



CLOUD COMPUTING APPLICATIONS

Ceph

Roy Campbell & Reza Farivar

Motivation

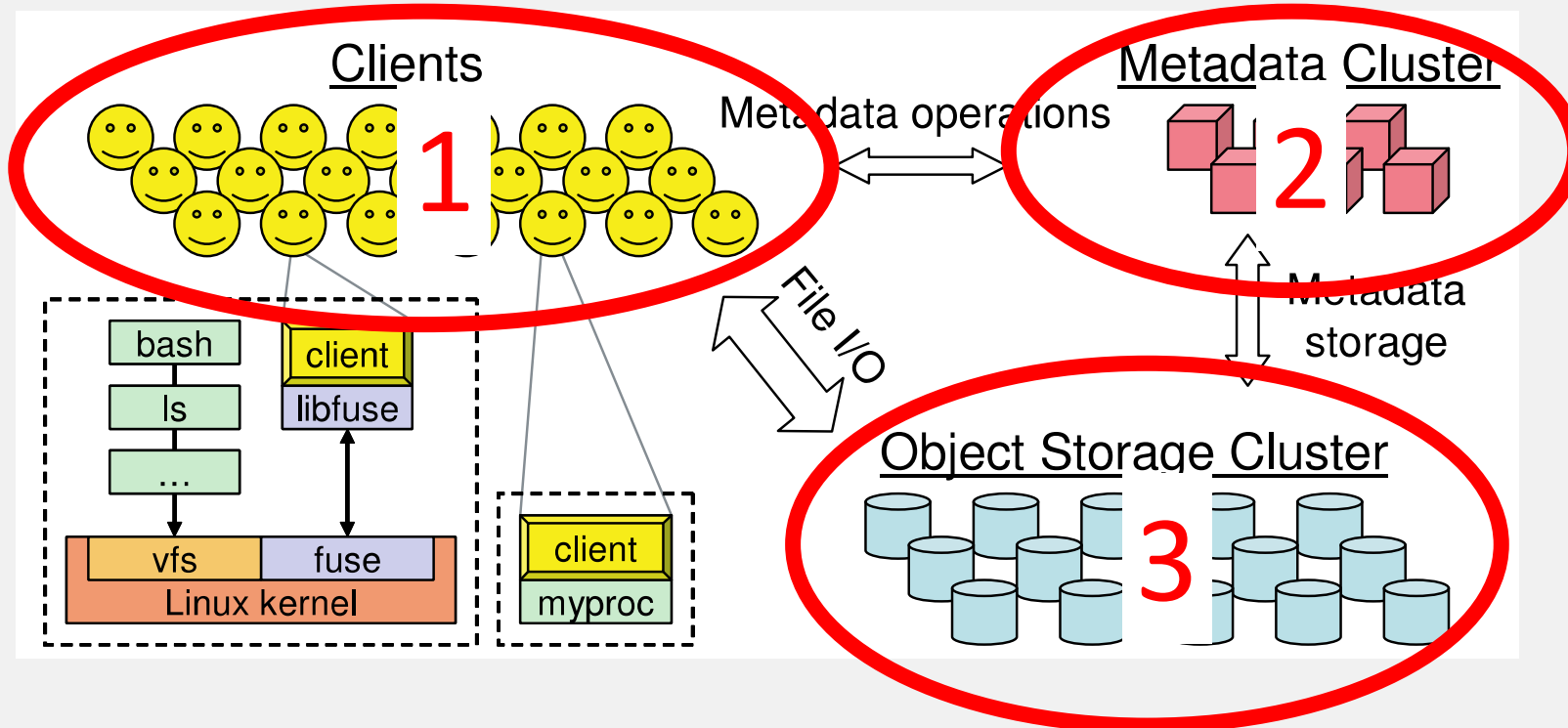
- Ceph is an emerging technology in the production-clustered environment
- Designed for:
 - Performance – Striped data over data servers
 - Reliability – No single point of failure
 - Scalability – Adaptable metadata cluster
 - More general than HDFS
 - Smaller files
- GlusterFS has more or less similar ideas
 - Uses ring-based hashing; Ceph uses CRUSH

Ceph Overview

- MDS – Meta Data Server
- ODS – Object Data Server
- MON – Monitor (now fully implemented)
- **Decoupled data and metadata**
 - I/O directly with object servers
- **Dynamic distributed metadata management**
 - Multiple metadata servers handling different directories (subtrees)
- **Reliable autonomic distributed storage**
 - ODS's manage themselves by replicating and monitoring

Ceph Components

- Ordered: Clients, Metadata, Object Storage



Decoupled Data and Metadata

- Increases performance by limiting interaction between clients and servers
- Decoupling is common in distributed filesystems: HDFS, Lustre, Panasas...
- In contrast to other file systems, Ceph uses a function to calculate the block locations

Dynamic Distributed Metadata Management

- Metadata is split among cluster of servers
- Distribution of metadata changes with the number of requests to even load among metadata servers
- Metadata servers also can quickly recover from failures by taking over neighbors' data
- Improves performance by leveling metadata load

Reliable Autonomic Distributed Storage

- Data storage servers act on events by themselves
- Initiates replication and
- Improves performance by offloading decision making to the many data servers
- Improves reliability by removing central control of the cluster (single point of failure)