SNIA DSI '16 DATA STORAGE INNOVATION CONFERENCE
Innovation in Storage Products, Services, and Solutions
June 13-15, 2016 | Marriott San Mateo | San Mateo, CA

Ceph For The Enterprise

David Byte
Sr. Technology Strategist
SUSE

Who is David Byte?

- Sr. technical strategist on the SUSE IHV Alliances & Embedded team
- Specializes in storage, HPC and ARM64
- Live in Jenks, OK – a suburb of Tulsa
- 20+ year veteran in the IT industry (15+ in storage and 20+ in Linux)

LinkedIn: http://LinkedIn/in/davidbyte

Blog: www.suse.com/communities/blog/author/davidbyte/

Agenda

A quick look at the market

Use cases & reference designs

Architecting a solution

Sizing a solution

Deployment notes

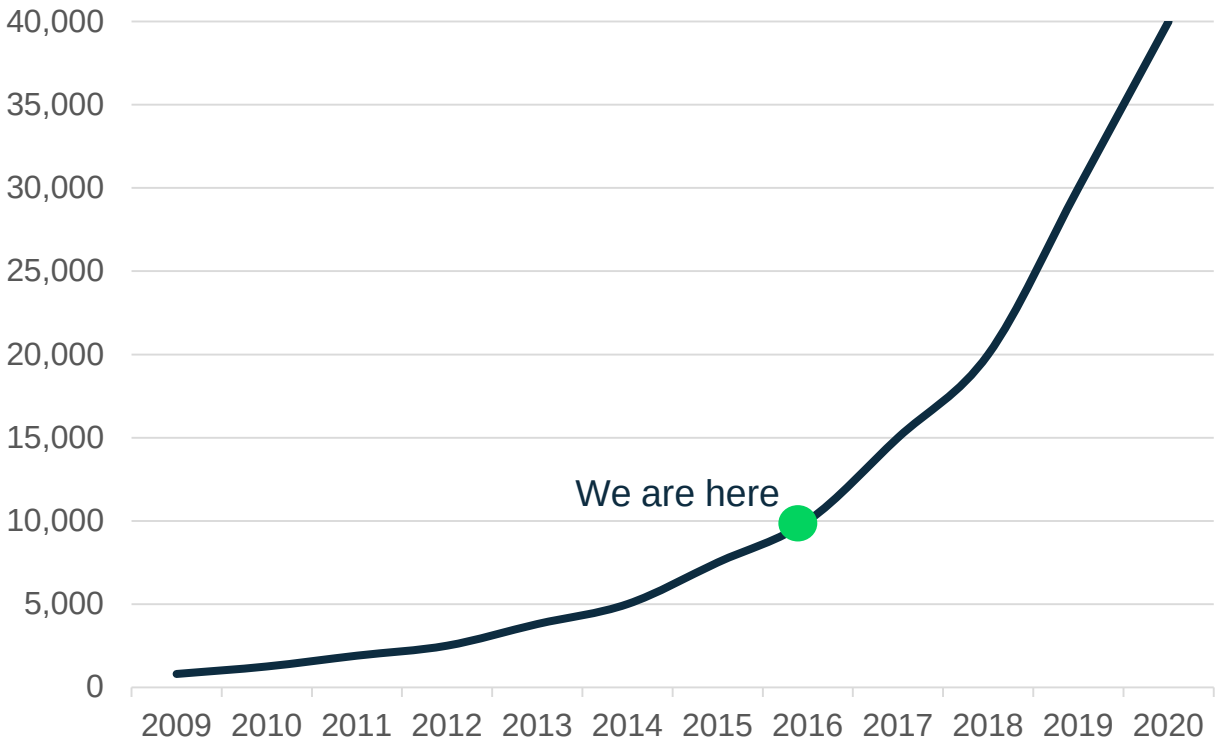
Tuning info

A few resources from SUSE

The Market

Business Has a Continuing Storage Challenge

Storage Doubles Every 18 Months



IT Budgets Are Flat

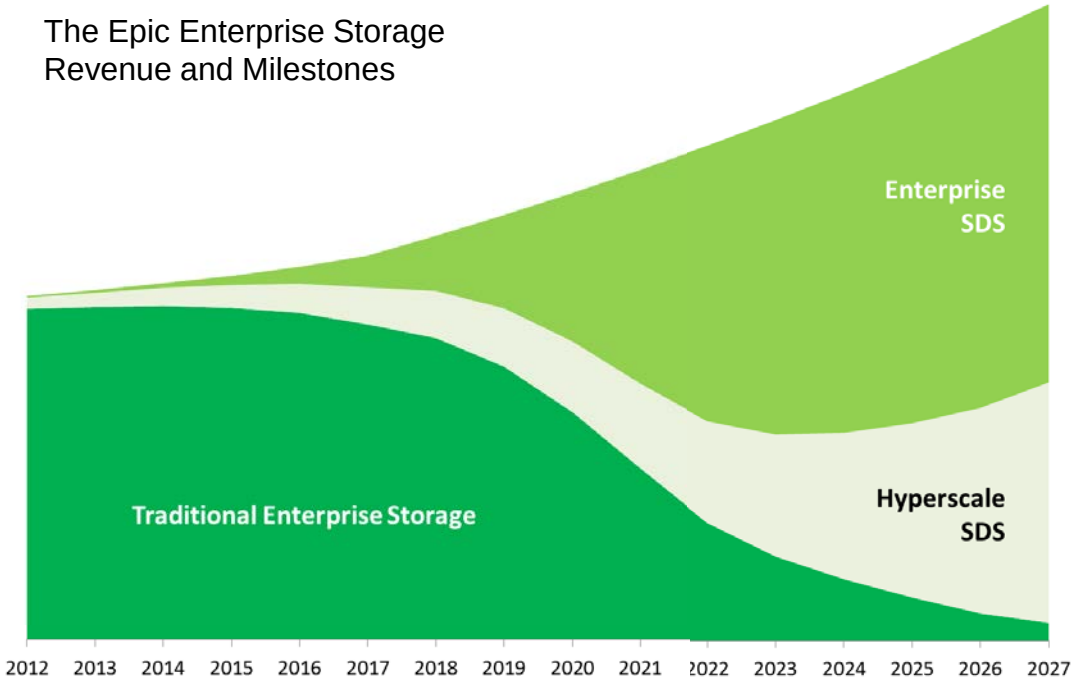
Storage costs decline **25%** annually

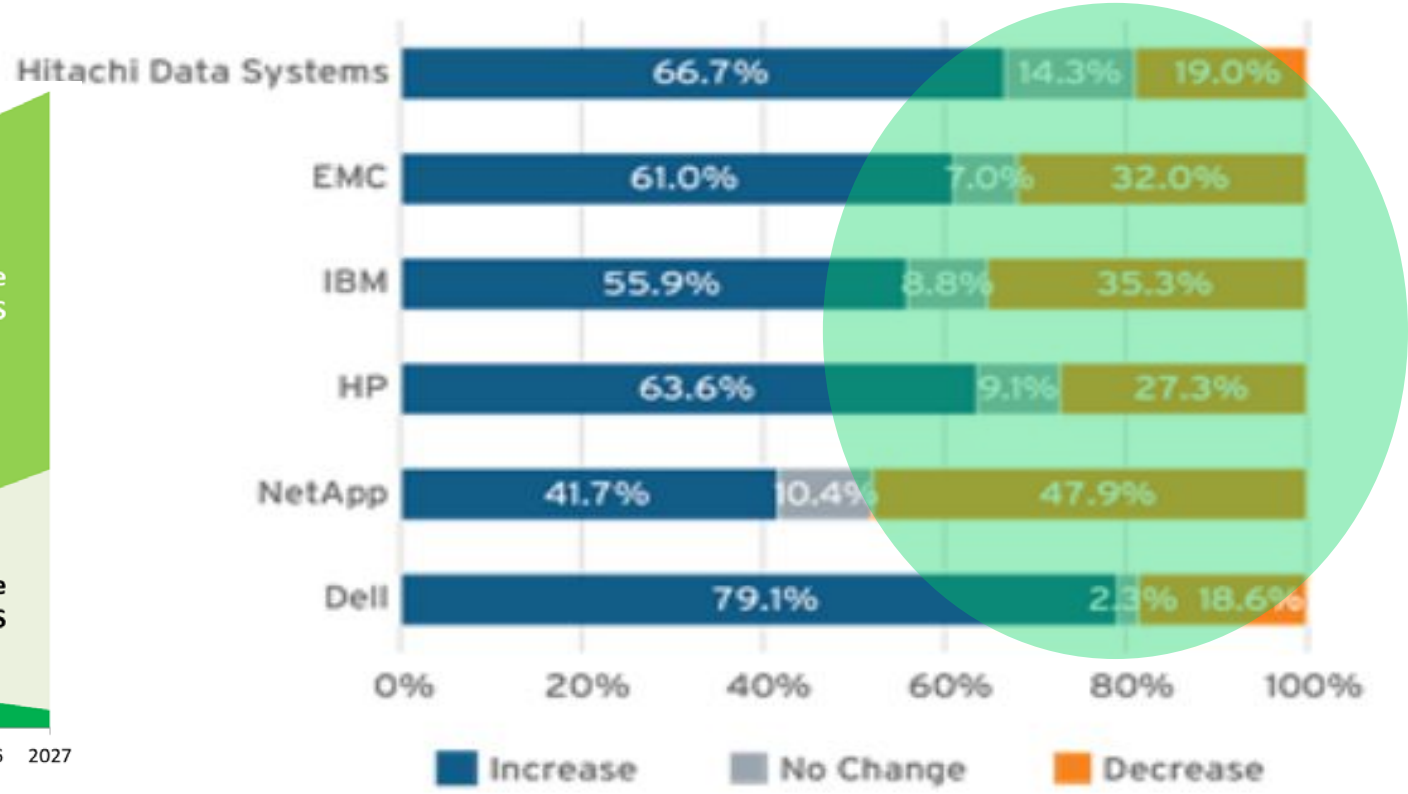
But data grows **40%** annually

Changes Are Coming

Market Shakeout - 2016 Spend vs 2015 Spend

The Epic Enterprise Storage
Revenue and Milestones





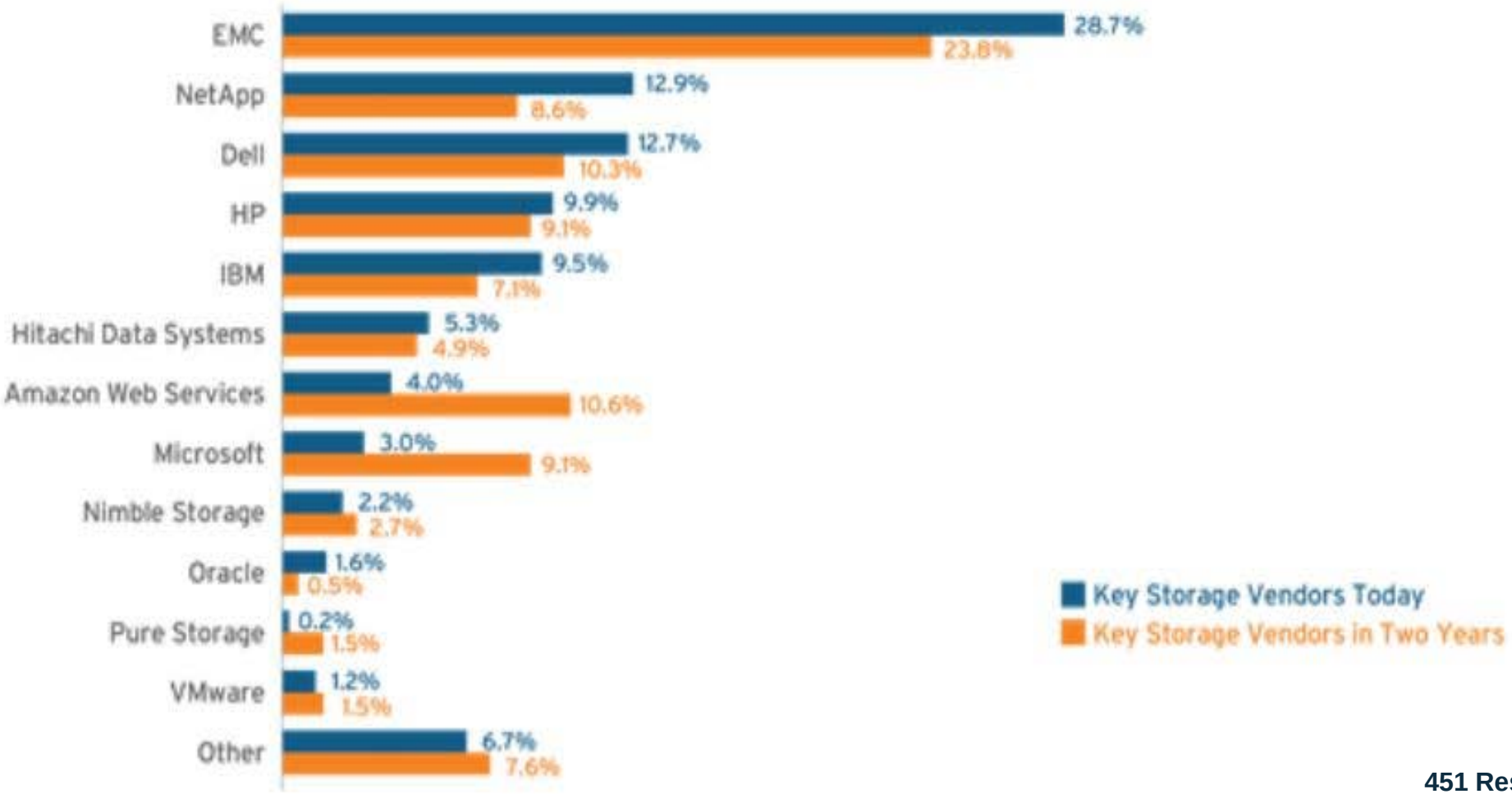
Source: 451 Research

By 2019, more than 50% of the storage capacity installed in enterprise data centers will be deployed with software-defined storage or hyperconverged integrated system architectures based on x86 hardware systems, up from less than 10% today

—Gartner

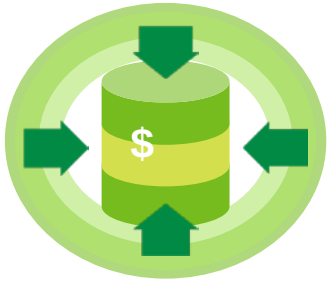
Who Would Have Thought?

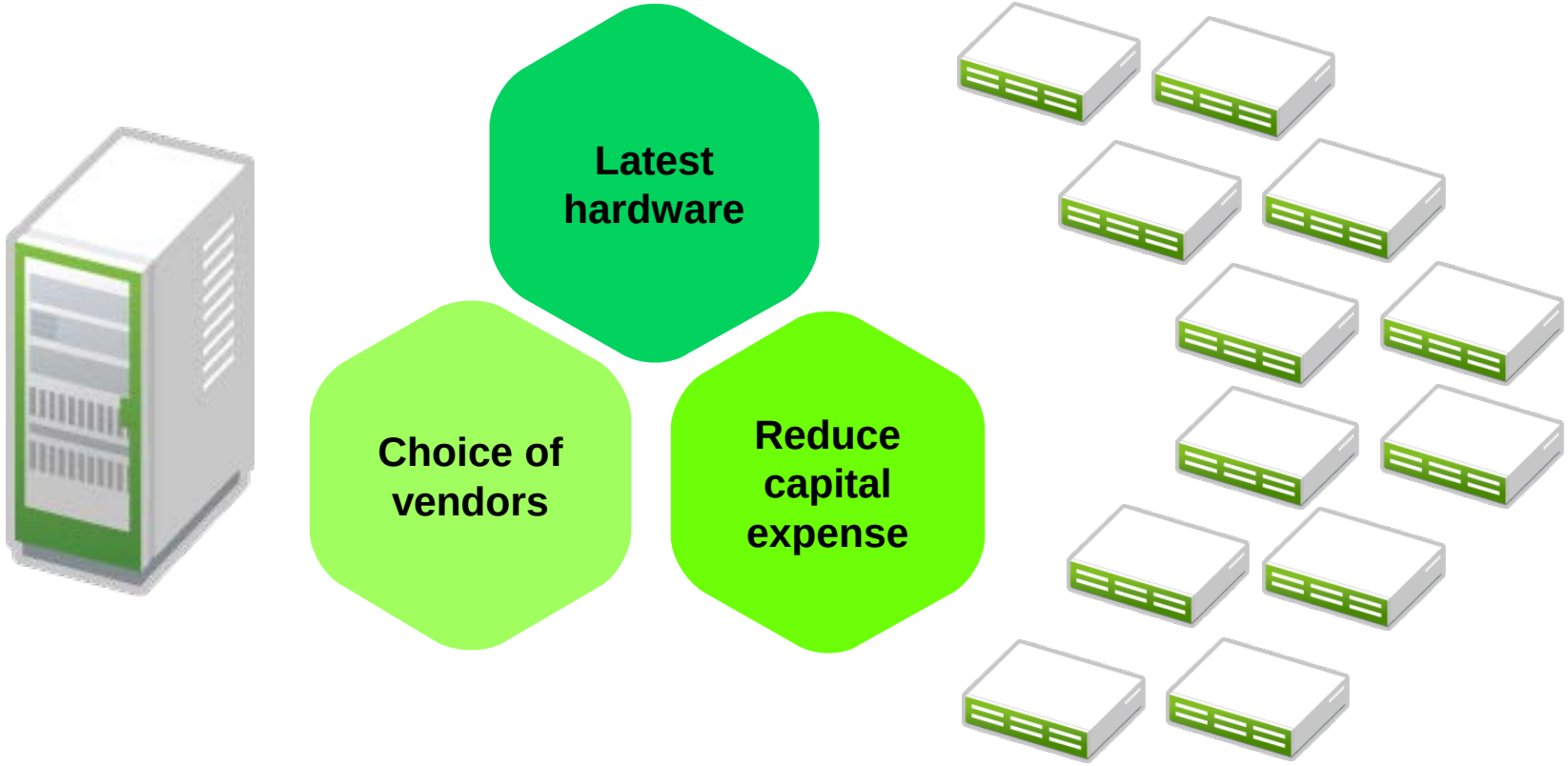
Amazon and Microsoft Move Into Top Six Storage Providers



451 Research

SDS Value





Making the Shift

Mode 1 – Gartner for Traditional

Legacy Datacenter

- Network, compute, and storage silos
- Traditional protocols – Fibre Channel, iSCSI, CIFS, NFS

Process-driven

- Slow to respond

This is what customers have today

Mode 2 – Gartner for Software Defined

Software Defined Data Center

- Software-defined everything

Agile infrastructure

- Supporting a DevOps model
- Business-driven

This is where customers want to go

Support today's investment

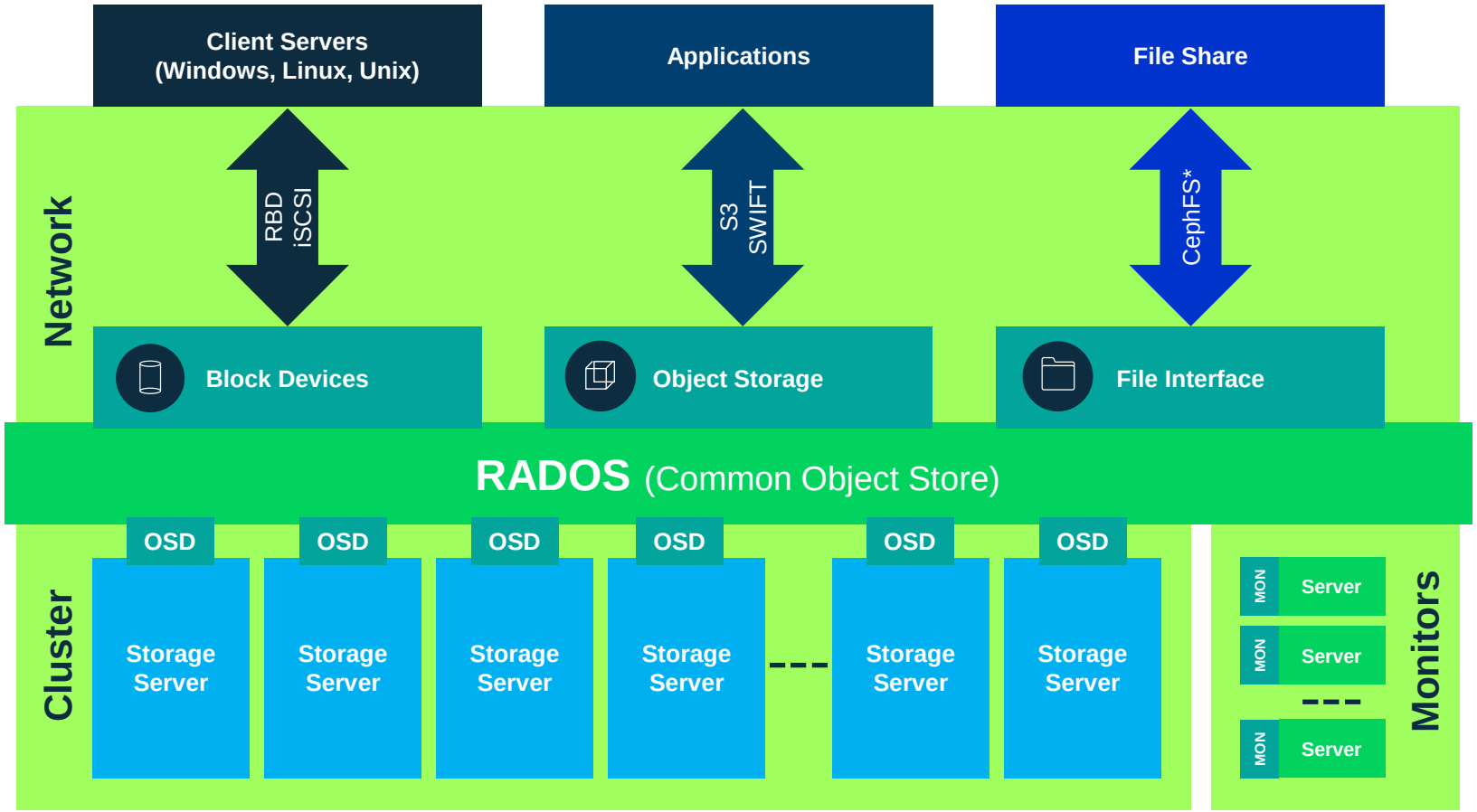
Adapt to the future

The SUSE Focus

SUSE Enterprise Storage powered by Ceph

- A highly scalable and resilient software-based storage solution.
- Enables organizations of any size to build cost-efficient, highly scalable storage.
- Utilizes off-the-shelf servers and disk drives.
- Self-managing and delivers storage functionality comparable to mid / high-end storage products at a fraction of the cost.

Open Source Ceph as the Base



Code Developers

782

Core	Regular	Casual
22	53	705

Total downloads

160,015,454

Unique downloads

21,264,047

Making Ceph Enterprise Consumable

Manageability

- Ease of install
- Centralized management, monitoring, reporting

Interoperability

- Unified block/file/object (heterogeneous OS access)
- Fabric interconnect

Efficiency

- Cache tiering
- Deduplication/compression
- Hierarchical storage management

Availability

- Backup/archive
- Continuous data protection
- Remote replication

Use Cases & Reference Designs

Really understand the problem(s) and drivers

Ask a lot of questions

- What business issues are driving this project?
- What applications will be interfacing with the storage?
- What kind of interface, block, object or file?
- What is the incumbent storage environment?

Content Store

Scientific Organizations

- Meteorological data
- Telescope recordings
- Satellite feeds

Media Industries

- TV stations
- Radio stations
- Motion picture distributors
- Web music/video content

Available Solutions

HPE – Apollo Series, 4200 and 4500 series

Thomas-Krenn – SES Appliance Capacity Optimized

Capacity Bundles via various reseller partners

Object or block bulk storage

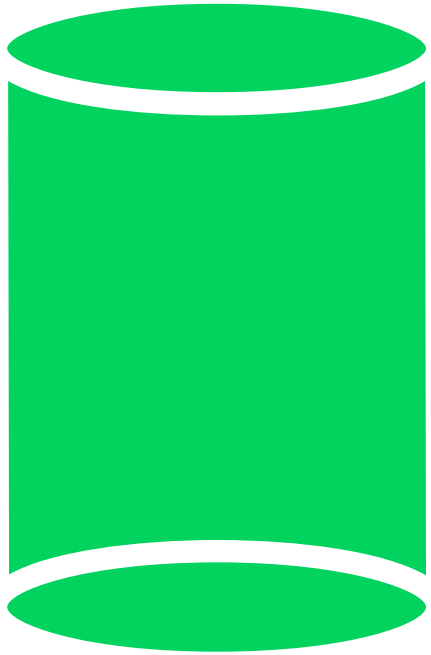
- Data that constantly grows during the course of business
- SharePoint data
- D2D Backup
  - HPE Data Protector and others
- Financial records

Video Surveillance

- Facility security surveillance
- Red light/traffic cameras
- License plate readers
- Body cameras for law enforcement
- Military/government visual reconnaissance

Virtual Machine (VM) Storage

- Ceph is already the leading storage choice for OpenStack environments
- Low and mid i/o virtual machine storage for major hypervisor platforms
  - kvm – native RBD
  - Hyper-V – iSCSI
  - VMware - iSCSI



Available Solutions For General Purpose And Performance Scenarios

HPE

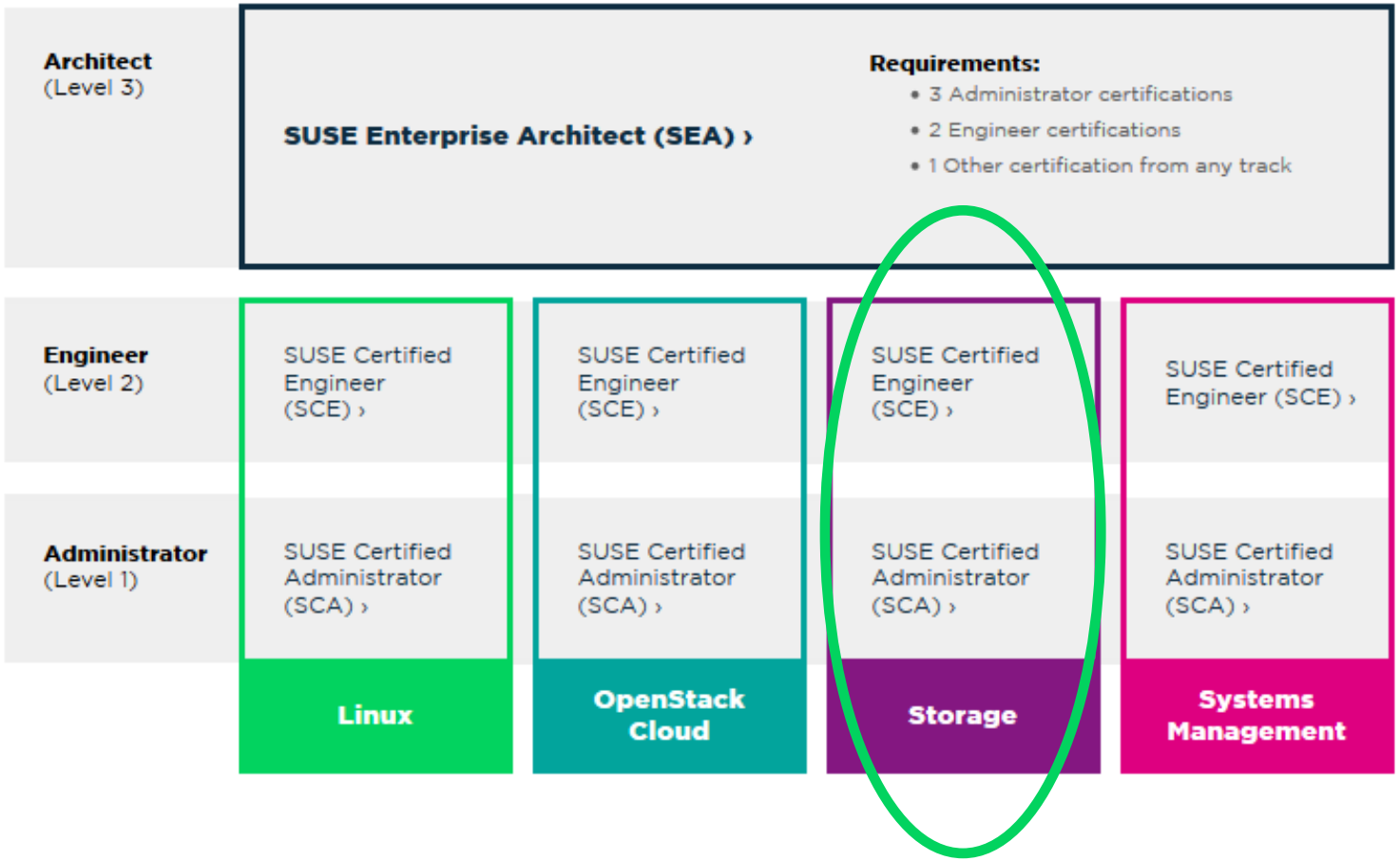
- Apollo – Apollo 4200
- Proliant – DL380/DL360

Thomas-Krenn – SES Appliance All-rounder

- Performance Optimized

General Bundles via various reseller partners

Training and Certification for Storage



SUSE Certified Administrator in Storage

SES101/201 – Introduction to Storage 2 Administration

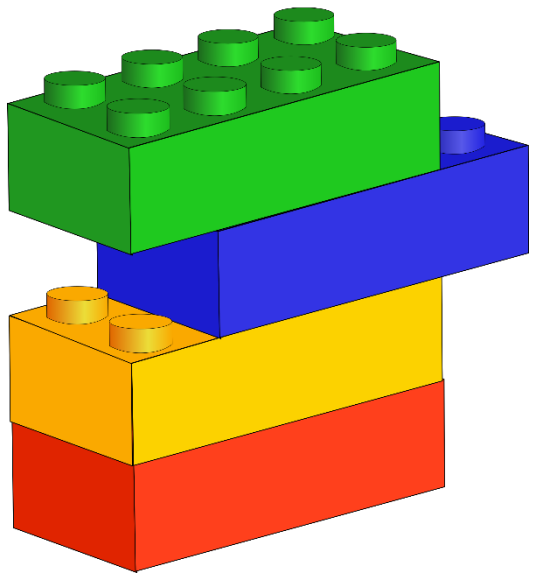
- Course currently in market
- Exam available end of June
- Update to version 3 in August (SES201)

SUSE Certified Engineer in Storage

SES301 – advanced course

- Course and exam to follow update to version 3 of admin level
- Second half of calendar 2016

Architecture



Plan for the highest resiliency

Bonded connections from multiple NICs to stacked switches

- Protect against NIC failure
- Protect against cable failure
- Protect against switch failure

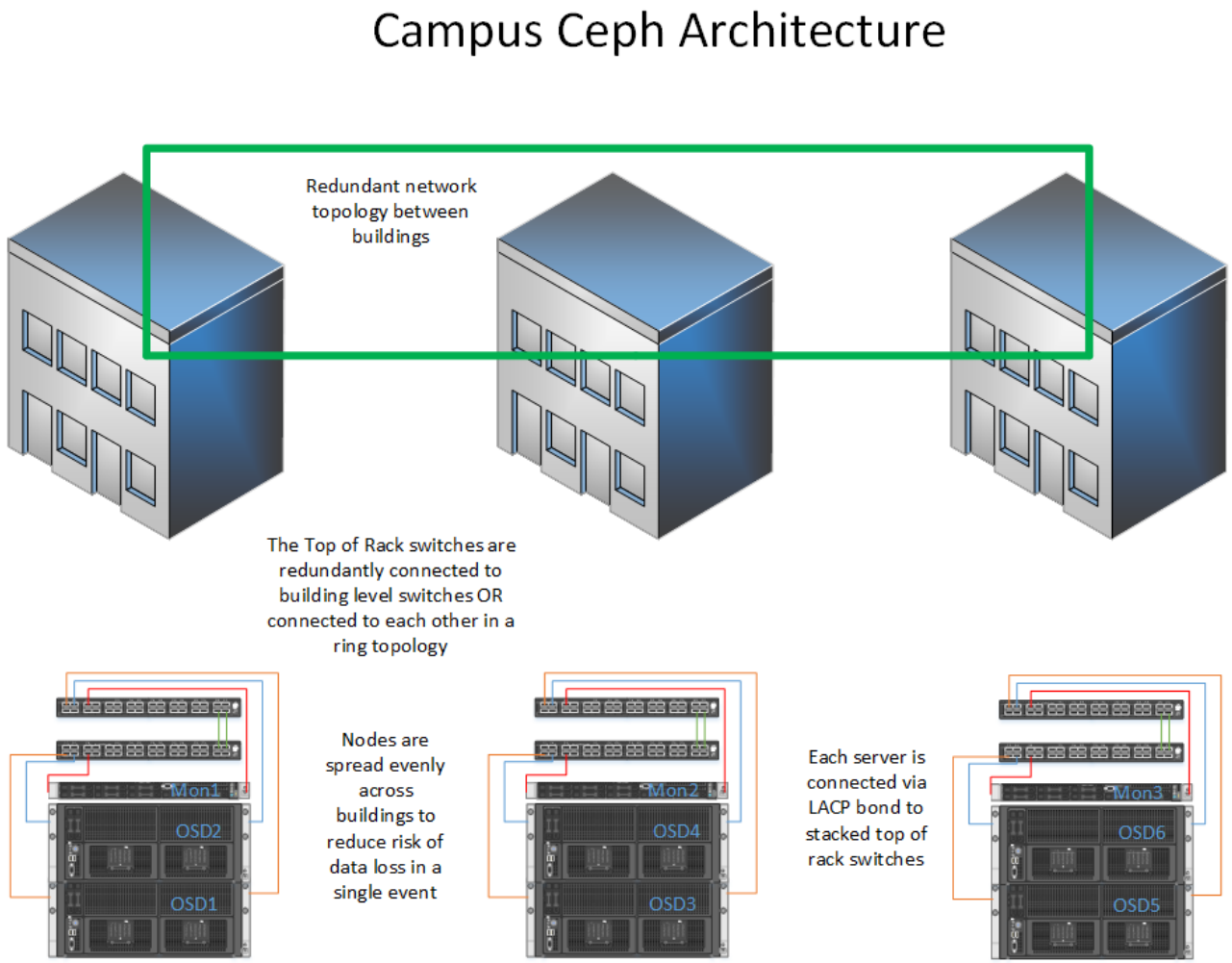
Stacked switches in ring or other redundant topology

- Bandwidth Aggregation
- Protect against switch failure

Spread across buildings on campus

- Protect against facility loss
- Provides ability to have gateway services (RGW, iSCSI) local to a building

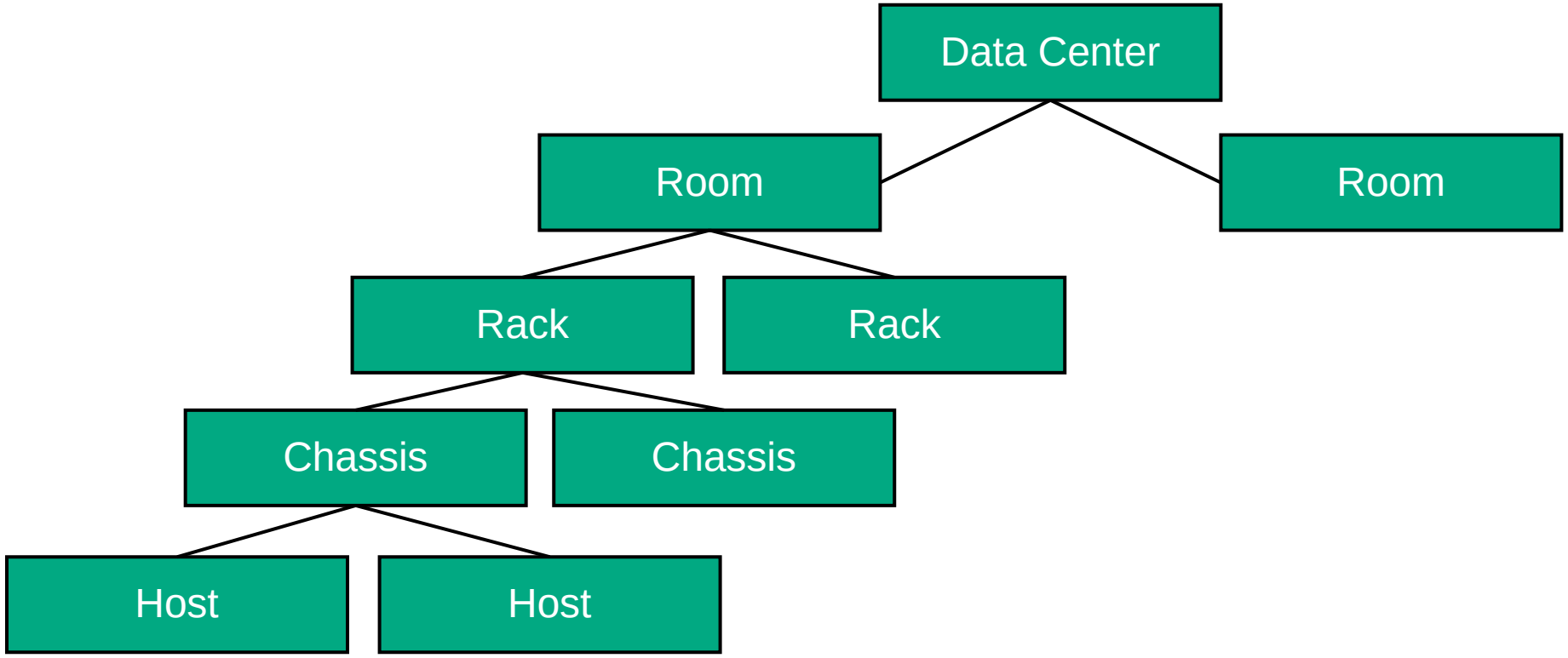
Plan for the highest resiliency



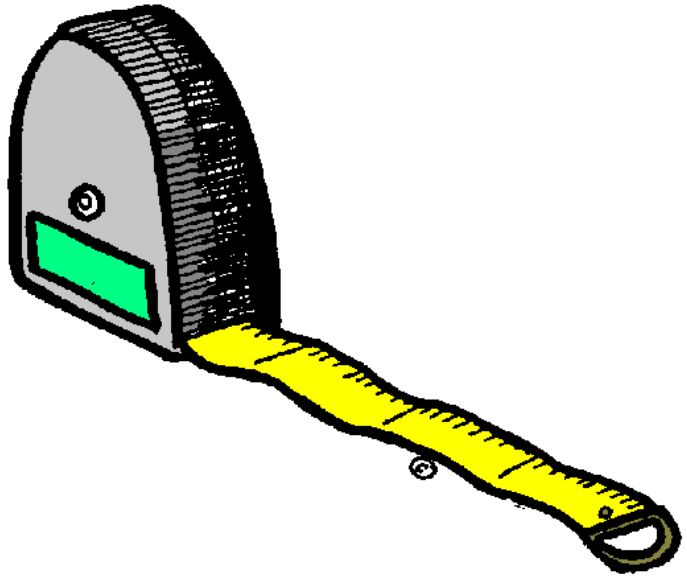
How do we make Ceph “aware”?

The Crush Map is the key!

Customize it for the site



Sizing a Solution



Use cases

Use Cases define design decisions

read heavy SSD read-only cache?

write heavy SSD/NVMe journals?

mix with reads on most recent 10% being >80% of read activity

- write-back cache tier on SSD

Block?

- SSD/NVMe journals

Rules of thumb

Storage Node (OSD)

- 2GB per TB of storage
- 2 Ghz per storage device
- 2 NICs, LACP to stacked switches

Monitor (Mon)

- 6 cores, 32GB, RAID1 SSD
- 2 NICs, LACP to stacked switches

iSCSI

- 6-8 cores, higher clock speed = better
- 2 NICs, LACP to stacked switches

Deploying The Solution

Base installation – Infrastructure

PXE

Autoyast

Custom kiwi image

SMT

NTP

DNS

Crowbar

Salt Coming Soon!

Naming Conventions

When possible, use naming conventions that are easy to identify purpose, protection, etc.

Nodes

- **By Function**
- **osd1, osdnode1, mon1, clu1-osd1**

Pools

- **doublereplica-rbd-mirror1**
- **openstack-images1**

RBD images

- **imgsvr-product-images.iscsi-lun1**

Validating your install

ceph status –w

rados benchmark

fio, iometer

COSbench – appliance image on susestudio

iscsi connect to host

Tuning

General information

Faster, Fatter Networks

- 40GbE +
- Jumbo Frames

Higher speed storage

- 7.2k <15k <SATA SSD < SAS SSD < NVMe (<NVDIMM?)

Intel Performance Portal for Ceph

- https://01.org/cephperf/ceph-performance-tunings

Resources

Available from SUSE

SUSE Enterprise Storage documents

- https://www.suse.com/documentation/ses-2/

Reference Architectures

- SUSE Enterprise Storage 2.0 Deployment for HPE ProLiant DL Series
- HP Apollo Series with SUSE Enterprise Storage

Blogs

- https://www.suse.com/communities/blog/category/suse-enterprise-storage/

Recorded webinars

- SUSE MOST – https://www.suse.com/partners/most/ (search for ceph)
- SUSE Chalk Talks – http://www.youtube.com/suse

ceph.com http://www.ceph.com

Questions and Answers


SUSE
We adapt. You succeed.