



CEPHALOCON APAC 2018

THE FUTURE OF STORAGE

22-23 March 2018 | BEIJING

Build HA NFS Cluster Based CephFS





-
- A vertical grey bar with four white circular markers, each connected to a list item by a thin grey line.
- 01 Why
 - 02 What
 - 03 How
 - 04 Future



Why

NFS(Network File System):

- 1) Widely used network protocol
- 2) Many enterprises still heavily depend on NFS to access their data from different operating systems and applications
- 3) Traditional storage system(SAN/NAS) is being replaced with SDS

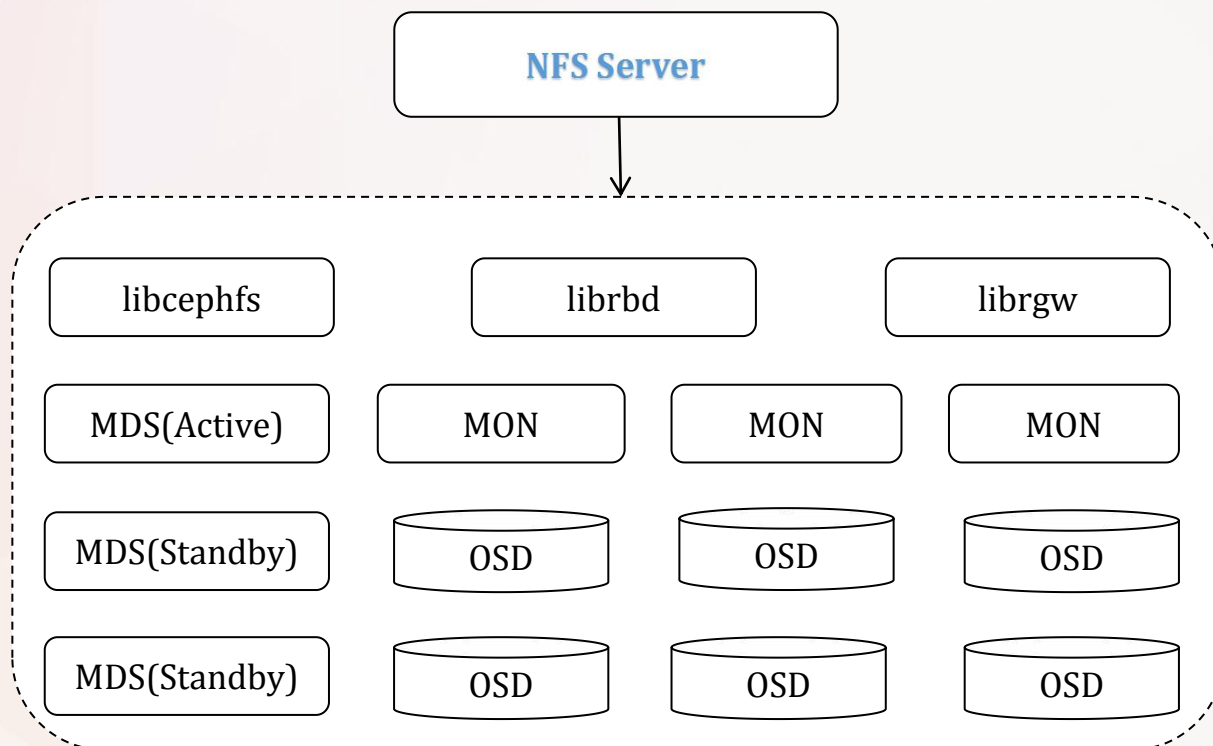


Why

CephFS(Ceph File System):

- 1) Ceph widely used
- 2) Ceph uniquely delivers object, block, and file storage in one unified system
- 3) Ceph-fuse need to be installed on client, which limits its application scope

What





What

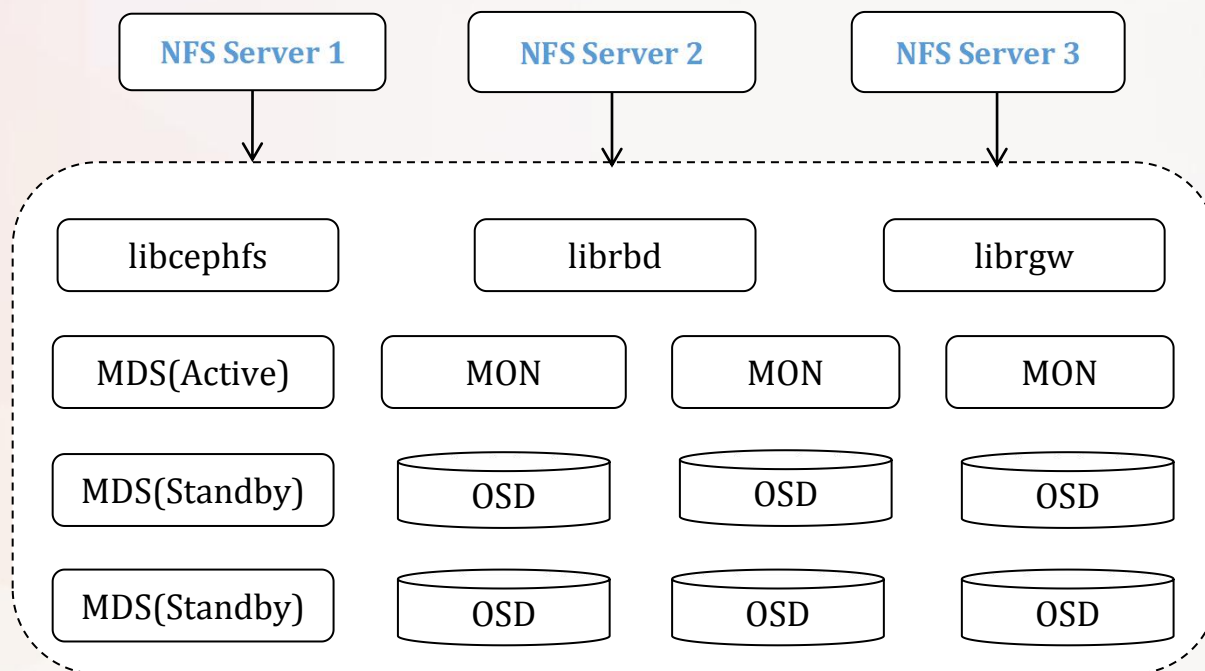
NFS Server mandatory requirements:

- 1) Reliability
- 2) Performance
- 3) Scalability

What

Reliability:

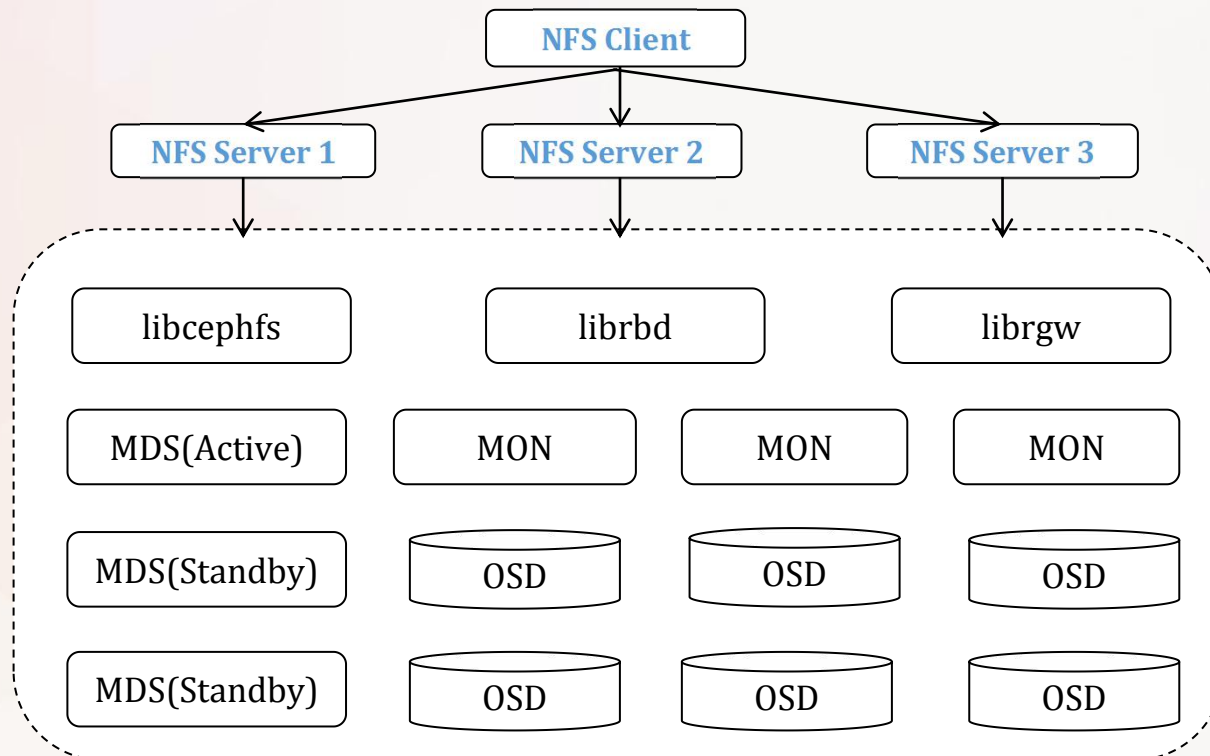
Multiple active NFS servers(Redundant NFS servers)



What

Performance:

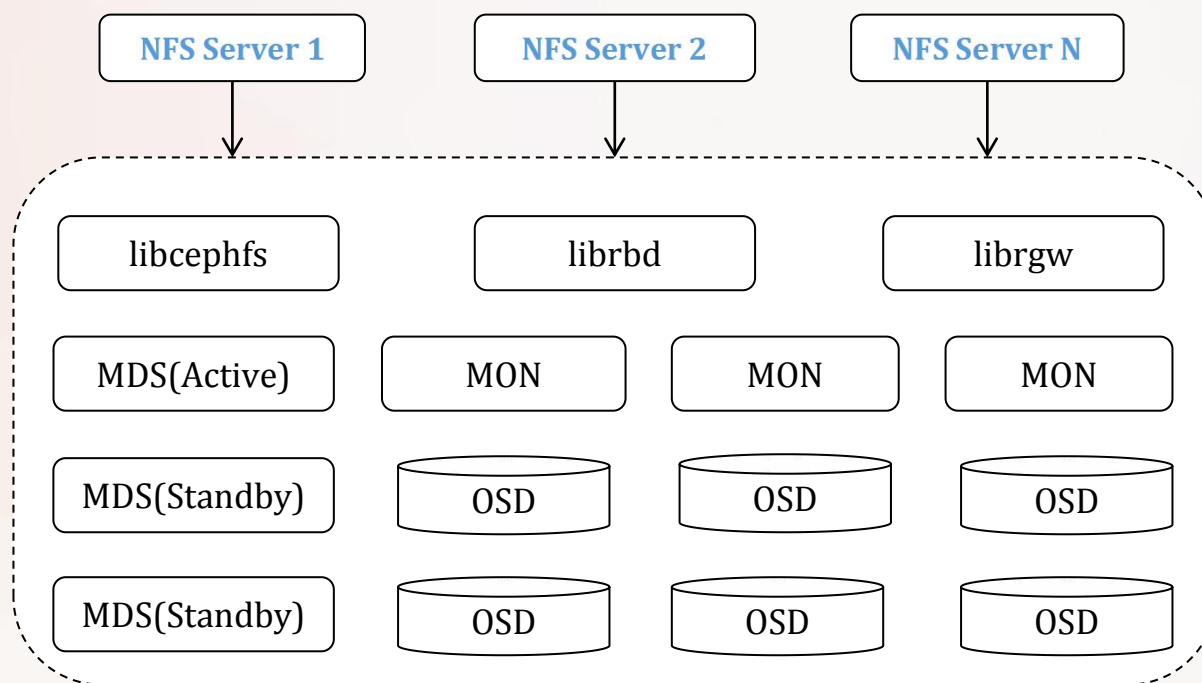
- 1) Ceph and CephFS
- 3) Multiple active NFS servers(Load Balance)



What

Scalability:

- 1) Ceph and CephFS
- 2) Add more NFS Servers





What

HA NFS Cluster:

- 1) NFS server based CephFS
- 2) HA
 - No SPOF(Single Point Of Failure)
 - LB(Load Balance)



How

Kernel NFS Server:

- 1) Performance – User mode can be slow
- 2) Hard to debug kernel module



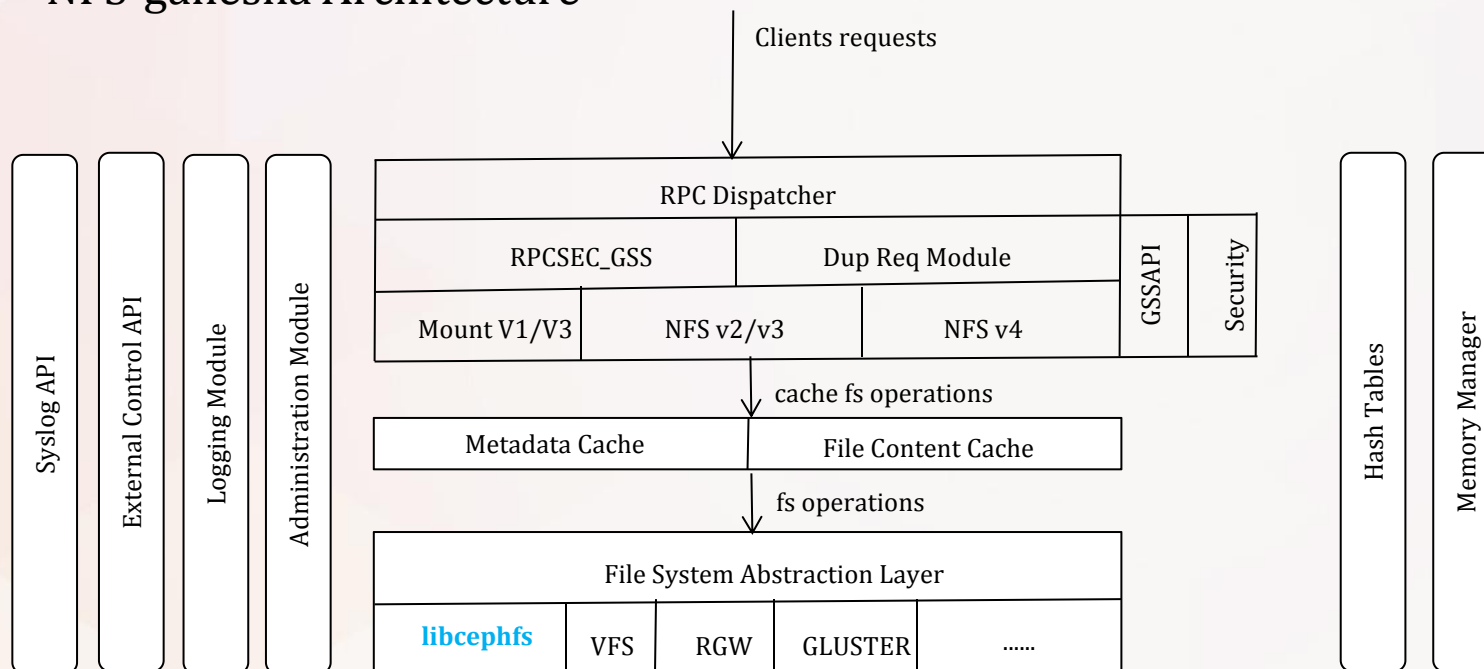
How

User Space NFS Server(NFS-ganesha):

- 1) Flexible and Plug-able FSAL
- 2) Clustering becomes natural and easy
- 3) Easy restarts, failover, failback implementation

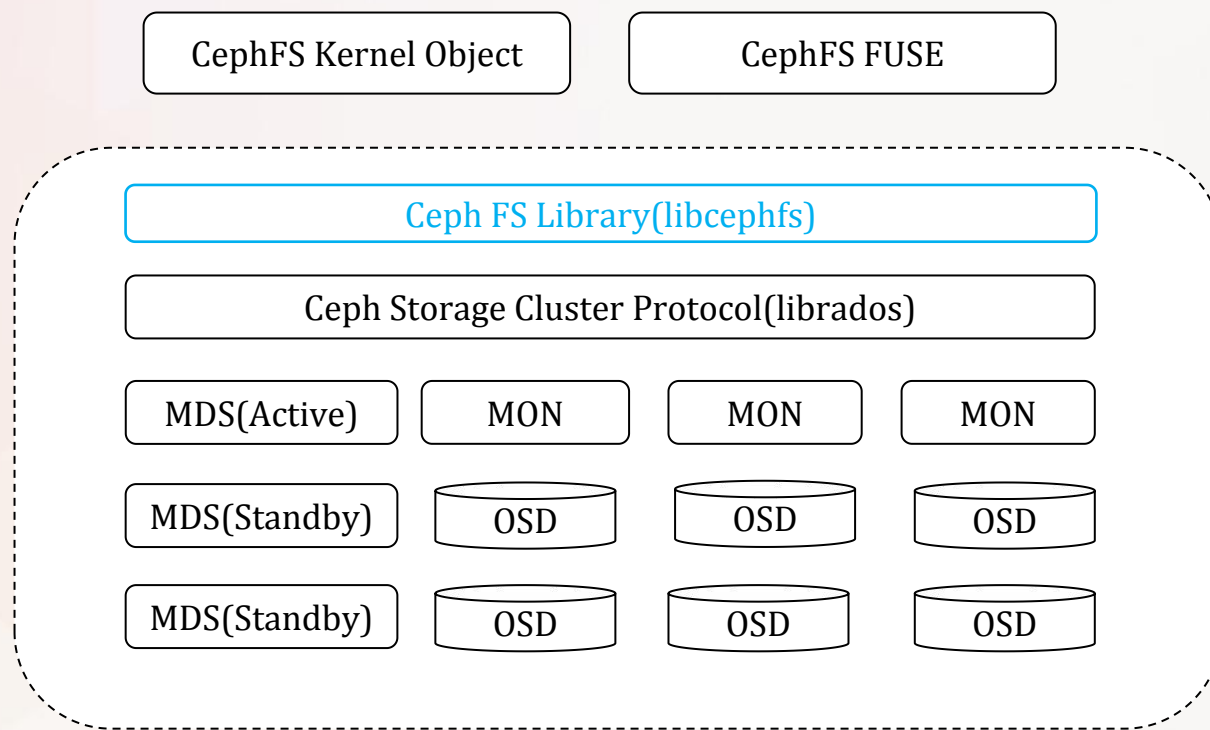
How

NFS-ganesha Architecture



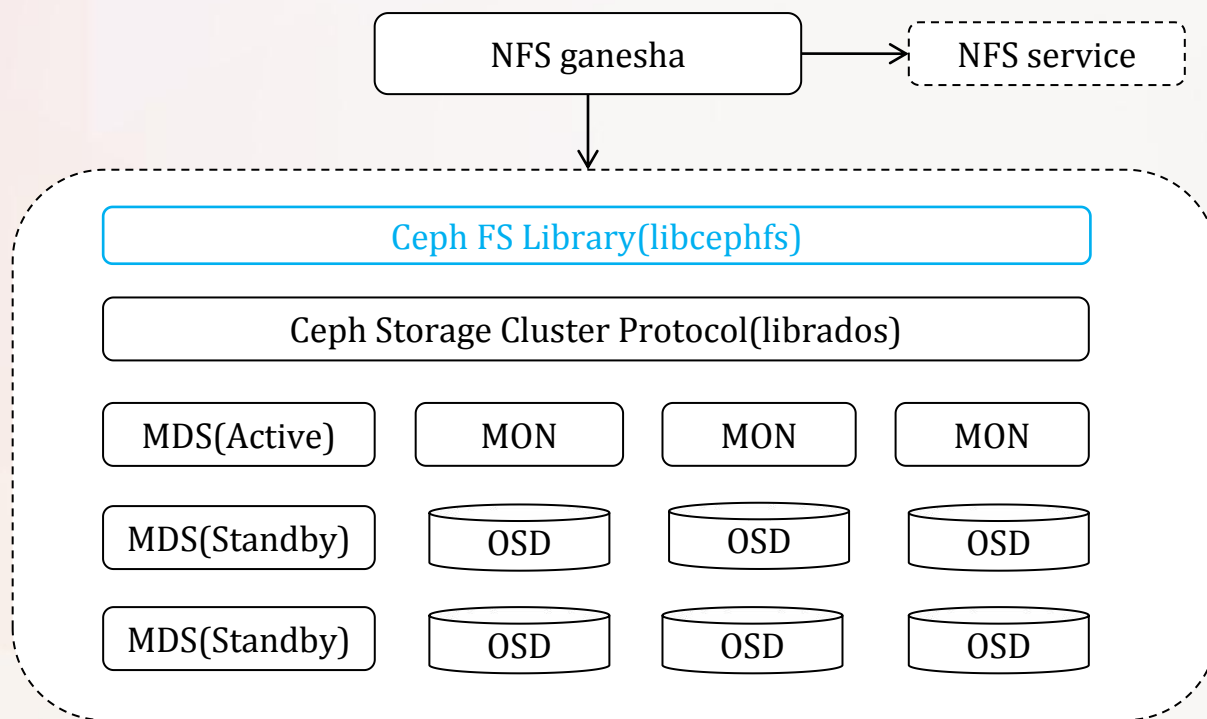
How

CephFS Architecture



How

NFS ganesha based CephFS





How

HA(High Availability) Solution:

- 1) Pacemaker + corosync
- 2) CTDB + LVS
- 3) HAproxy
- 4) Piranha(RedHat)
- 5) Heatbeat + mon
- 6) Keepalive

.....



How

CTDB:

CTDB is a cluster implementation of the TDB database used by Samba and other projects to store temporary data.

CTDB features:

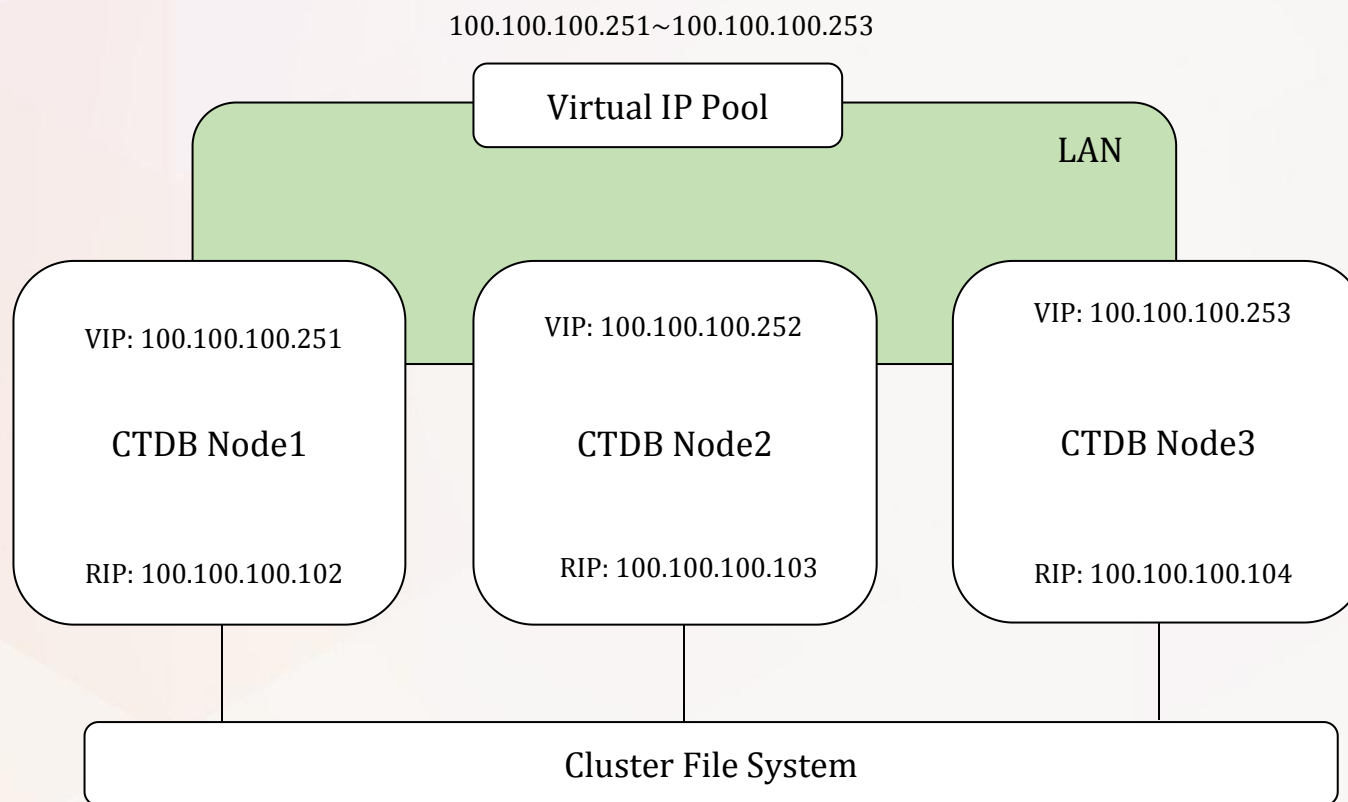
- 1) provides a TDB that has consistent data and consistent locking across all nodes in a cluster.
- 2) CTDB is very fast.
- 3) In case of node failures, CTDB will automatically recover and repair all TDB databases that it manages.
- 4) provides HA features such as node monitoring, node failover, and IP takeover.

CTDB requirements:

a clustered filesystem, it must be mounted and available on all nodes.

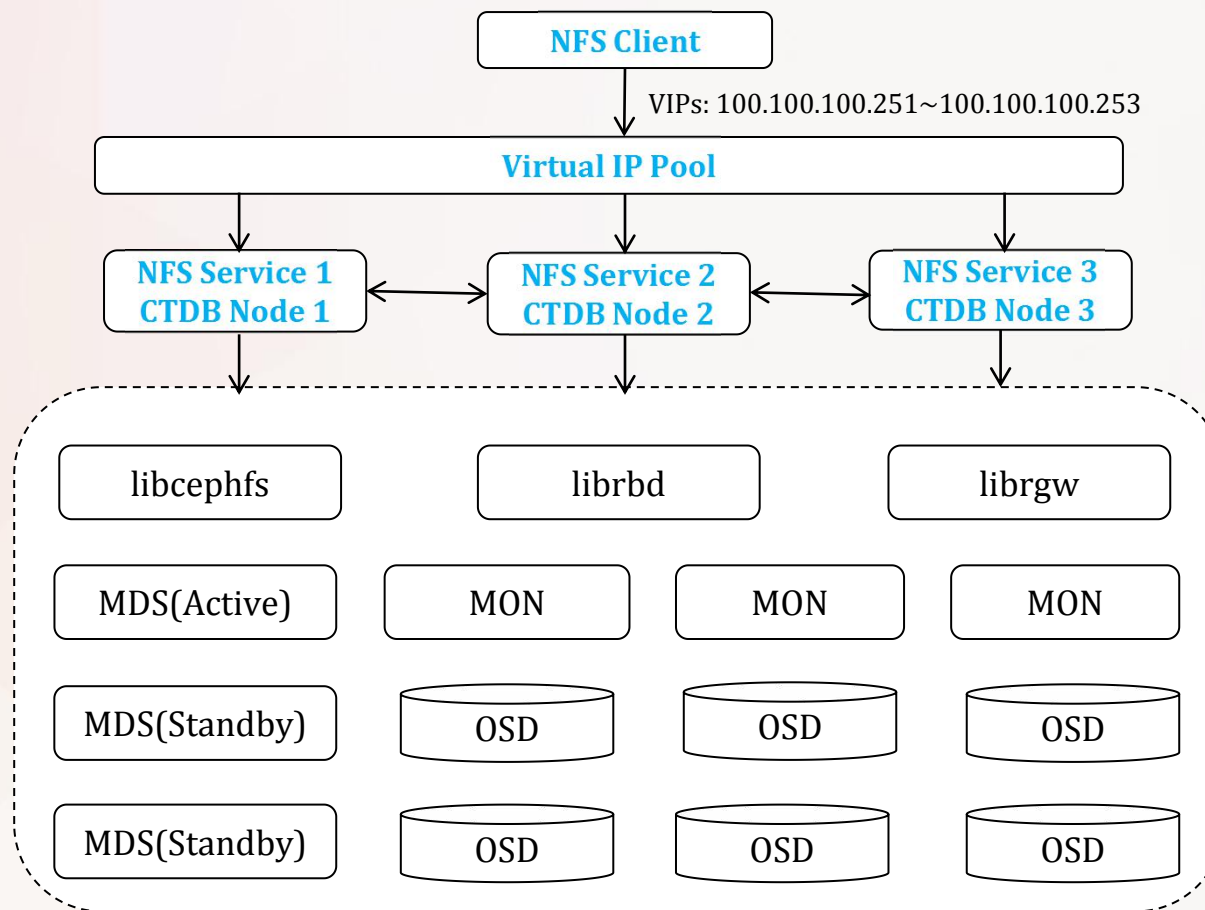
How

CTDB Architecture



How

NFS Service based CTDB





How

NFS Service based

CTDB:

- 1) Multiple VIP
- 2) Load Balance

Solutions:

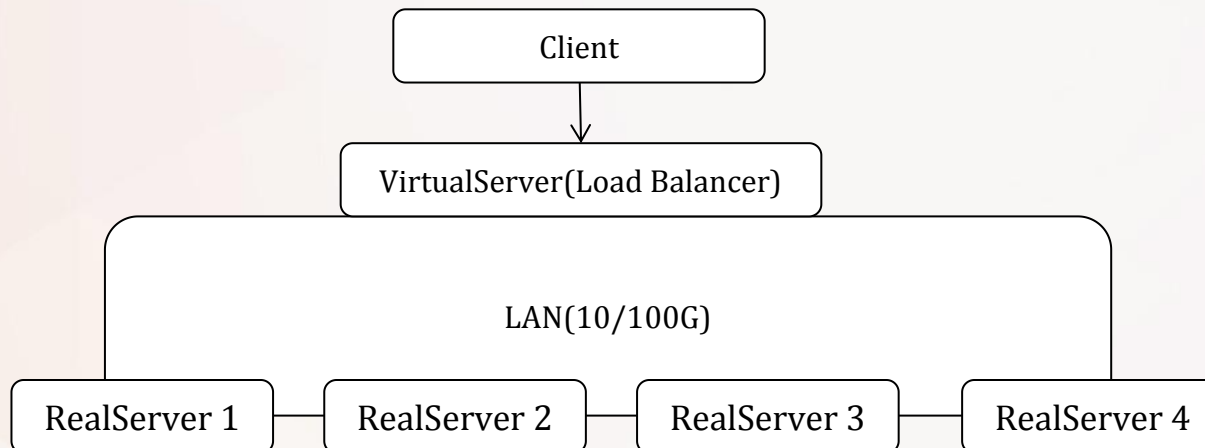
- 1) Round Robin DNS
- 2) LVS

How

LVS(Linux Virtual Server):

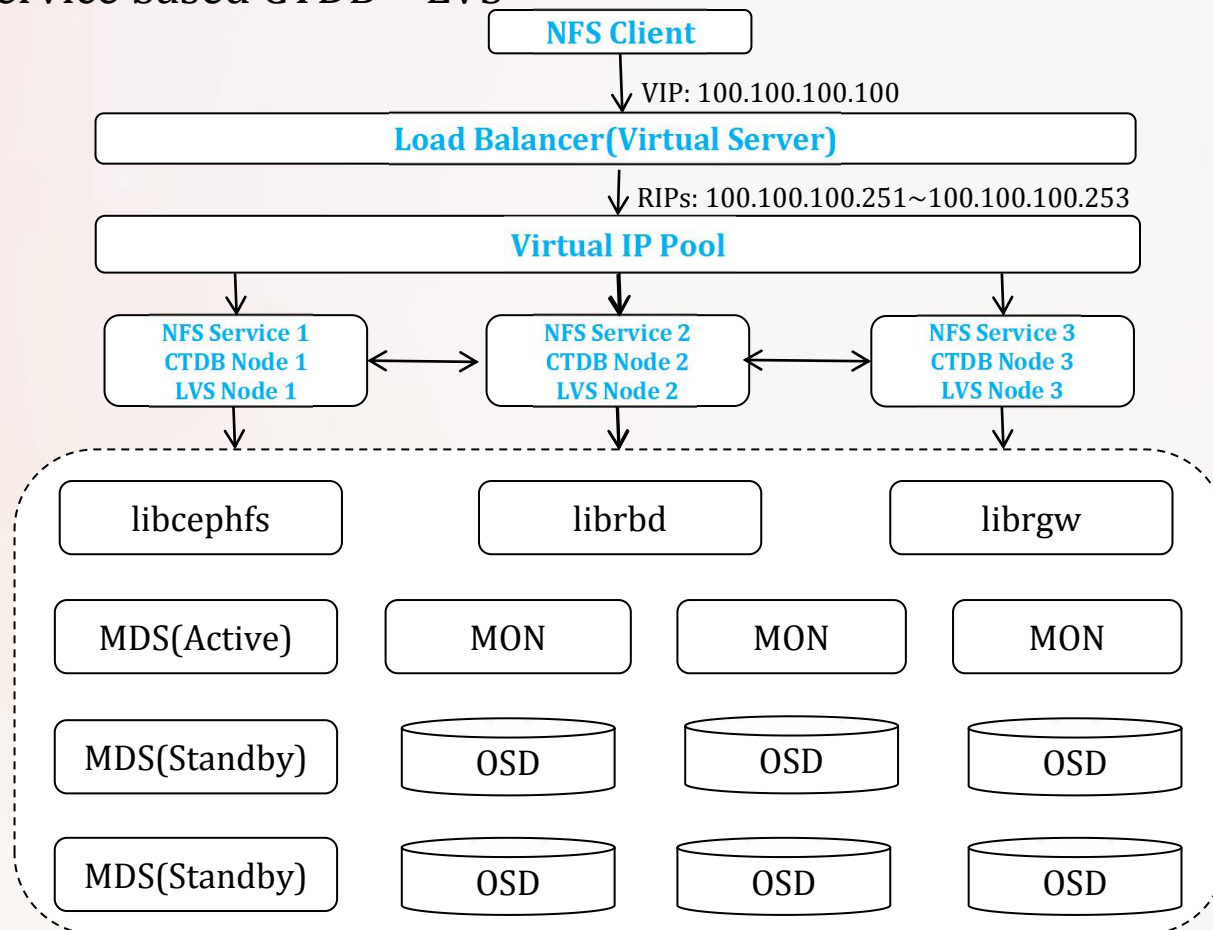
Virtual server is a highly scalable and highly available server built on a cluster of real servers.

It was in the official kernel 2.6.10 released on December 25, 2004.



How

NFS Service based CTDB + LVS





Future

- 1) Performance
- 2) CTDB depends on a clustered file system(CephFS).