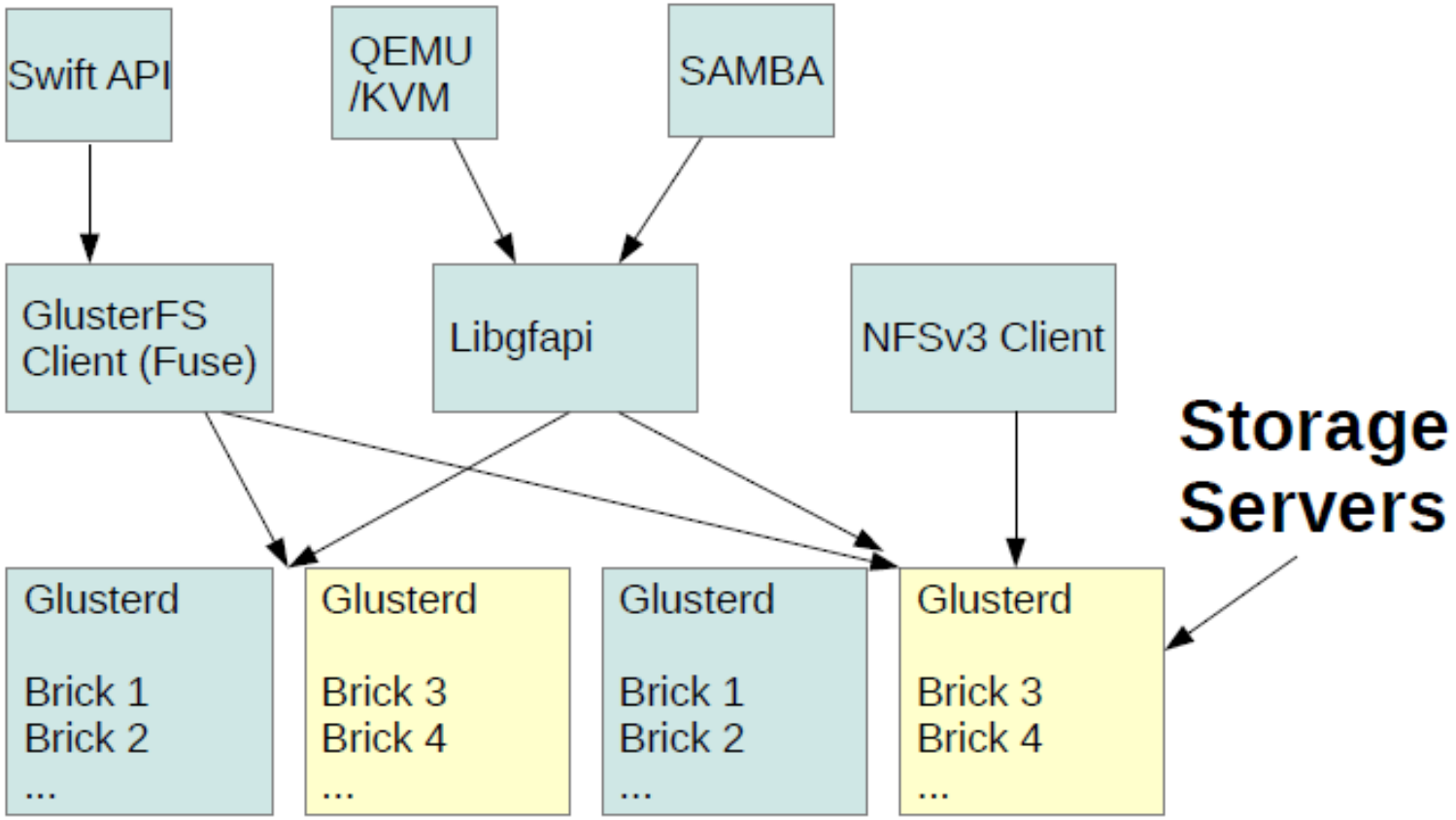
Terminology

- **Trusted Storage Pool** is a trusted network of storage servers.
- **Brick** is the basic unit of storage, represented by an export directory on a server in the trusted storage pool.
- **Volume** is a logical collection of bricks. Most of the gluster management operations happen on the volume.

Key features

- GlusterFS uses a native FUSE-based client to export the file system
  - Filesystem in Userspace (FUSE) is a loadable kernel module for Unix-like OS that lets non-privileged users create their own file systems without editing kernel code.
- No data silos - *files, objects and block devices all in the same namespace*
- No single point of failure
- Global namespace (logical grouping of Ids)

Client Access Overview



Key features

- **Elasticity:** Storage volumes are abstracted from underlying hardware and can be grown, shrunk, or migrated across physical systems as necessary.
- **High availability:** Synchronous *n*-way file replication ensures high data availability and recovery, access from anywhere.
- **Scalability:** 1 machine ->thousands of systems

Key features

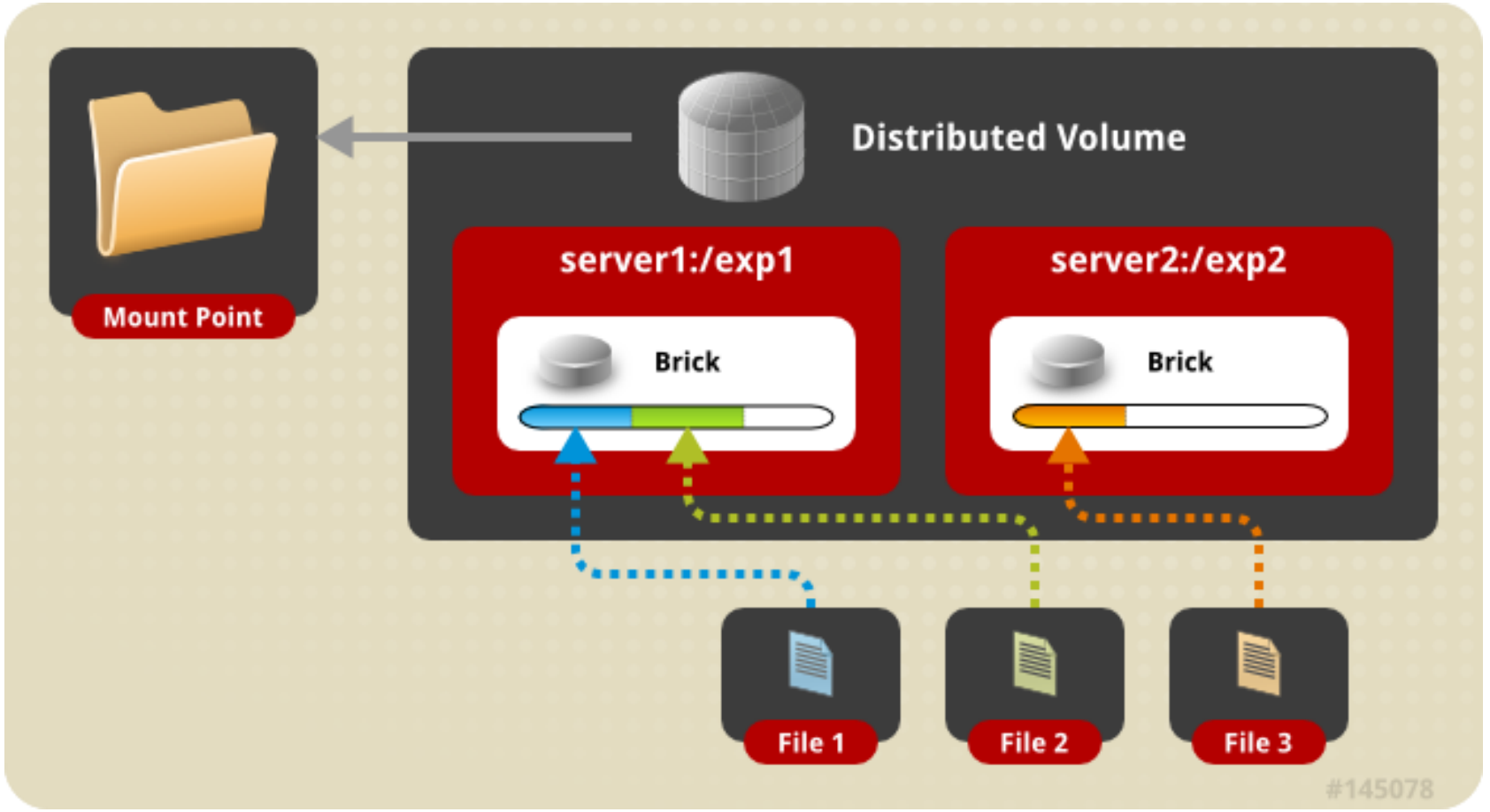
- **Flexibility:** GlusterFS runs in userspace, so there is no need for kernel patches, custom modules, and so on, reconfigurability.
- **Geo-replication:** GlusterFS enables you to replicate the whole storage system between different datacenters or geographic locations.
  - Master-slave (mirroring), asynchronous cascading

Cluster and Performance translators Distributed Hash Table Translator

- No centralized metadata storage concept, it is stored with the data itself, use of **Elastic Hash** *(DHT xlator)*
  - DHT is the real core of how GlusterFS aggregates capacity and performance across multiple servers. to place each file on exactly one of its subvolumes.
  - It's a routing function, not splitting or copying.
- AFR(Automatic File Replication) Translator

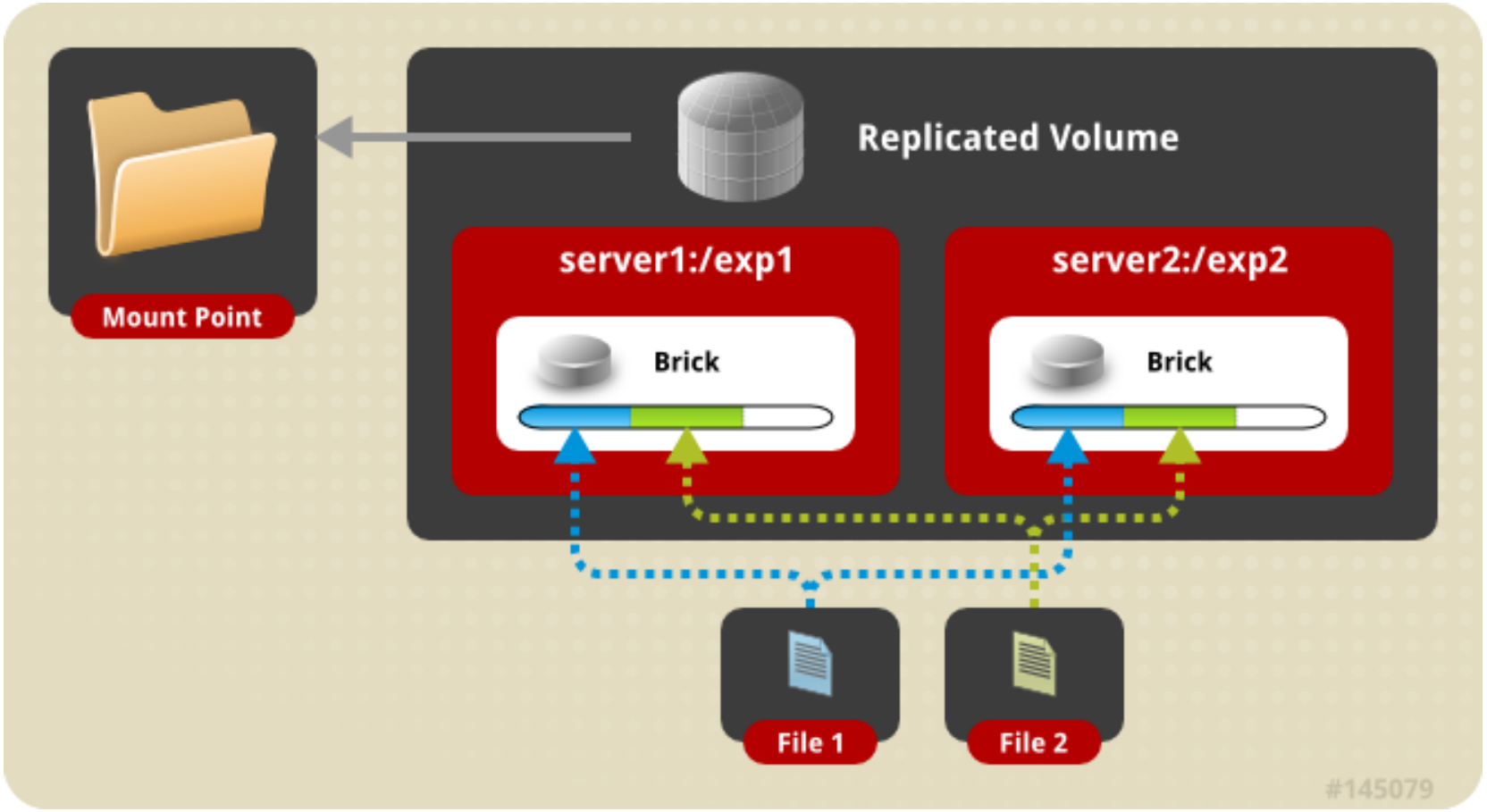
Architecture

- Distributed Glusterfs Volume



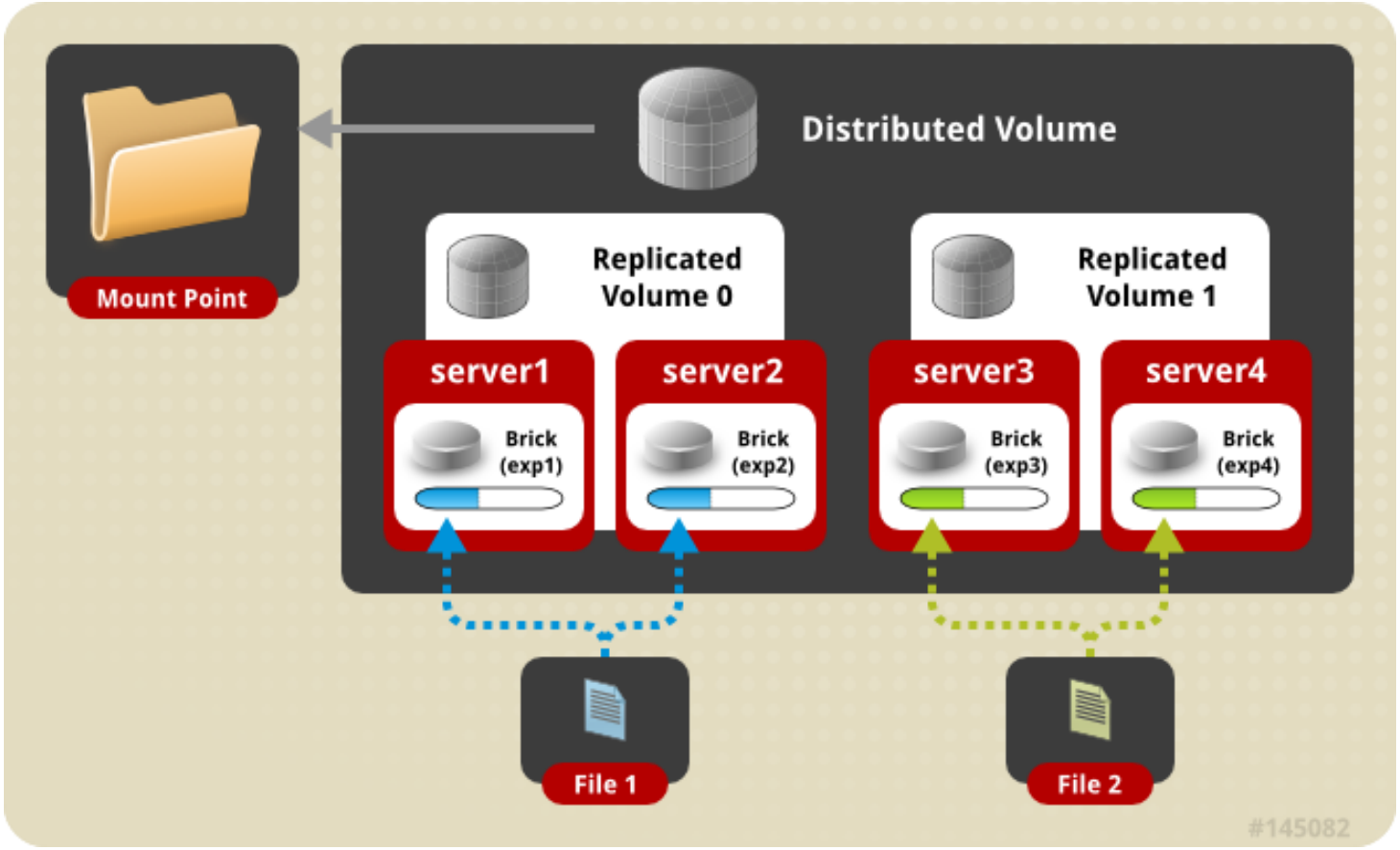
Architecture

- Replicated Glusterfs Volume



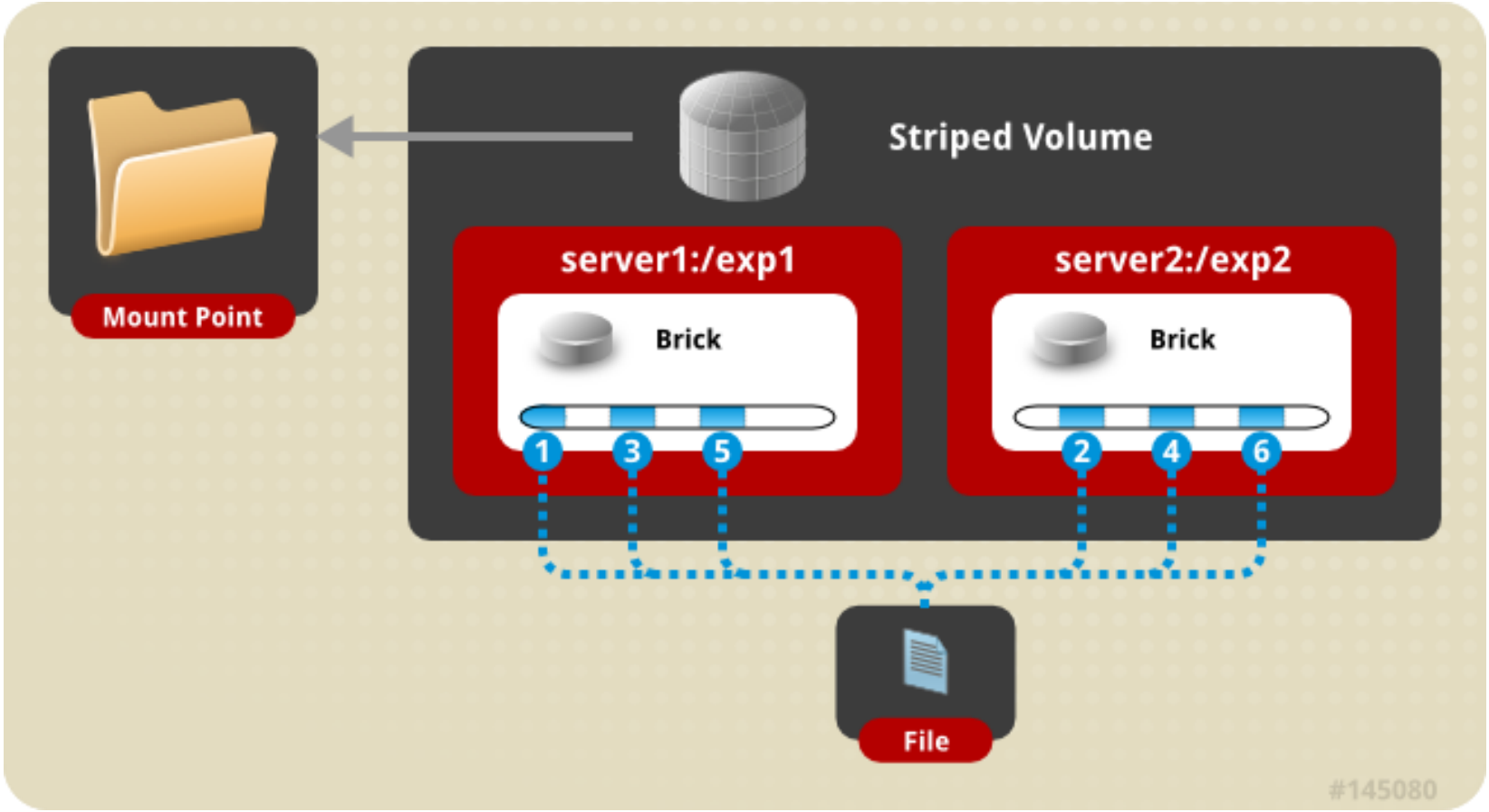
Architecture

- Distributed Replicated Glusterfs Volume



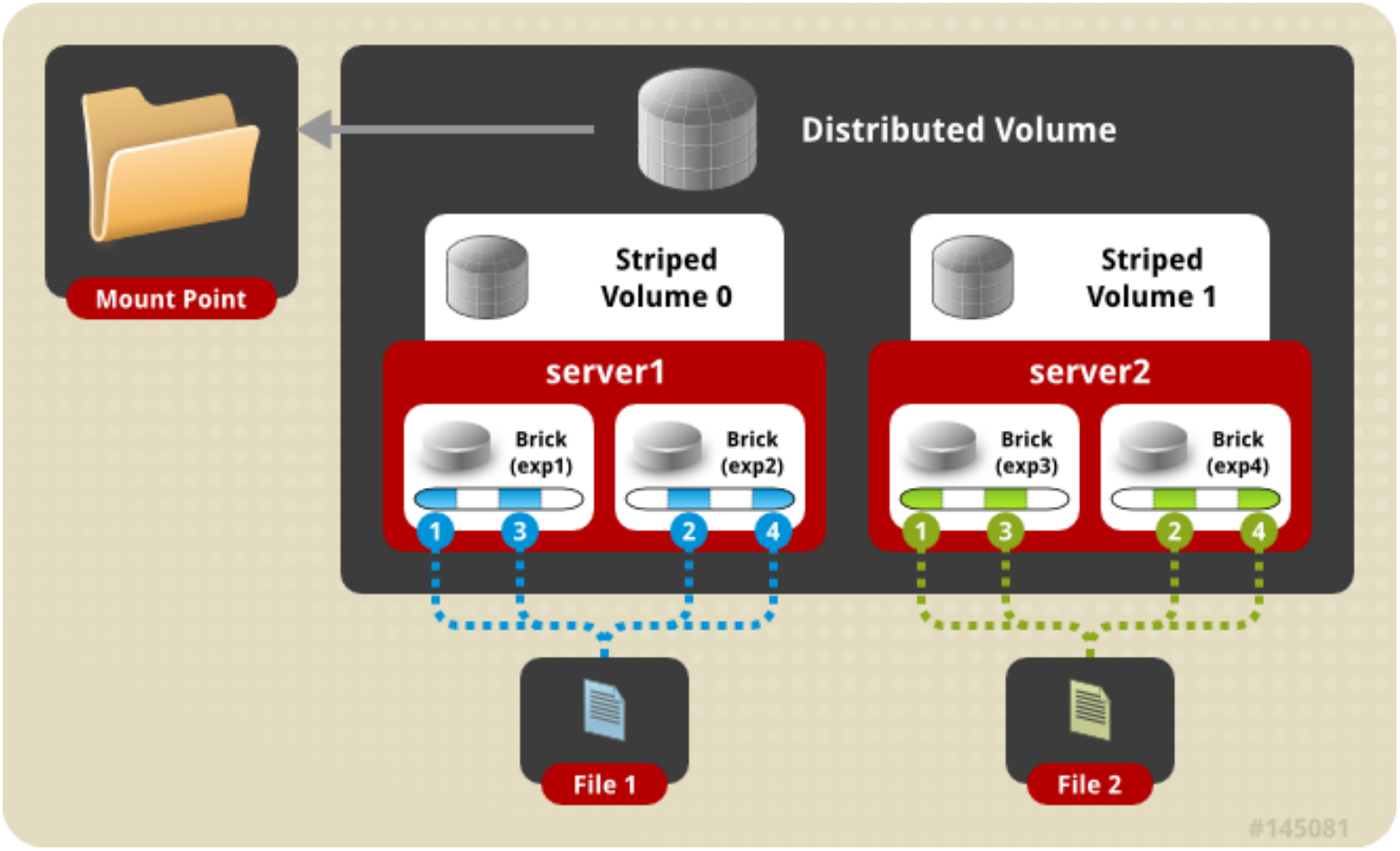
Architecture

- Striped Glusterfs Volume



Architecture

- Distributed Striped Glusterfs Volume



To sum up..

- Glusterfs allows enterprises to combine large numbers of commodity storage and compute resources into a high performance, virtualized and centrally managed pool.
- Architecture based on needs
- Capacity and performance can scale independently

➢Create the world's largest and most dynamic community for open software-defined storage

Thank you

Q&A.