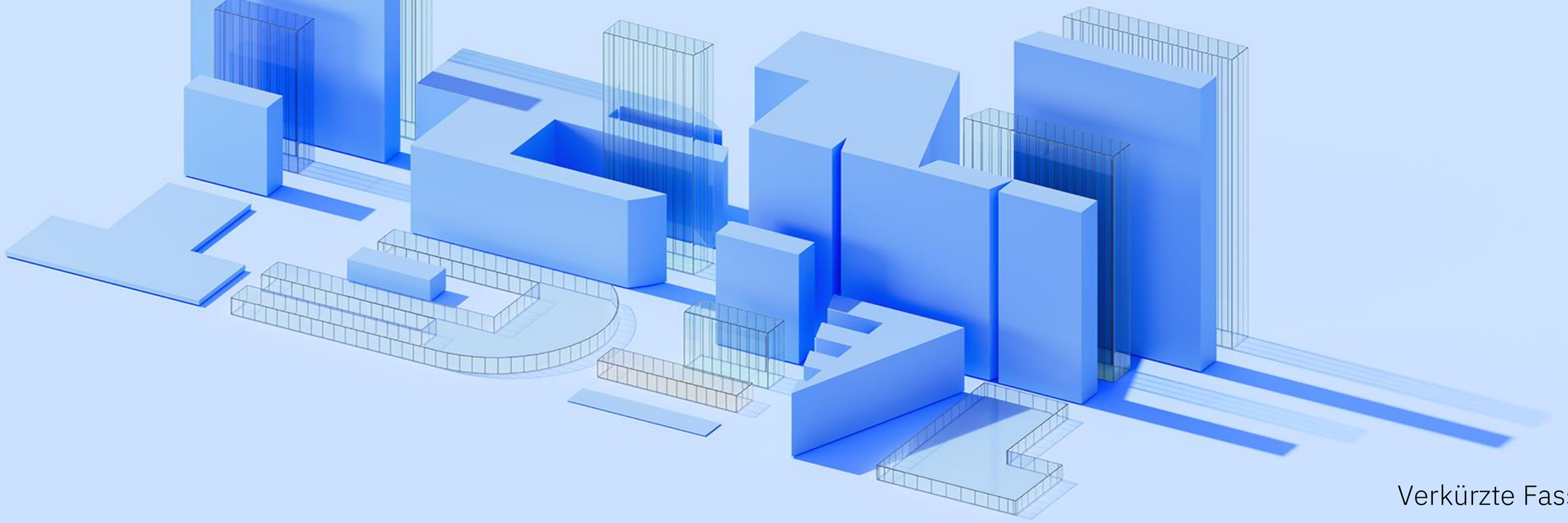
Wie ein deutscher Krankenversicherer AI zur Optimierung von Arbeitsabläufen nutzt

Red Hat OpenShift.AI zur Verbesserung von Workflows



Verkürzte Fassung

Hafid Haddouti
Principal Architect, Tech Lead Red Hat Services EMEA

Daniel Danner
Lead Architect for IBM Fusion

IBM

IBM Fusion HCI

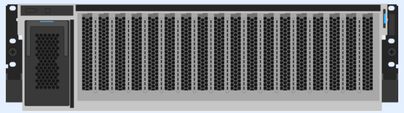
“Red Hat OpenShift in a Box”

- OpenShift installed by IBM
- Bare Metal Performance
- Scalable
- Resilient
- Highly Available

What is Fusion HCI?

An engineered infrastructure platform optimized for OpenShift, VM and AI workloads

GPU nodes



GPU nodes
Up to 8x NVIDIA RTX6000,
H100 or H200 NVL GPUs per node

200 GbE OpenShift network

High speed switches
200 GbE OpenShift network

Management network

Management switches
Used for internal appliance management

Hyper-converged servers with NVMe

Compute/storage nodes
Customizable Openshift worker and control nodes

AI

Accelerate Generative AI Solutions

Inferencing, summarization, fine tuning

Data Lakehouse modernization

OPENSHIFT

Deploy in days vs months

App Modernization at your own pace

IBM Cloud Paks, ISV, or custom apps

Built-in HA/DR, Backup/Recovery

OPENSHIFT

Virtualization

VM Migration

Removing vSphere/vSAN dependency with container native storage

Built-in HA/DR, Backup/Recovery

More flexibility for AI-centric solution designs

Flexible primary rack removes the requirement for six storage nodes

1

2

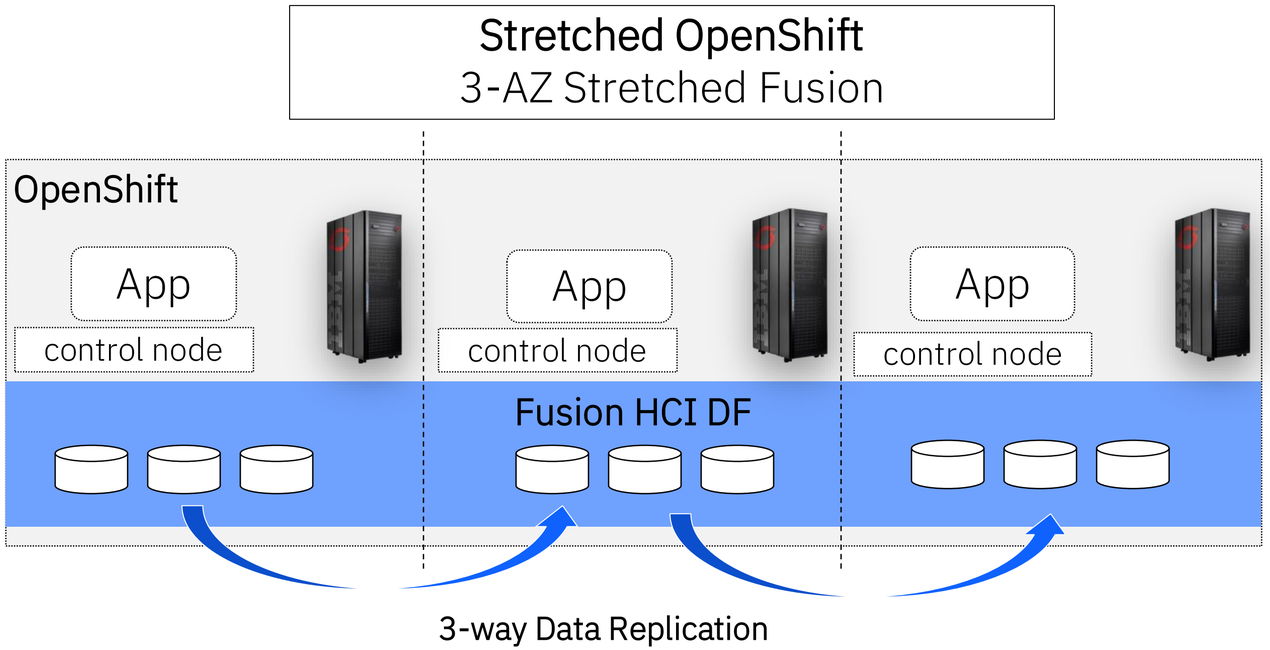
Dell GPU nodes enable AI-solutions on non-Lenovo hardware

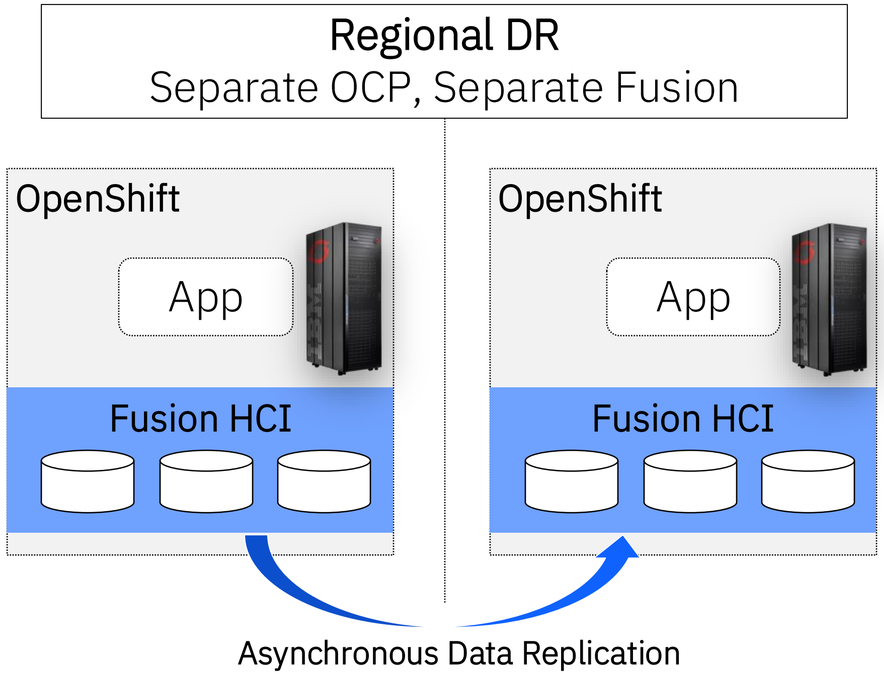
3

Flexible expansion rack enables racks of GPU nodes for inferencing at scale

Disaster Recovery

Data Foundation with internal storage



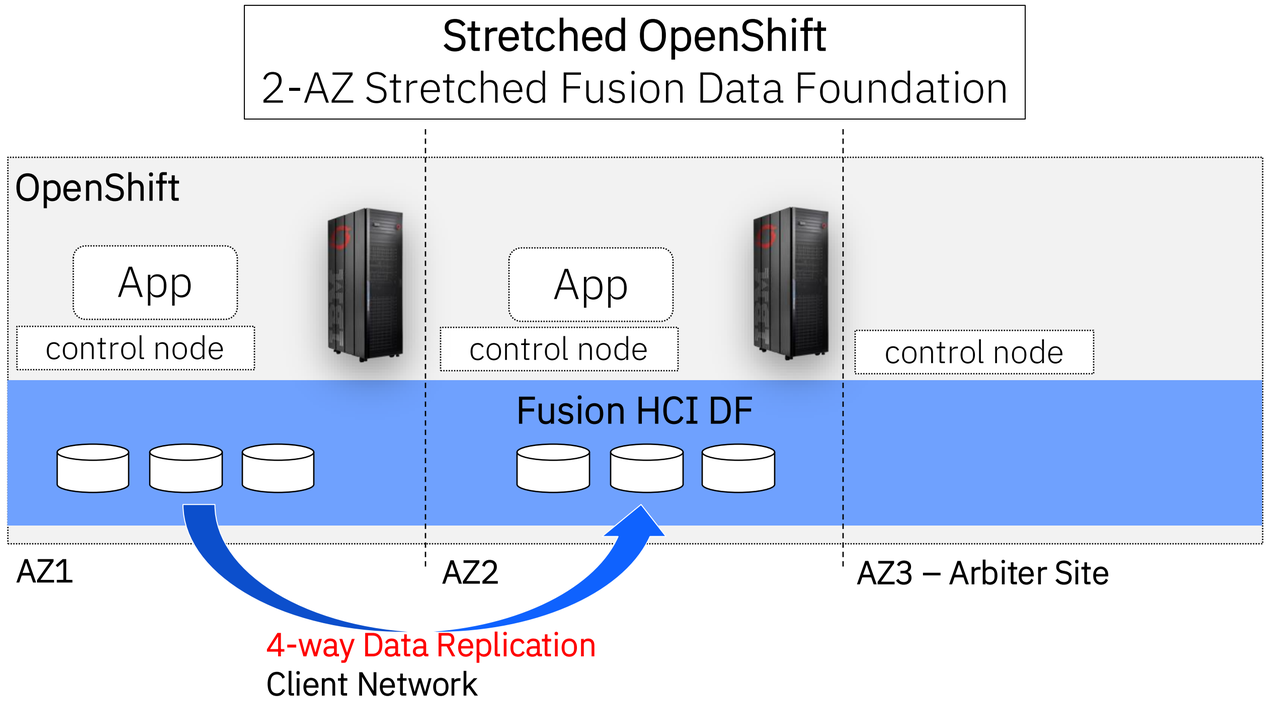


Topology	One OCP cluster, one DF cluster	RHACM orchestration required. 1 Primary replicates to 1 DR.
RTO (downtime)	RTO=0 (continuous)	RTO=variable
RPO (data loss exposure)	RPO=0	RPO=5 minutes
Infra requirements	On-prem; usually <10ms latency between sites	No network latency limits

Disaster Recovery

Data Foundation with internal storage

Fusion 2.11



- 3 AZs in a metropolitan area
- Fusion HCIs connected over client network

- AZ3 contains only a single control node
  customer rack and networking required

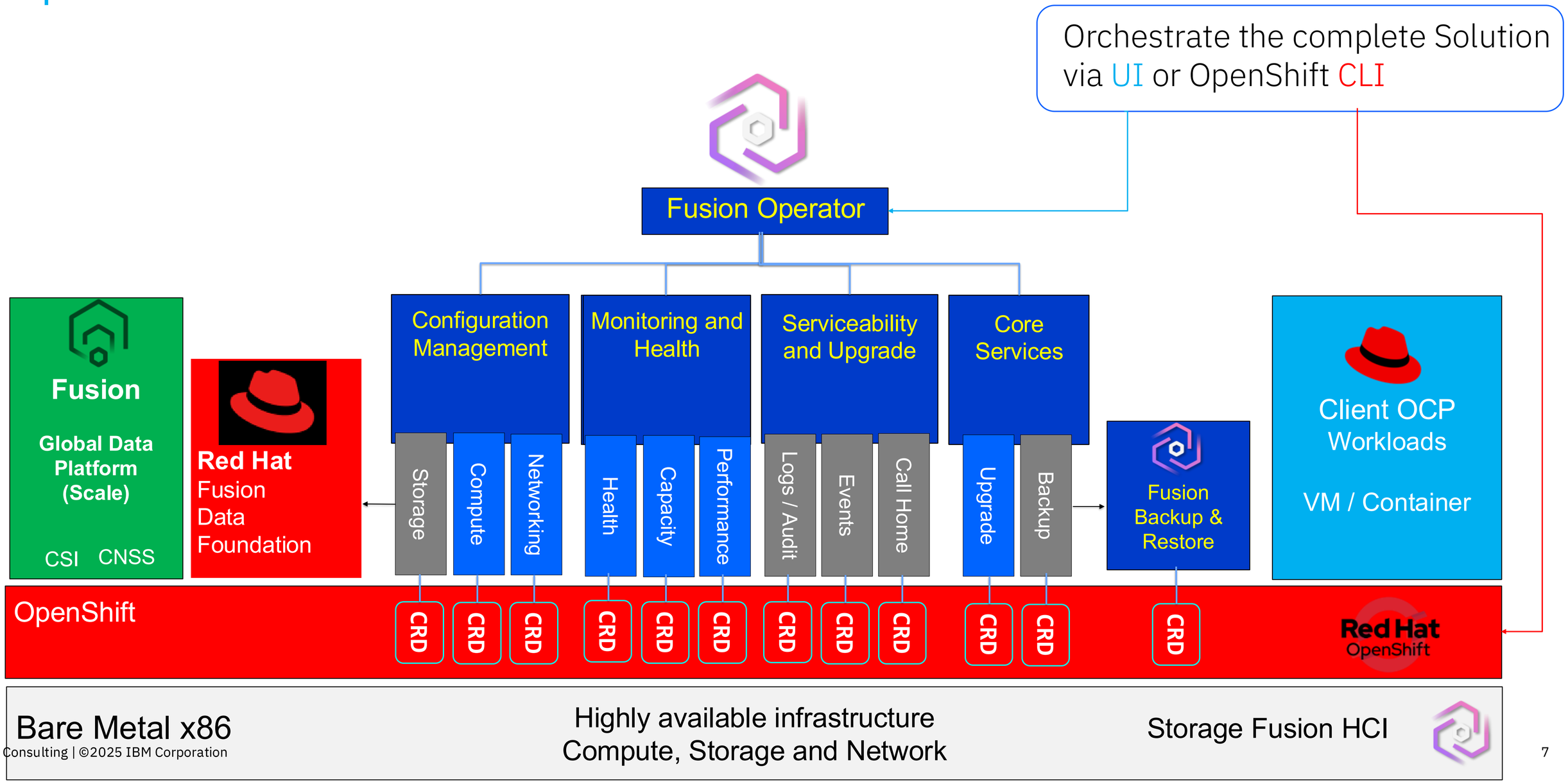
- Data synchronously replicated across AZ1 and AZ2
- Two replicase of the data in AZ1 and AZ2

- Control plane distributed over all three zones
- This is NOT OpenShift 2AZ + Arbiter

Topology	One OCP cluster, one DF cluster
RTO (downtime)	RTO=0 (continuous)
RPO (data loss exposure)	RPO=0
Infra requirements	On-prem; usually <10ms latency between sites

IBM Fusion

Complete Container Native Software Architecture

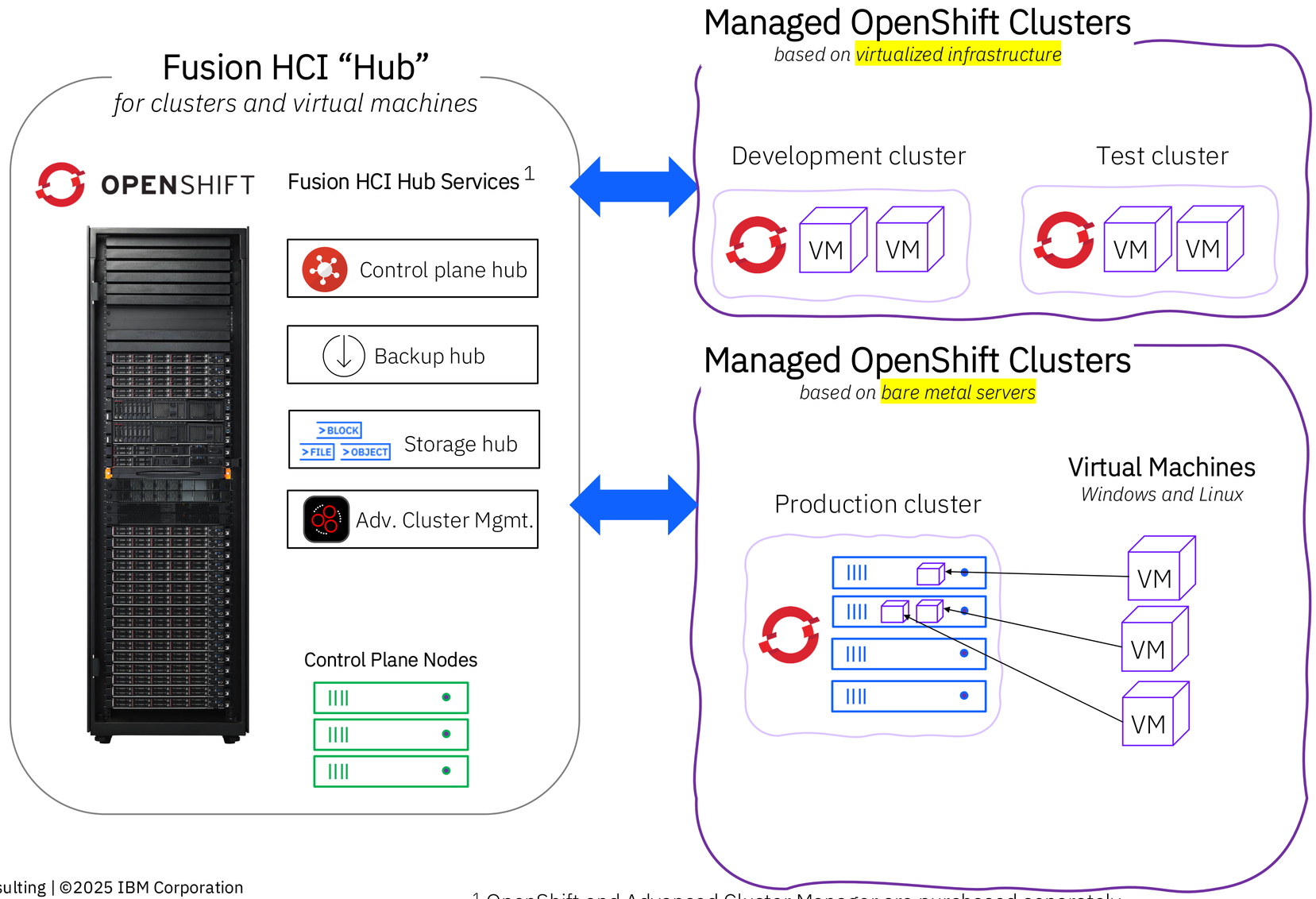


Fusion HCI can host multiple Openshift Clusters

based on virtual machines or bare metal

Control Plane Nodes

Worker Nodes



Key Concepts

- Leverages Hosted Control Planes
- Red Hat Advanced Cluster Manager to deploy hosted clusters
- Fusion Data Foundation provider mode: Clients to consume storage from the hub
- Fusion Backup & Restore hub and spoke architecture
- Strong storage tenant isolation between hosted clusters (Data Foundation)
- Allows to deploy different OpenShift versions within one Fusion HCI Rack.

[1] OpenShift and Advanced Cluster Manager are purchased separately

The IBM Fusion HCI business case

Reduce Complexity & Costs

- Slash infrastructure overhead with a turnkey solution—no DIY headaches.

Value – Faster time to market, reduce costs to integrate, deploy and maintain.

Streamline Operations

- Simplify deployment, updates, and maintenance through automated, centralized management.

Value – Sustain higher levels of availability and increase client satisfaction.

Future-Proof Investments

- Support containers, VMs, and AI workloads as technologies evolve

Value – Easily adapt to emerging technologies and evolving workloads.

Accelerate Time to Value

- Deploy in days, delivering immediate ROI and faster innovation cycles.
- Spend more time on applications … less time on infrastructure

Value – Accelerate Time to Revenue with an integrated, well tested, and flexible platform.

Reduce Risk

- Leverage advanced HA/DR capabilities to safeguard critical workloads and prevent downtime.
- One company for delivery, one number for support

Value – Reduce risk of missed schedules, and performance objectives.

Empower Developers

- Provide on-demand clusters for rapid prototyping, seamless testing, and continuous delivery.

Value – Free up time and effort to focus on higher value services and competitive differentiation.