



数据驱动 存储创新


2017企业存储技术峰会·北京站

2017 ENTERPRISE STORAGE TECHNOLOGY SUMMIT BEIJING

2017.01.18 · 北京东方美爵酒店

主办方

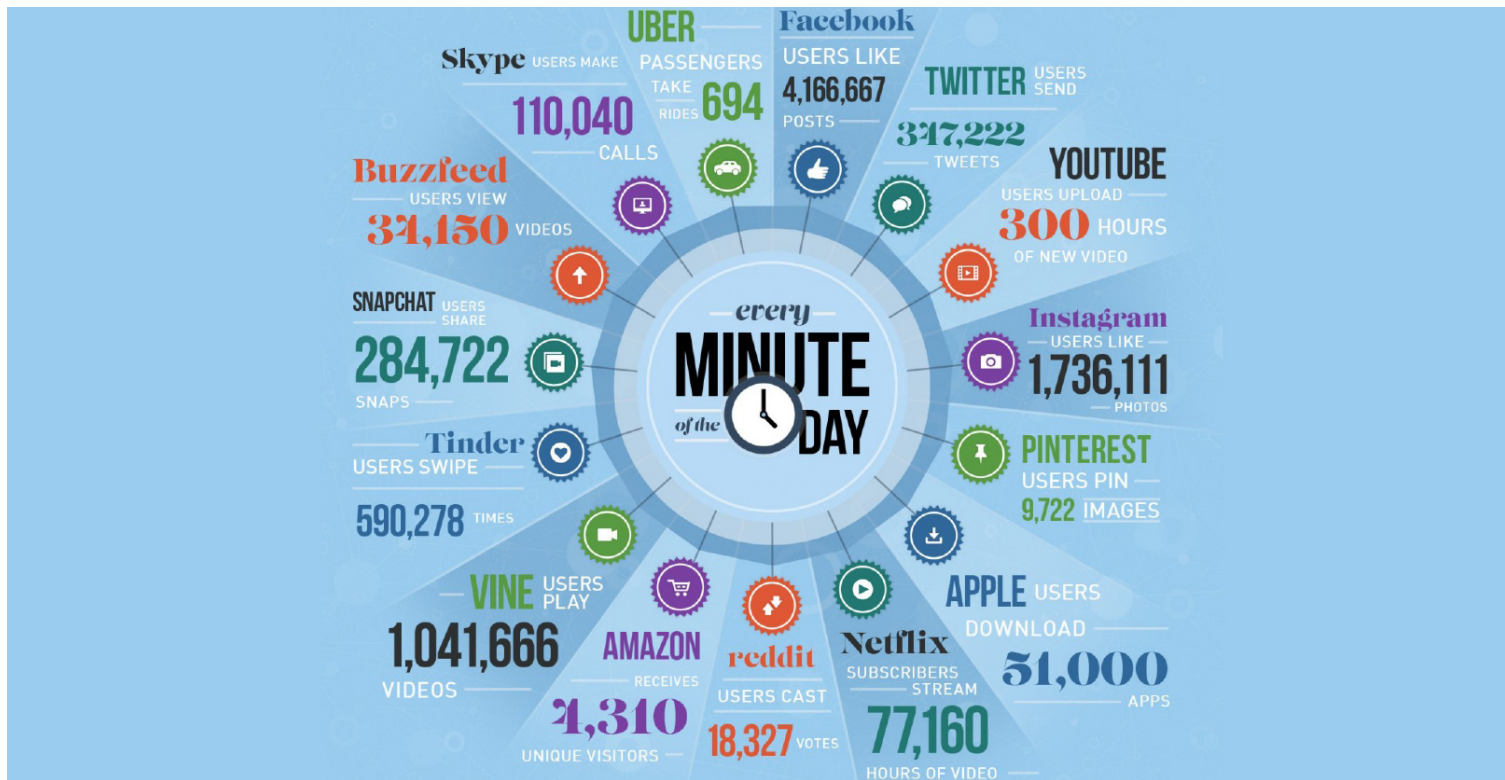
日知录技术社区





数据中心的演进与分布式 存储的崛起

张家驹
Red Hat



数据中心的演进

开发模式



瀑布式



敏捷



DevOps

应用架构



一体化



N层



微服务

部署与封装



裸机



虚拟服务



容器

应用基础架构



数据中心



托管



混合云

存储



纵向扩展



横向扩展



软件定义的存储

数据中心的演进

开发模式



瀑布式



敏捷



DevOps

应用架构



一体化



N层



微服务

部署与封装



裸机



虚拟服务



容器

应用基础架构



数据中心



托管

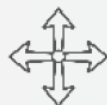


混合云

存储



纵向扩展



横向扩展



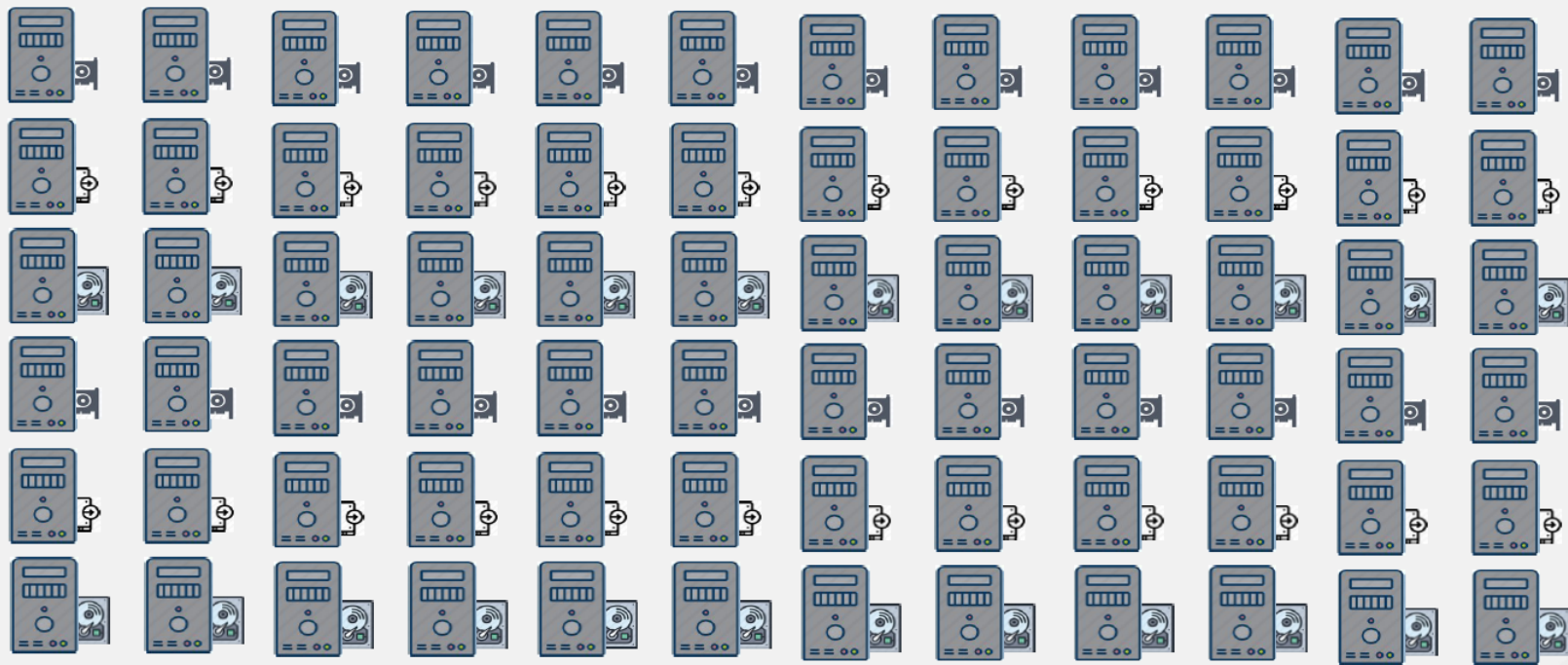
软件定义的存储

存储的演进

传统 → 开放、软件定义



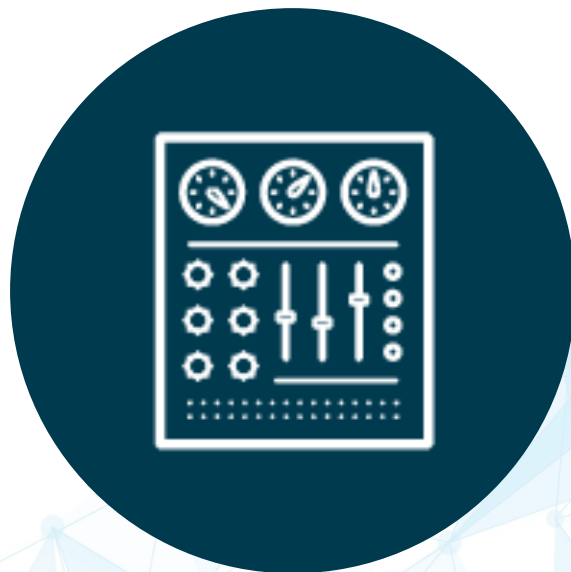
分布式大规模横向扩展



两个基本问题

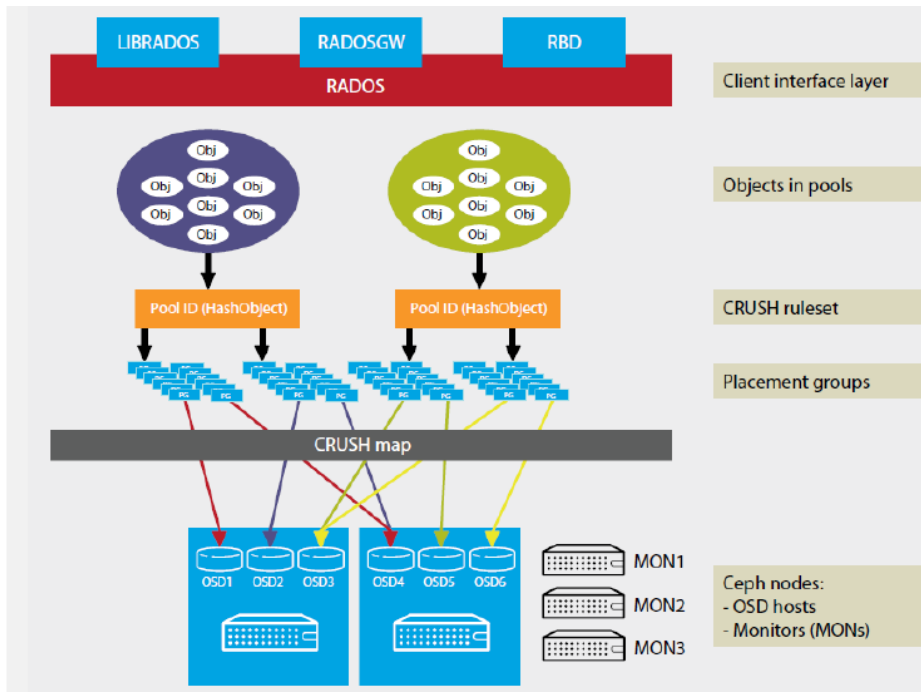
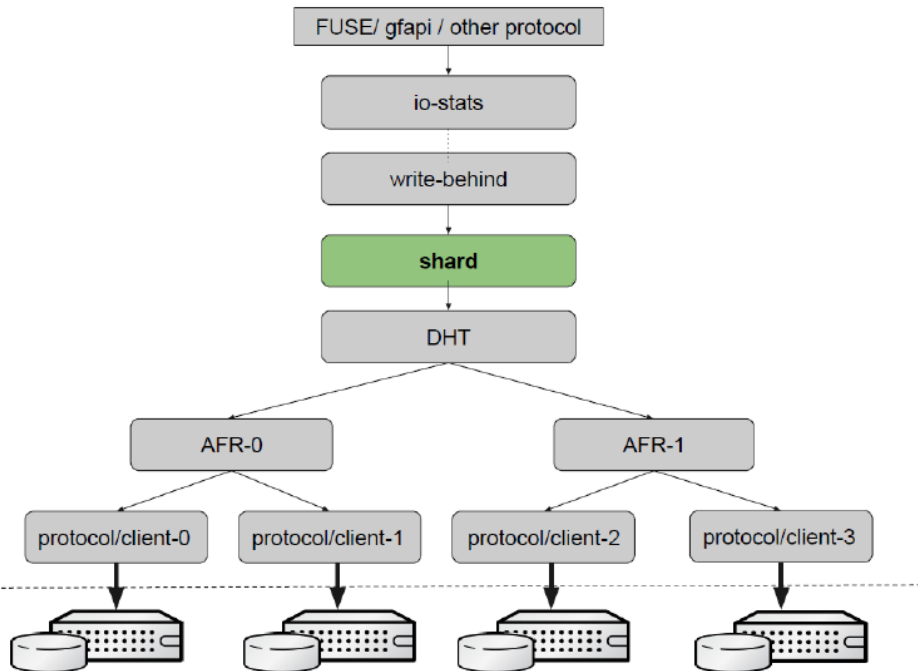


存储池化

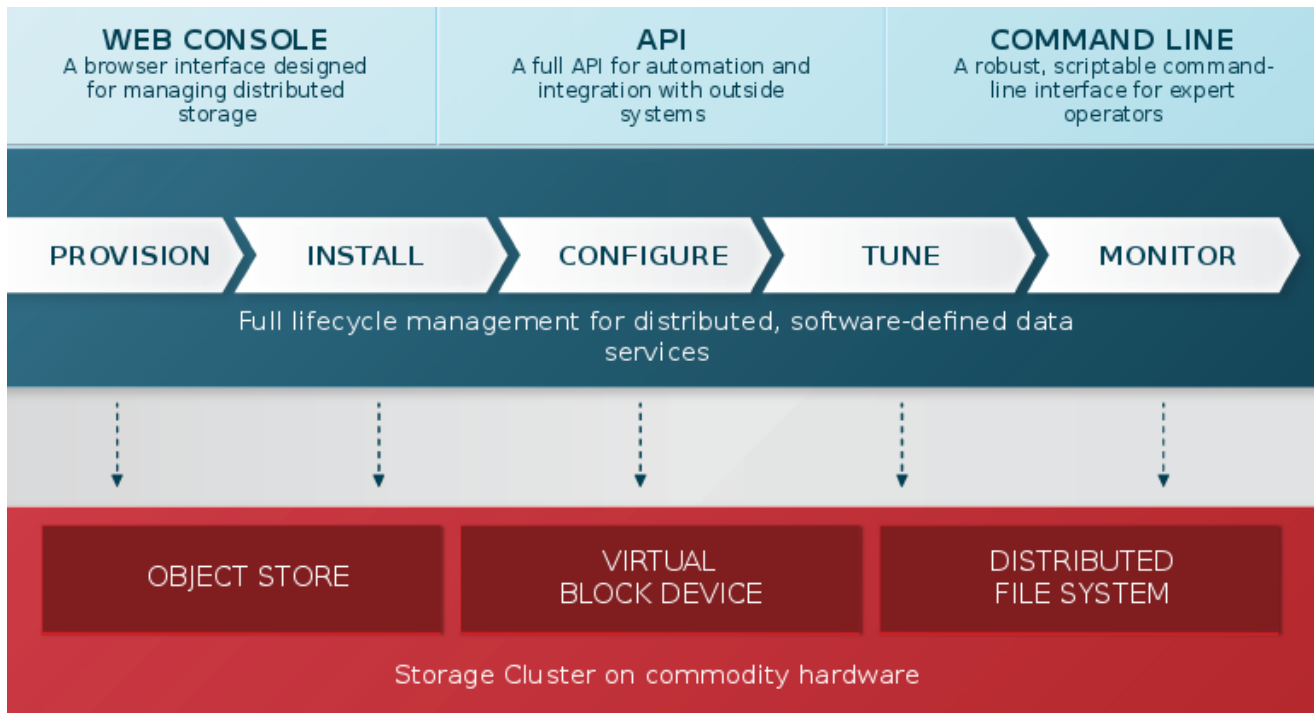


存储编排

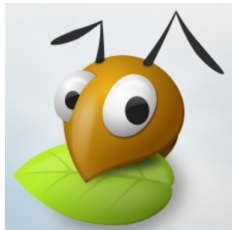
存储池化



存储编排



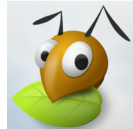
持续改进和创新



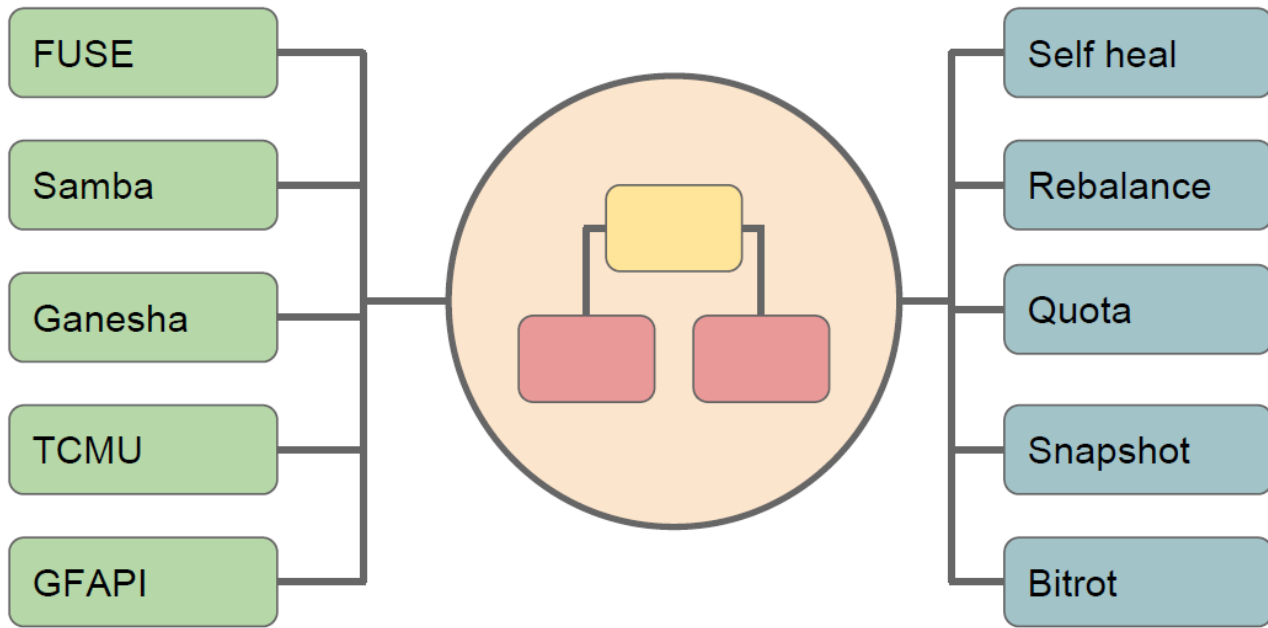
- ✦ CTDB / NFS-Ganesha ...
- ✦ GlusterD 2.0
- ✦ AFR → NSR
- ✦ Gluster as block storage
- ✦ Heketi
- ✦ GFProxy

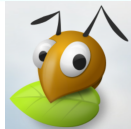
- ✦ RADOS QoS
- ✦ EC Overwrites
- ✦ Ceph MGR
- ✦ Dedup
- ✦ Multi-MDS
- ✦ BlueStore, ZetaScale



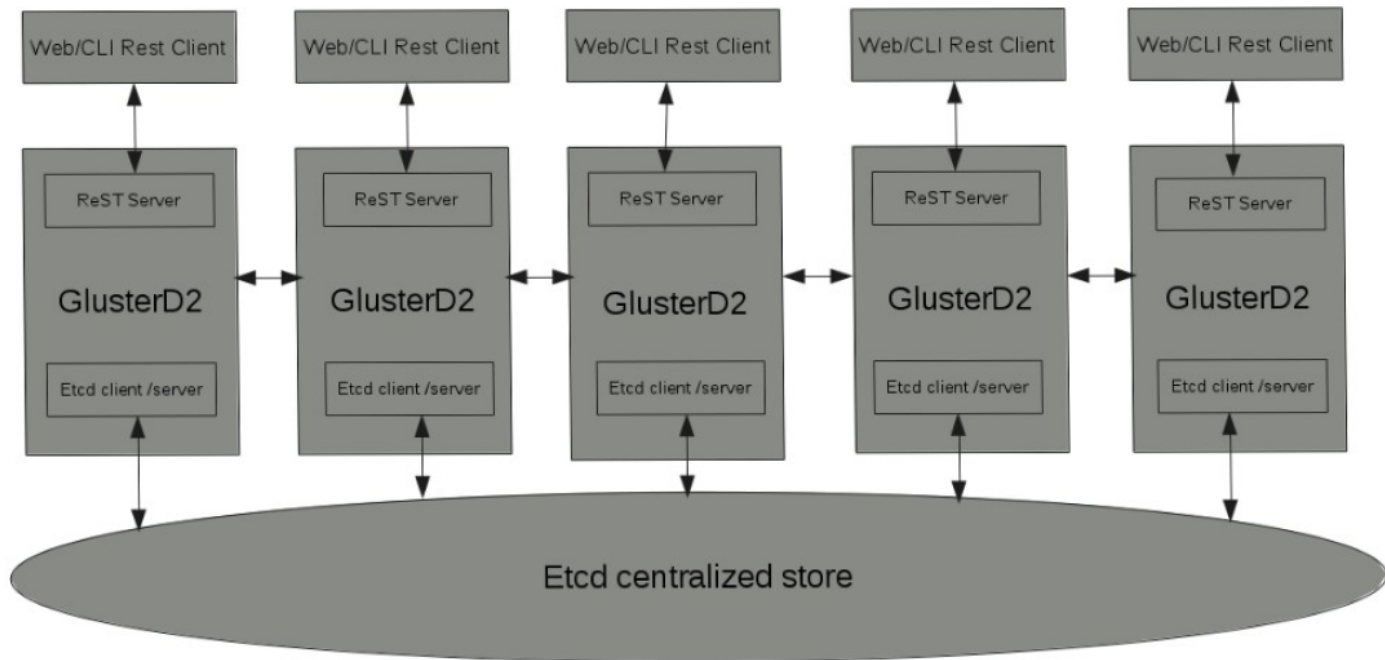


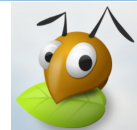
CTDB / NFS-Genesha ...



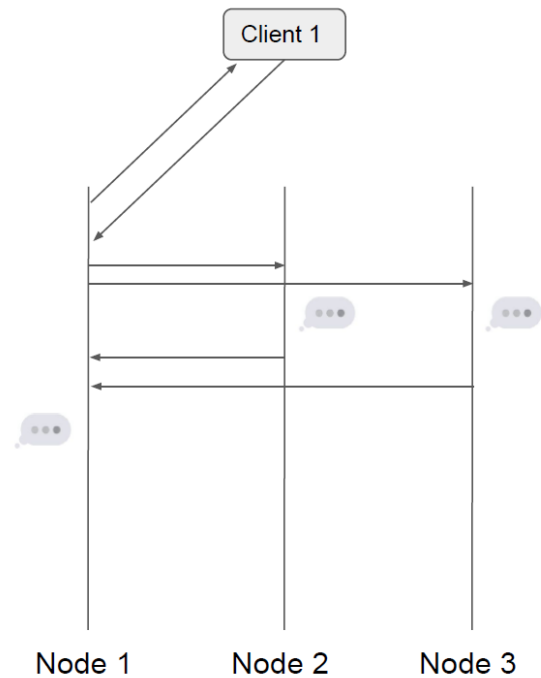
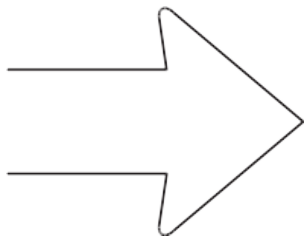
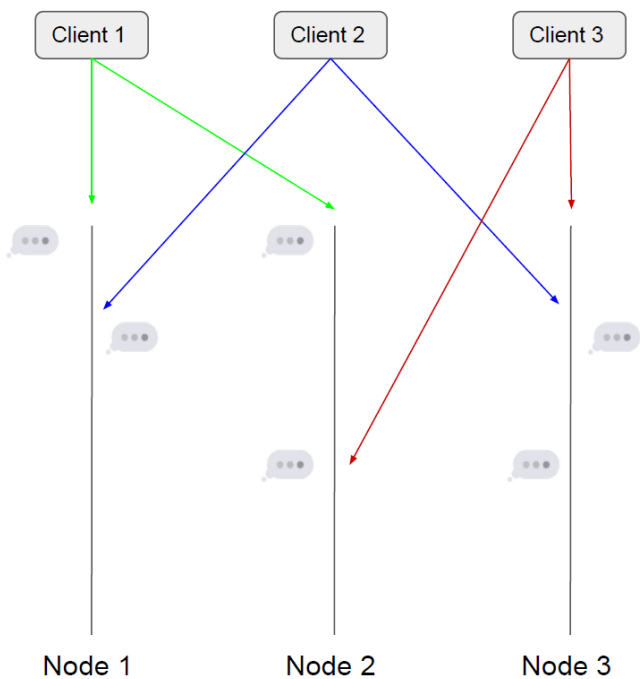


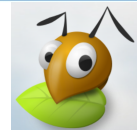
GlusterD 2.0





AFR → NSR

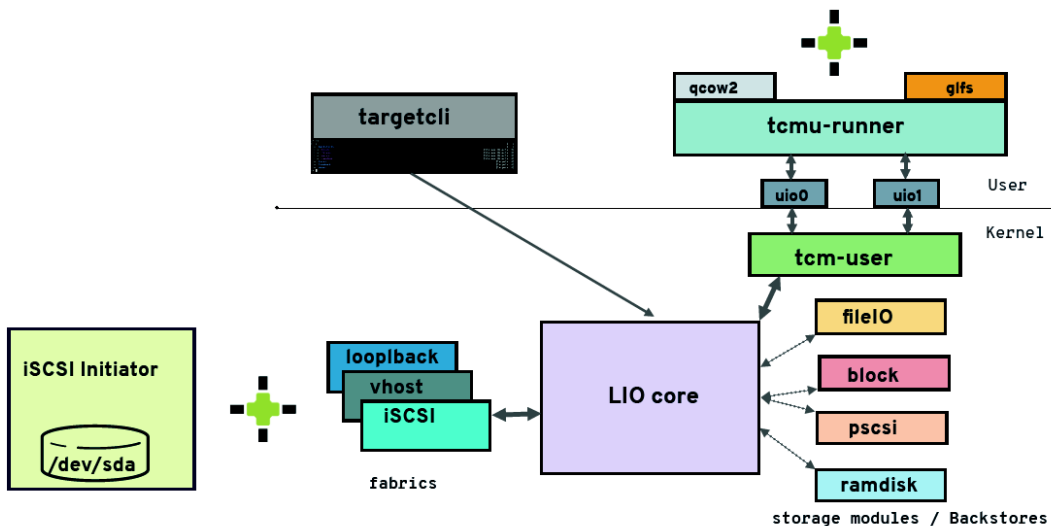
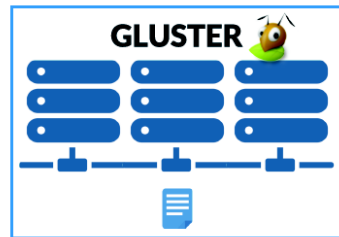


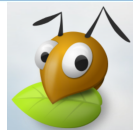


Gluster as Block Storage

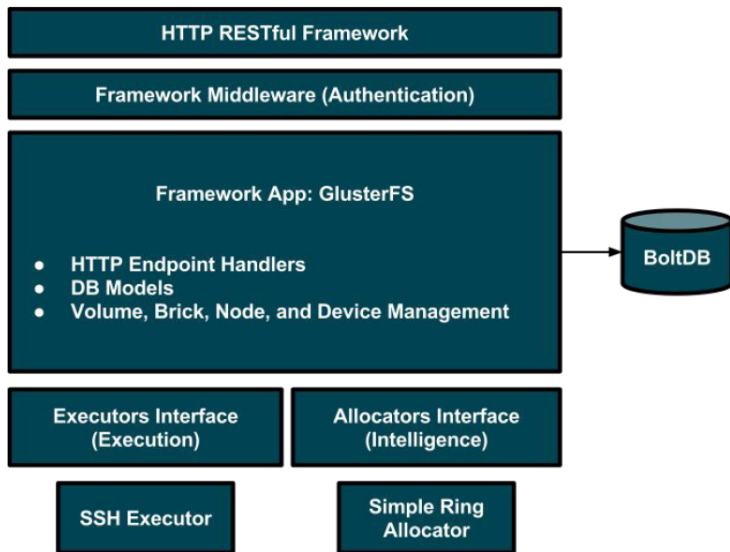
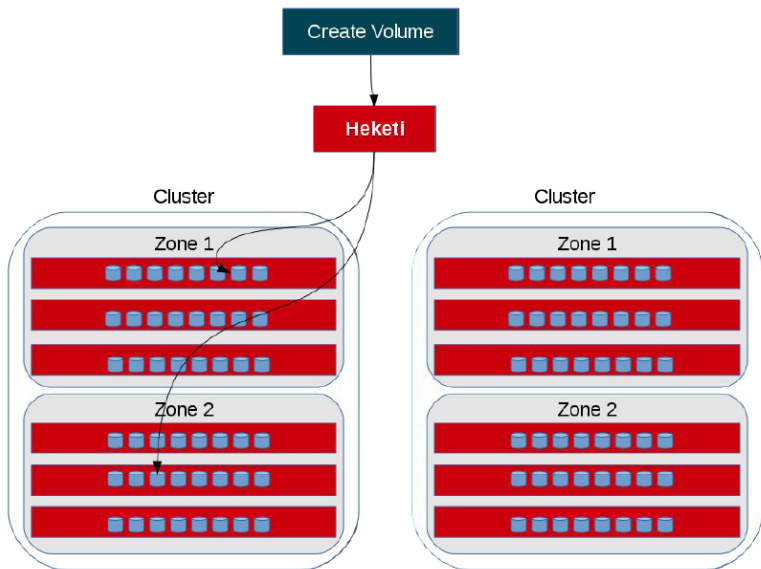


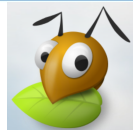
ARCHITECTURE



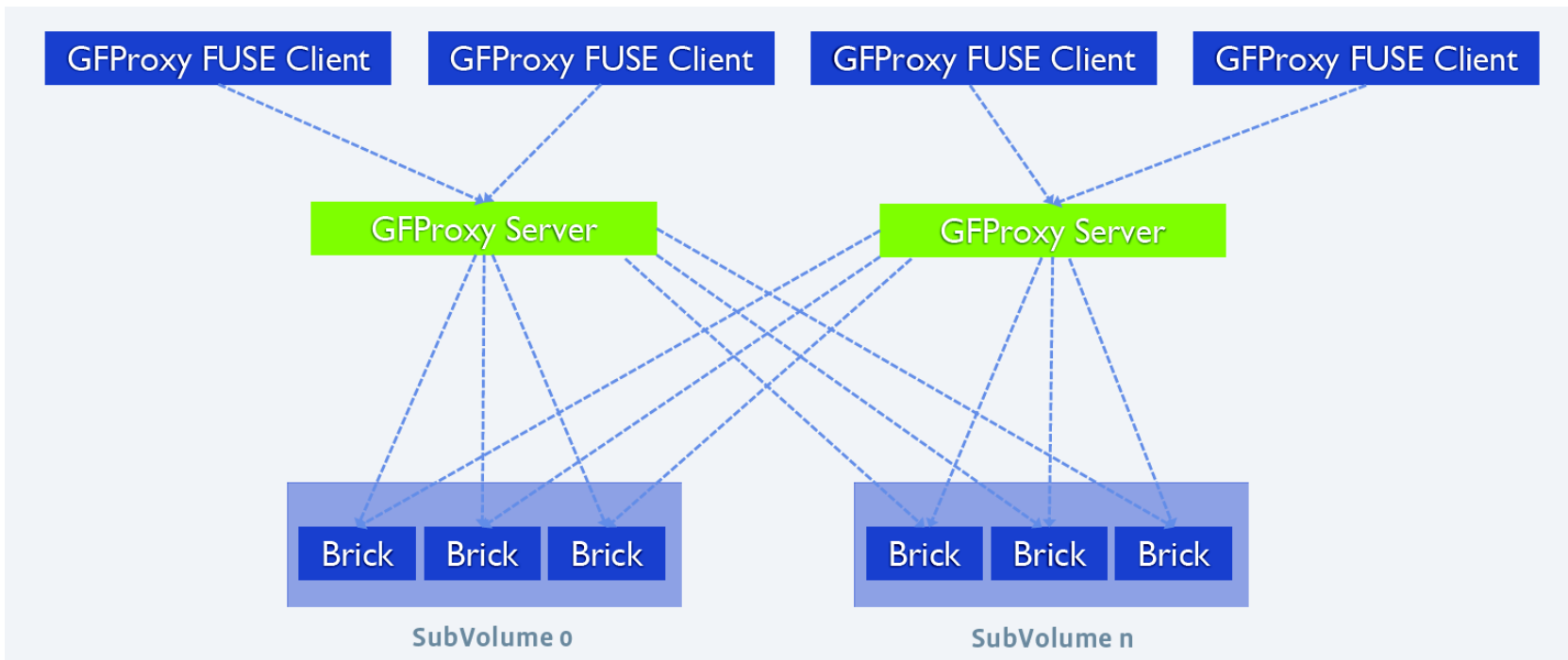


Heketi





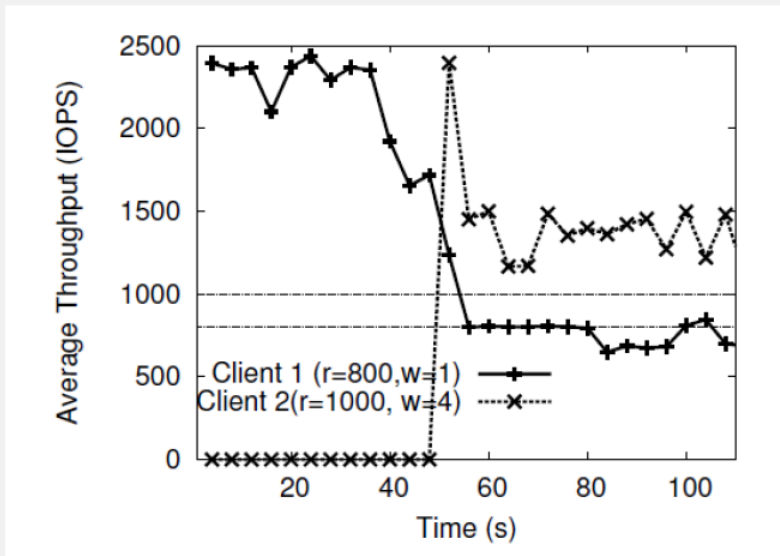
GFPProxy





RADOS QoS

- 基于 dmClock 算法实现
- QoS 参数
 - 预留 IOPS
 - 上限 IOPS
 - 权重
- QoS 策略可基于不同维度设置
 - IO 类型 (recovery, scrub,...)
 - 存储池
 - 客户端





EC Overwrites

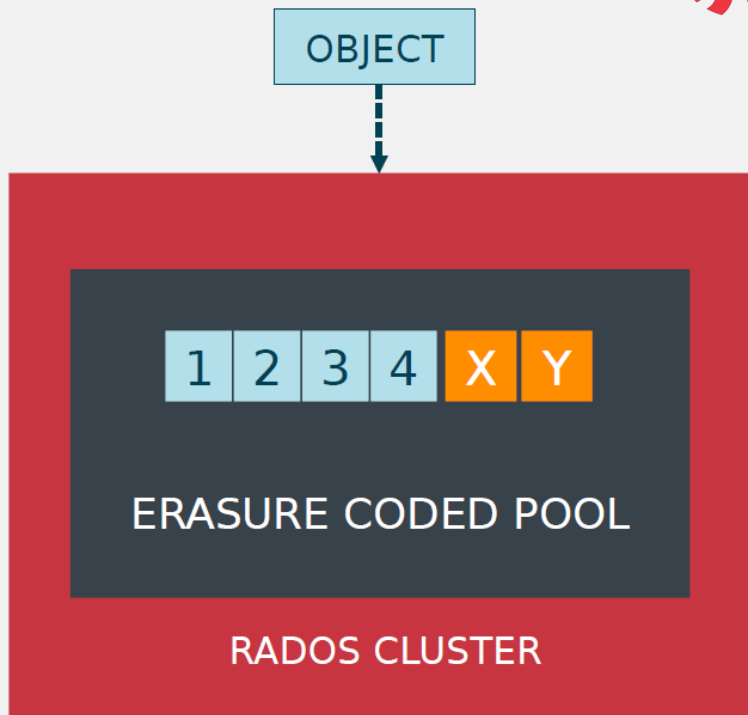
- 之前的纠删码池的实现只支持追加写
优点：简单、稳定；
缺点：不能直接给块设备和文件系统来用

- 最新的纠删码池的实现支持覆盖写，可直接为块设备和文件系统所用

关键技术点：

- 两阶段提交
- 小 IO 写的性能优化
- BlueStore 的高效的“move ranges”操作

优点：大幅降低 TCO





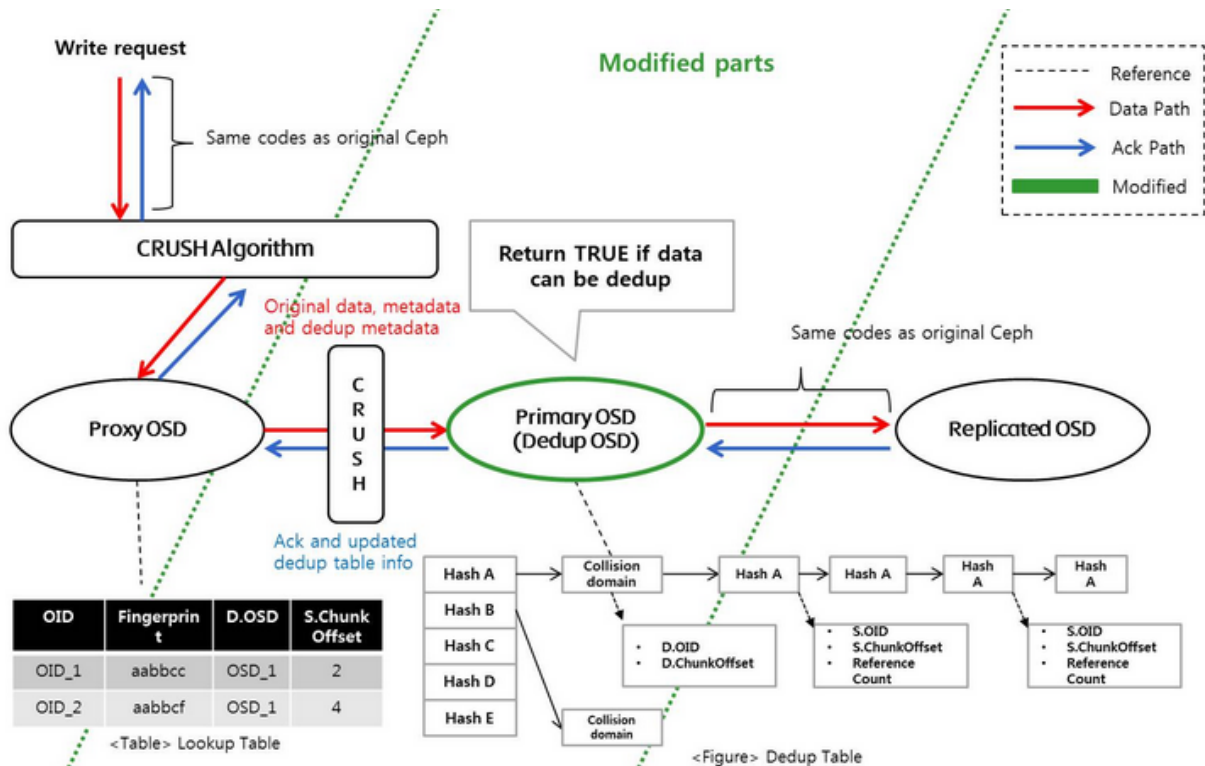
Ceph MGR

- 将存储管理面的功能从 MON 中剥离出来，由单独的进程 MGR 来实现，MON 只实现最核心的集群功能
- MGR 中进一步增强运维管理功能，利用 Ceph 内部机制提取信息提高效率，并支持运维人员扩展自定义插件
- 支持高级管理功能，支持 Rest API, 便于与第三方管理平台对接



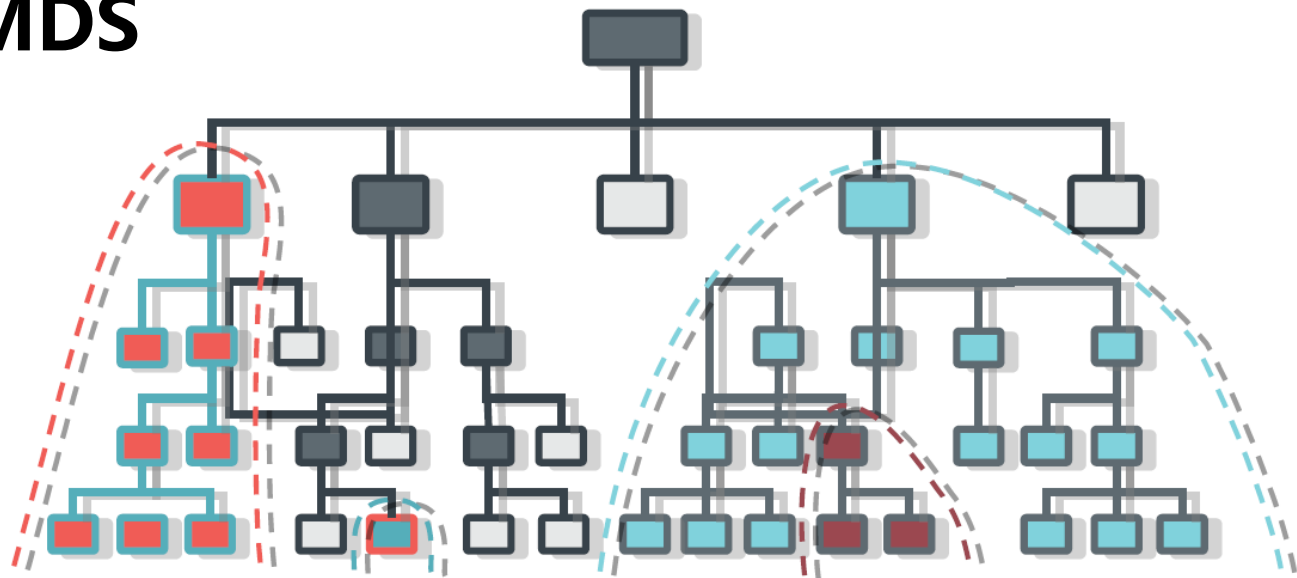


Dedup



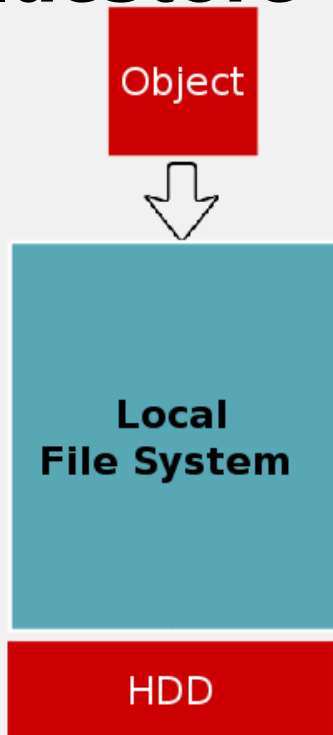


Multi-MDS





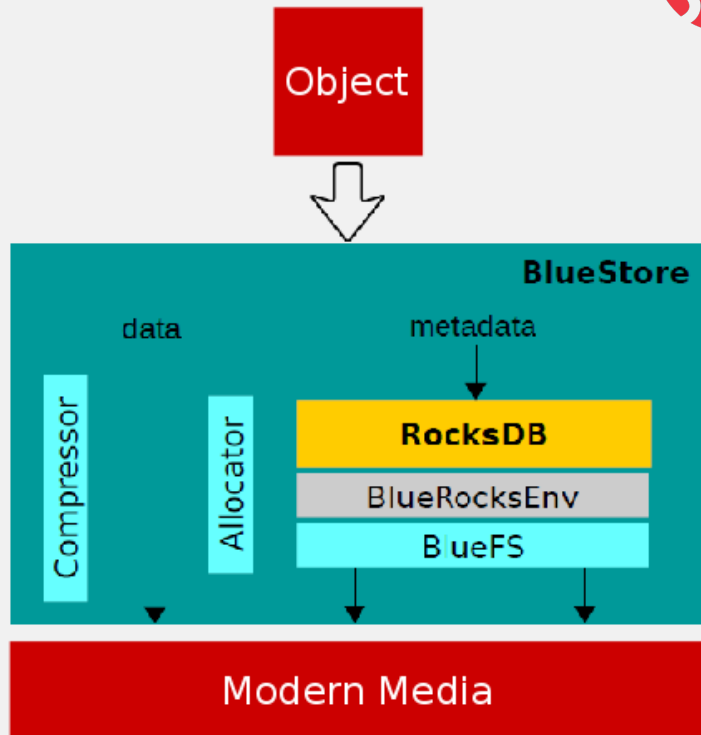
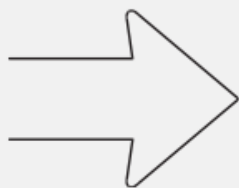
BlueStore



所有对象存储在平坦的命名空间里面，每个对象包括：

- 唯一标识
- 二进制数据
- 元数据

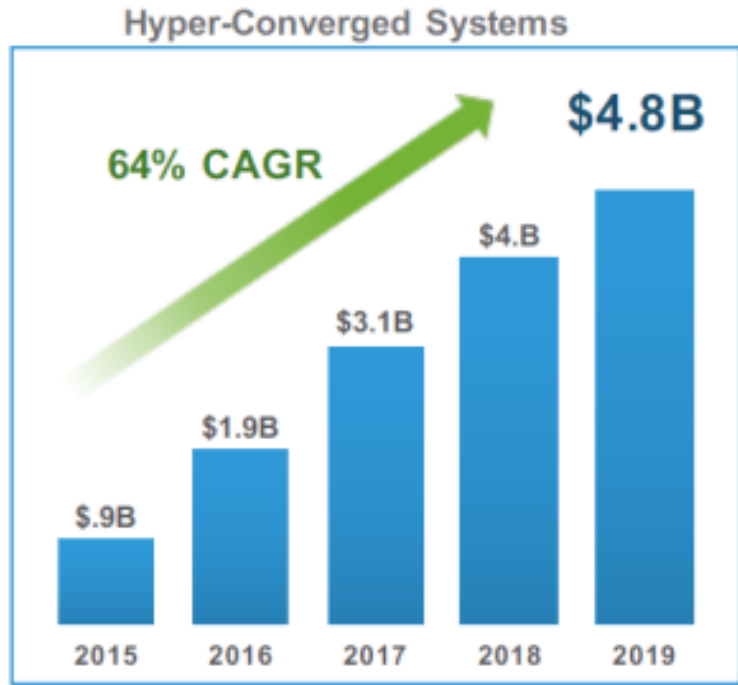
通常以 XFS 作为本地文件系统



超融合走向白热化

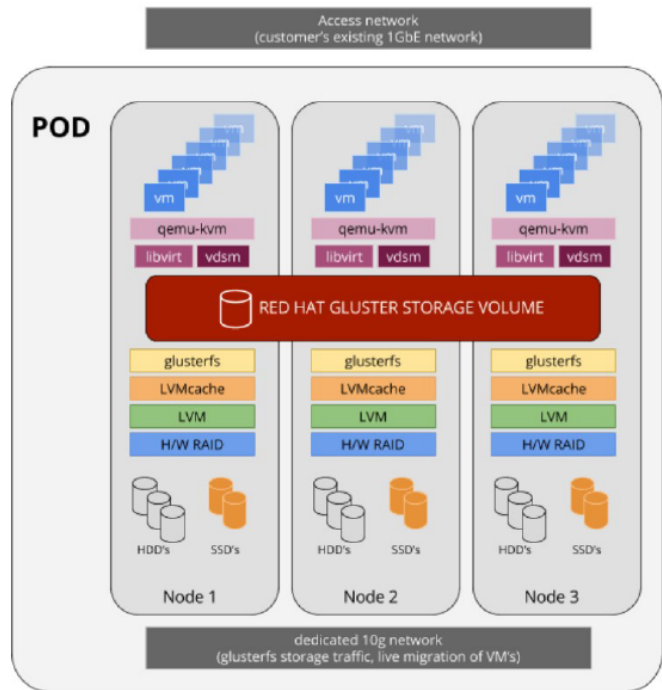
“Hyperconverged Systems collapse core storage and compute functionality into a single, highly virtualized solution. A key characteristic of hyperconverged systems that differentiate these solutions from other integrated systems is their ability to provide all compute and storage functions through the same server-based resources”

31 Mar 2016 - IDC



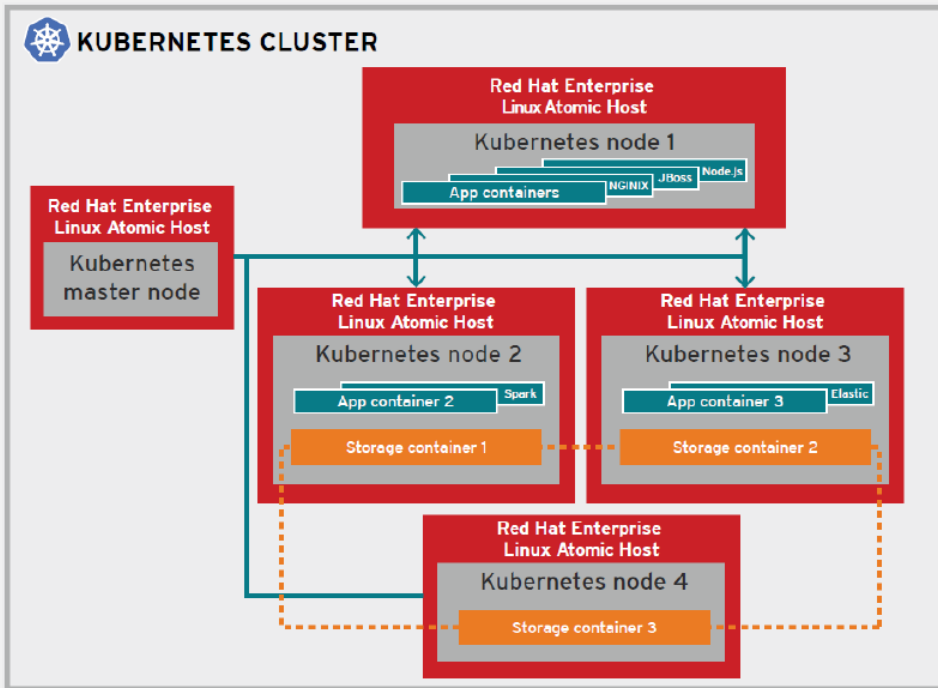
Source: IDC Hyperconverged Systems 2015-2019 Forecast, February, 2016

基于KVM虚拟化的超融合架构



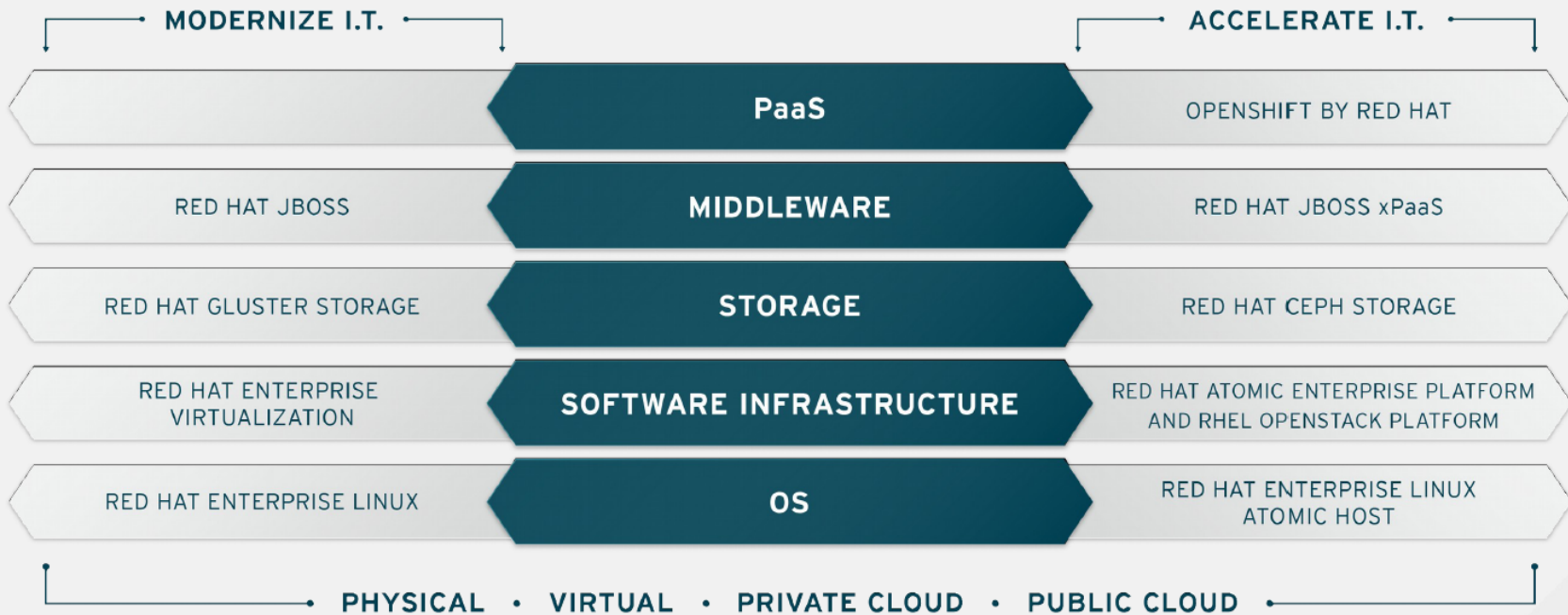
- Integrated management and monitoring
- At-a-glance dashboard
- automated(ansible) or manual deployments
- Systemd based resource control
- Direct-io support
- More granular virtual machine images (sharding)
- Data locality
- Multi-threaded self heal
- UI integration
- Sharding aware geo-replication
- DR support with geo-replication
- Cache acceleration with lvmcache

容器原生存储



- “Containerized” Gluster or Ceph
- Seamless, reliable file storage managed by Kubernetes
- No rewrite for POSIX compatibility
- Dev “owns” storage - can bypass Ops to deploy and maintain it alongside apps

红帽开源技术堆栈





THANK YOU



plus.google.com/+RedHat



facebook.com/redhatinc



linkedin.com/company/red-hat



twitter.com/RedHatNews



youtube.com/user/RedHatVideos