Peanut butter and jelly
Mapping the deep Integration between Ceph and OpenStack

Sean Cohen
Associate Manager, OpenStack Product Team

Sébastien Han
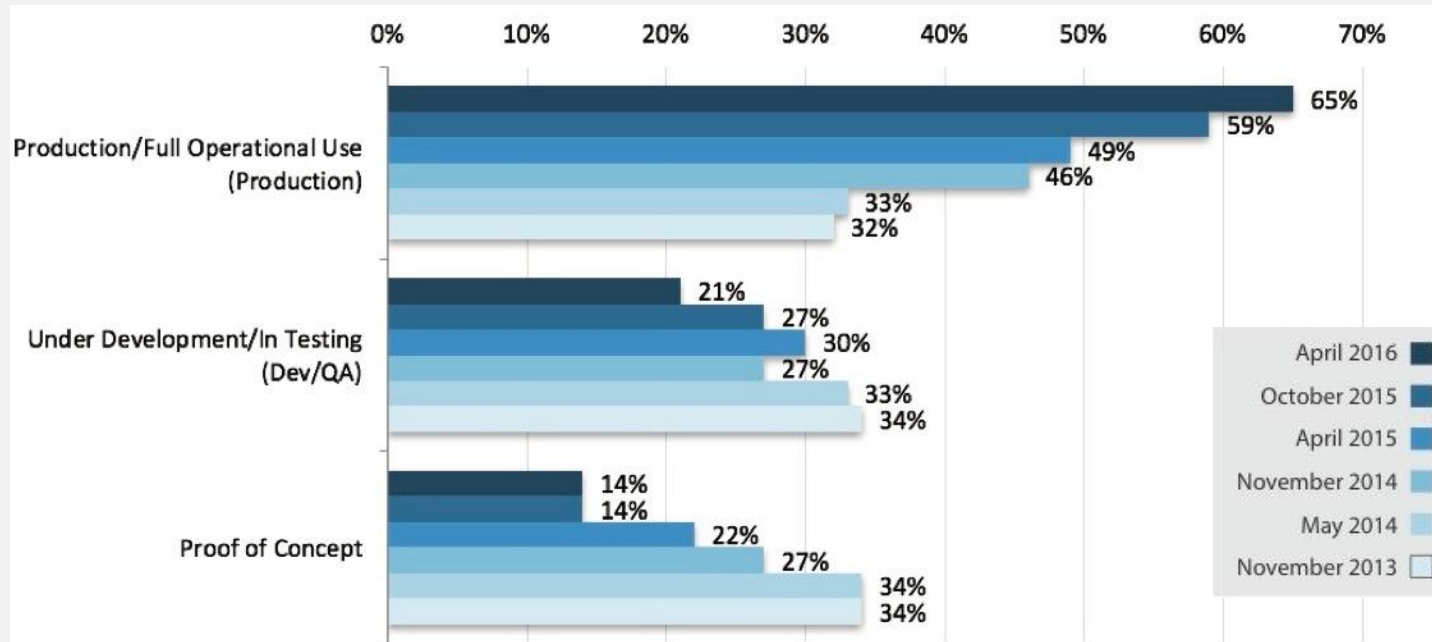Principal Software Engineer, Storage Architect

Federico Lucifredi
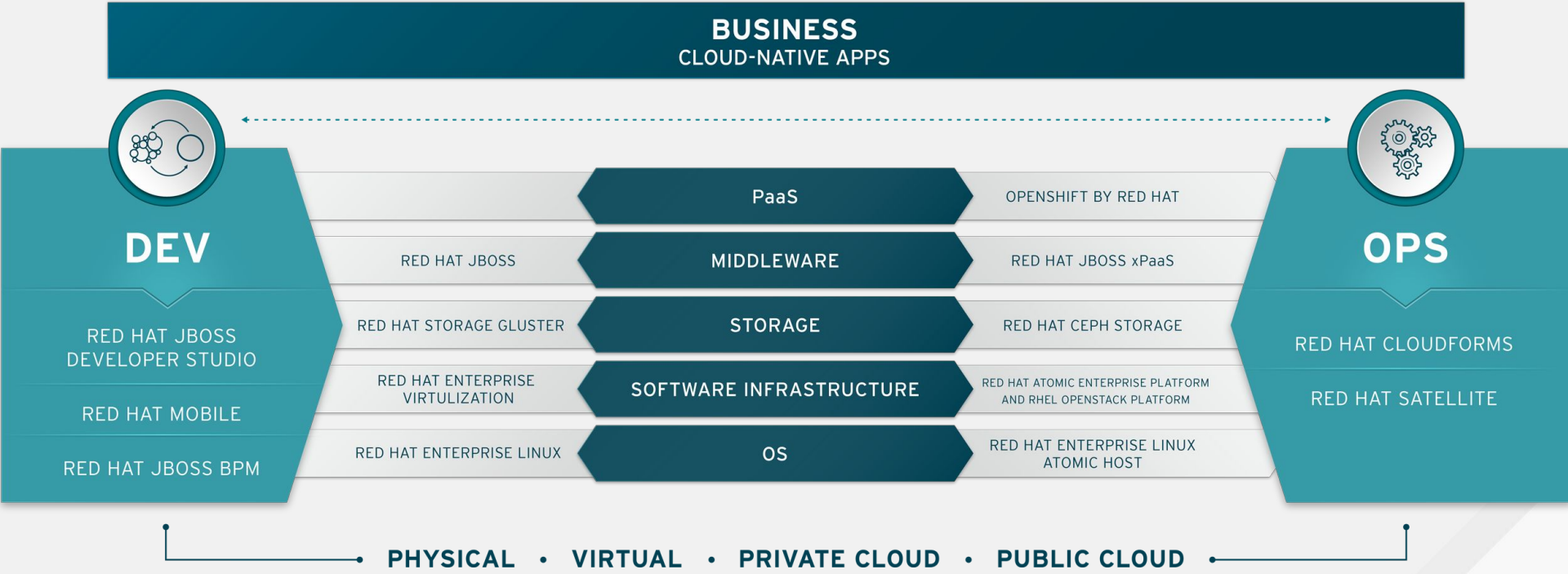Product Management Director, Red Hat Ceph Storage

#redhat #rhsummit

OPENSTACK?

OpenStack adoption

65% OF CLOUDS ARE IN PRODUCTION

Source: OpenStack Foundation April 2016 User Survey

Red Hat Open Hybrid Cloud

BUSINESS
CLOUD-NATIVE APPS

DEV

OPS

	PaaS	OPENSHIFT BY RED HAT
RED HAT JBOSS	MIDDLEWARE	RED HAT JBOSS xPaaS
RED HAT STORAGE GLUSTER	STORAGE	RED HAT CEPH STORAGE
RED HAT ENTERPRISE VIRTULIZATION	SOFTWARE INFRASTRUCTURE	RED HAT ATOMIC ENTERPRISE PLATFORM AND RHEL OPENSTACK PLATFORM
RED HAT ENTERPRISE LINUX	OS	RED HAT ENTERPRISE LINUX ATOMIC HOST

RED HAT JBOSS DEVELOPER STUDIO

RED HAT MOBILE

RED HAT JBOSS BPM

RED HAT CLOUDFORMS

RED HAT SATELLITE

PHYSICAL • VIRTUAL • PRIVATE CLOUD • PUBLIC CLOUD

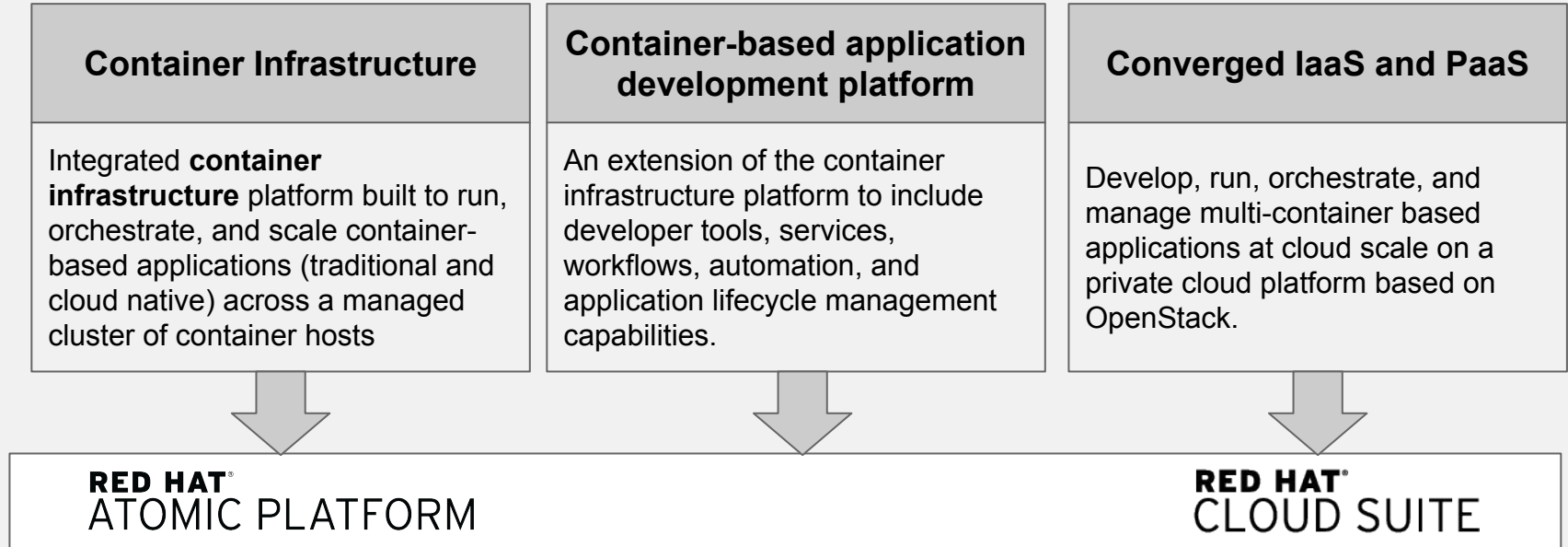redhat.

Red Hat OpenStack Platform Overview

- Red Hat OpenStack Platform delivers an integrated and open foundation to create, scale, and manage a secure and reliable public or private OpenStack cloud.

- Customer benefits:
  - Delivers a production-ready cloud platform that combines Red Hat's hardened OpenStack infrastructure
  - Co-engineered and integrated with Red Hat Enterprise Linux
  - Offers a telco-grade, massively scalable platform
  - Delivers the highest levels of OpenStack performance

redhat

OpenStack provides elastic and scalable platforms

Managed Private Cloud	Software Defined Storage	Telco/NFV	Hybrid Cloud
A Managed Private cloud is a compute platform implemented on premise under the control of the IT department, with a management framework to bring flexibility across various resource pools.	**Software** to manage policy-based provisioning and management of data **storage** independent of the underlying hardware. Can use industry standard servers and disks rather than purpose-built or proprietary appliances.	Provides a stable, robust and scalable means for Telco providers to detach from current networking equipment at a lower cost that allows them greater flexibility for future growth needs	Hybrid cloud is a cloud computing environment which uses a mix of on-premises, private cloud and third-party, public cloud services with orchestration between the two platforms.

RED HAT®
OPENSTACK
PLATFORM

**RED HAT®**
CEPH STORAGE

RED HAT®
OPENSTACK
PLATFORM

redhat.

OpenStack enables agile and responsive Apps

Container Infrastructure	Container-based application development platform	Converged IaaS and PaaS
Integrated **container infrastructure** platform built to run, orchestrate, and scale container-based applications (traditional and cloud native) across a managed cluster of container hosts	An extension of the container infrastructure platform to include developer tools, services, workflows, automation, and application lifecycle management capabilities.	Develop, run, orchestrate, and manage multi-container based applications at cloud scale on a private cloud platform based on OpenStack.

RED HAT®
ATOMIC PLATFORM

RED HAT®
CLOUD SUITE

Largest Certified Partner Ecosystem

- Over 350+ members since launch in April 2013
- Over 900 certified solutions in partner Marketplace
- Over 4,000 RHEL certified compute servers
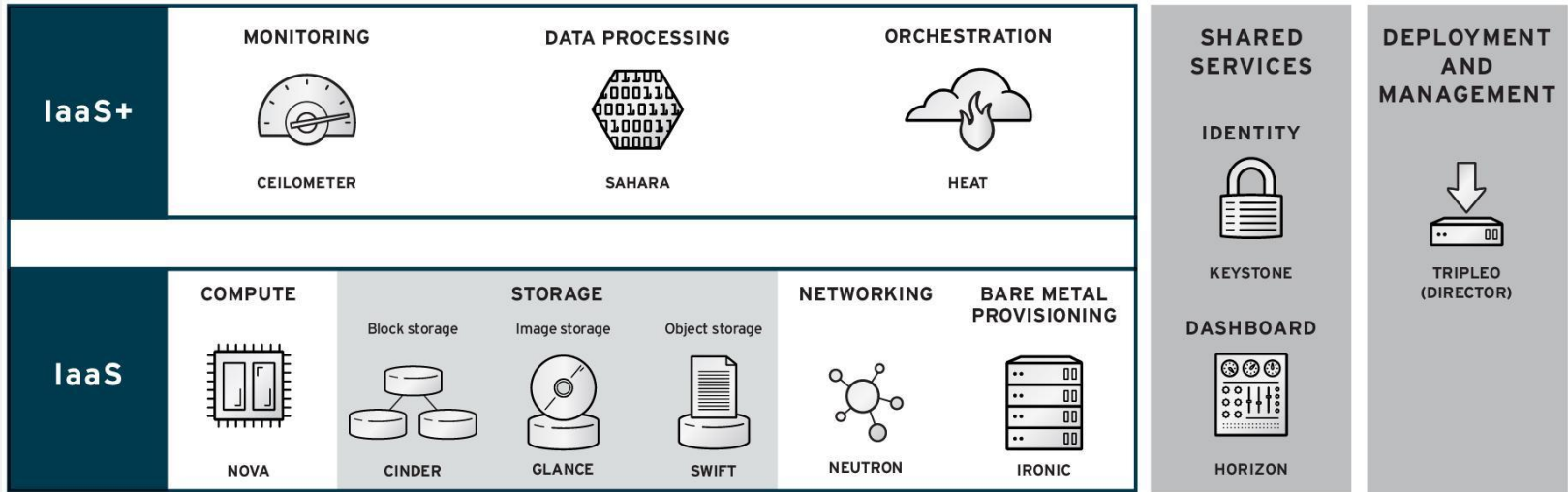
OEMs, IHVs, ISVs
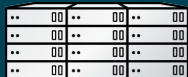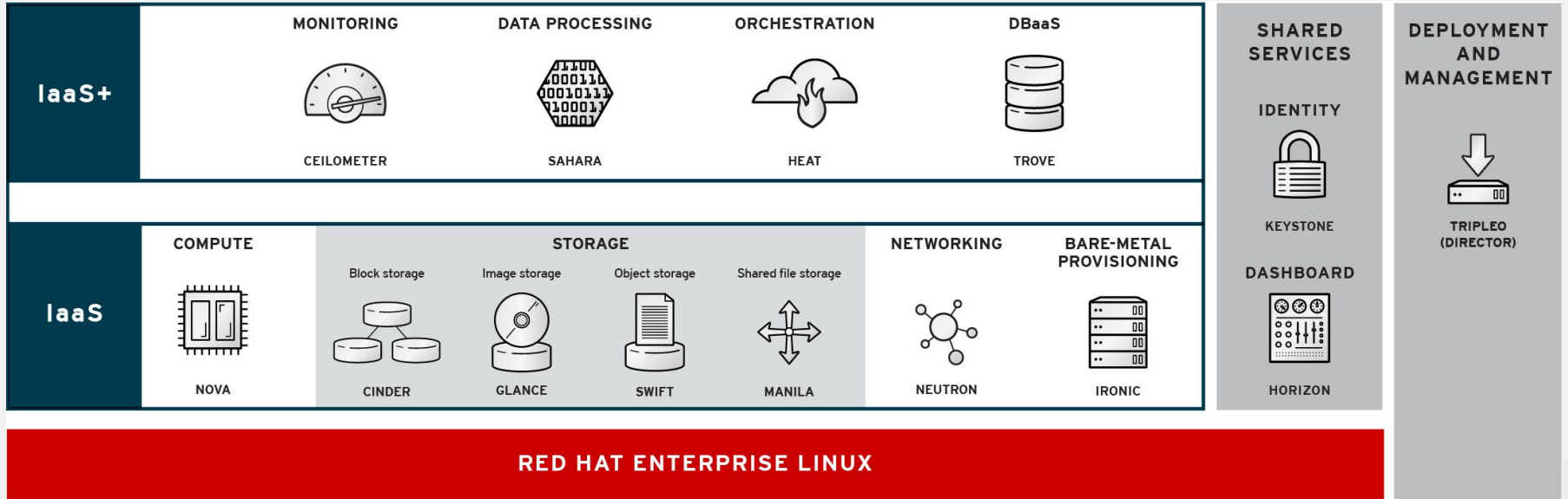
Channel Partners

System Integrators

Cloud Service Providers
Managed Service Providers

redhat.

OpenStack: Framework for the Cloud

- Needs to access hardware resources
- Needs an operating environment, hypervisor, services
- Leverages existing code libraries for functionality

Main components

		MONITORING	DATA PROCESSING	ORCHESTRATION		DBaaS			SHARED SERVICES	DEPLOYMENT AND MANAGEMENT
IaaS+		CEILOMETER	SAHARA	HEAT		TROVE			**IDENTITY**	
	COMPUTE	**STORAGE**				**NETWORKING**	**BARE-METAL PROVISIONING**		KEYSTONE	TRIPLEO (DIRECTOR)
		Block storage	Image storage	Object storage	Shared file storage				**DASHBOARD**	
IaaS	NOVA	CINDER	GLANCE	SWIFT	MANILA	NEUTRON	IRONIC		HORIZON	

RED HAT ENTERPRISE LINUX

redhat.

OpenStack Storage

OpenStack supports four types of persistent Storage:

- **Block storage (Cinder)** Provides persistent block storage to running instances. Its pluggable driver architecture facilitates the creation and management of block storage devices.
- **Image Storage (Glance) -** Stores and retrieves virtual machine disk images. OpenStack Compute makes use of this during instance provisioning.
- **Shared File Systems (Manila) -** provides a set of services for management of shared file systems such as CIFS, NFS, CephFS & HDFS in a multi-tenant environment.
- **Object storage (Swift)** - Stores and retrieves arbitrary unstructured data objects via a RESTful, HTTP based API. It is highly fault tolerant with its data replication and scale out architecture. Its implementation is not like a file server with mountable directories.

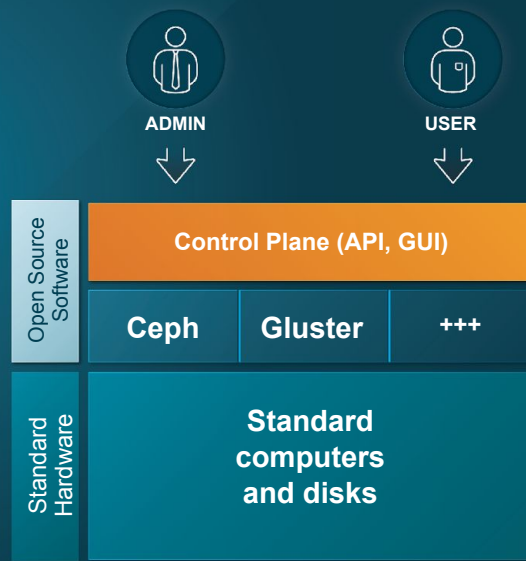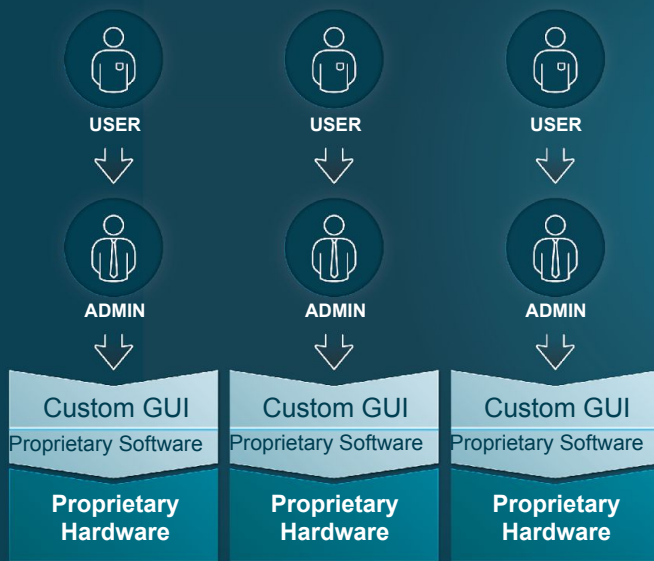# Storage challenges in OpenStack

- Multiple storage interfaces to consume
  - File
  - Block
  - Object
- They don't all scale at the same pace
- Mixing up storage technologies is painful to manage, needs more talents/skills

redhat.

Rising tide of software-defined storage

"By 2016, server-based storage solutions will lower storage hardware costs by 50% or more."

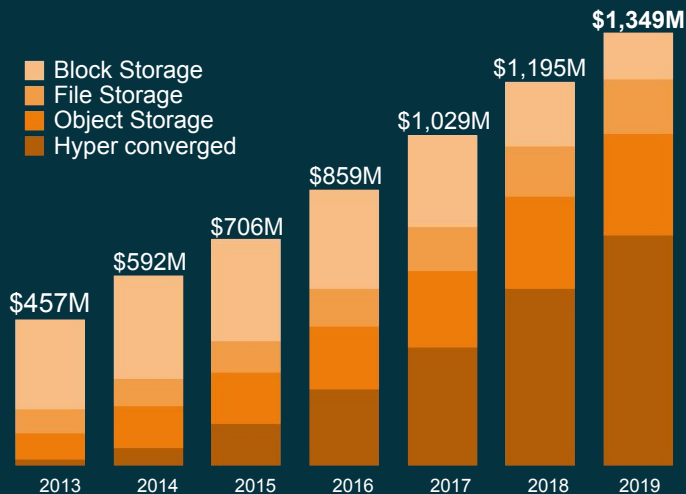Gartner: "IT Leaders Can Benefit From Disruptive Innovation in the Storage Industry"

"By 2020, between 70-80% of unstructured data will be held on lower-cost storage managed by SDS environments."

Innovation Insight: Separating Hype From Hope for Software-Defined Storage

"By 2019, 70% of existing storage array products will also be available as software only versions"

Innovation Insight: Separating Hype From Hope for Software-Defined Storage

SDS MARKET SIZE BY SEGMENT

- Block Storage
- File Storage
- Object Storage
- Hyper converged

$457M — 2013
$592M — 2014
$706M — 2015
$859M — 2016
$1,029M — 2017
$1,195M — 2018
$1,349M — 2019

Source: IDC

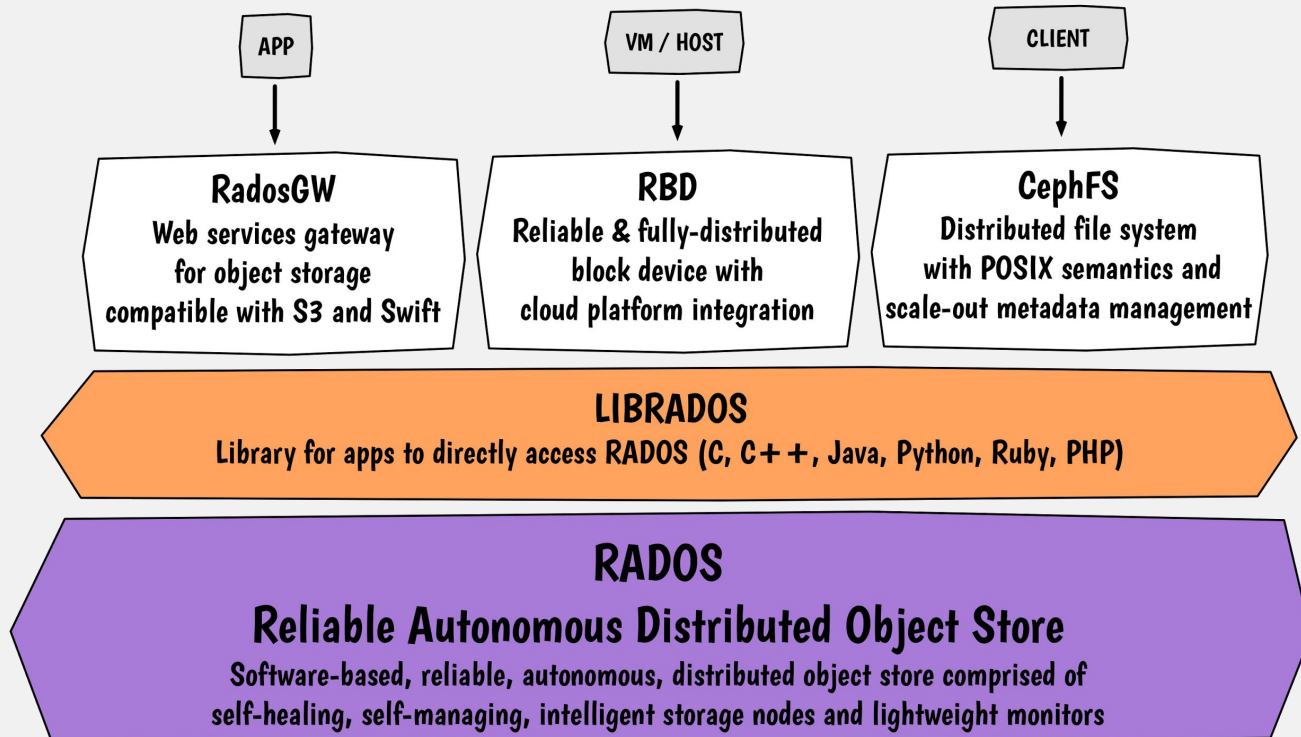~20% CAGR between 2015 and 2019

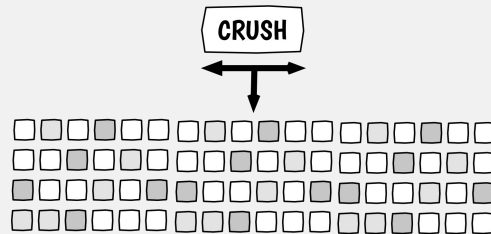redhat.

CEPH?

What is Ceph?

- Open, massively-scalable, software-defined
- Flexible, scale-out architecture on clustered commodity hardware
- Single, efficient, unified storage platform
- User-driven storage lifecycle management with 100% API coverage
- Integrated, easy-to-use management console
- Designed for cloud infrastructure and emerging workloads

CEPH OVERVIEW

APP

VM / HOST

CLIENT

RadosGW
Web services gateway
for object storage
compatible with S3 and Swift

RBD
Reliable & fully-distributed
block device with
cloud platform integration

CephFS
Distributed file system
with POSIX semantics and
scale-out metadata management

LIBRADOS
Library for apps to directly access RADOS (C, C++, Java, Python, Ruby, PHP)

RADOS
Reliable Autonomous Distributed Object Store
Software-based, reliable, autonomous, distributed object store comprised of
self-healing, self-managing, intelligent storage nodes and lightweight monitors

CRUSH

Controlled **R**eplication **U**nder **S**calable **H**ashing:

- Pseudo-random placement algorithm
- Statistically uniform distribution (hash based)
- Rule-based configuration
- Topology aware

RED HAT®
CEPH STORAGE 2

Powerful, production-grade, distributed storage for OpenStack

- Open, massively scalable, and software-defined
- Flexible, scale-out architecture on clustered commodity hardware
- Specifically designed for cloud infrastructure and emerging workloads
- Seamlessly integrated: Block (ephemeral and persistent), Object, and file storage on COTS
- 64TB included in Red Hat OpenStack Platform

redhat.

RED HAT® CEPH STORAGE 2

"DevOps" deployment

ceph-ansible (with Ansible 1.9)

"Apple" deployment

Red Hat Storage Console 2

RGW

Multisite v2
AWSv4
LDAP/AD
KeyStone v3
Swift API updates (Multi-tenant, Expiration, SLO, Bulk Delete)

RBD

Volume Mirroring

Tech Previews

BlueStore

CephFS

NFS Gateway for RGW

RADOS

Improved Security
(uid 'ceph' for all processes)

Reduced impact of scrubbing on client I/O
(Unified Queue)

3x small write performance on SSD with RHEL 7.2

redhat.

OPENSTACK AND CEPH INTEGRATION
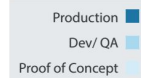
Ceph adoption in OpenStack
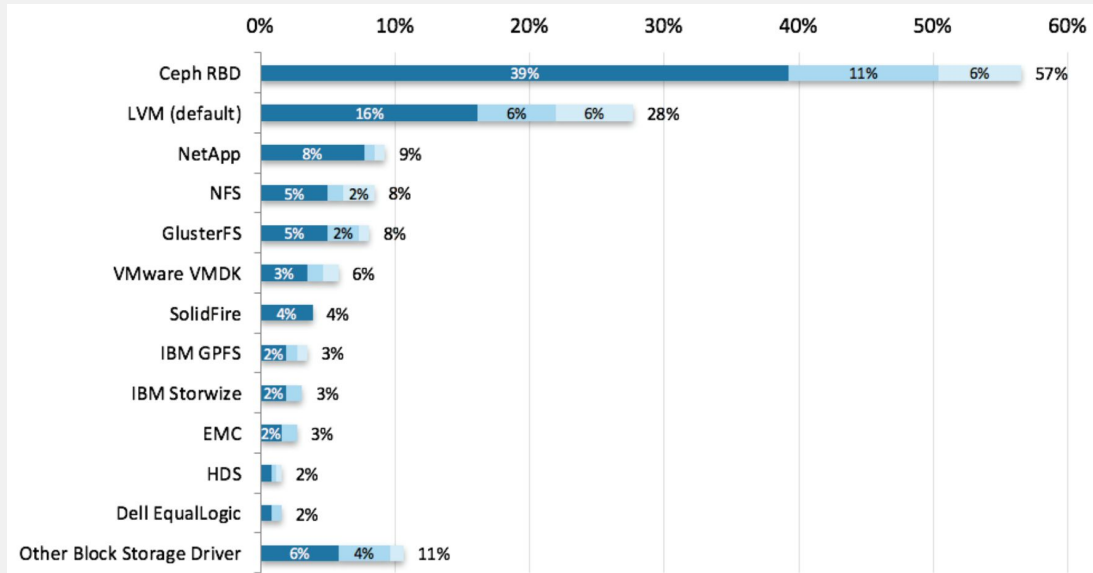
Figure 5.7 n=260

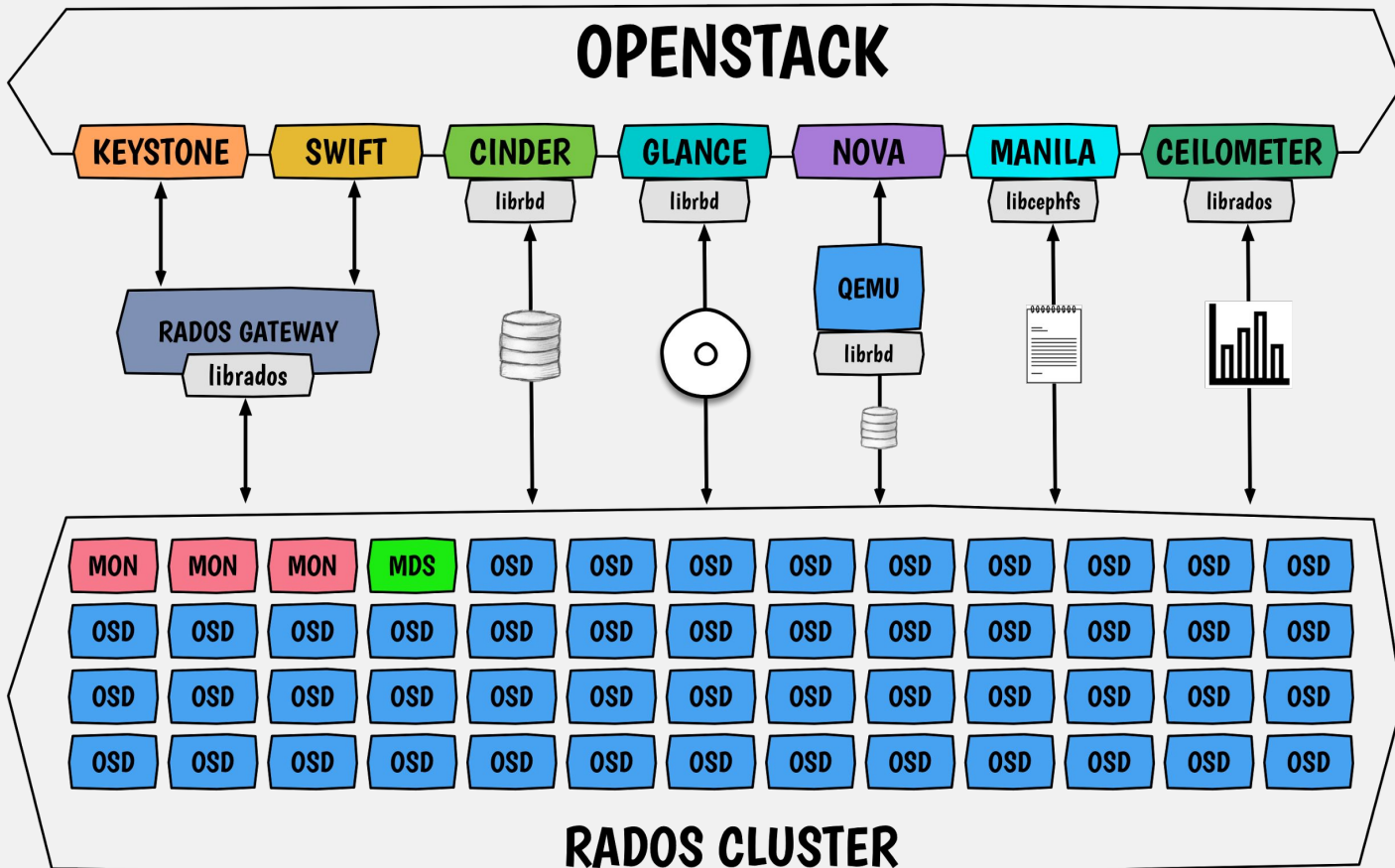Percentages are rounded to the nearest whole number; bar length shows fractions.
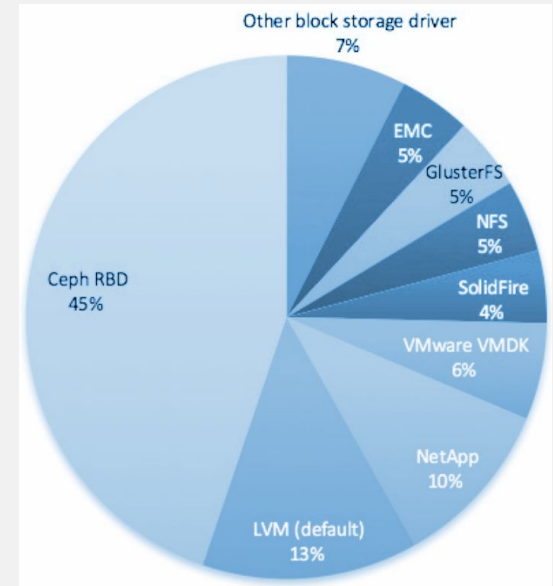
The unified story

- **Ongoing effort** since OpenStack's creation
- **Ceph is the foundation** of the storage stack in OpenStack
- Provides a single and flexible storage layer
- **Integrates** perfectly with all **OpenStack's components**

Tight product integration

As of Red Hat OpenStack Platform 8 you get:

- 64TB capacity of Red Hat Ceph Storage unrestricted use with Red Hat OpenStack Platform storage services.
- Red Hat OpenStack Platform Director (deployment tool) integration - RHCS is the default block storage for RHOSP.
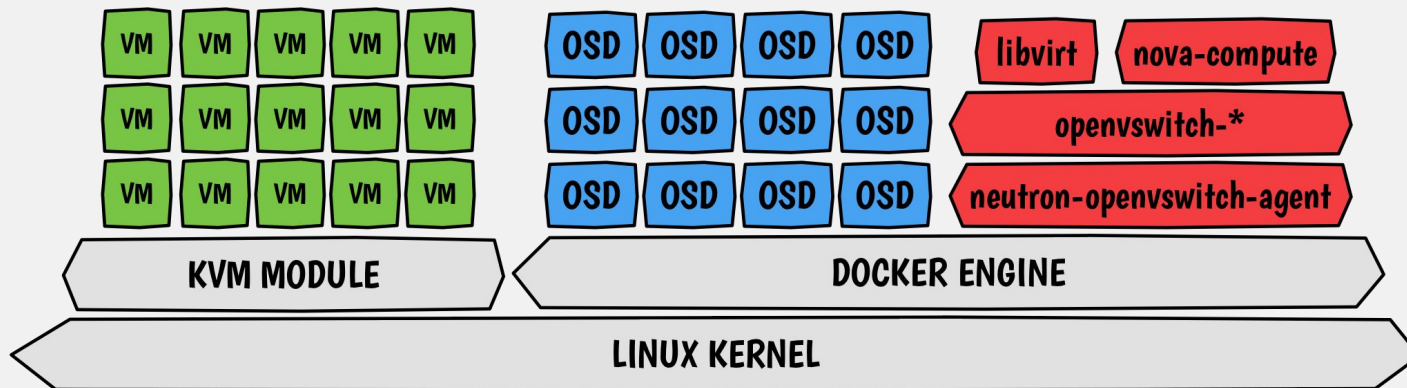
OPENSTACK ARCHITECTURE EXAMPLES WITH CEPH

Hyperconverged

Hypercon... what?

- **Co-locate compute** and **storage** resources on the same machine
- **Fine control of resources** using cgroups, NUMA and CPU pinning
- Enabler for **container-based infrastructure**
- Can benefit from a local hit when performing an IO from a guest
- Component upgrades made easy with containers

HYPERCONVERGED NODE IN-DEPTH

VM qemu-kvm

VM	VM	VM	VM	VM
VM	VM	VM	VM	VM
VM	VM	VM	VM	VM

KVM MODULE

OSD	OSD	OSD	OSD
OSD	OSD	OSD	OSD
OSD	OSD	OSD	OSD

libvirt	nova-compute
openvswitch-*	
neutron-openvswitch-agent	

DOCKER ENGINE

LINUX KERNEL

redhat.

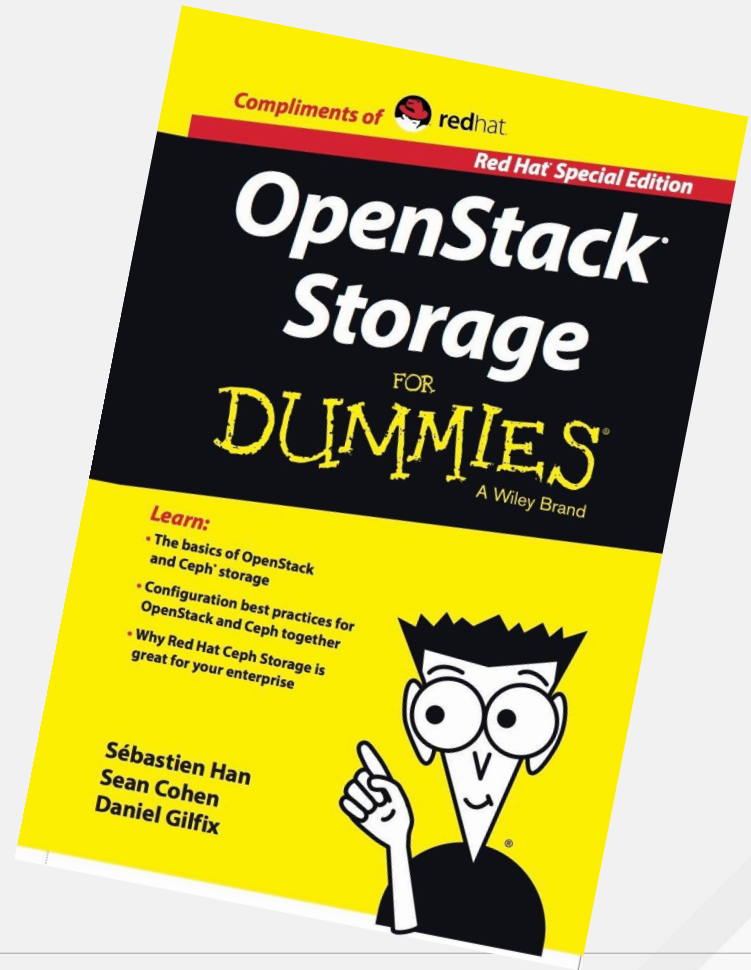# Multi-site with replicated storage

- Re-using our basic robust stack
- Multiple **isolated** OpenStack environments
- With **replicated storage** for disaster recovery using Ceph
- Each site has each other's data
- Live synchronization
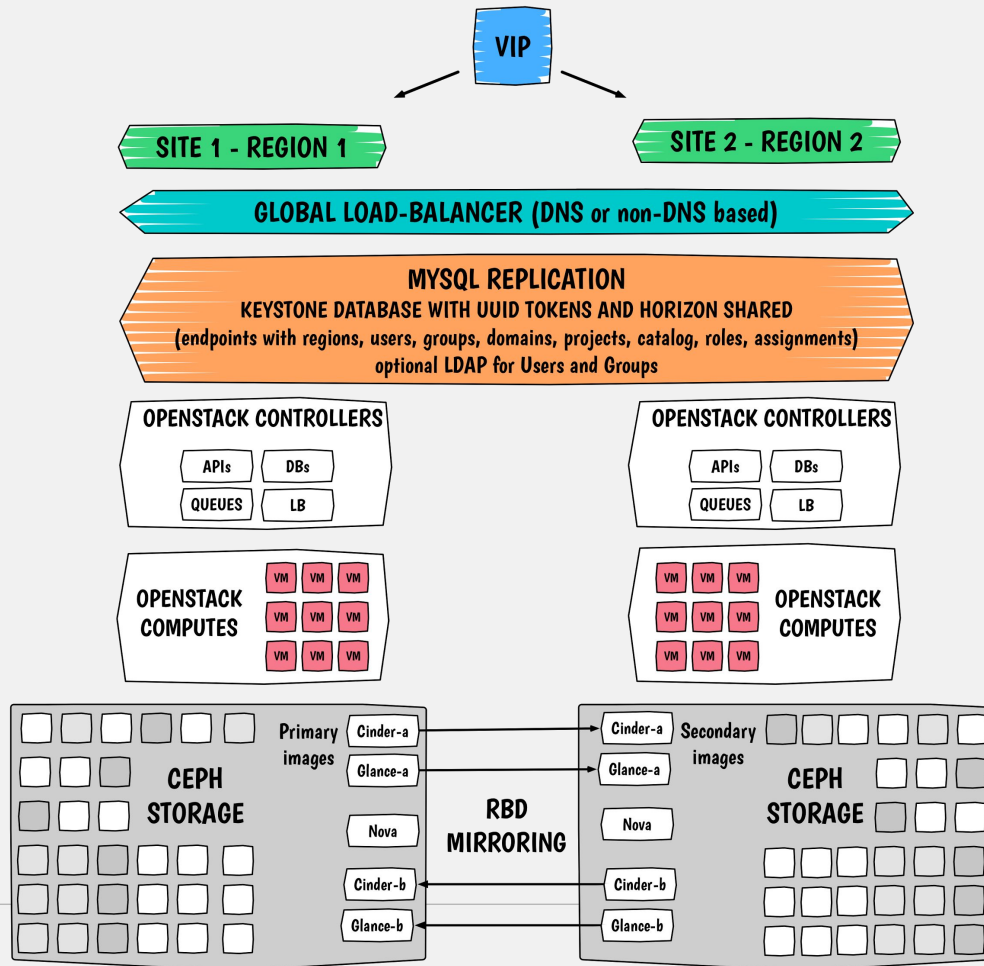
Thanks!

Sean Cohen - scohen@redhat.com
Sébastien Han - seb@redhat.com
Federico Lucifredi - federico@redhat.com