

redhat.

# State of OpenShift on Bare Metal

## OpenShift Commons Gathering - Seattle

Jose Palafox, Technical Program Manager for CNCF, Intel

Jeremy Eder, Senior Principal Performance Engineer, Red Hat

Dave Cain, Senior Architect, Red Hat

December 10th, 2018

# Agenda

1. Background and why
2. Futures and where we're going
3. Coming soon in OpenShift
4. What you can consume today



# Background

- Bare Metal Cloud Market

- \$26.21 billion by 2025; CAGR: ~40%
- AWS, IBM, Oracle, and other cloud providers offering BM
- Driven by demand for performance and latency sensitive applications that take advantage of low-level hardware features



## Announcing General Availability of Amazon EC2 Bare Metal Instances

Posted On: May 17, 2018

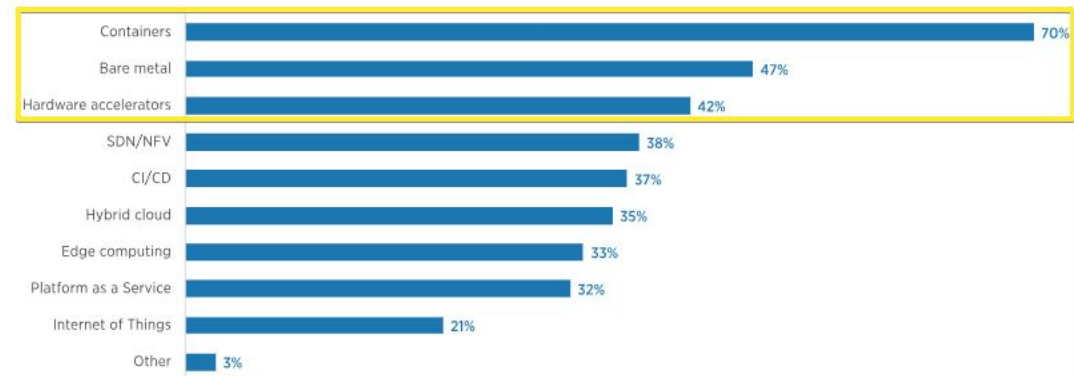
Amazon Elastic Compute Cloud (EC2) bare metal instances provide your applications with direct access to the processor and memory resources of the underlying server. These instances are ideal for workloads that require access to the hardware feature set (such as Intel® VT-x), or for applications that need to run in non-virtualized environments for licensing or support requirements. Bare metal instances are built on the Nitro system, a

- <https://www.marketsandmarkets.com/PressReleases/bare-metal-cloud.asp>
- <https://www.grandviewresearch.com/press-release/global-bare-metal-cloud-market>

IBM brings the ease of containers to complex workloads with managed Kubernetes on bare metal

March 14, 2018 | Written by: Jason McGee, VP and IBM Fellow, IBM Cloud

## Which emerging technologies interest OpenStack users?



<https://www.openstack.org/user-survey/2018-user-survey-report>

# Growing interest and adoption

- Majority of Containers and Kubernetes environments run on/in VMs
  - OpenStack, AWS/Azure/Google, vSphere
- OpenShift on bare metal interest growing, driven by multiple factors
  - Reducing VM sprawl and cost (software/infra expense)
  - Application workload & performance requirements driven (accelerators)
    - Device Manager

# Intel and Hardware Accelerators

**ALTERA**  
now part of Intel

Deep Learning  
Multi use FPGA and Smart NIC FPGAs  
QuickAssist for TLS and Compression offload  
Discrete GPU



Intel®  
QuickAssist  
Technology

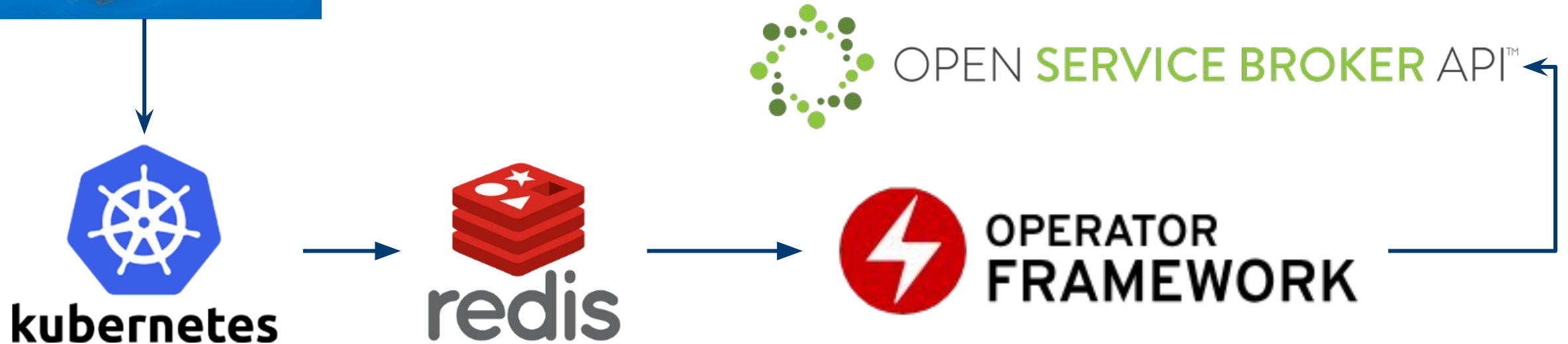


# Futures and where we're going

# Hardware Accelerated Services

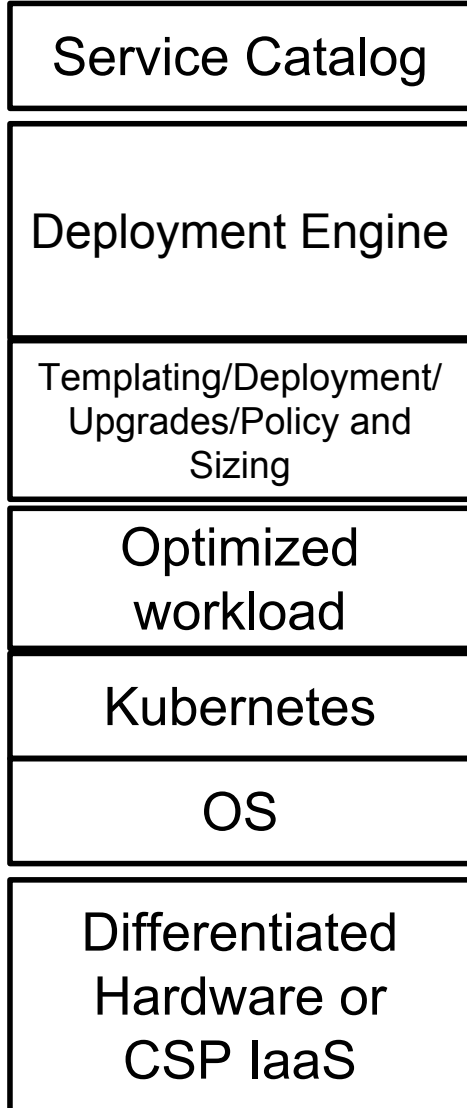


- Expose Accelerator or Data Center Device to Kubernetes
- Identify target use cases
- Create provisioning and management tools
- Enable Developer Self Service





# End to End Enablement



Publish Optimized Service Catalog Extension



Create service broker for on premise services



Drive community standardization around Redis Operator



redis

Modify Redis Upstream to use Persistent Memory



RED HAT®  
OPENSSHIFT

Expose Persistent Memory to the scheduler via CSI-Driver

Ensure Persistent Memory is exposed to the Kernel in RHEL and RHEL Variants

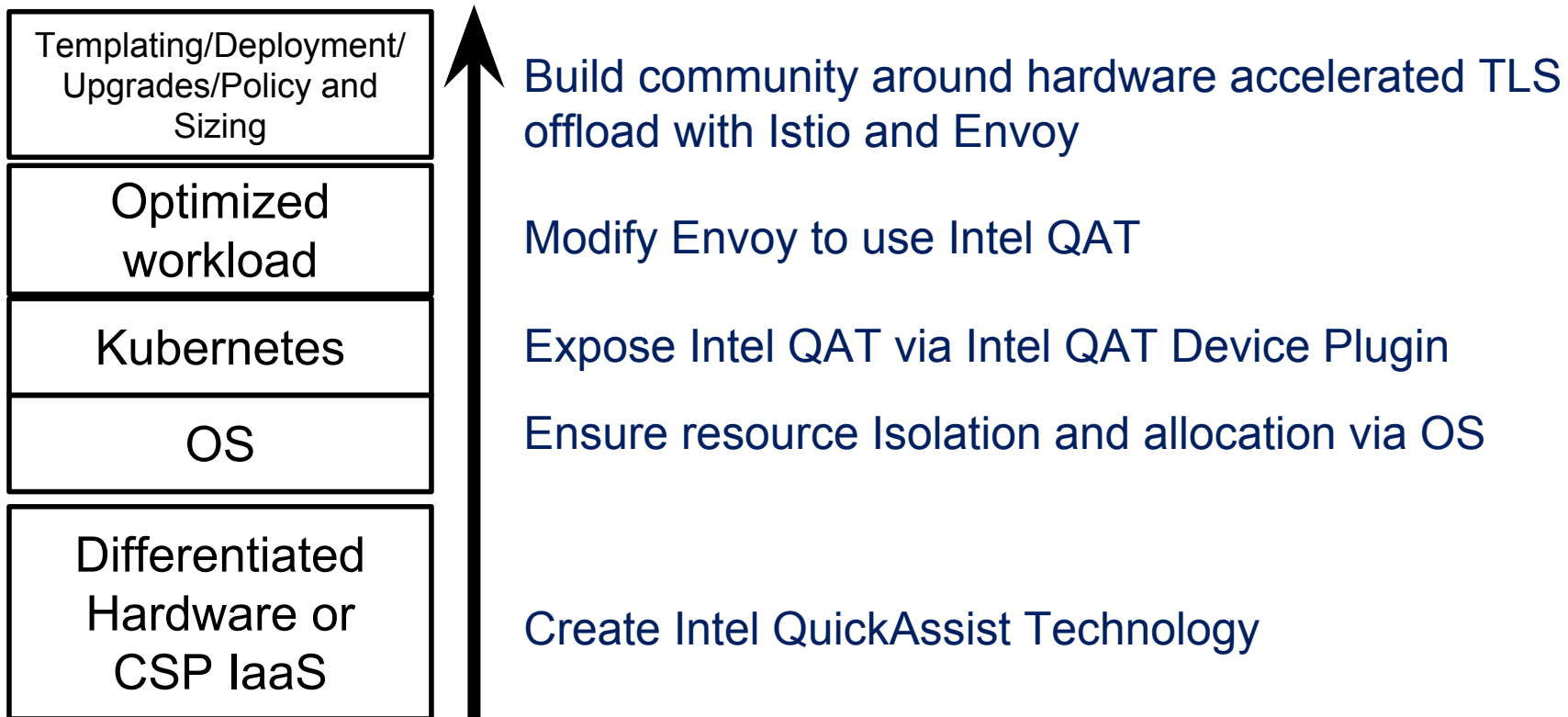
RED HAT®  
ENTERPRISE LINUX®

Create Optane for DC and Ensure OxM Support





# End to End Enablement

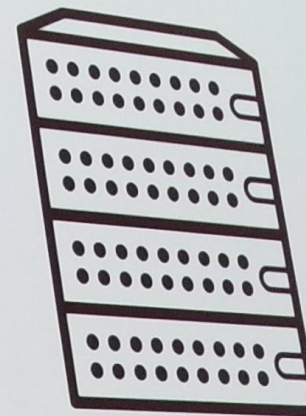


Current PR: <https://github.com/envoyproxy/envoy/pull/5161>

# Coming soon in OpenShift Container Platform

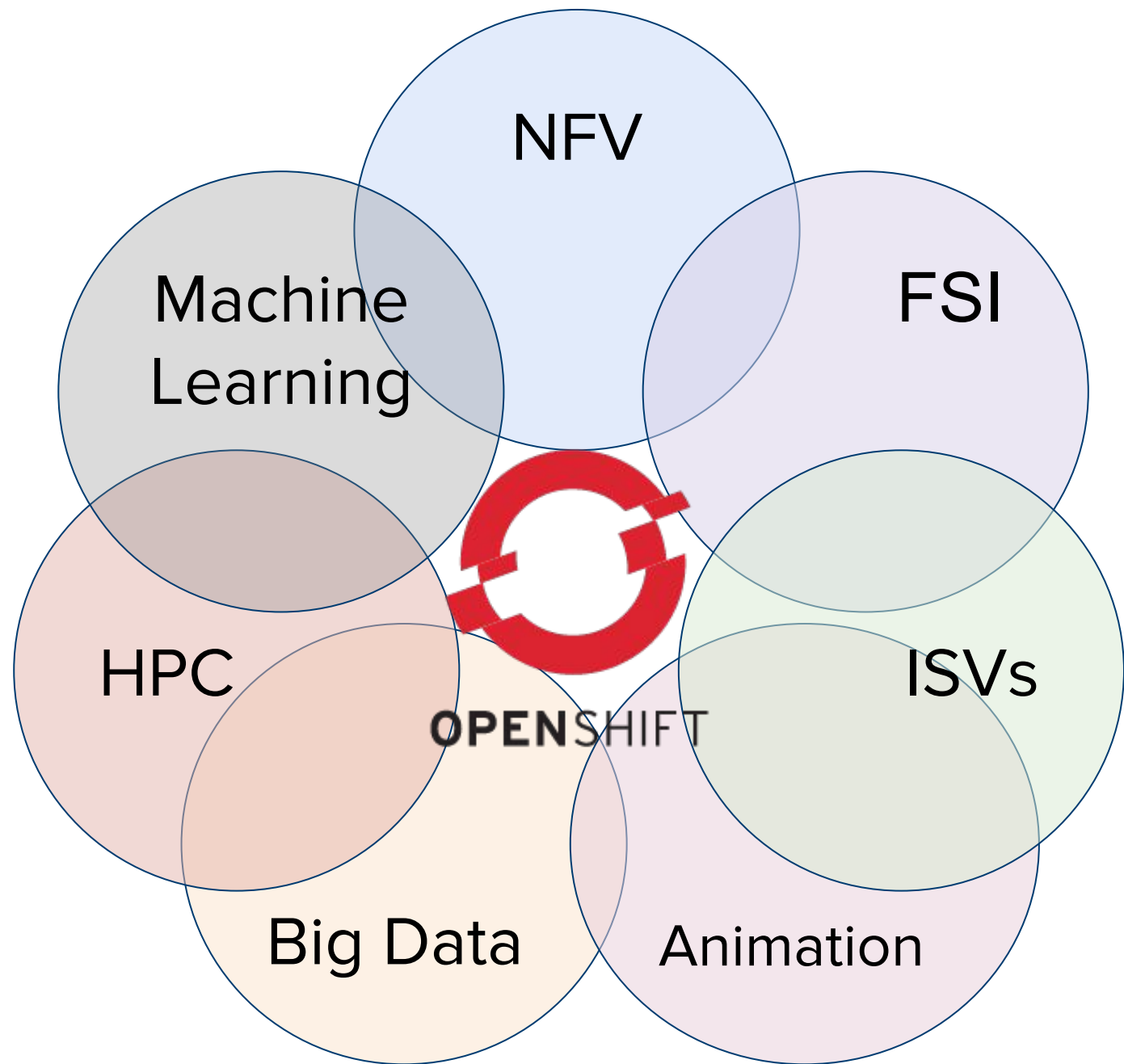
# Bare Metal

- OpenShift 4 brings more simplicity and automation
  - Applied to bare metal clusters, as well
- Bare metal infrastructure management that
  - Is automated
  - Works with Kubernetes native interfaces
- Vision of a converged infrastructure platform
  - including VMs with KubeVirt
- Enabling key environments and use cases
  - Performance sensitive applications
  - Edge architectures



So much overlap...

Coordinate, and plumb  
these generically

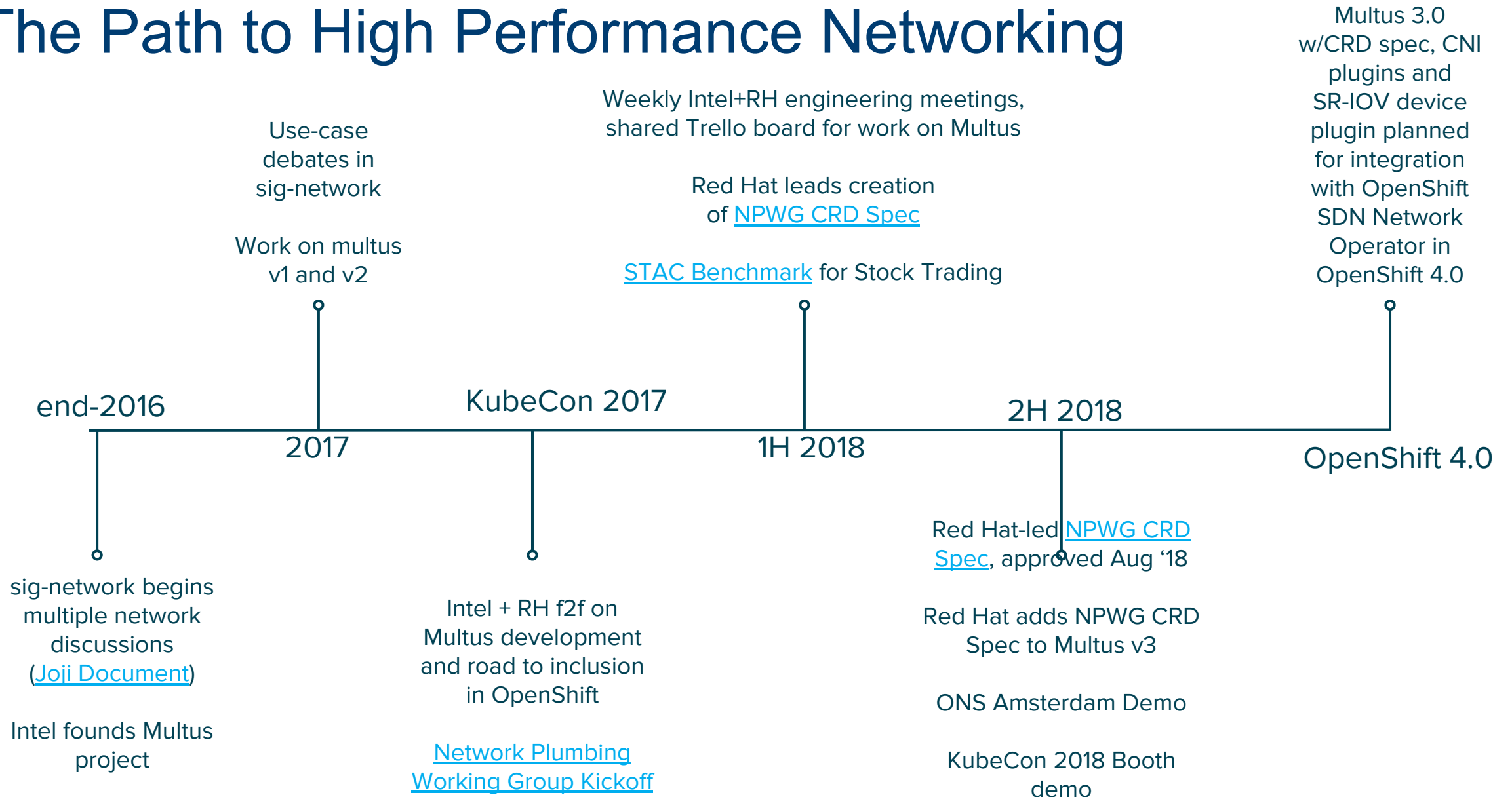


# Use-cases for High Performance Networking

Use-case/Vertical	Separate Control/Data	Fast Data Plane (no overlay)	SR-IOV	Kernel Bypass
Telco/NFV	Yes	Yes	Yes	Yes
Financial Services	Yes	Yes	Maybe	Yes
Video Streaming	Yes	Yes	Yes	Maybe
Software Defined Storage	Yes	Yes	Maybe	No
KubeVirt	Maybe	Yes	Yes	Maybe

Technology Availability	Bare Metal	AWS	Azure	GCE	OpenStack	VMware
Separate Control/Data	Yes	Yes	Yes	Yes	Yes	Yes
SR-IOV	Yes	Yes	Yes	No (virtio only)	Yes	Yes
Additional NICs	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Bypass	Yes	Maybe	No	No	Yes	Yes

# The Path to High Performance Networking



# Intel + Red Hat Open Source Collaboration

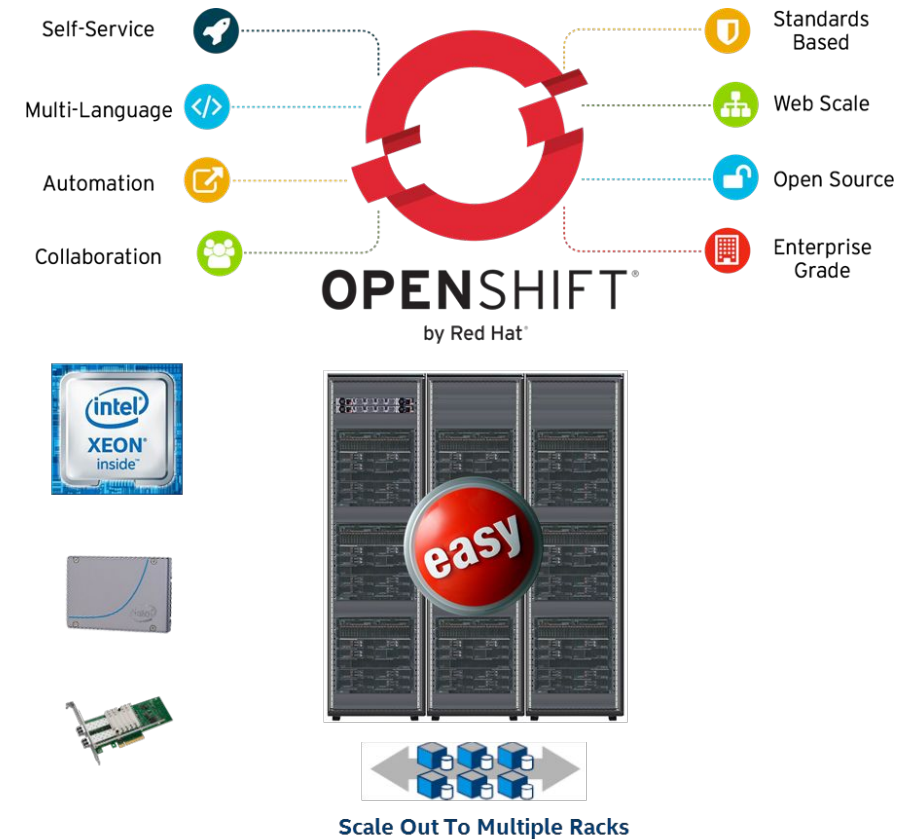
- sig-node
  - Resource Management Working Group
    - Device Plugins, CPU Pinning, NUMA/Topology
  - Node Feature Discovery
- sig-network
  - Network Plumbing Working Group
    - Multi-Network CRD Specification
    - Multus (multiple networks)



# What you can consume today

# Intel / Red Hat Solution Goals

- A fully integrated container application platform (CaaS/PaaS) with enterprise grade Kubernetes and infrastructure management, simplified for on-premises deployment, management, and scale
- Capable of deployment in about an hour, with high availability and persistent storage, automated with Red Hat Ansible, and configured to be horizontally scalable to multiple racks
- Turnkey, easy to order solution (pre-loaded hardware and software), interconnected via a leaf-spine network topology
- Differentiation: A bare-metal on-premises solution purpose-built for dev-test use cases, optimized for high performance and durability, utilizing hardware infrastructure including Intel® Xeon® Scalable processors, Intel® Solid State Drive (SSD) technology, and Intel® Network Interface Cards (NICs)
- A fast method of releasing cloud-based, stateful (persistent) containerized applications to a target audience, providing customers with an enterprise grade solution



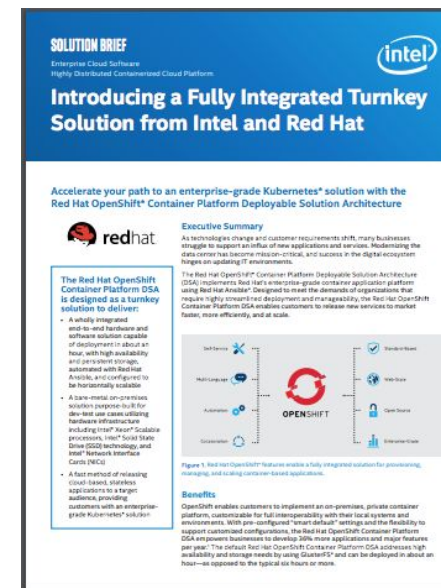
# Solution Launch (December 2017)

<https://red.ht/intelopenshift>

Reference Architecture (24 pages)

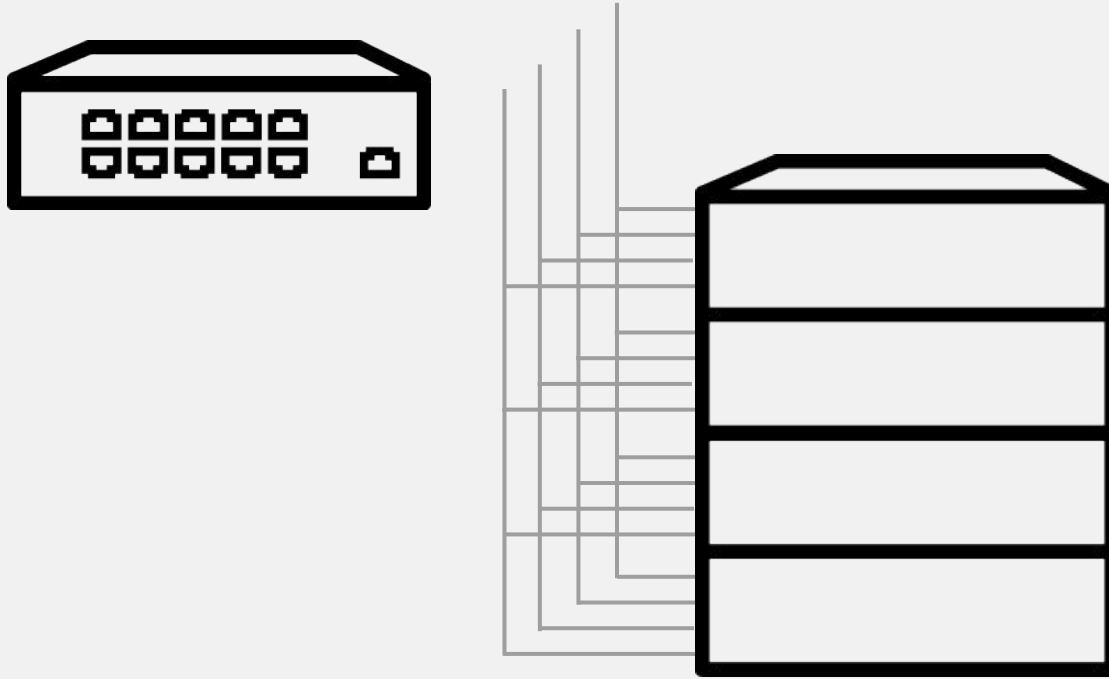


Solution Brief (2 pages)



# Solution Deployment Workflow

New hardware, racked and wired



# Solution Deployment Workflow

Identify management node for the bastion



# Solution Deployment Workflow

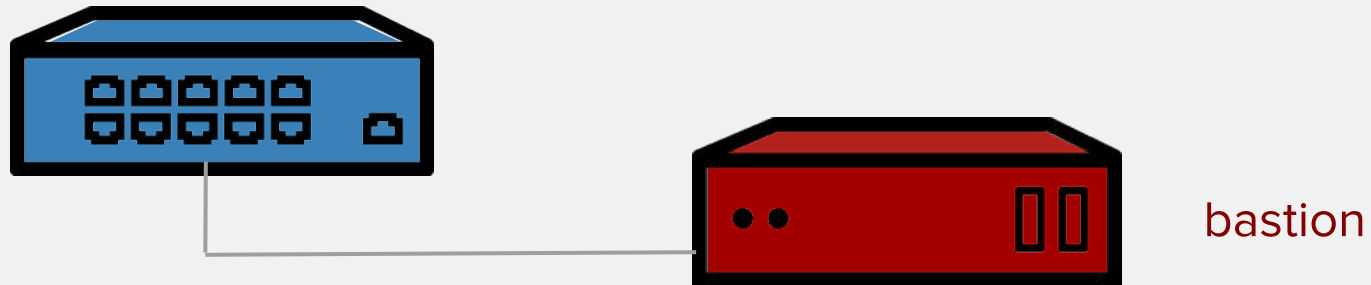
Configure and provision bastion node, subscribe, and download ansible playbooks



bastion

# Solution Deployment Workflow

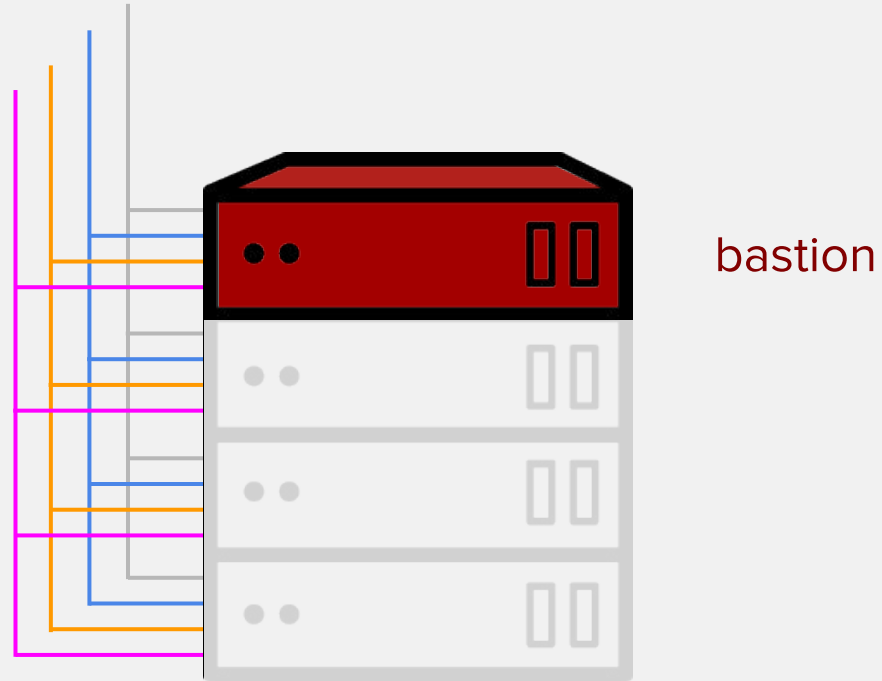
Provision networking infrastructure





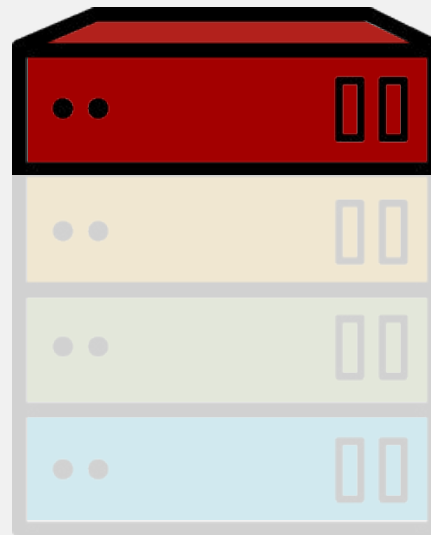
# Solution Deployment Workflow

Baremetal server preparation with RHEL deployed



# Deployment Workflow

Define OpenShift Nodes: Master, Infrastructure, Application, and Storage



bastion

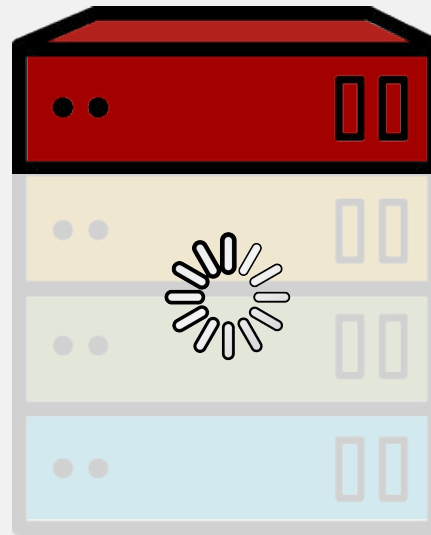
Master Nodes

Infrastructure Nodes

Application Nodes

# Deployment Workflow

Launch openshift-ansible playbooks



bastion

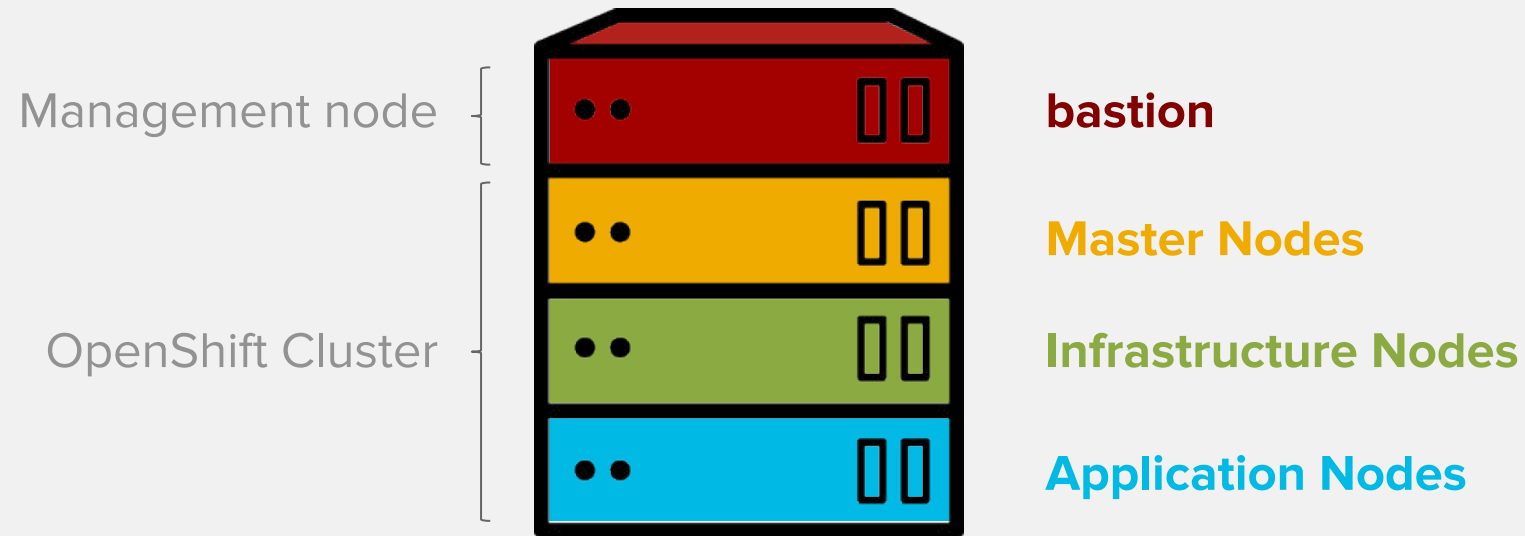
Master Nodes

Infrastructure Nodes

Application Nodes

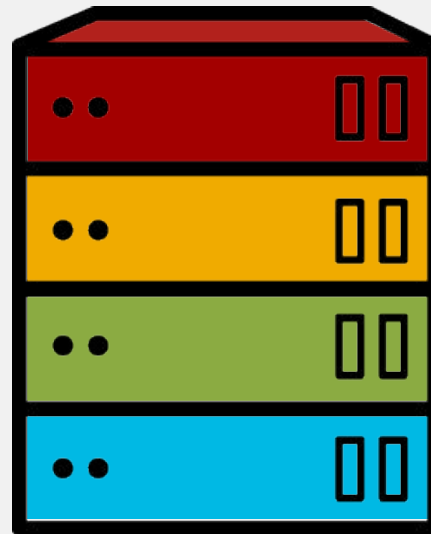
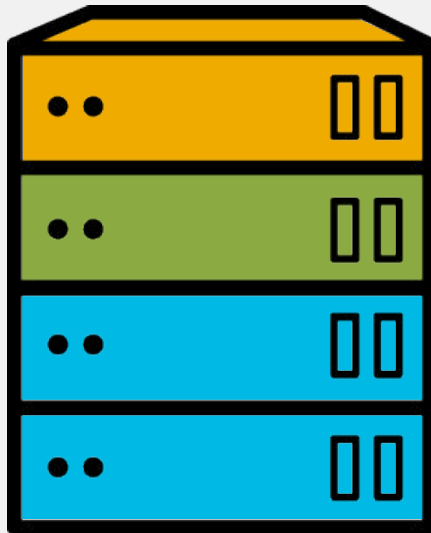
# Deployment Workflow

Deployed OpenShift Container Platform



# Deployment Workflow

Deployed OpenShift Container Platform



# Available Today

Lenovo Reference Architecture: <https://bit.ly/2RDSyKR>



Cisco CVD Design Guide: <https://bit.ly/2G3YW8e>

Cisco CVD Deployment Guide: <https://bit.ly/2MraKbW>



Dell EMC: <https://dell EMC.com/openshift>

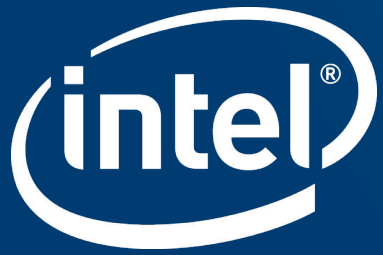


# Drive Use Case Based Optimized Hardware Solutions

- Validated best known hardware configurations
- Available for build by OEM of choice
- Fully integrated with use case based benchmarks, deployment tooling, user guides, etc.







# Thank you

[intel.com/redhat](https://intel.com/redhat)  
[redhat.com/intel](https://redhat.com/intel)

 **slack**  
[openshiftcommons.slack.com](https://openshiftcommons.slack.com)