SUSE Enterprise Storage 4

Powered by Ceph

전철민 과장
Technical Specialist
chris.chon@suse.com

Agenda

- Market Status
- SUSE Solution Offering
- SES Architecture
- What SUSE provides in detail
- Partnerships
- Where to use
- Sizing Guide
- Success story

Market Status

Enterprise Data Storage Needs

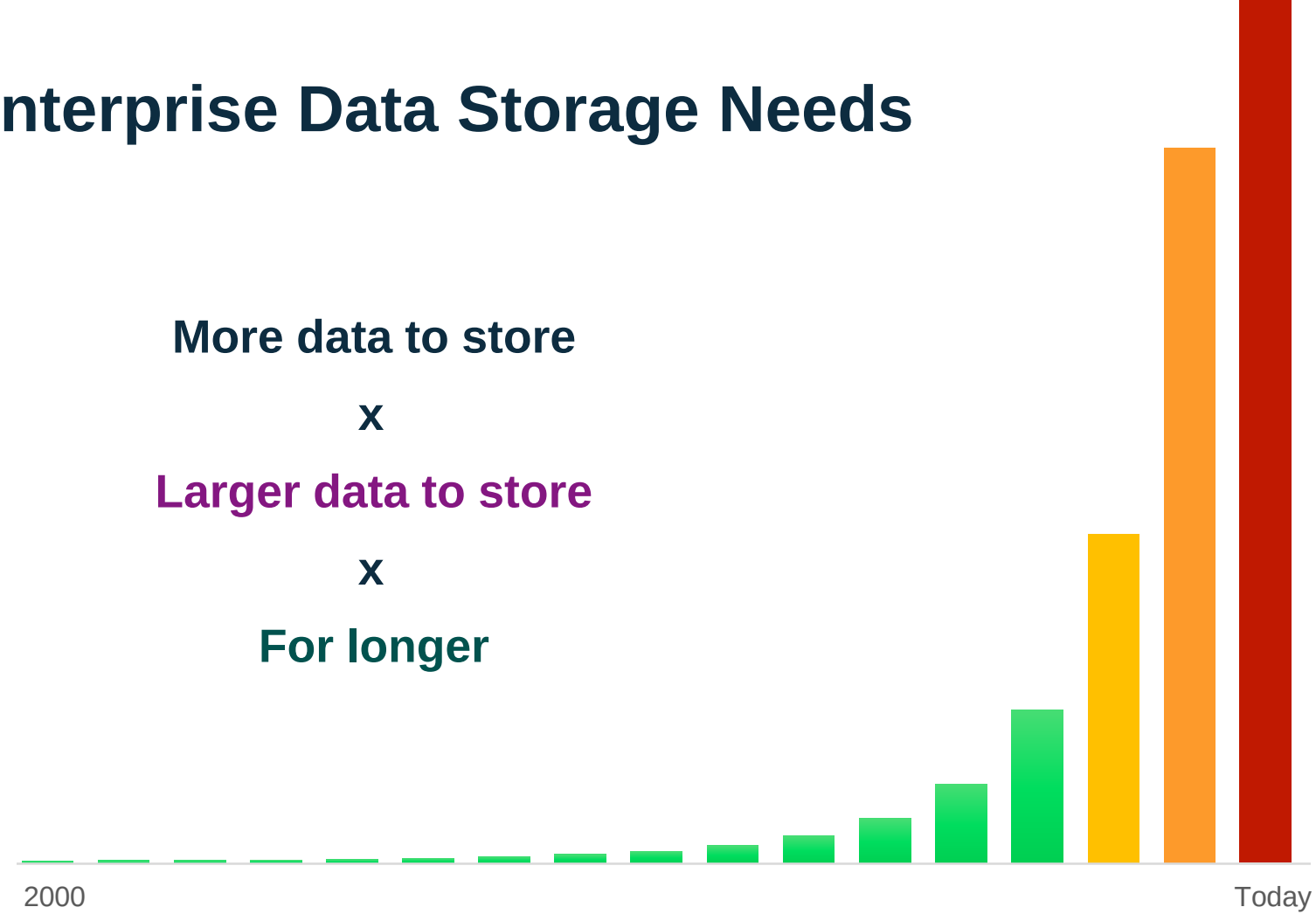
More data to store

x

Larger data to store

x

For longer



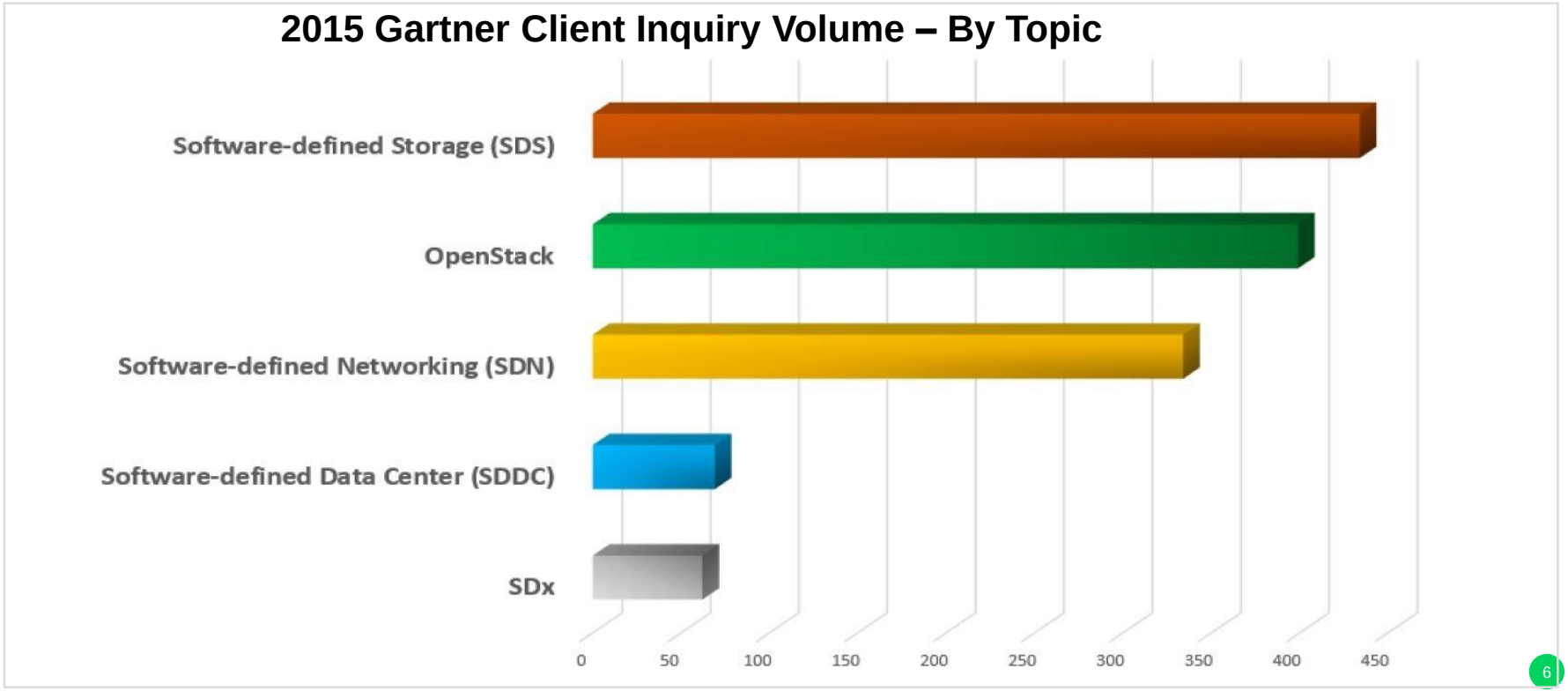
Emerging Market:

Software-defined Storage is Pricing Traditional Storage Out of the Market

Based on data from Wikibon and IT Brand Pulse

Know Your Market

Are Your Customers Asking YOU About Software-Defined Storage?



Making the Case for Software-Defined Storage in 2016 – Roger Cox/Gartner

Strategic Planning Assumption

By 2021, more than 80% of enterprise data will be stored in scale-out storage systems in enterprise and cloud data centers from 30% today.

© 2016 Gartner, Inc. and/or its affiliates. All rights reserved.

Top Reasons for Implementing SDS

Improving agility	• Faster time to value • Ability to manage across locations and classes of storage
Increase flexibility	• Multiple SDS alternatives to choose from • Select hardware platforms and hypervisors of choice
Decouple SW from HW	• Adopt industry standard HW platforms • Increases innovation in each layer
Reducing capex/opex	• Lower acquisition costs without sacrificing SLOs • Lower ongoing HW maintenance and SW support costs

© 2016 Gartner, Inc. and/or its affiliates. All rights reserved.

Gartner

SUSE Solution Offering

Limitations of Traditional Storage

Unable to Scale and Manage Data Growth

Expensive

Hard to adopt innovation

SUSE Enterprise Storage Lowers Costs

Lower Acquisition Cost

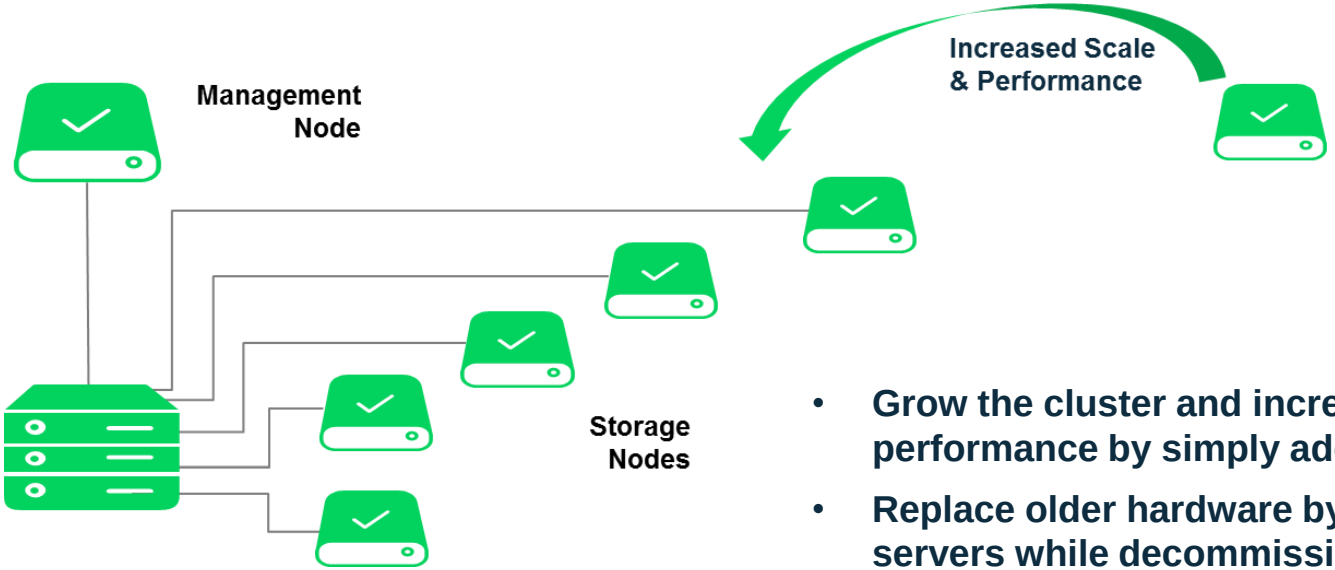
- Open source software running on **x86 hardware**
- Software is **priced by the node** instead of capacity, achieving price parity with public cloud

Lower Management Cost

- Self-managing and self-healing software reduces management overhead
- Rolling Upgrade

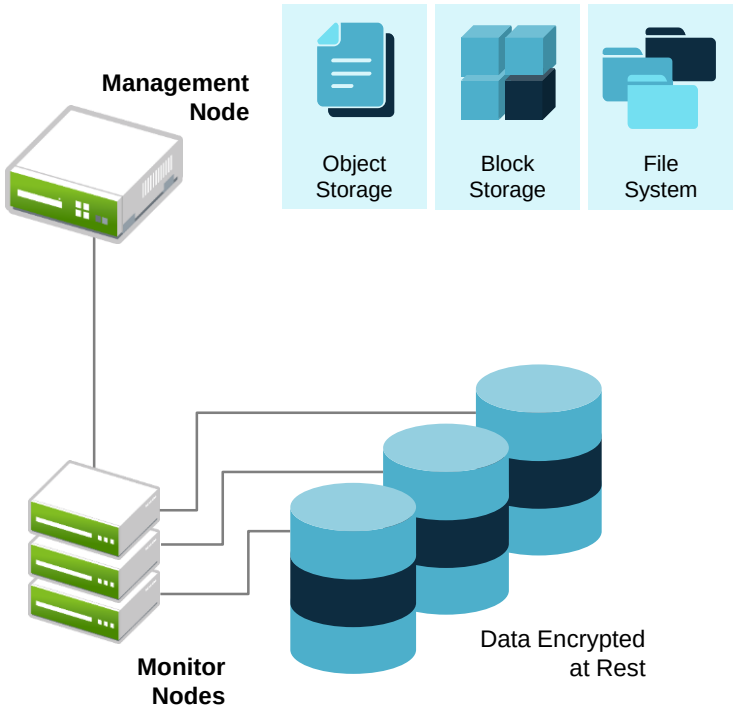
Accommodate Growth Easily

No More Forklift Upgrades



- **Grow the cluster and increase its performance by simply adding new servers.**
- **Replace older hardware by adding new servers while decommissioning older out-of-warranty servers.**

Truly Unified – Supports Object, Block and File System Storage in the Same Cluster



Automated management

Heterogeneous OS Access

Why SUSE and Software Defined Storage

• Open source Software defined storage platforms are **based on the Linux operating system**.

• **SUSE is a Linux OS pioneer and successful software vendor** with thousands of installations. Customers should expect to receive nothing less than expert support from SUSE for their software based storage.

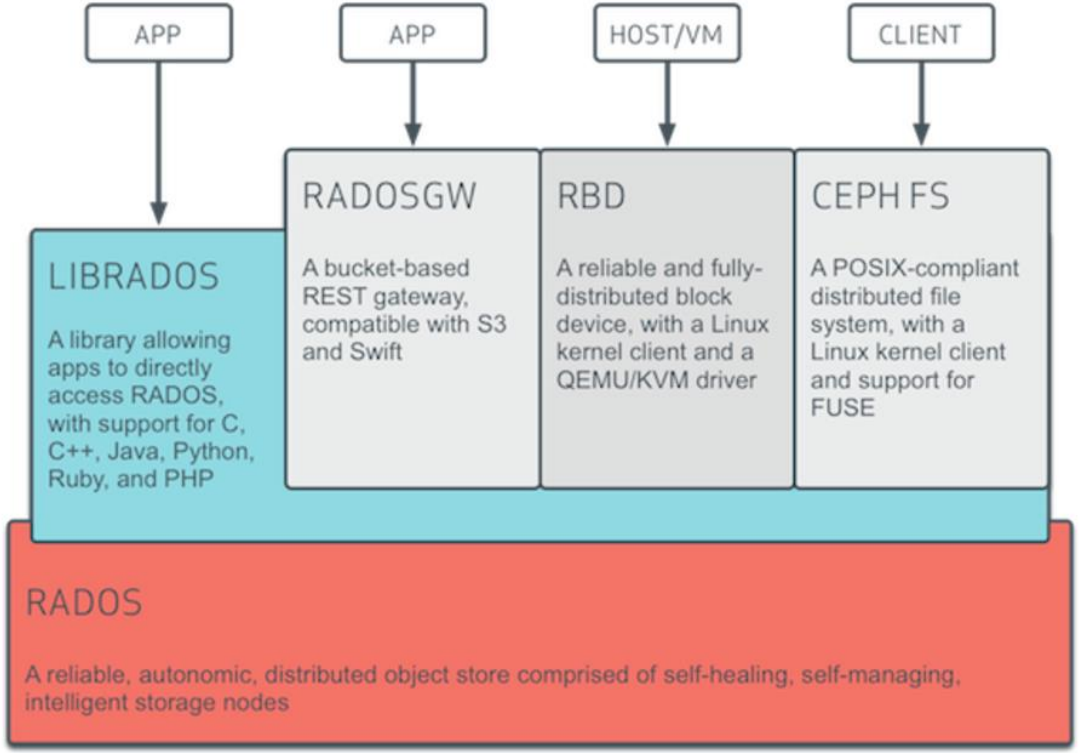
A Long History and leading innovation in Data

SUSE was the first to:

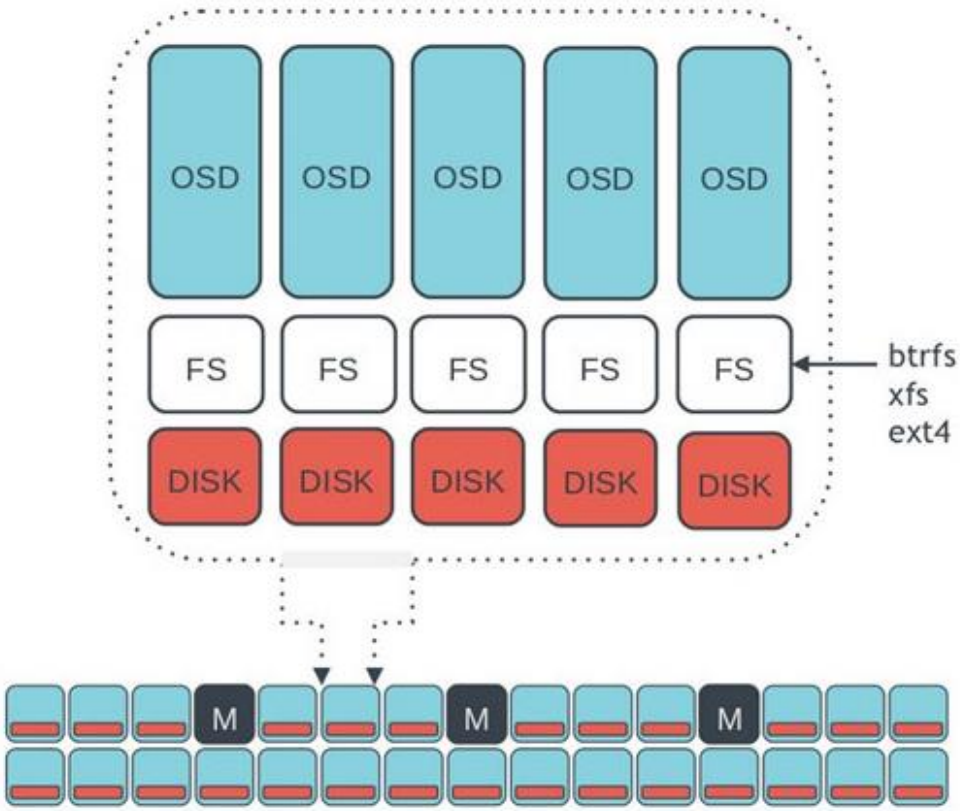
- Ship supported XFS to customers
- Ship supported ReiserFS
- Ship supported OCFS2
- Ship supported btrfs
- Ship SCALE-OUT ISCSI for Ceph
- Ship supported CephFS

SUSE Enterprise Storage Architecture

Ceph Architecture

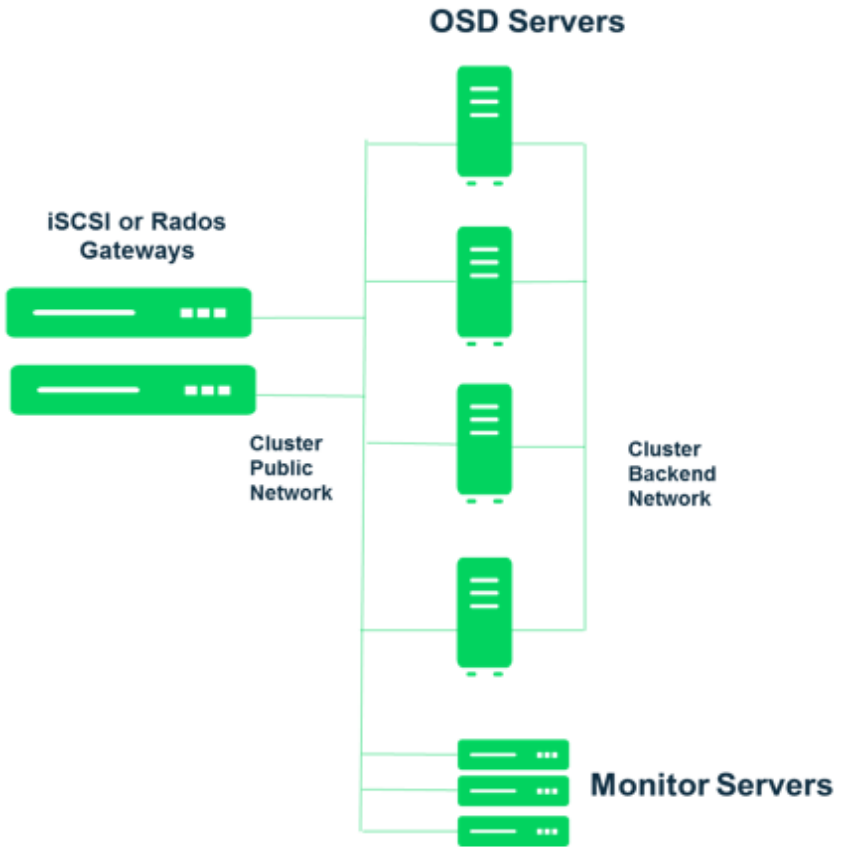


Ceph Architecture



Basic Network Diagram

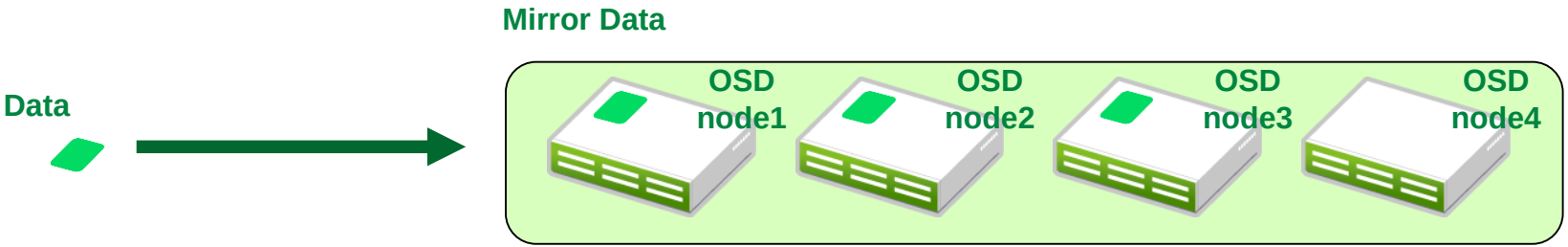
- Based on Ceph Jewel
- 4 OSD servers
- 3 Mon servers
- 1 or 2 Gateways
- 2 Network devices



What SES provides in detail

Lower Management Cost

- High Available Storage
- Scale-out without performance degrade
- Easy Upgrade
- Self-healing / Self-reparing

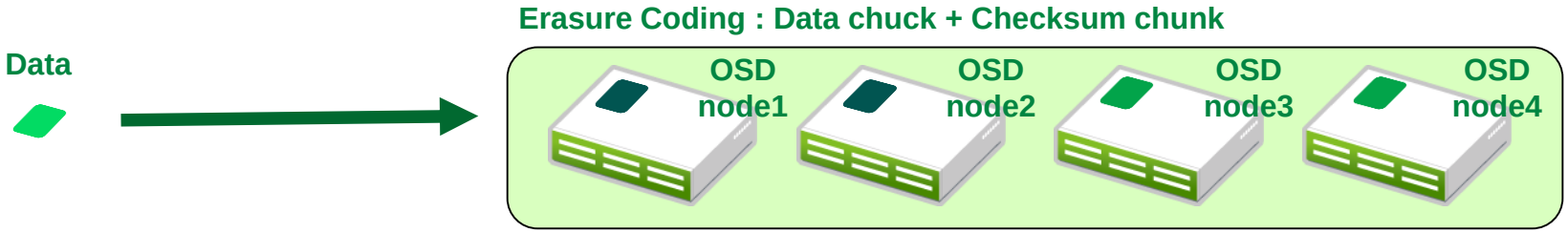


Various functions

- Snapshot : pool, rbd
- Thin provisioning(rbd)
- Cache tier

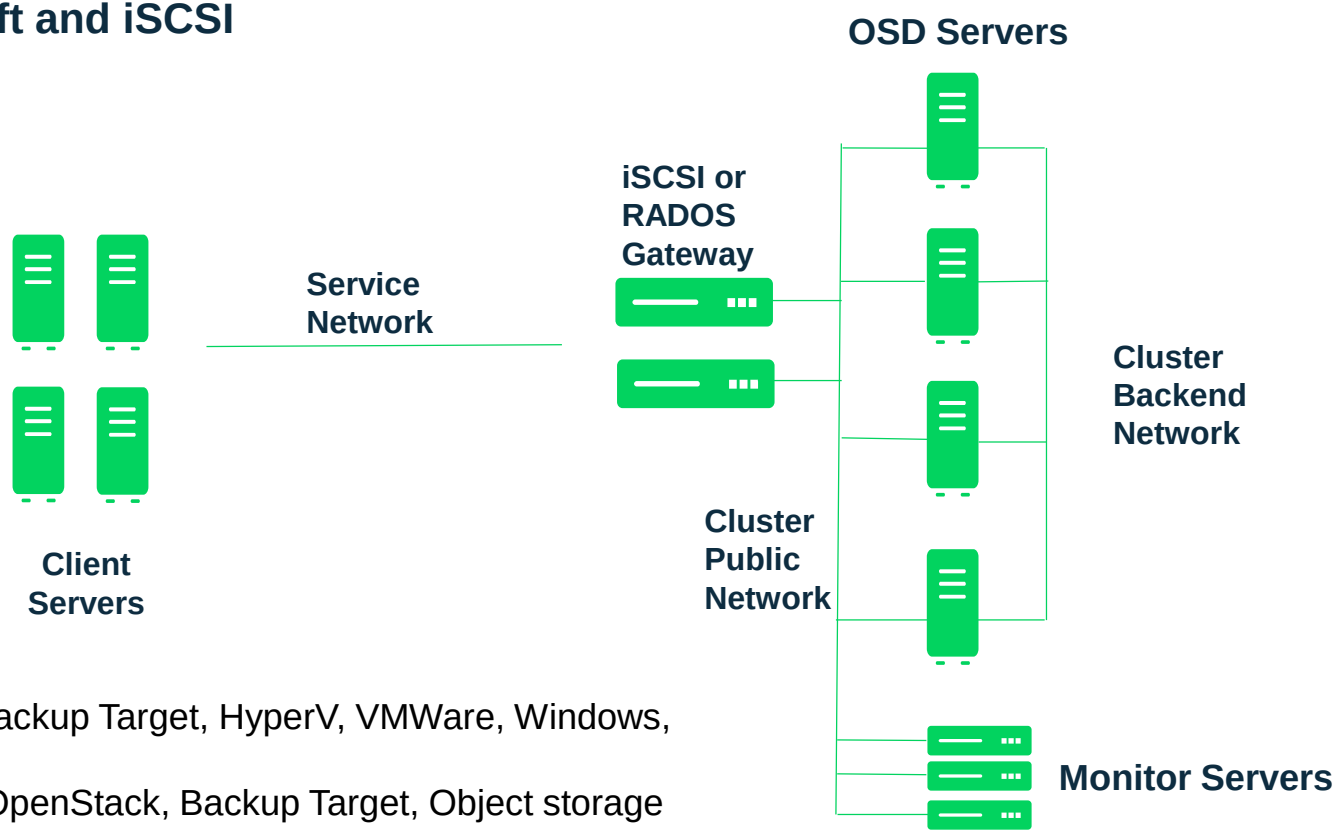
  - Erasure coded pool via RBD
  - Limited number of high performance Storage

- Erasure coding



Various Usage with one solution

S3, Swift and iSCSI

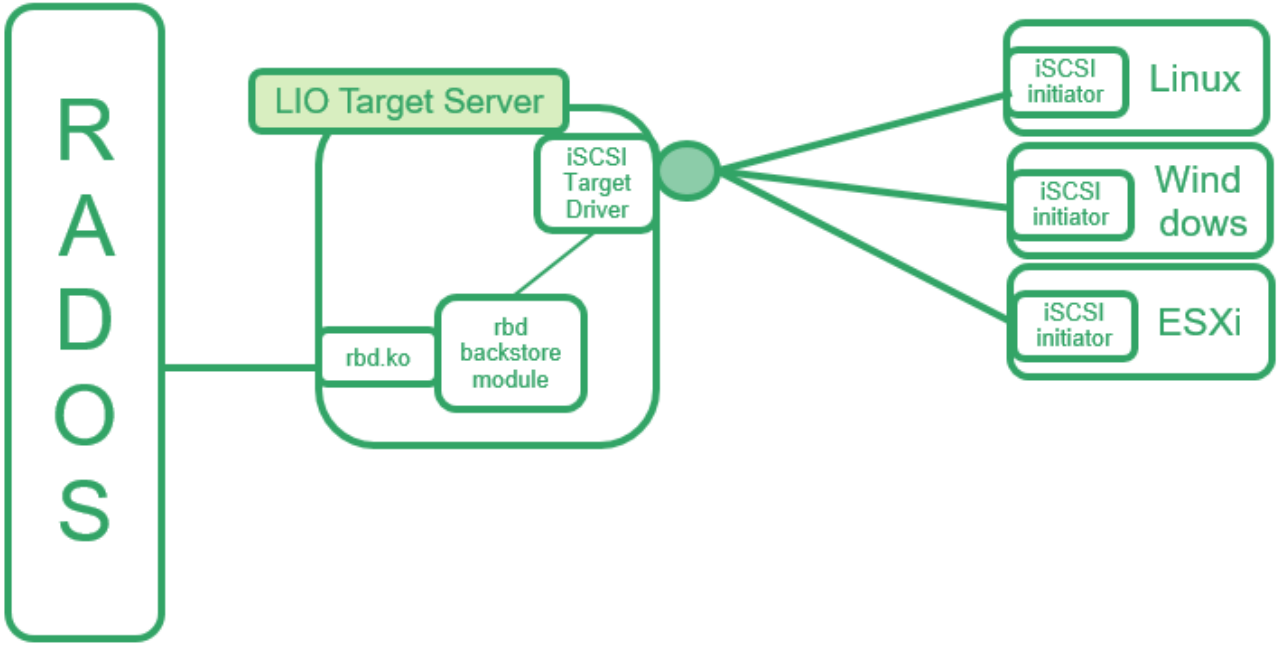


- iSCSI : Backup Target, HyperV, VMWare, Windows, Linux
- Object : OpenStack, Backup Target, Object storage

Various Usage with one solution

iSCSI Gateway

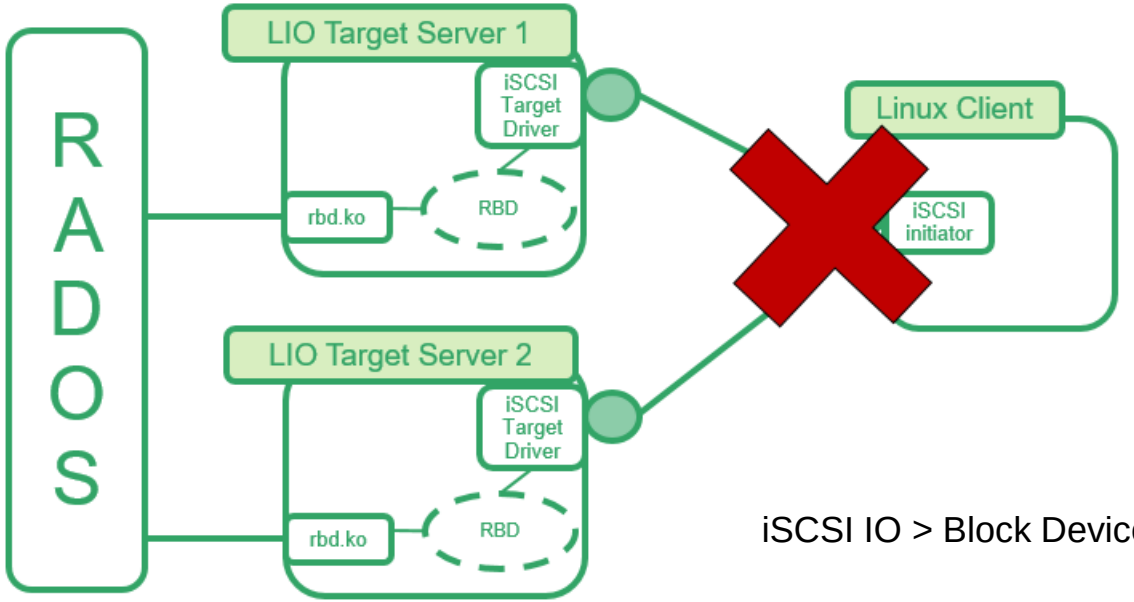
- RBD backstore module
- lrbd : Multi-node configuration utility
- LIO Target package
- Kernel RBD module



Various Usage with one solution

iSCSI Gateway without SES(SUSE Enterprise Storage)

- **Active-Active Multipath is not available**

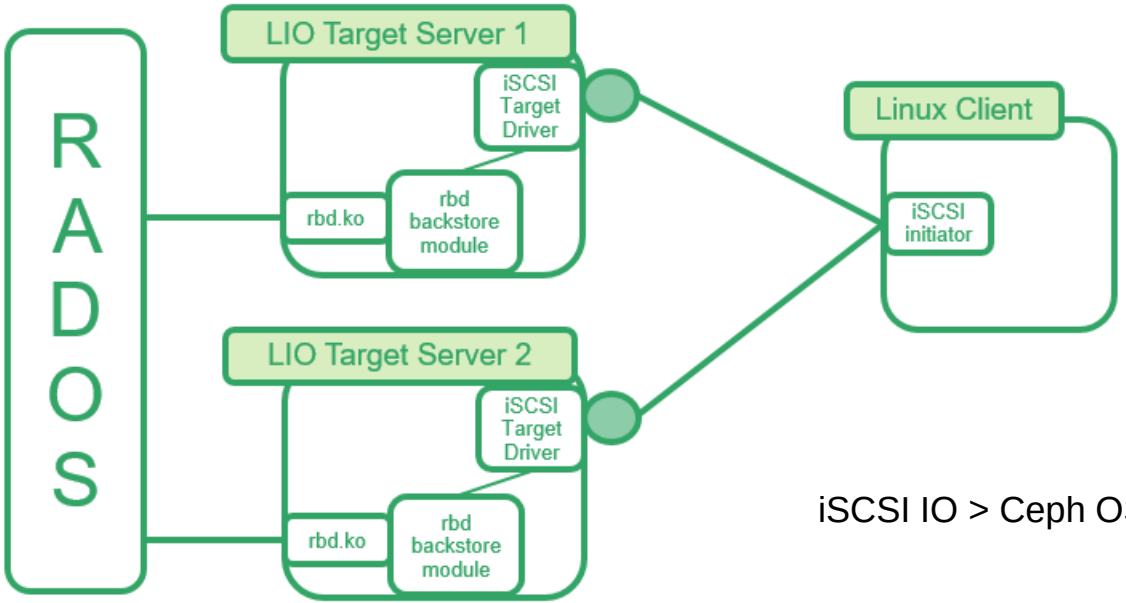


iSCSI IO > Block Device > Ceph OSD Requests

Various Usage with one solution

iSCSI Gateway with SES

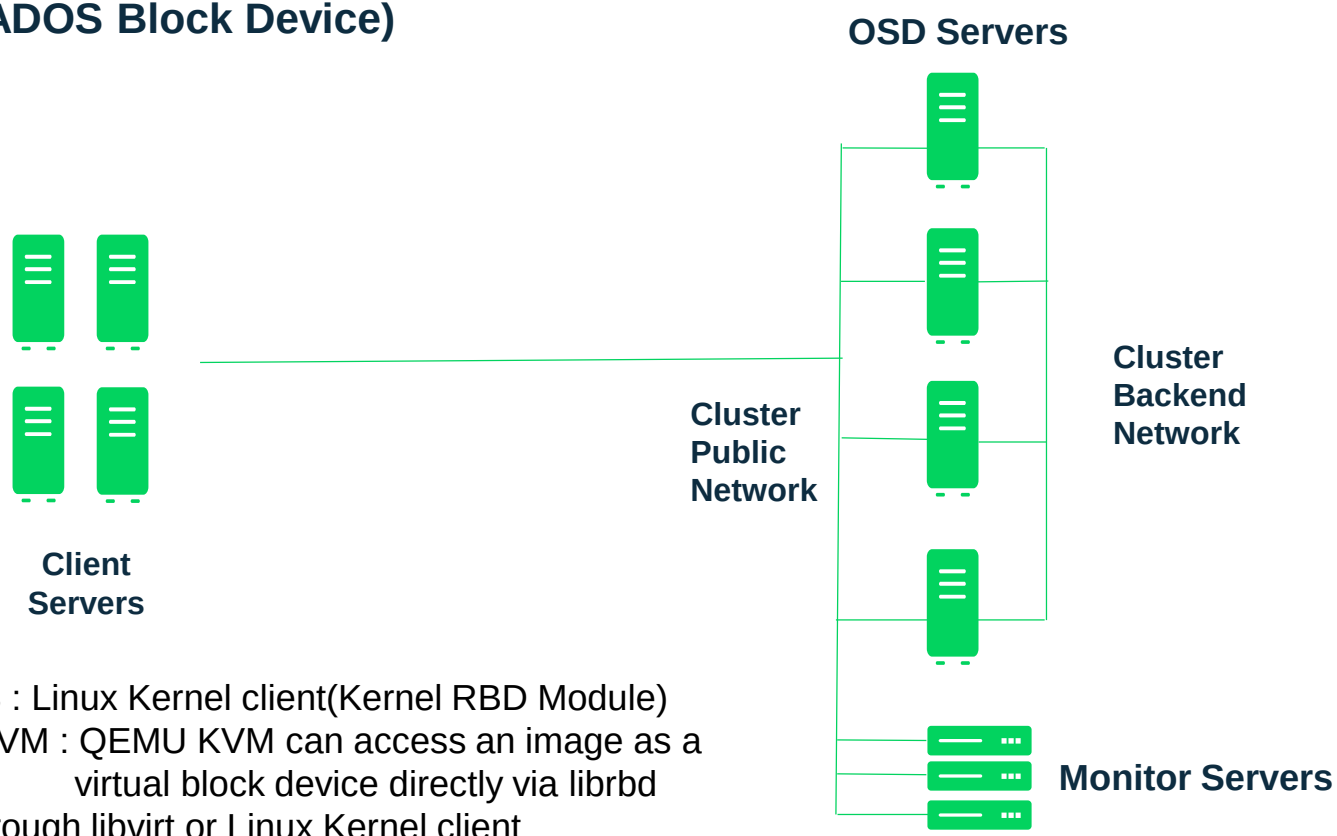
- **Active-Active Multipath is available**



iSCSI IO > Ceph OSD Requests

Various Usage with one solution

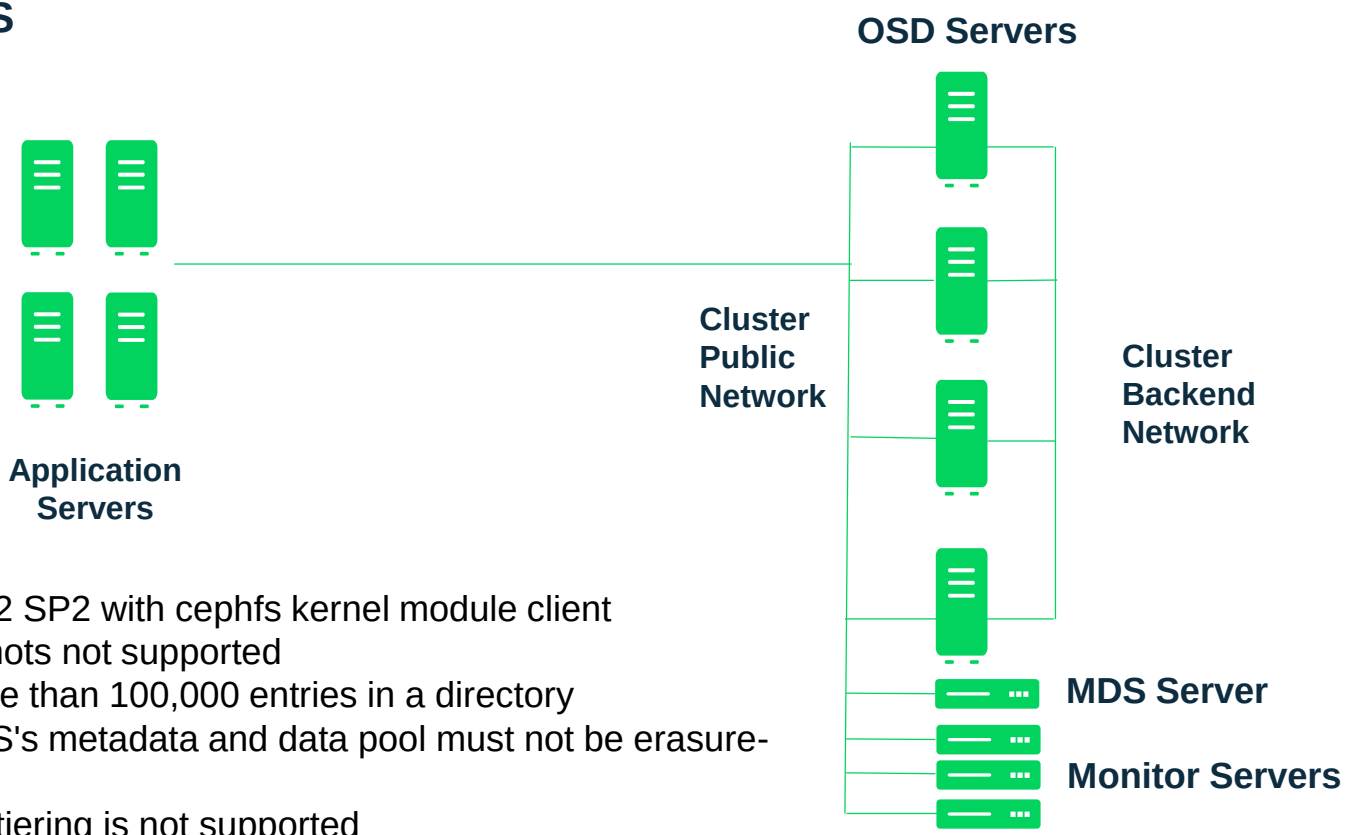
RBD(RADOS Block Device)



- Linux OS : Linux Kernel client(Kernel RBD Module)
- QEMU KVM : QEMU KVM can access an image as a virtual block device directly via librbd
- Xen : Through libvirt or Linux Kernel client

Various Usage with one solution

CephFS

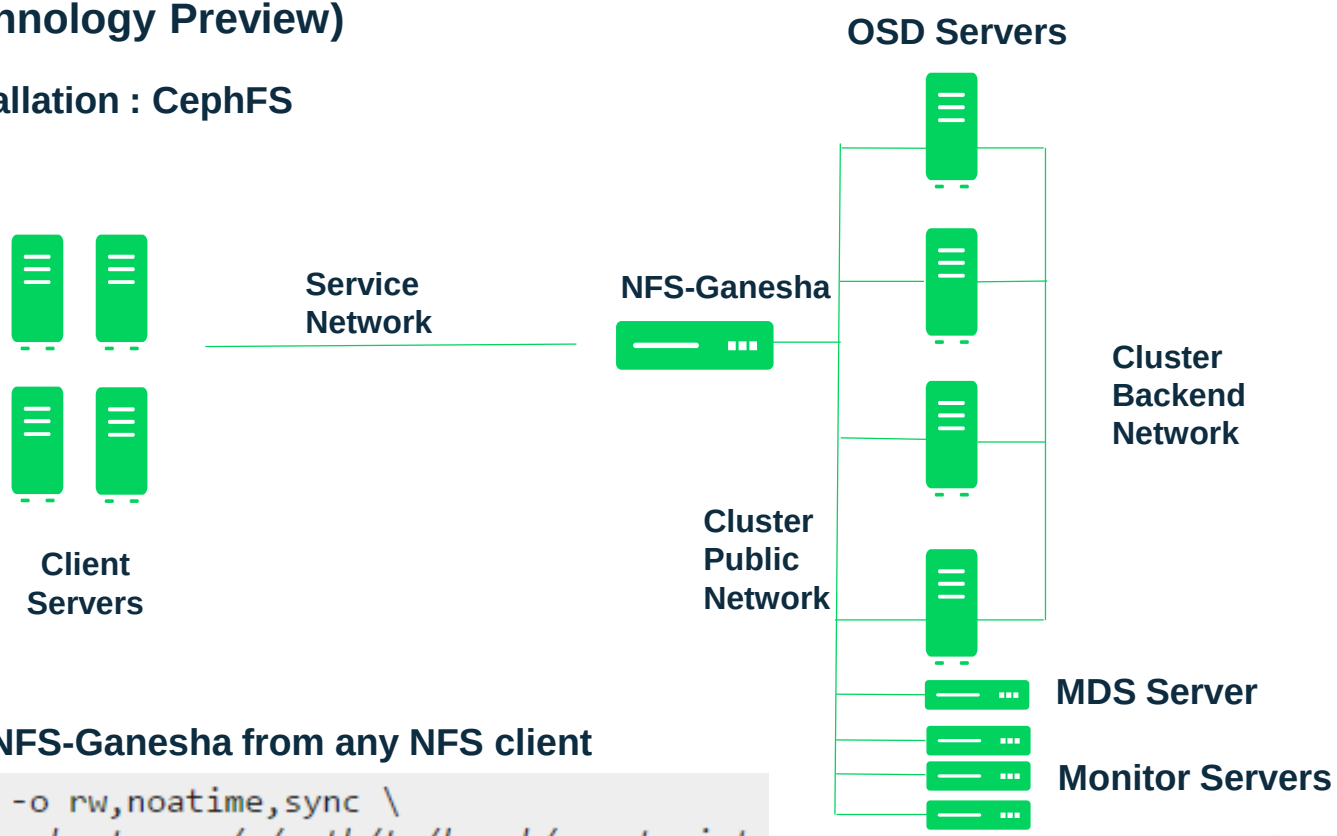


- SLES12 SP2 with cephfs kernel module client
- Snapshots not supported
- No more than 100,000 entries in a directory
- CephFS's metadata and data pool must not be erasure-coded.
- Cache tiering is not supported

Various Usage with one solution

NFS-Ganesha(Technology Preview)

- Two type of installation : CephFS



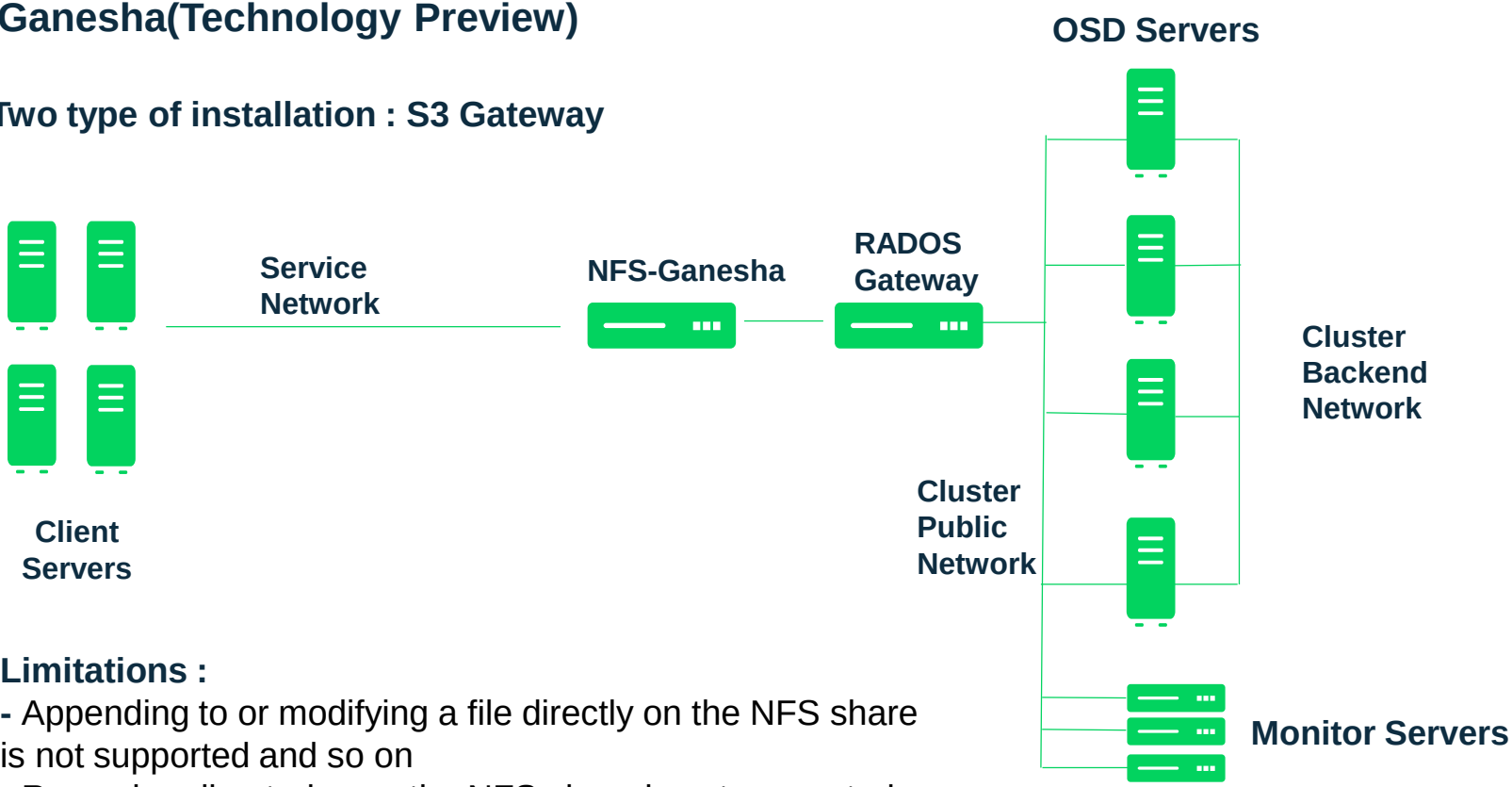
- Clients access NFS-Ganesha from any NFS client

```
sudo mount -t nfs -o rw,noatime,sync \
 nfs_ganesha_server_hostname:/ /path/to/local/mountpoint
```

Various Usage with one solution

NFS-Ganesha(Technology Preview)

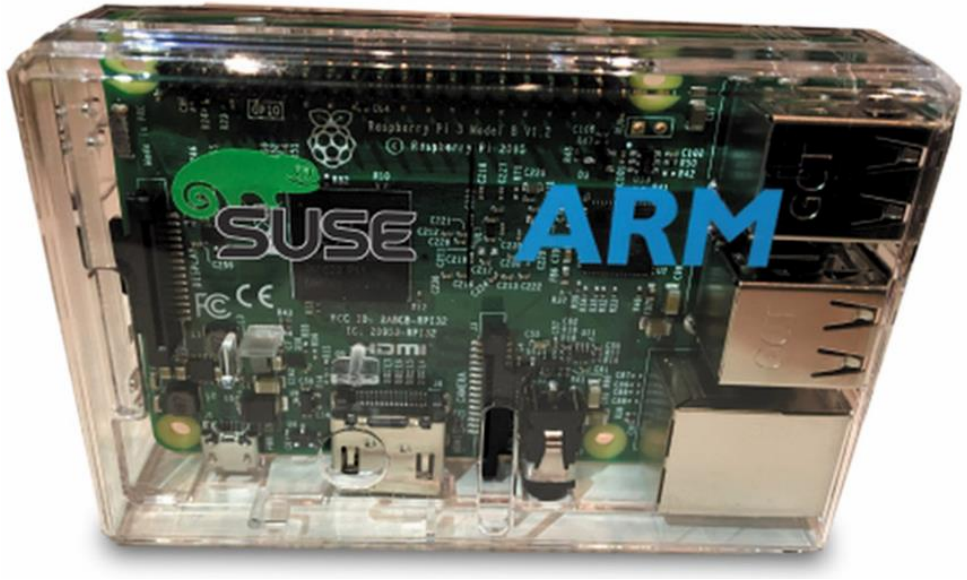
- **Two type of installation : S3 Gateway**



- **Limitations :**
  - Appending to or modifying a file directly on the NFS share is not supported and so on
  - Removing directories on the NFS share is not supported

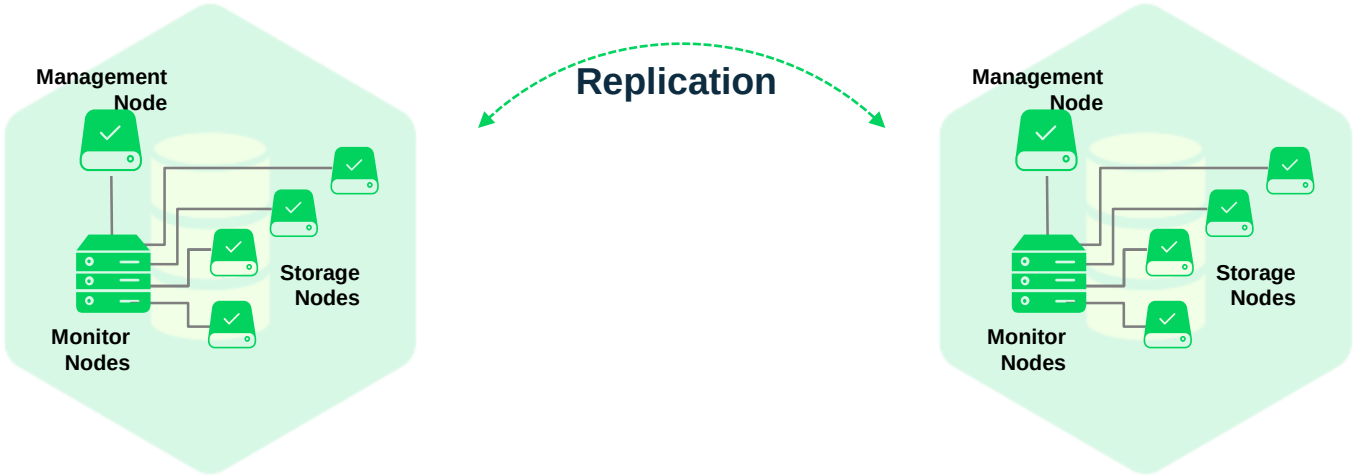
Various Usage with one solution

Support Aarch64 architecture



Help Disaster Recovery

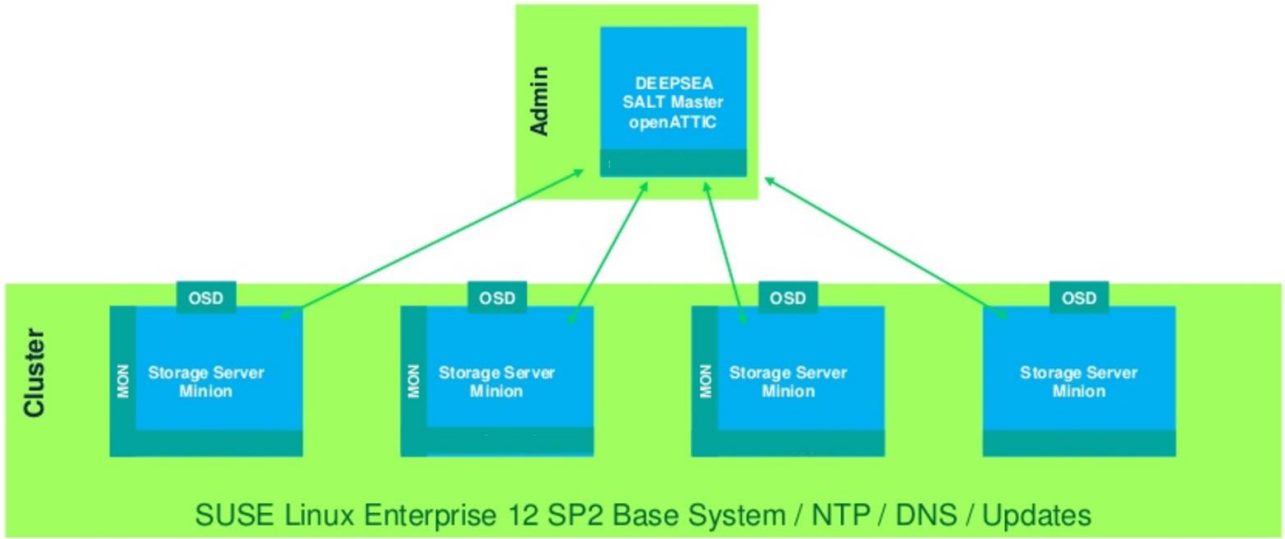
Mult-site replication



- Block : Only for native RBD, Not iSCSI
- Object : Active-active replication
- CephFS : Not supported

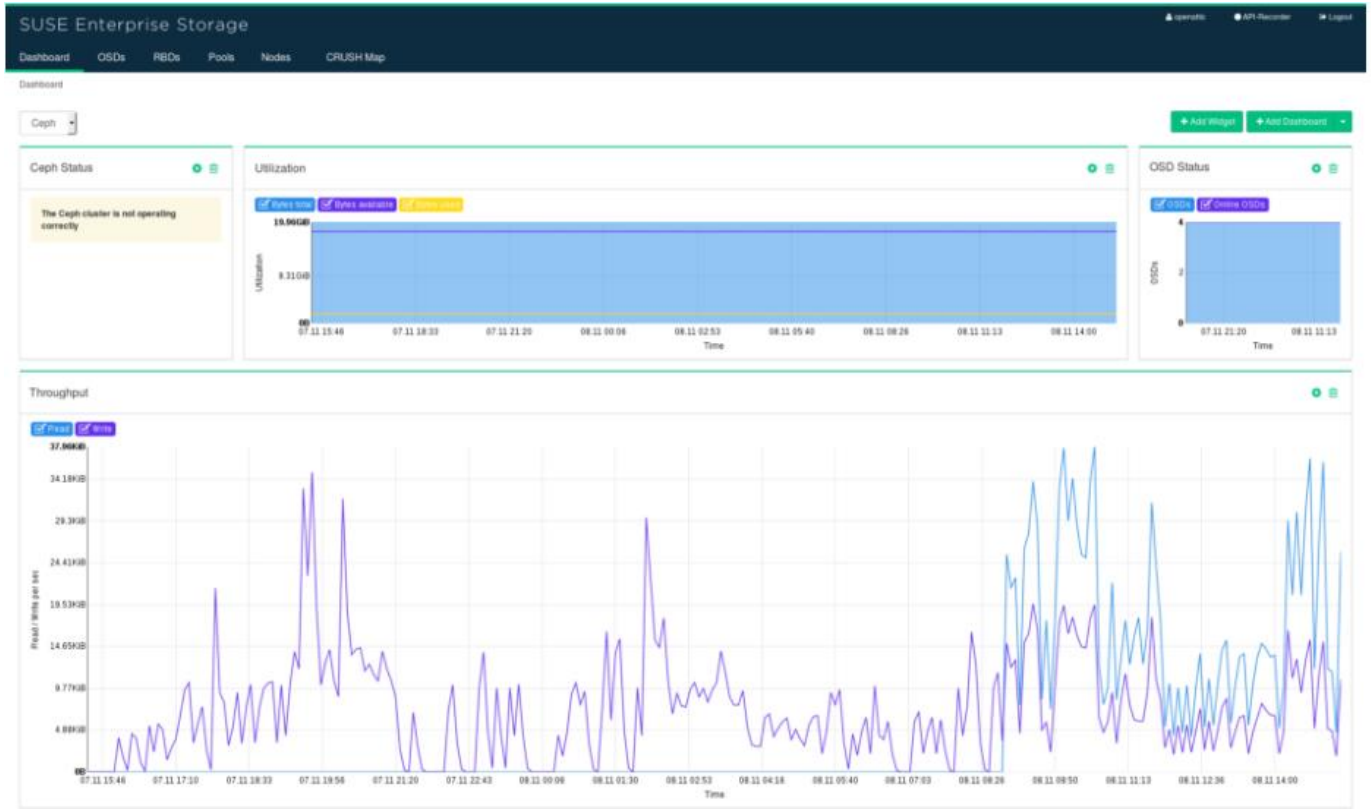
Advanced Management

Deepsea



Advanced Management

openATTIC



Advanced Management

openATTIC

- Dashboard
  - Ceph status, Utilization, OSD status and Throughput
- OSDs
  - List(hostname, status, Crush Weight)
- RBDs
  - List(pool, size, used), Create and Delete
- Pools
  - List(Usage, PG, Replica size, type, crush_ruleset), Create and Delete
- Node
  - List(hostname, public address, cluster, rolls, key status) and Edit
- CRUSH Map
  - View CRUSHMAP and Edit

Partnership

Partnerships

Backup to Disk Solution

Veeam · Veritas · Commvault · Hewlett Packard Enterprise

SAP Hana Storage Solution

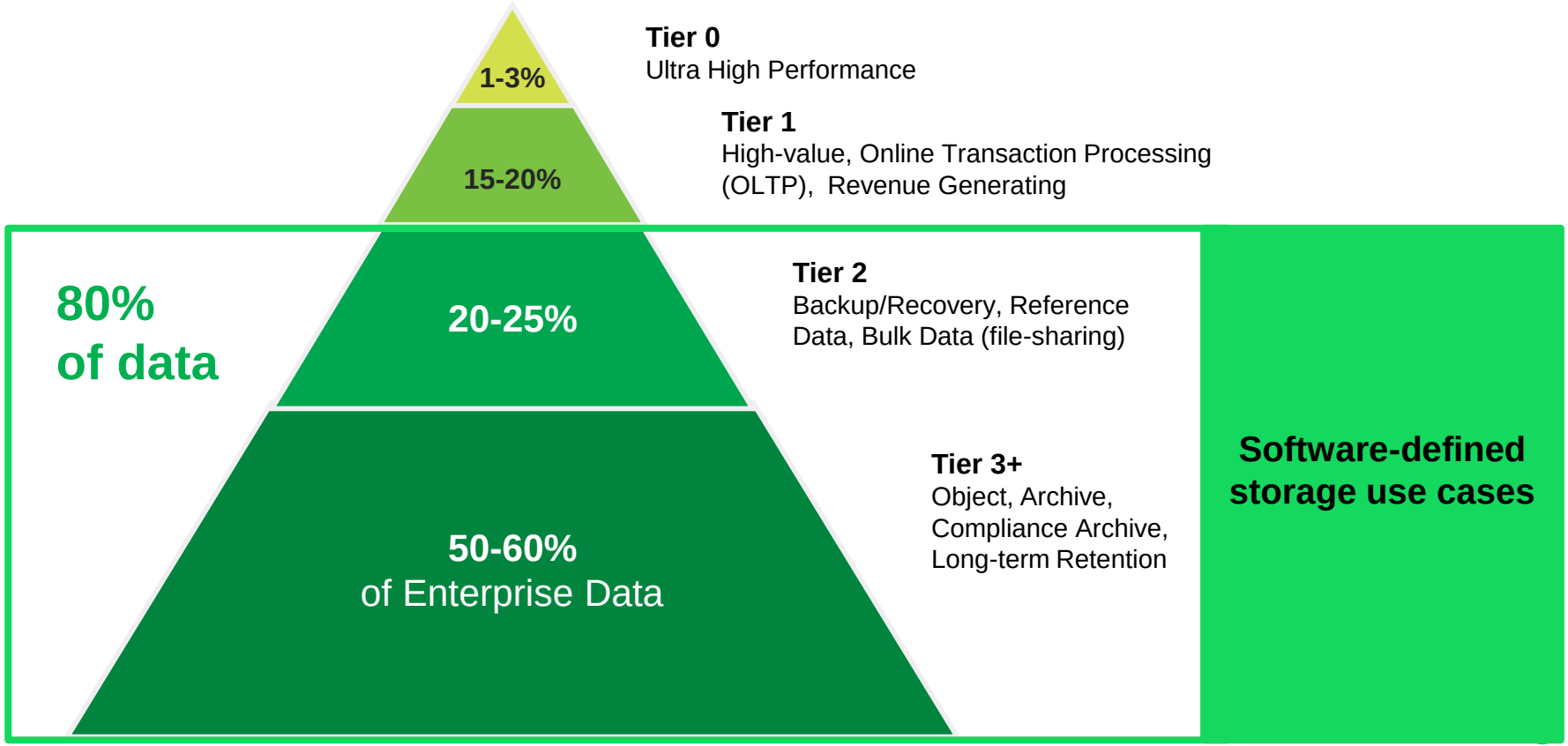
Lenovo

Compliant Archives

iTernity

Where to use

Enterprise Data Capacity Utilization

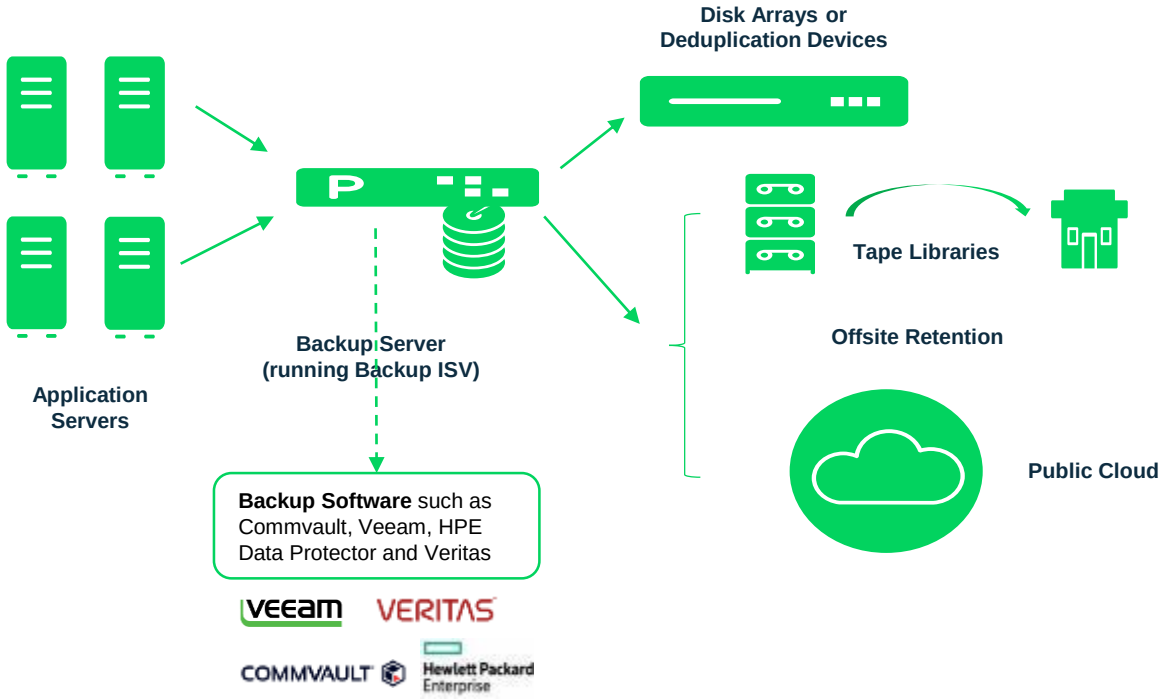


Source: Horison Information Strategies - Fred Moore

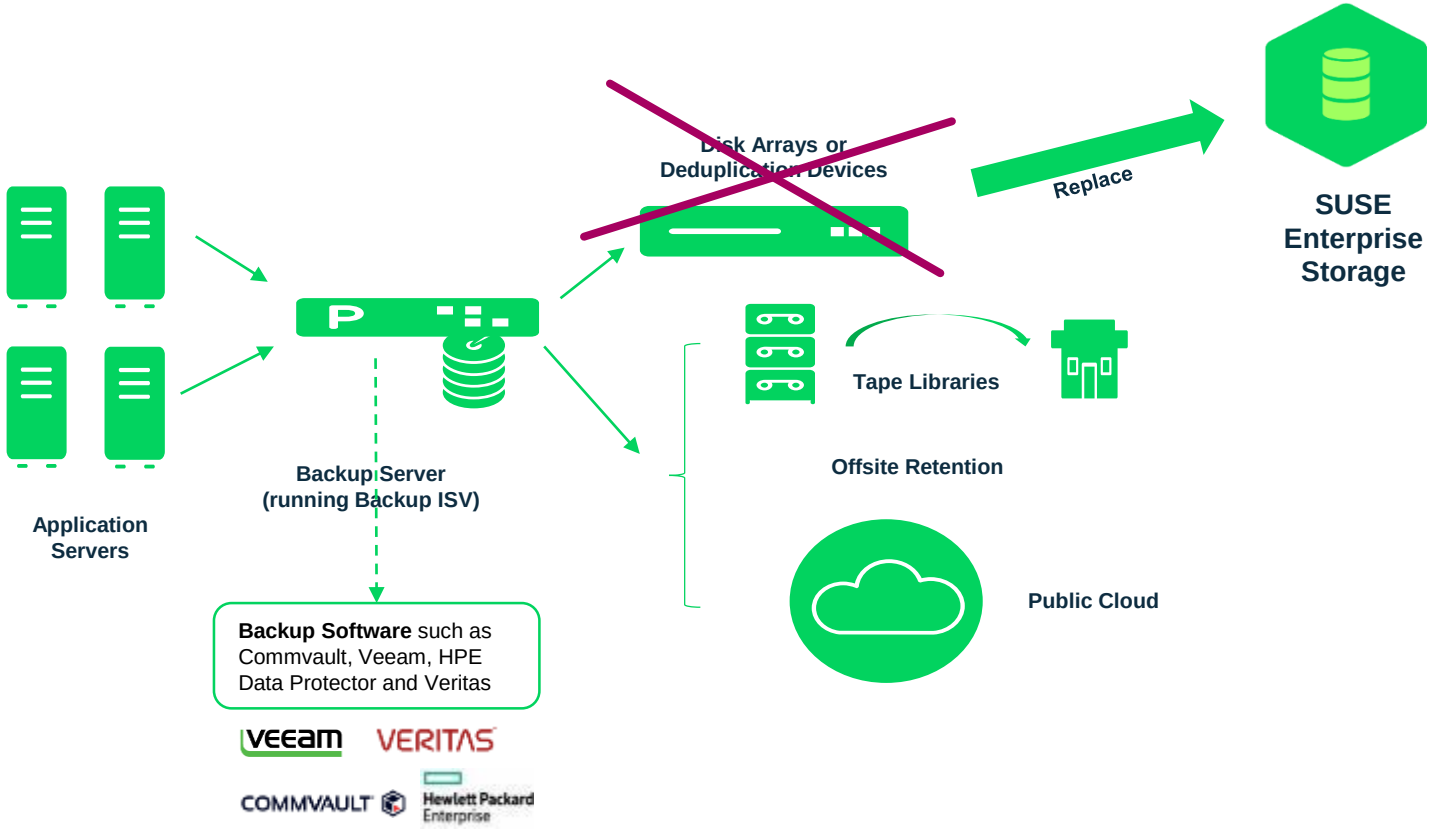
Large Data Use Cases

Content Storage & Sharing	Backup & Archival Use Cases	Compliance Storage & Archives	Cloud Storage
Video Surveillance & Bodycam	Backup & Archive Target	Health Records Archives	OpenStack Cloud
Media Asset Management Solution	HPC Storage Archives	PACS Modality Archives	Public Cloud Storage
Streaming Video Repository	Geological Archives	Financial Data Archives	Private Cloud Storage
		Legal Archives	
		Genomics Archives	

Current Solution Architecture

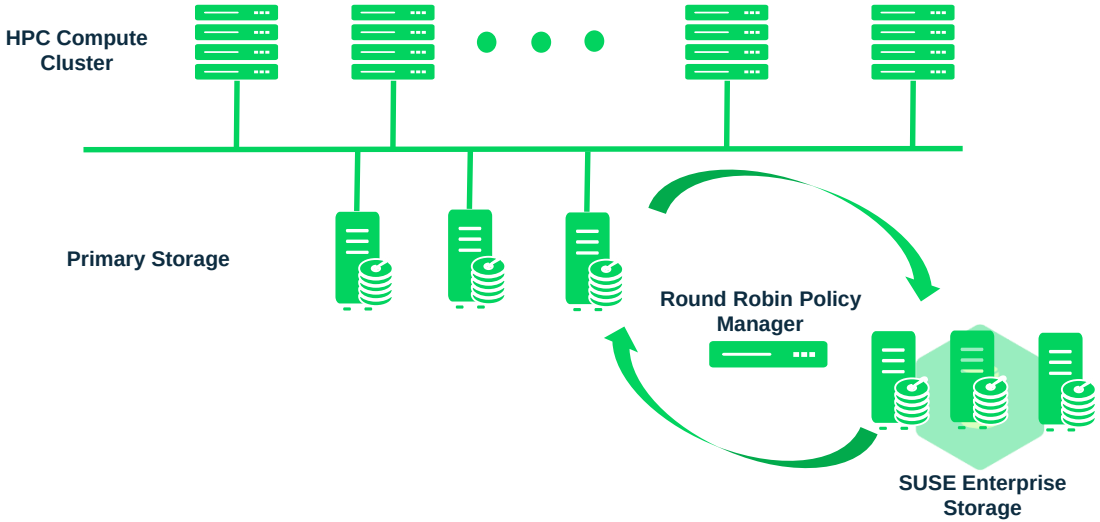


Current Solution Architecture



HPC Storage Archive Use Case

Archive Data not Needed Immediately to Secondary Tier



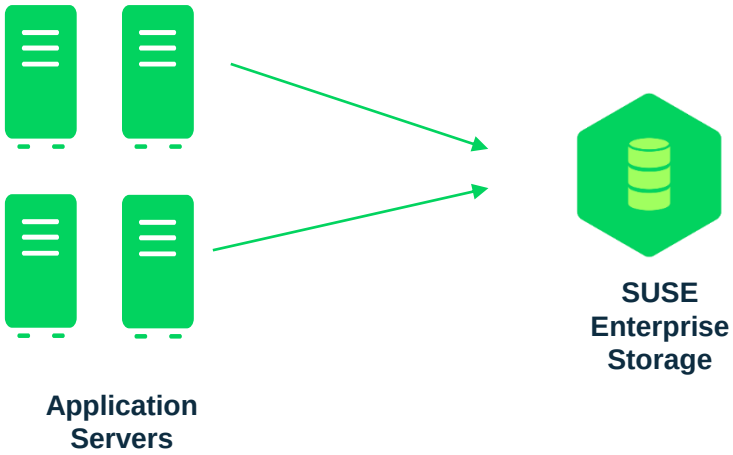
Lower TCO

Easy to Grow

Reduce Footprint

Never Migrate

Bulk Storage



Sizing guide

Minimum requirement

- Monitor node
  - RAM : 2G
  - OS Disk : SSD
  - Network : Bonded network
  - Minimum 3 monitor nodes

- RADOS Gateway node
  - CPU : 6-8 CPU cores
  - RAM : 32G (64G recommended)

- Storage(OSD) node
  - CPU : 1.5GHz of a CPU core per OSD
  - RAM : 2G per TB,
    4G additional RAM for cache tiering
  - OS disk : SSD
  - Network : 10GbE networks
  - Minimum 4 storage nodes

- iSCSI node
  - CPU : 6-8 CPU cores
  - RAM : 16G

Refer to the link below for more detail
https://www.suse.com/documentation/ses-4/book_storage_admin/data/cha_ceph_sysreq.html

80TB Storage with minimum requirement

with an iSCSI Gateway

- SES1
  - Role : Admin, iSCSI, OSD
  - CPU : 2 Xeon E5-2683v4(16 cores, 2.1Ghz)
  - RAM : 136(120+16)G
  - Network : 2 Inter 530SFP+
  - OS disks : 2 200G SSDs
  - Journal disks : 2 250G SSDs
  - OSD disks : 10 6TB nearline SAS HDDs

- SES2, SES3, SES4
  - Role : Monitor, OSD
  - CPU : 2 Xeon E5-2680v3(12 cores, 2.5Ghz)
  - RAM : 122(120+2)G
  - Network : 2 Inter 530SFP+
  - OS disks : 2 200G SSDs
  - Journal disks : 2 250G SSDs
  - OSD disks : 10 6TB nearline SAS HDDs

- 2 x 10G Switches

Success Story

Medical Research Customer in Canada

Use Case:

- High-powered telescope with archival data that will all go to SUSE Enterprise Storage.
- Some virtual machine storage

Solution Design:

- 1PB Cluster
- ISCSI – Primarily
- Testing Object, CephFS, NFS
- Eventual Multi-site

Network:

- 802.3ad 40GbE connections
- Redundant switches

European Telco

Use Case:
Helion OpenStack

Solution Design:
- RBD for OpenStack Cinder
- Swift for Glance
- Started with minimal configuration with dedicated Monitor nodes

Network:
- 802.3ad 10GbE connections
- Redundant switches

S2

https://www.suse.com/success/stories/s2/

- **Application: SUSE OpenStack Cloud**
- **Used for: 100% backend storage for glance, cinder and nova**
- **Critical success factors: integration with Openstack**

University Hospital Essen

https://www.suse.com/success/stories/university-hospital-essen/

- **Application: Medical imaging + file sharing**
- **Used for: 100% iSCSI block (for Windows)**
- **Total capacity: 300TB**
- **Critical success factors: deployment & expansion speed, cost, commitment to customer in the face of challenges**

Orchard Park Police

https://www.suse.com/success/stories/orchard-park-police-depa

- Application: Body cam archive
- Used for : 100% iSCSI block (for Windows)
- Total capacity: 100TB+
- Critical success factors: cost, heterogeneity, complete solution

Phact B.V.

- **Application: VMware**
- **Used for : managing external voice engines, connection servers, webservers, payment services, small databases**
- **Total capacity: 240 TB**

Q&A

Thank you!

Unpublished Work of SUSE LLC. All Rights Reserved.

This work is an unpublished work and contains confidential, proprietary and trade secret information of SUSE LLC. Access to this work is restricted to SUSE employees who have a need to know to perform tasks within the scope of their assignments. No part of this work may be practiced, performed, copied, distributed, revised, modified, translated, abridged, condensed, expanded, collected, or adapted without the prior written consent of SUSE.
Any use or exploitation of this work without authorization could subject the perpetrator to criminal and civil liability.

General Disclaimer

This document is not to be construed as a promise by any participating company to develop, deliver, or market a product. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. SUSE makes no representations or warranties with respect to the contents of this document, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. The development, release, and timing of features or functionality described for SUSE products remains at the sole discretion of SUSE. Further, SUSE reserves the right to revise this document and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes. All SUSE marks referenced in this presentation are trademarks or registered trademarks of Novell, Inc. in the United States and other countries. All third-party trademarks are the property of their respective owners.