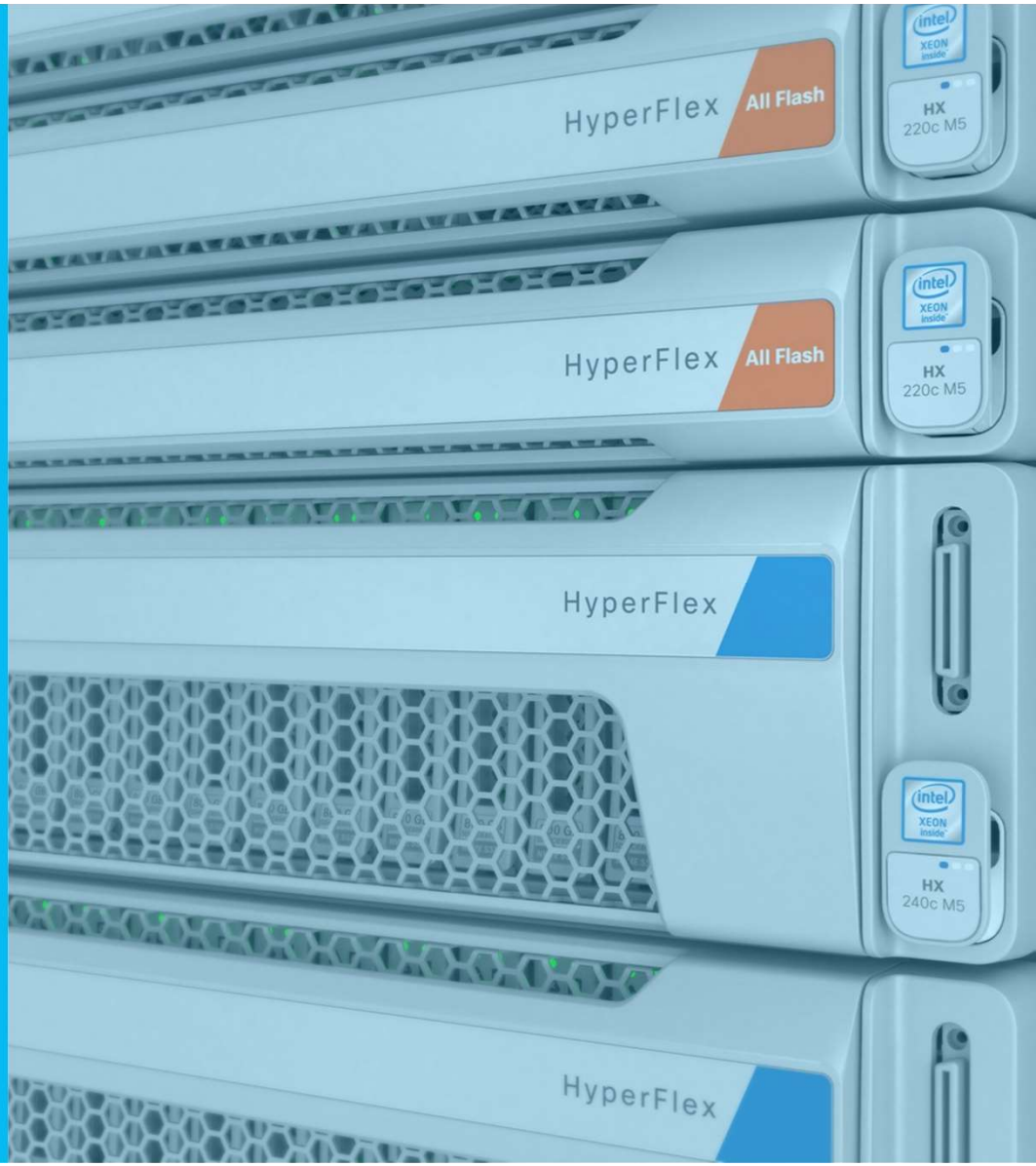
HyperFlex M5 & HyperFlex Data Platform 3.0 Overview and Update

Phil Lowden
Consulting Systems Engineer – Service Provider Data Center
Cisco Global Service Provider

This version: Feb. 27th 2018

Agenda

- Introduction
- UCS and HyperFlex Review
- High Availability
- Data Protection
- Security
- HyperFlex Edge
- Use Cases
- Summary



Cisco's Hyper-convergence Journey So Far...

Pumping Innovations with Rapid Pace

All Flash | SED | Enterprise Workloads | Profiler, Sizer | Edge | Connect UI | Native Replication/DR | Scale | RBAC | Localization | M5 Skylake

Fastest Growing HCI

2500+

New Customers

Ready for Biz. Critical

#1

in Performance and consistency

Customer Success Tools

15

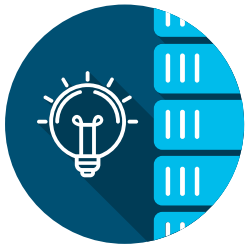
CVDs, Ref. Architectures

25

White papers, Soln. Guides

VDI, VSI, SQL, Oracle, SAP, Exchange, Splunk, Veeam, …

HYPERFLEX 3.0 NEW



Any App.

Microsoft Complete

NEW Enterprise Kubernetes Containers

Mission Critical Applications

Application Performance Monitoring

Any Cloud.

NEW Cisco-Google Hybrid Cloud Solution

Multi-Hypervisor, Multi-Cloud

IaaS and Private Cloud

Intelligent Workload Placement

Any Scale.

64 Node & Capacity Scaling

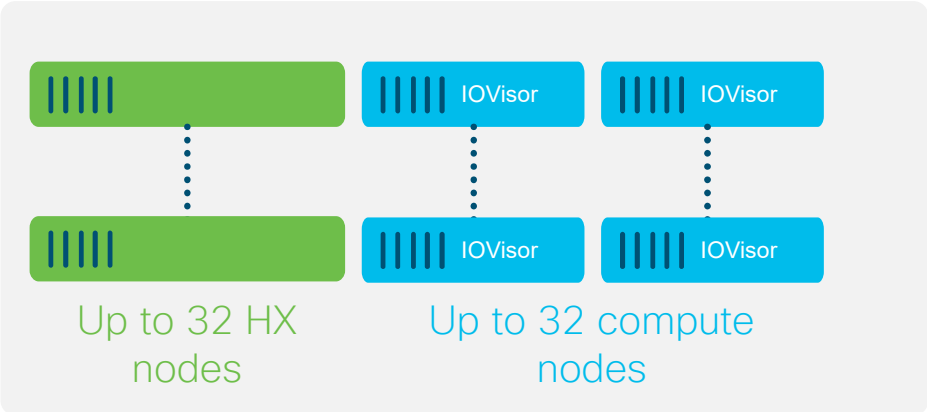
NEW Stretched Clusters

Larger Form Factors

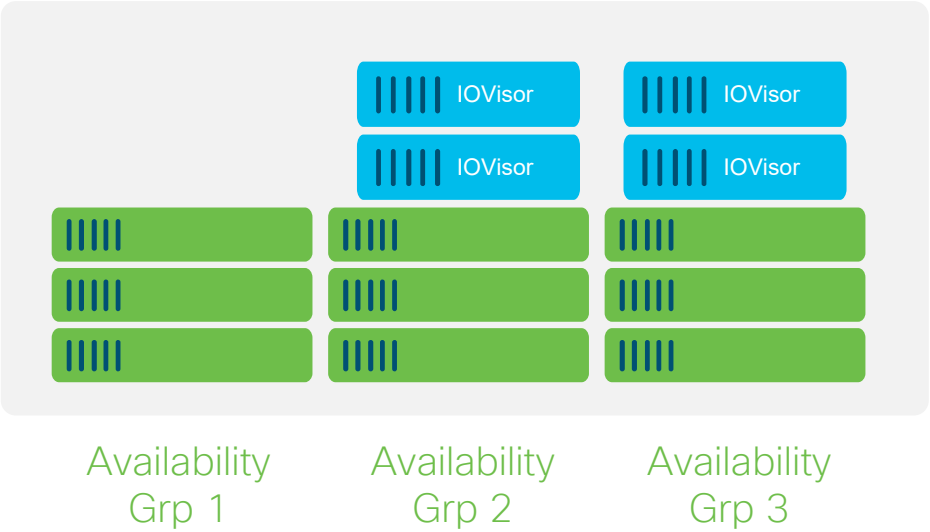
Cisco Intersight Managed

64 Node Scale with Resiliency
Cloud Scale Data Platform

64 Node & Capacity Scalability

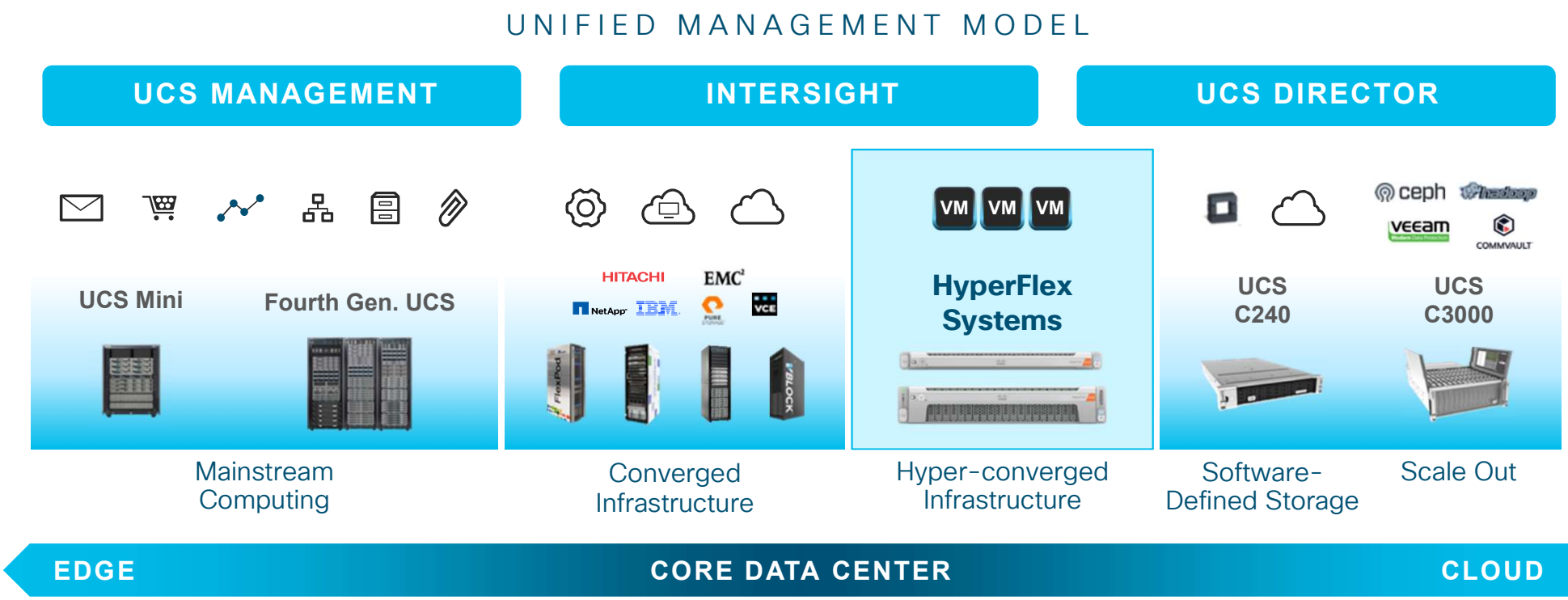


Automated Availability Zones



Based on Cisco UCS

One Architecture for Operational Simplicity



Form Factors

HyperFlex 220 M4/M5
Small Form Factor (SFF)

- 6 – 8 HDD/SSD
- 1.2/1.8 TB HDD, 960GB SSD, 3.84TB SSD

HyperFlex 240 M4/M5
Small Form Factor (SFF)

- 6 – 23 HDD/SSD
- 1.2/1.8 TB HDD, 960GB SSD, 3.84TB SSD

HyperFlex 240 M5 Large Form Factor (LFF)

- 6 – 12 HDD's
- 6 & 8 TB 7.2K HDD capacities

HyperFlex Data Platform

The Cisco HX Data Platform

HX LOG STRUCTURED FILE SYSTEM
DESIGNED SPECIFICALLY FOR HYPERCONVERGENCE

DISTRIBUTED

Object-Based File System Architected for Scale-Out, Distributed Storage

ADVANCED DATA SERVICES

Built Into File System Architecture

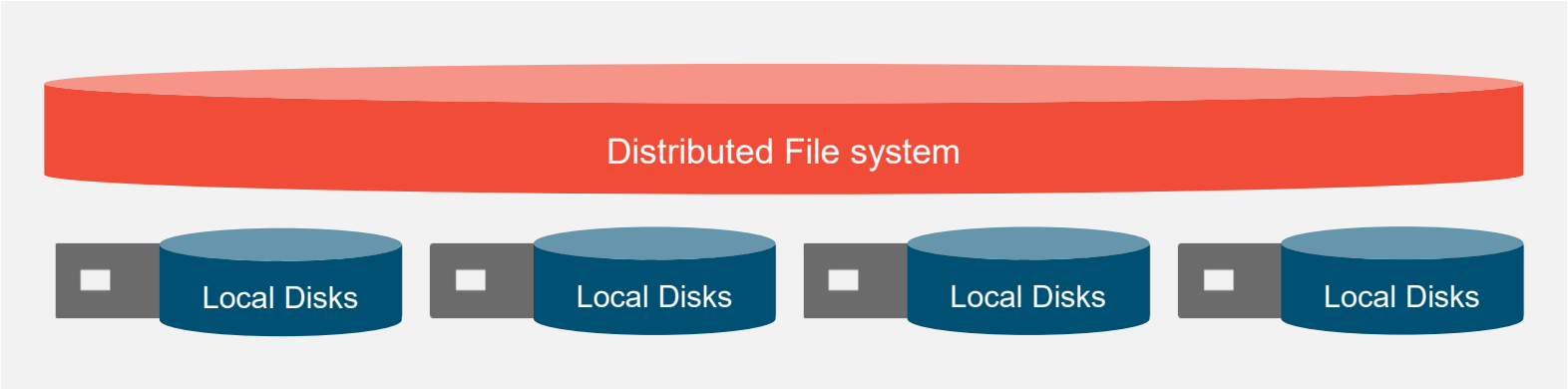
FUTURE READY

Designed for Containers and Next-generation Applications

Building on the Right Foundation

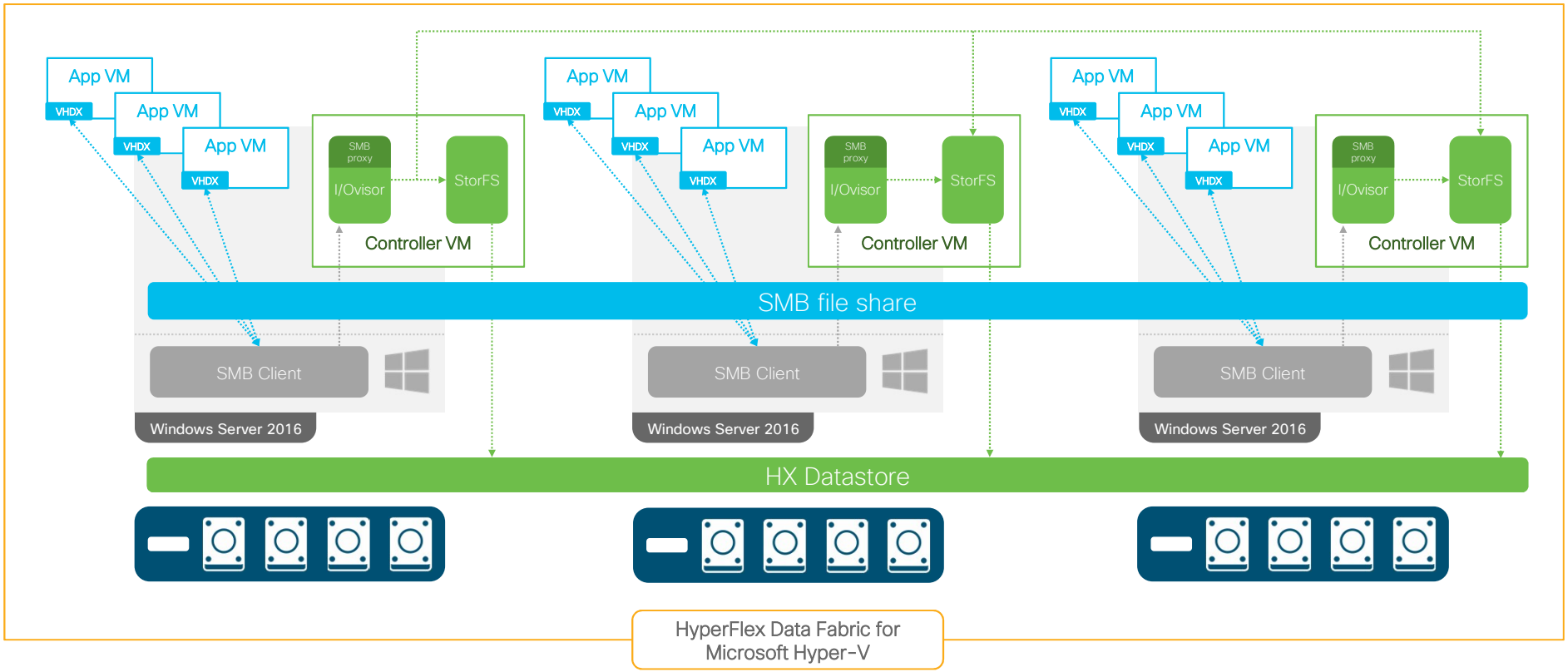
Cisco HX Data Platform

Unique Architecture

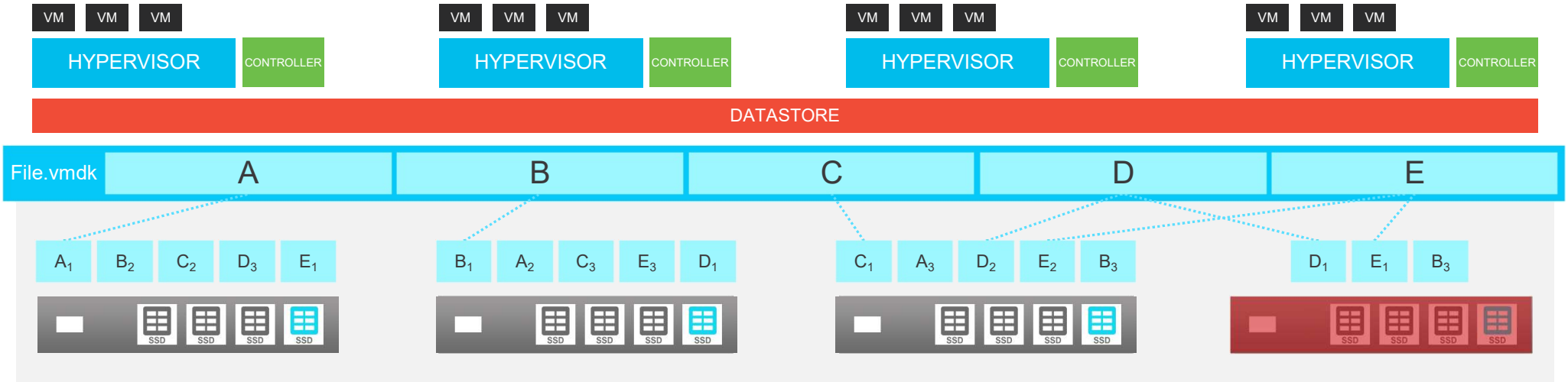


Built From the Ground Up for Hyperconvergence	Distributed Log-Structured File System Designed for Scale-out, Distributed Storage	Advanced Data Services (Snapshots, Clones, Replication) and Data Optimization (Inline Dedupe, Compression) Without Trade-offs	Better Flash Endurance and Disk Performance	Computing, Storage, Networking, and Hypervisor Integration	No Reliance on Legacy Filesystems or Technology

Controller VM Architecture with Hyper-V

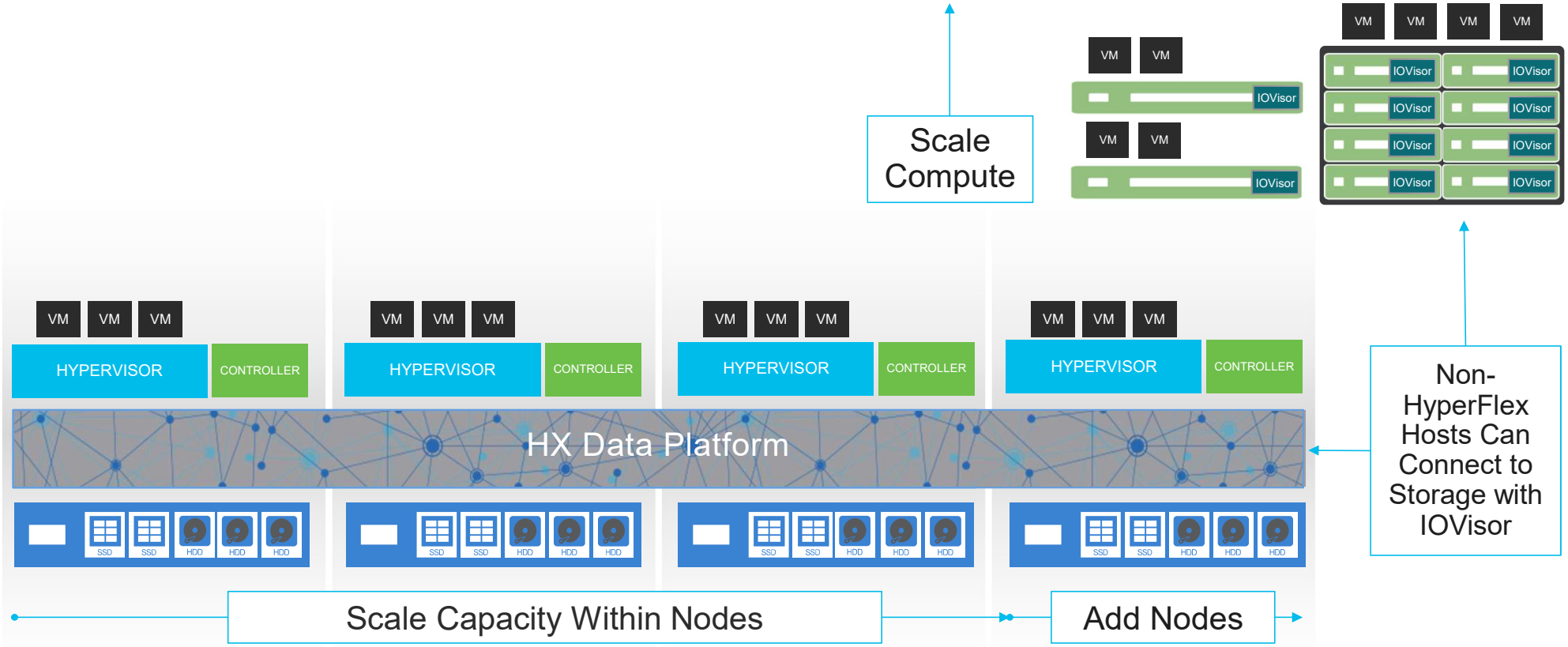


Data Distribution and Non-Disruptive Operations



- Stripes blocks of virtual disk files across servers using a hashing algorithm
- Replicate one or two additional copies to other servers
- Handle entire server or disk failures
- Restore back to original number of copies
- Rebalance VMs and data post replacement
- Rolling “one-click” software upgrades

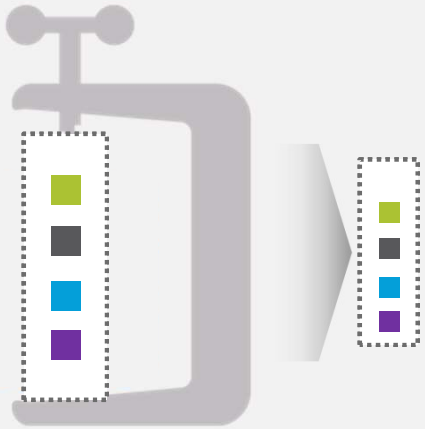
Independent Scaling of Compute and Capacity



Data Services

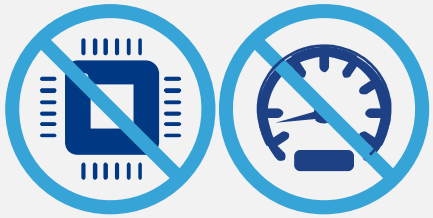
Continuous Data Optimization

Log-Structured File System Yields More Efficient Data Optimization



Inline Compression

Inline Deduplication

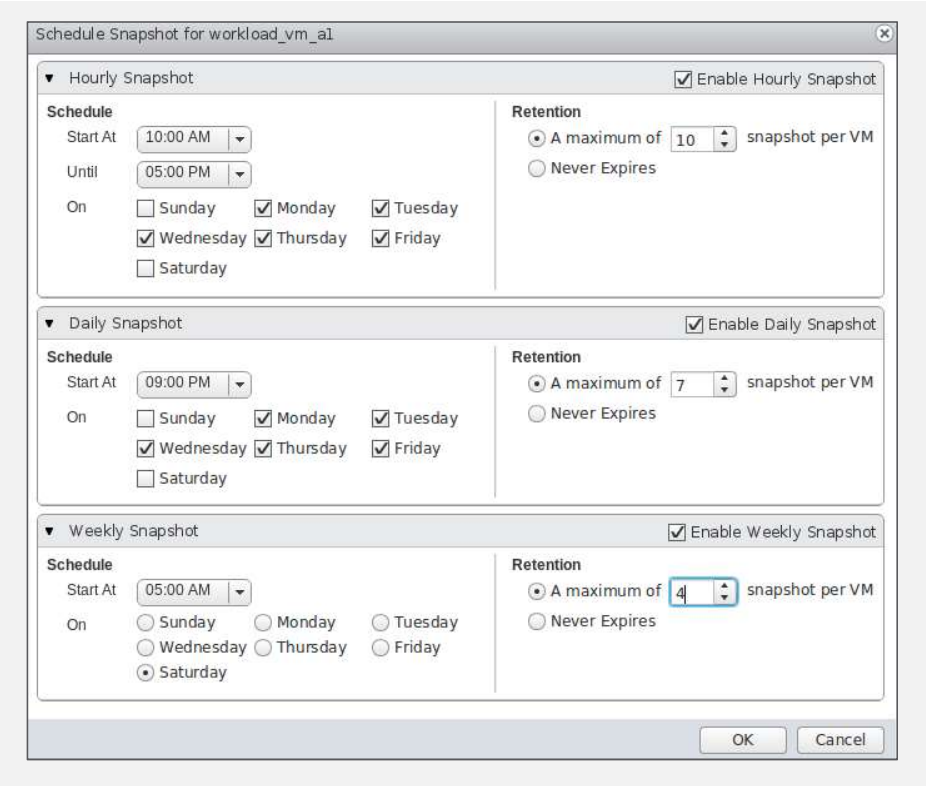


No Special Hardware
No Performance Impact
No Config. lock-in
No Additional License

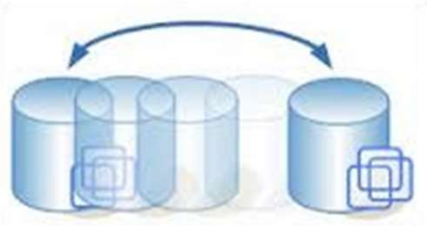
30–50% space savings	20–50% space savings	Lower Cost

Fast and Flexible Native Snapshots

- Pointer-based snapshots
  - Space-efficient with no performance penalty vs. VMware Redo Log Snaps
  - Fast creations and deletions
- Fine-grained or coarse-grained
  - VM-level or VM folder-level
- VAAI-integrated
  - Quiesced and crash-consistent
- Use vCenter Snapshot Manager
- Policy-based schedules and retention

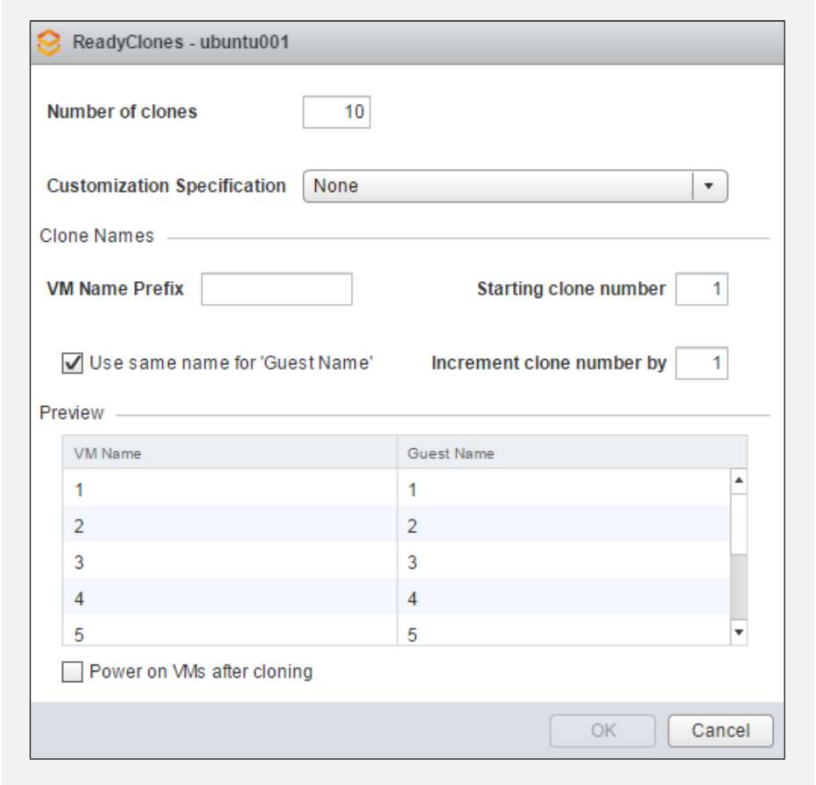


Native VM Clones for Rapid Provisioning



- Pointer-Based Writeable Snapshots (Instantaneous Clones)
- VAAI integrated
- VM-level granularity

- Batch creation GUI
  - Apply unique names
  - Use customization spec to apply IP
  - Powerful tool to rapidly setup a large set of VMs using just VC (without scripting or View composer); Up to 256 clones in parallel per job
  - Golden/Base VM can be a template, powered on or powered off



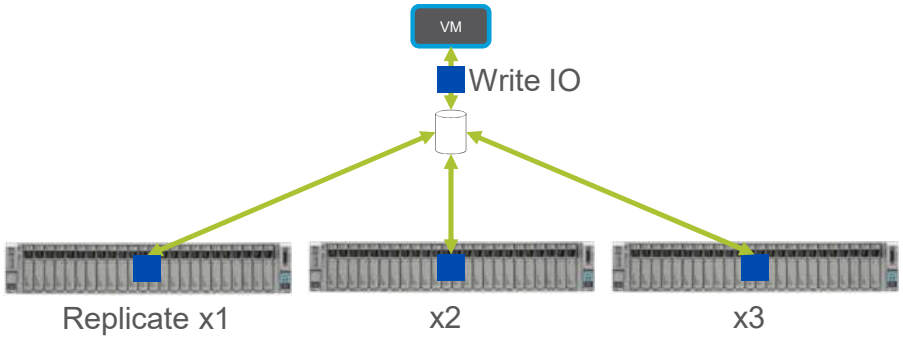
Built-in, Multi-layer High Availability

Level 1: HXDP Architectural HA

Built-in Data Protection and High Availability

Data Protected by Replication of Data Across the Cluster Nodes

Replication Factor 3 (RF3)
Default and Recommended is Replication Factor = 3
Every block is written to 3 different nodes in the cluster
Higher availability to survive multi-point failures; Higher device protection
Reduces raw disk capacity to 33%



Note: RF3 is strongly recommended for high availability

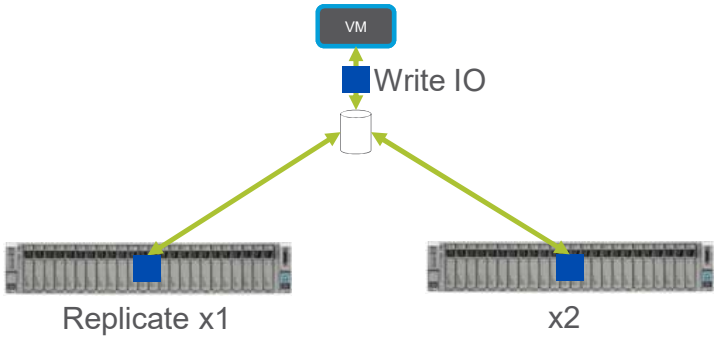
Replication Factor	3 or 4 Node Cluster	5+ Node Cluster
3	Simultaneous Failures Supported: 1 node / 2 drives*	Simultaneous Failures Supported: 2 nodes / 2 drives*

*drives across different nodes

Data Protection and High Availability

Data Protected by Replication of Data Across the Cluster Nodes

Replication Factor 2 (RF2)
Default and Recommended is Replication Factor = 3
Every block is written to 2 different nodes in the cluster
Lower availability to survive failures; Lower device protection
Reduces raw disk capacity to 50%



Disclaimer: While RF2 is available, **RF3 is strongly recommended** for high availability

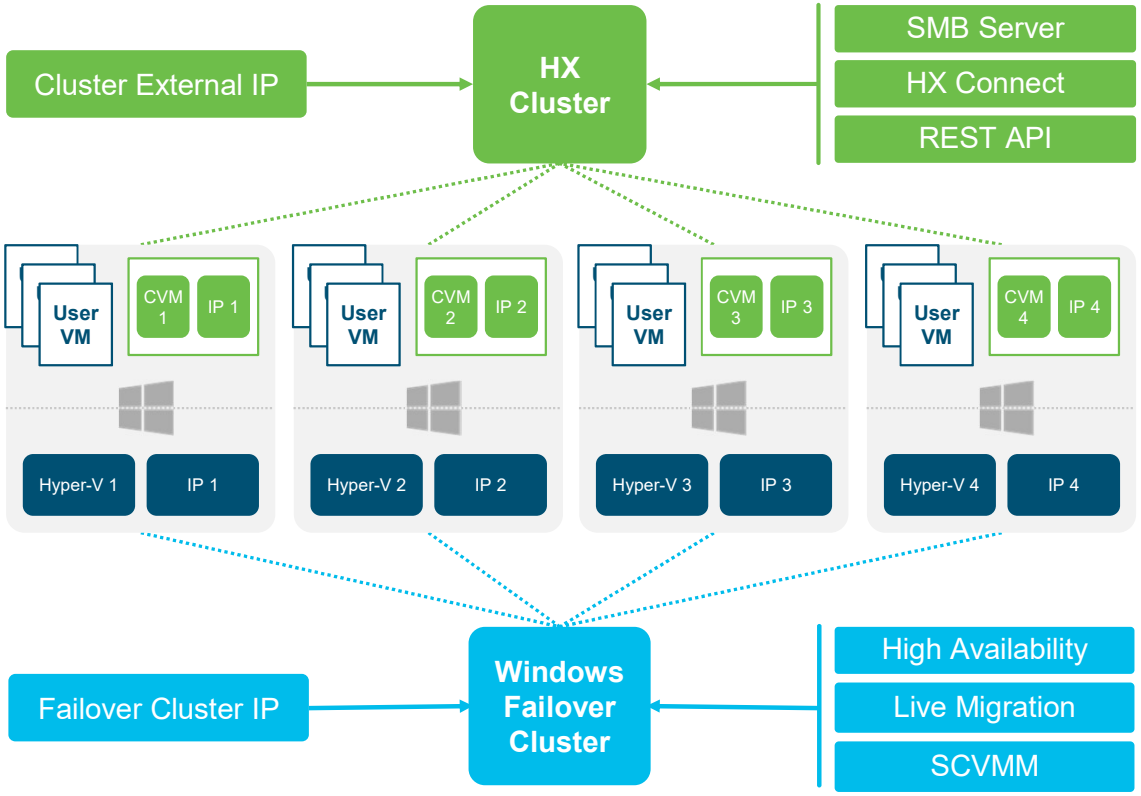
Replication Factor	3 or 4 Node Cluster	5+ Node Cluster
2	Simultaneous Failures Supported: 1 node / 1 drive	Simultaneous Failures Supported: 1 node / 1 drive

Cluster Self-Healing

Self-healing	
Disk Failure	In 1 min
Node Failure	2 hours

- Node failure:
  - Index replay optimizations decrease reboot time
  - VMware HA moves VMs to active nodes.
  - Delay timeout can be modified (SE/Support only)
  - Self-healing delay to account for transient conditions such as maintenance mode
  - Node rebalance command available for on-demand rebalancing on failure
- Disk failure:
  - If a disk fails after node failure then the larger timeout value is used
  - Rebalance is more granular and latency spikes with disk removal eliminated
- SSD failure:
  - If caching SSD fails, VMs have no availability hit. Caching SSD replacement will rebalance cluster cache.
    - Slight impact to latency as aggregate read cache diminished until SSD replaced and rebalanced.
  - Housekeeping SSD will impact SCVM but will not cause VMs to failover.

Windows Failover Clustering with HyperFlex

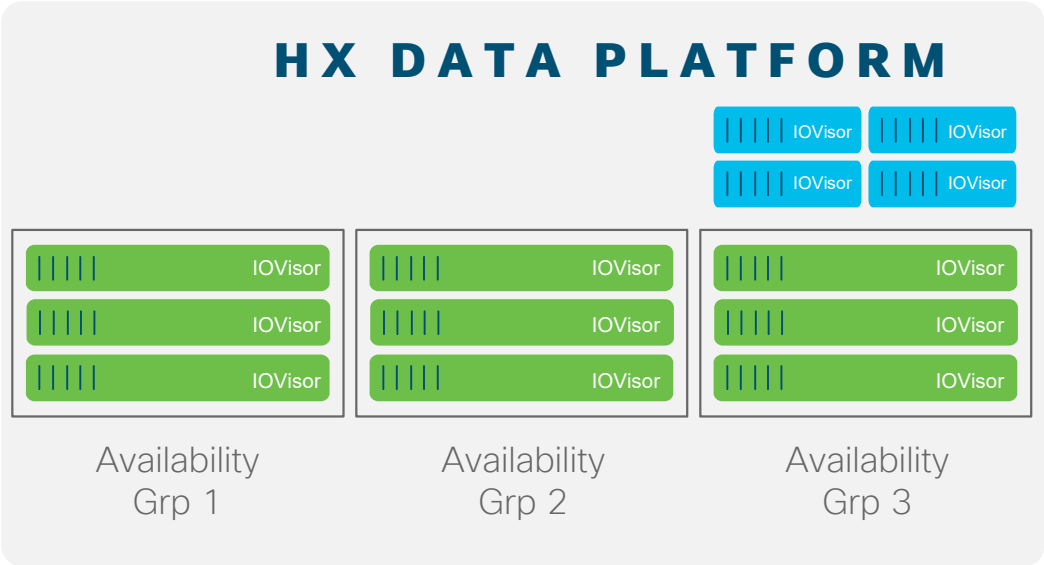


1. VM Load Balancing or Dynamic Optimization with SCVMM
2. VM-level HA, Advanced Resiliency
3. Virtual Machine Start Order
4. Host resource protection (from noisy VMs)
5. Application-level HA, Resiliency
6. Cloud witness (Azure-based)

Level 2: Logical Availability Zones

Logical Availability Zones (LAZ)

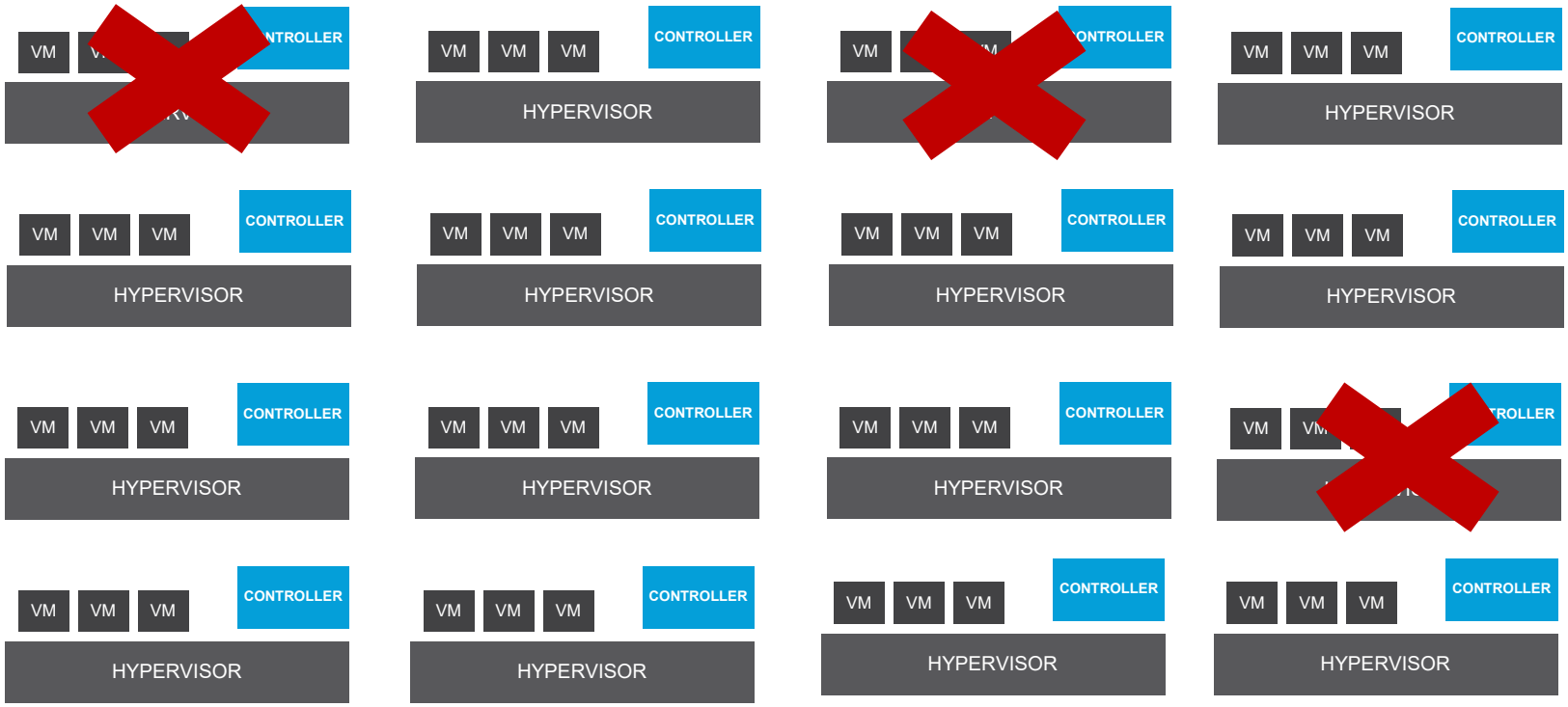
- Cluster Scale With High Availability
- Increased resiliency without added manageability overhead
- How does it work?
  - HX nodes grouped into logical “availability groups” (N/A for compute nodes)
  - HXDP never places 2 copies of the data in the same availability group
  - Clusters with LAZ can survive > 2 simultaneous node failures without data loss or loss of availability
  - Tolerate more independent failures



LAZ Failure Scenario

Cluster State: **Offline**

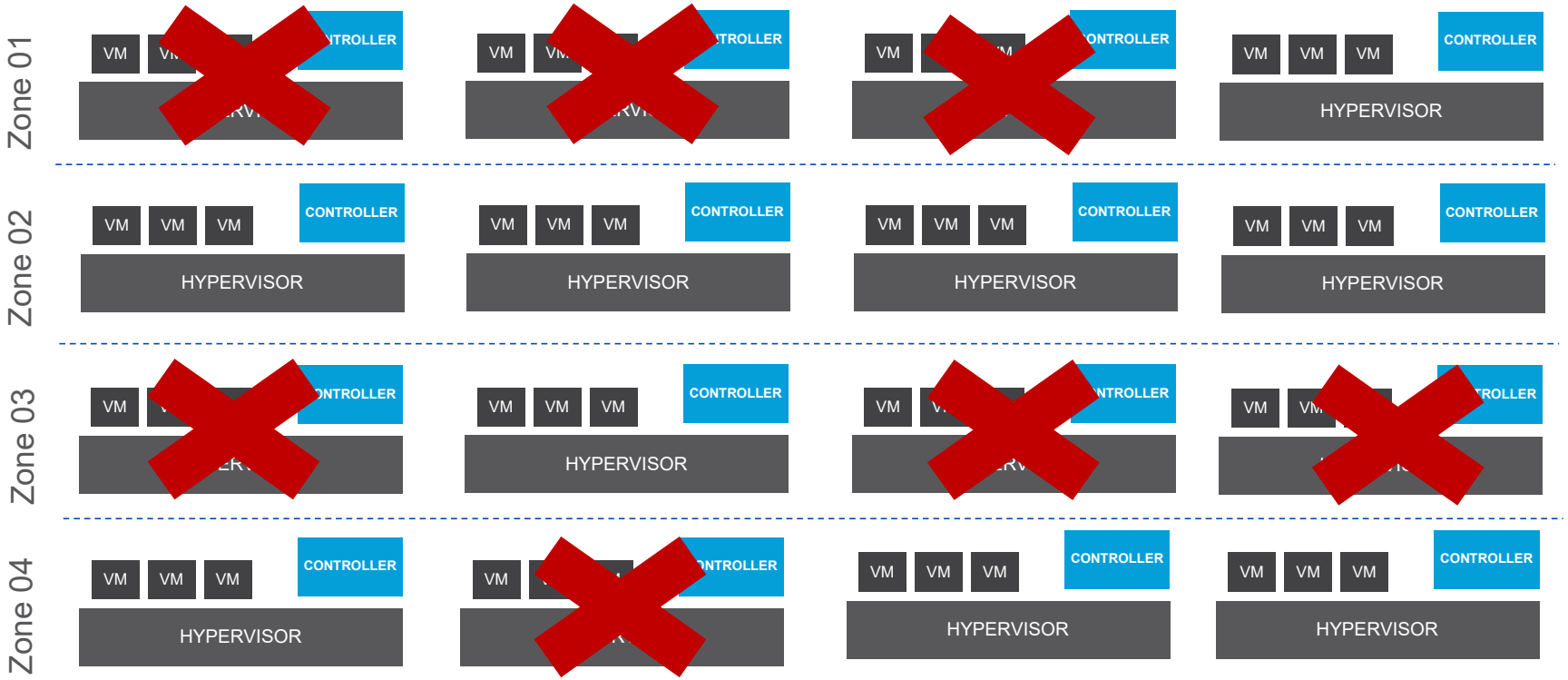
LAZ: Off



LAZ Failure Scenario

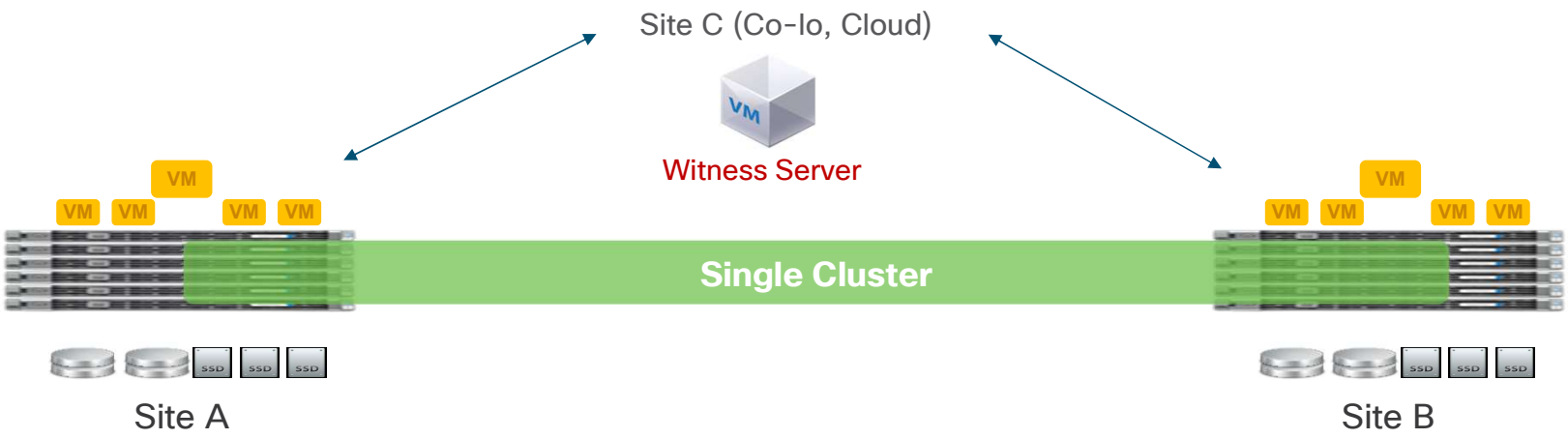
Cluster State: **Offline**

LAZ: **On**



Level 3: Stretched Cluster

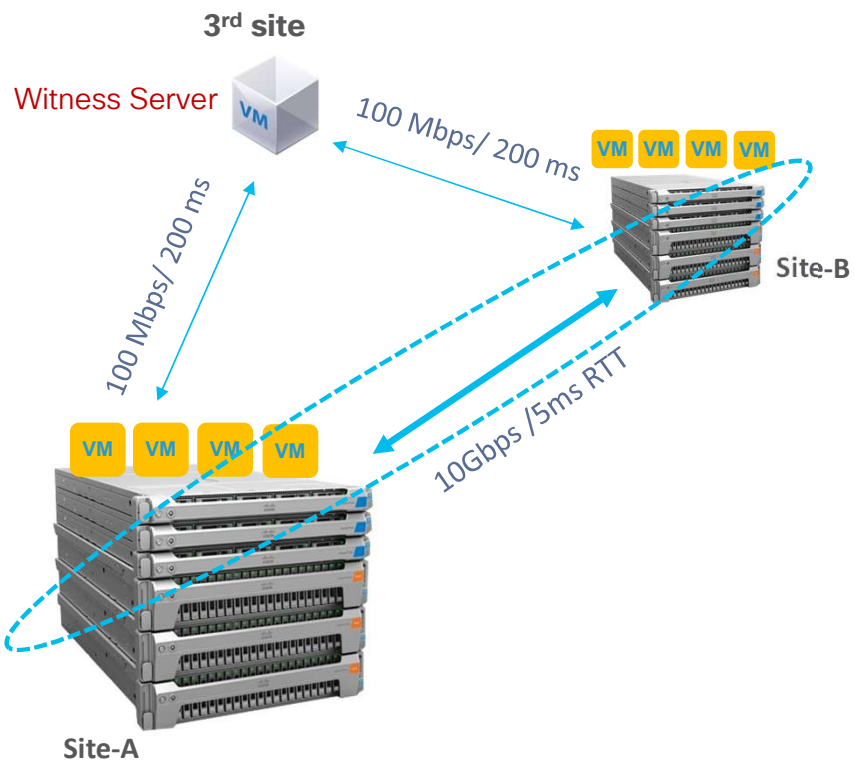
What is a Stretched Cluster?



1. A stretched cluster is a **single cluster** with nodes geographically distributed
2. Storage is mirrored across each sites
3. Sites need to be connected over low latency network
4. Geo-failover (VM) is like failover in a single cluster
4. "Split-Brain": Condition when nodes on either sites cannot see each other
   - Network failure
   - Site failure
5. "Witness" : An entity hosted on a 3rd site responsible for deciding the which site survive after a split-brain

HyperFlex Stretched Cluster

ZERO RPO ! NEAR ZERO RTO!



Configuration Support

- ✓ Single Stretched Cluster across 2 sites
- ✓ Symmetric Configuration
- ✓ Site to host a "Witness Server" (small VM)
- ✓ 8 +8 nodes on each site

Management

- ✓ Cross site Cluster creation
- ✓ Non disruptive online rolling upgrade
- ✓ Site awareness in HX Connect
- ✓ Site specific Alarm and Events on a single Dashboard

IO Path

- ✓ Active-Active sites – VMs Active on each site
- ✓ VM Read IOs served locally
- ✓ VM Write IOs Sync-Writes across sites
- ✓ 2x copies on each site

HA Operations

- ✓ Recover from a Site failure
- ✓ Recover from a Local failure
- ✓ Failover of VM
- ✓ vMotion of VM
- ✓ Split Brain handling

HyperFlex Stretched Cluster

Network & Infrastructure: Supported Configurations & Prerequisites

- ✓ 10Gbps (dedicated), 5ms RTT latency between the 2 active sites
- ✓ 100 Mbps, 200ms RTT latency between the active sites & witness site
- ✓ FI based configuration
- ✓ Existing FIs are supported

vCenter
- Single vCenter for both sites
- Need credential during cluster creation
- Can be on either site or 3rd site
- Can be a VM
- Configure independently for HA as required

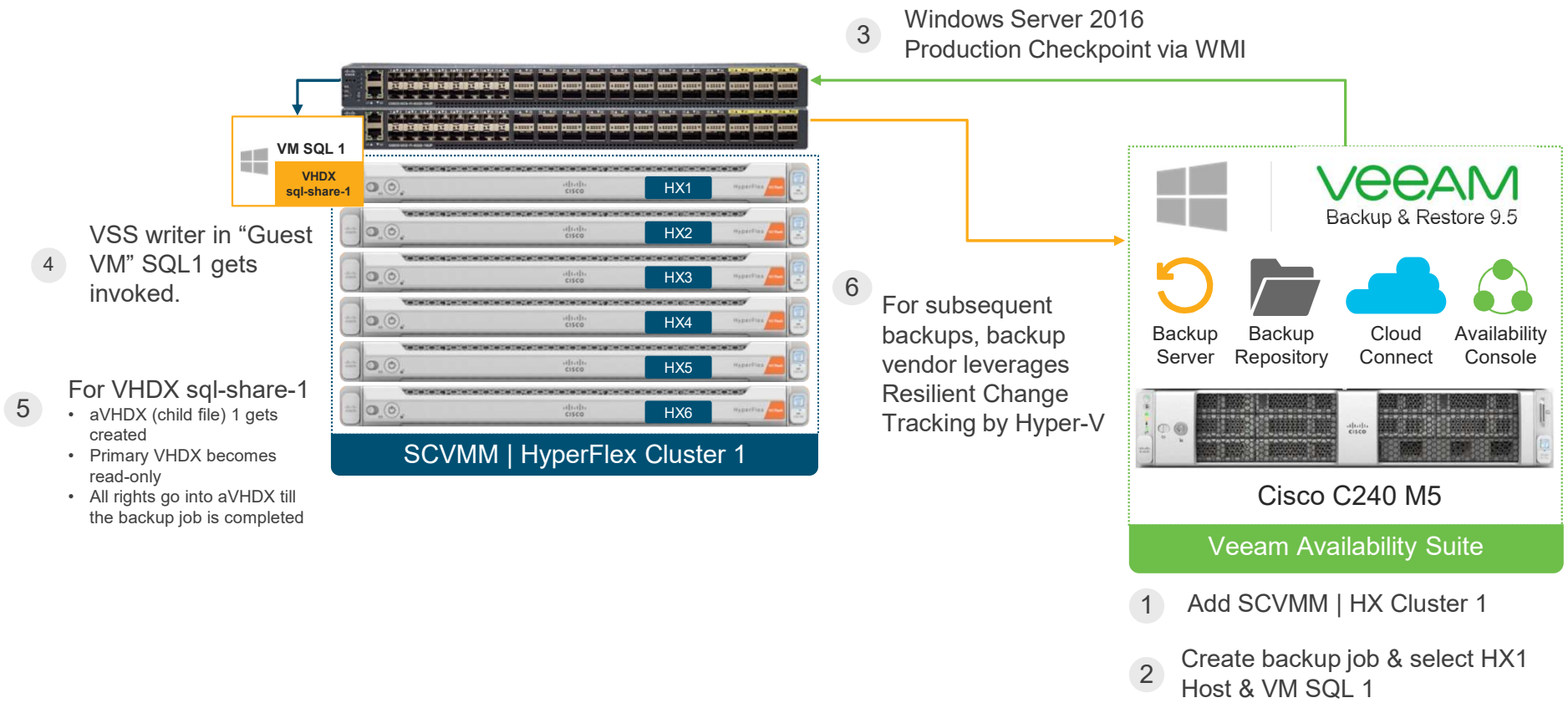
UCSM:
- Need UCSM credentials for both sites
- Two separate FI Domains (not automatically synchronized)

VLAN
- IP addresses for nodes on both sites
- Stretched VLANs across both sites

Witness
- Access to 3rd Site to host witness VM
- IP address and connectivity for the witness VM

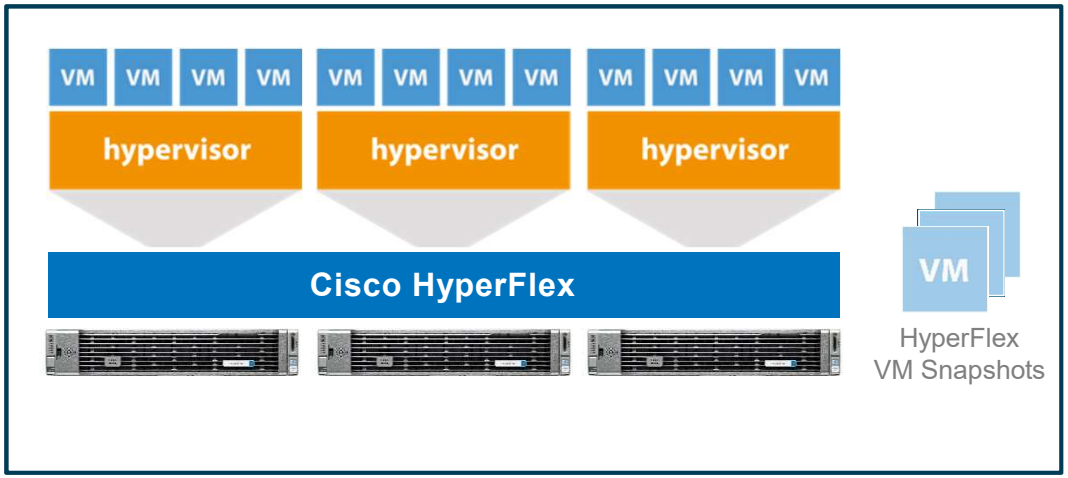
Data Protection

Data Protection with Hyper-V



Veeam’s HyperFlex Snapshot integration

Backup from Cisco HyperFlex Snapshot

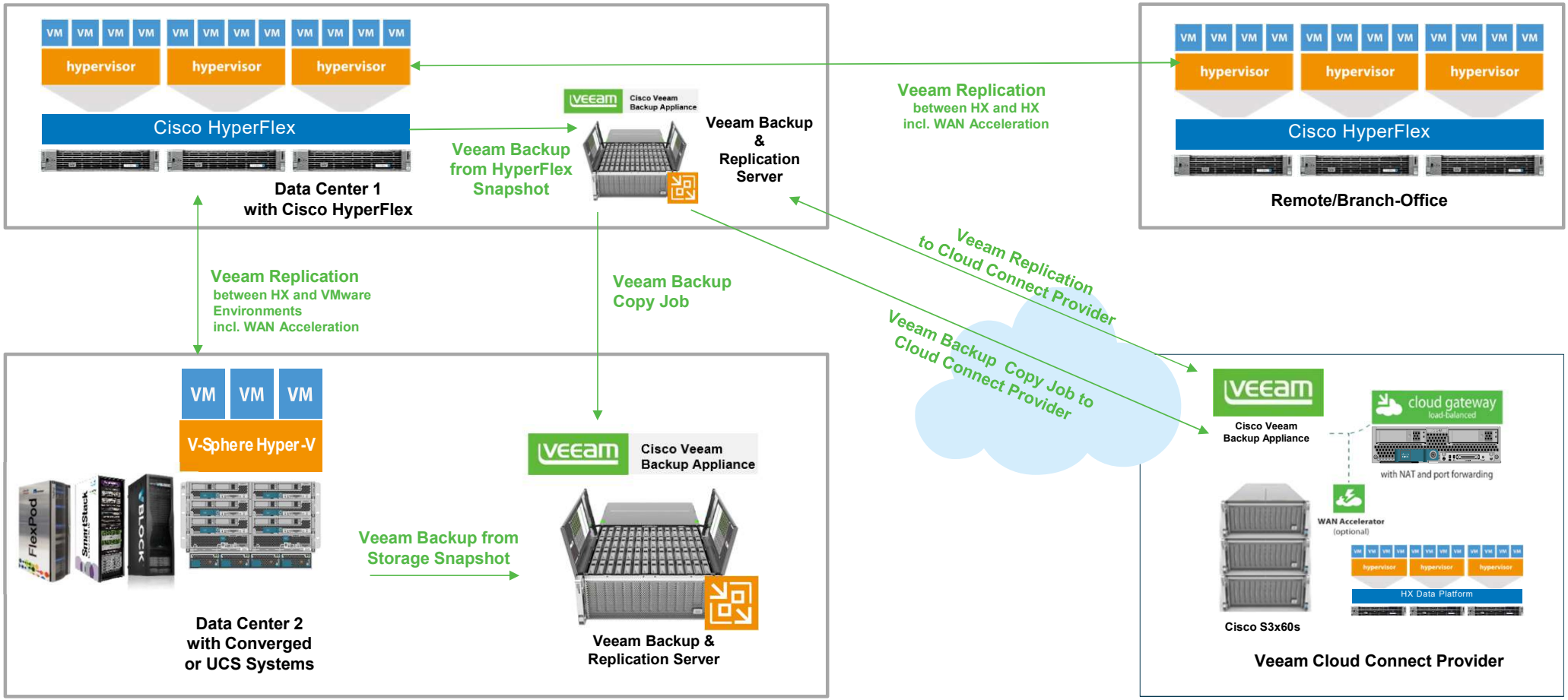


Backup/Restore
NFS Backend

Veeam leverages **HyperFlex Snapshot** based Backups to achieve:

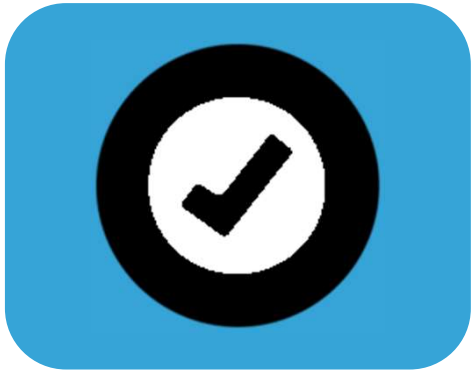
- Agentless Application Aware Backup (AD/Exchange/SharePoint/SQL/Oracle/…)
- Minimizes the typical VMware backup API performance impact on production VMs during backup
- Improves RTO and RPO times with technologies like Instant VM Recovery out of Veeam backups
- More than doubles the possible backup speed compared with standard VMware API backups

Disaster Recovery Solutions with Veeam and Cisco



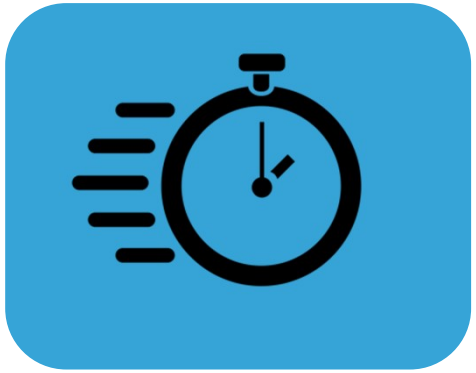
A Replication Solution for Mission Critical workloads

Reliable



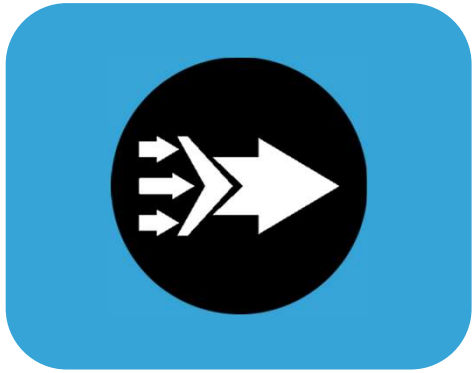
- TCP based reliable transmission
- Internal checksums to protect against network corruptions
- Robust fault handling- retries and timeouts
- Crash consistent snapshots using VMTools

Performant



- Replication streams ‘scale-out’ to all nodes to maximize network utilization and lower CPU overheads
- Coalesced large IO reads and transfers

Optimized



- Data on the wire compressed, incremental snapshots
- Intelligent pattern detection
- Different Primary and Target side configurations
- HX 3.0 does not support dedupe on-wire

HX Replication: Recommended Configurations

Options for Cost Effective DR

- Dissimilar cluster sizes on both sites – Primary site can have more number of nodes than the DR site
- Mix of HX220 and HX240 Clusters
- Need to plan for the storage capacity based on the number of VMs to be protected

Optimal DR setup

- Active-Active DR setup – Clusters on both sites host active VMs
- Configure the clusters for Bi-directional replication

HX Replication: Best Practices

Cluster Configuration

- Replicate one cluster to one cluster only
- You can mix clusters with HX240 & HX220 but be aware of capacity limits
- For optimal usage of resources consider Active-Active Cluster

Point-in-Time Copies

- Replica is current as of last update
- Only one PIT
- May affect recovery if corruption has happened farther in the past than the last incremental
- Use VMTools to quiesce VMs during replication snapshots for better crash consistency

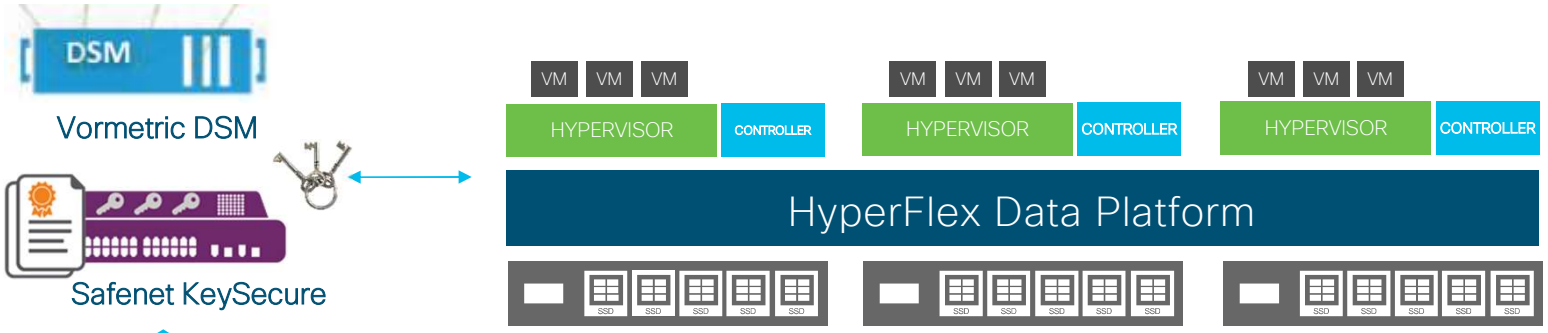
HyperFlex Security

Security

- Data-at-rest encryption, with self-encrypting drives
- Enterprise key-management support
- Vulnerability assessments and hardening
- Role-based access control (RBAC)
- Certifications: FIPS, Common Criteria
- Compliance for data privacy: HIPAA, PCI-DSS, NIST, FedRAMP, FISMA, Europe GFPR, UK Data Protection, etc.

HyperFlex Data-at-rest encryption

- ✓ Familiar easy HX Installer
- ✓ Simplicity of HX Connect – easy to manage
- ✓ Power of UCS – policy enforcement



Enterprise Key-Management Support

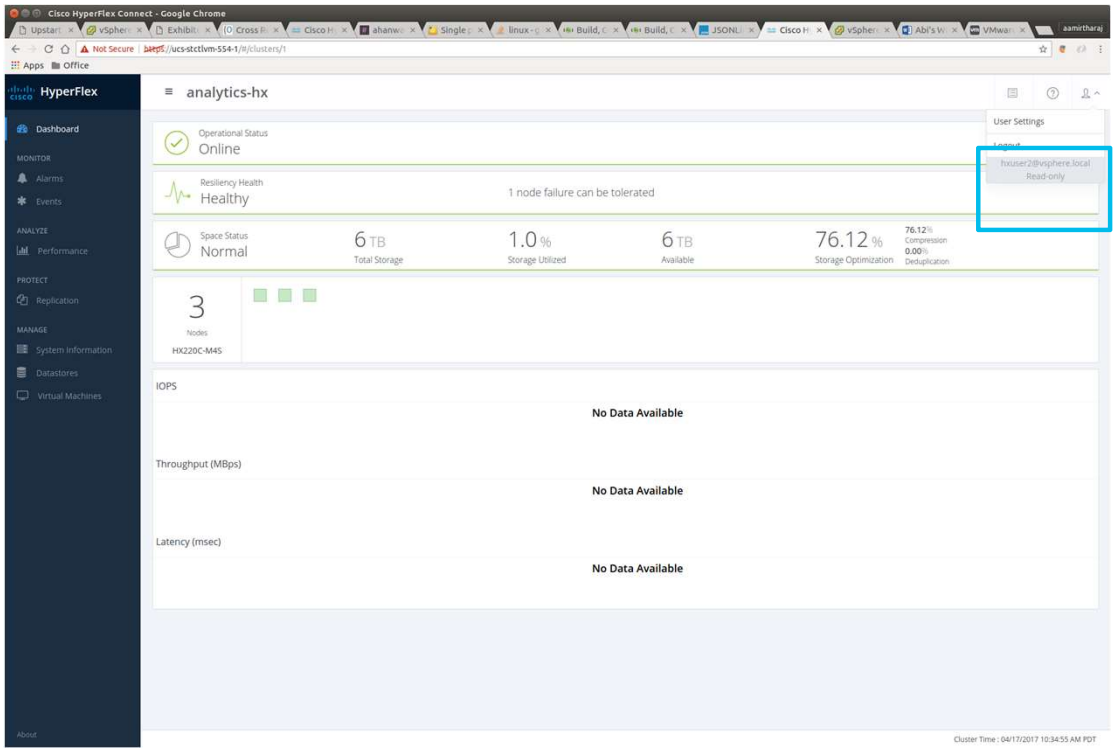
- Secure Key Lifecycle Management
- Highly Available
- FIPS Certified

Self-Encrypting Drives (SEDs)

Security without performance penalty

Secure Management – RBAC/Audit

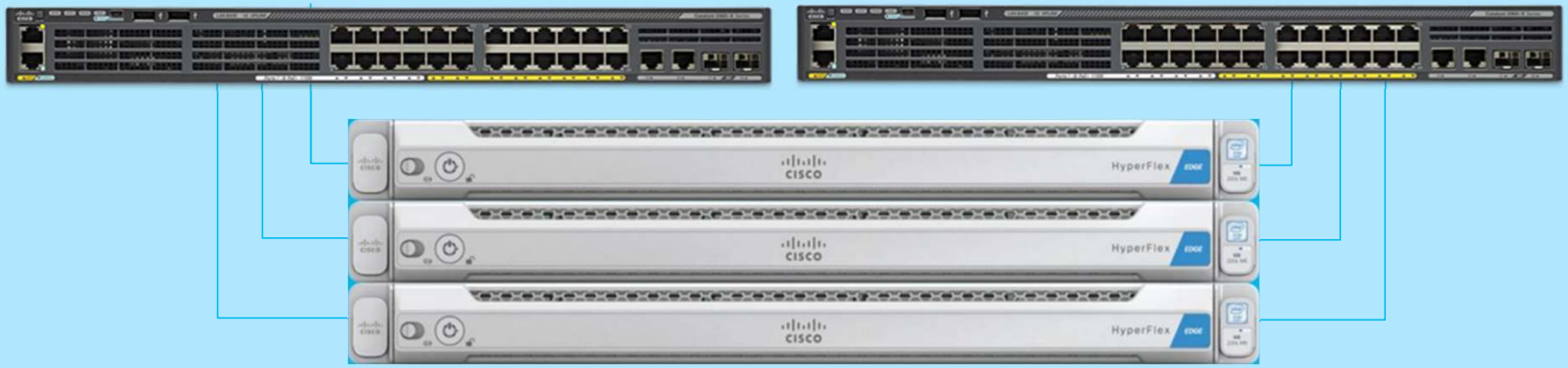
- Two Roles
  - Admin Role
  - Monitoring Role (no actions permitted)
- User management thru vCenter
  - User in Administrator group has Admin role
  - User with read only role has Monitoring role
  - VC supports Active Directory and Groups
  - Note: Local admin user supported but can not perform vCenter operations
- RBAC enforced for HX Connect UI, REST APIs and CLIs
  - Note:VC plug-in not recommended for monitoring role
- All actions are audited in audit log file



HyperFlex Edge

Remote Office Branch Office Solution

Designed for customers with no plans to upgrade or refresh their network at the remote office



- Single or dual Intel Xeon processor

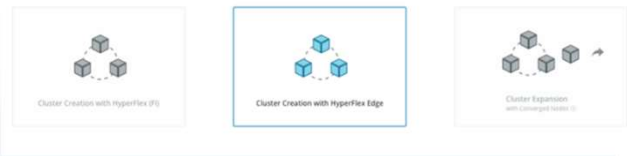
- 3 – 8x HDD
- 128GB – 3TB RAM
- 2 – 4x 1Gbps ports

Existing Network

Tuned for 1Gb/sec network

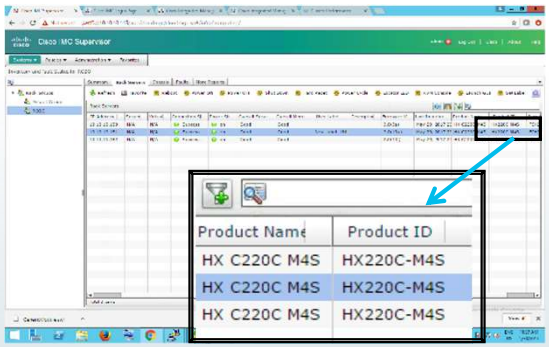
End-to-end HCI Automation

Extending Reach of Edge



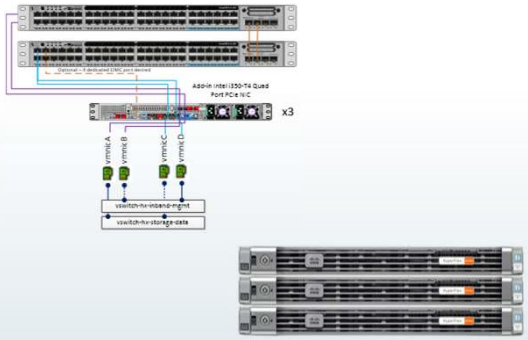
Automate Deployment

- CIMC automation for settings and policies
- ESXi configuration
- Pre-configured settings for HX Edge



IMC Supervisor support

- Centralized monitoring and firmware management and server config. for HX Edge
- Supported starting with IMCS v.2.1.0.2

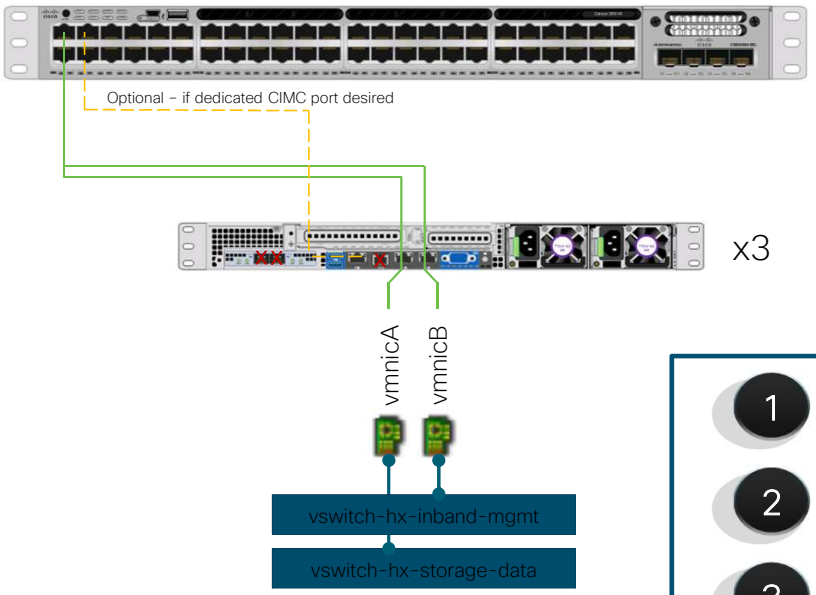


New Configurations

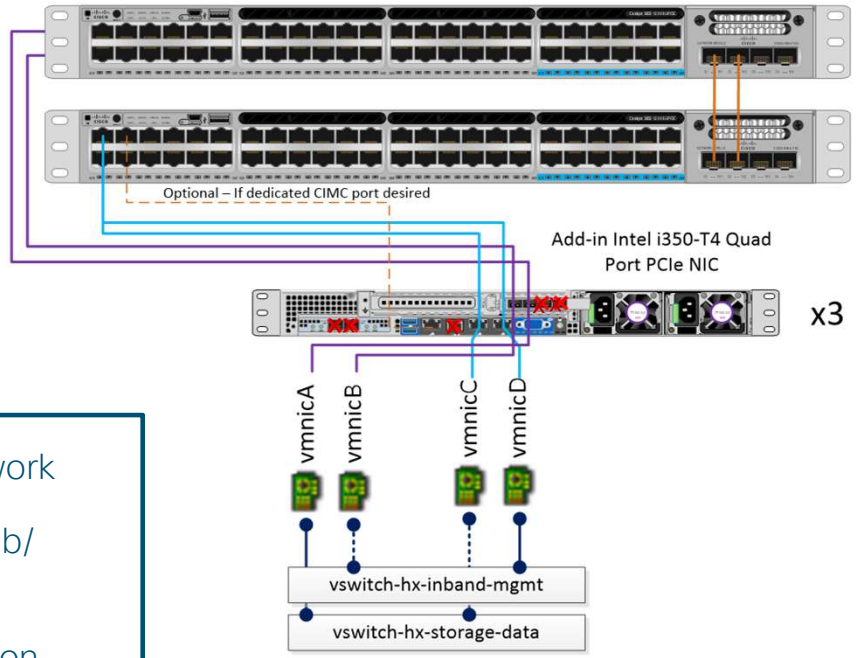
- Dual Switch (1Gb)– protect against network failure
- Evaluating new configs. – All-Flash HX220

HyperFlex Edge Configurations

Single Switch
HX Edge Configuration



Dual Switch
HX Edge Configuration



1. Existing Network
2. Tuned for 1Gb/ sec network
3. HCI Automation

HyperFlex Use Cases

Primary HyperFlex Use Cases

Virtual Desktop Infrastructure

- Low upfront costs
- Consistent performance
- Predictable scaling

Server Virtualization

- Reduce operational complexity
- Adaptive scaling
- Always-on resiliency

Test and Development

- Agile provisioning
- Frequent iterations
- Instant cloning and snapshots

Large Remote Branch Office

- Simple deployment
- Centralized management
- No “fly-and-fix” missions

Databases

- Consistent, low-latency
- High IOPS
- All-flash nodes

HyperFlex Solutions / Workloads Portfolio

WORKLOADS / APPLICATIONS			
Citrix XenDesktop	VMWare Horizon	Oracle Single Instance	Cisco UC
VSI on HX	MS SQL	SAP (Non – Prod / Prod)	ACI
Oracle RAC	Splunk	MS Exchange	N1K

INTEGRATION / INTEROP	
Veeam	Commvault
Cloud Center	UCSD
3rd Party Arrays	NVIDIA

Cisco HyperFlex: HX Data Platform

Data Services and Storage Optimization

HX Tools (for sizing)

Enterprises Run Mission Critical Apps on HyperFlex

CITRIX
Exchange
Cisco webex
Microsoft SQL Server
ORACLE DATABASE
ORACLE E-BUSINESS SUITE
Microsoft
SharePoint
Cisco Unified Contact Center
MySQL
SIEBEL
PeopleSoft
vmware
McKESSON
Epic
MUREX
SAP

- Sustained Low Latency
- Availability & Resiliency
- Enterprise Data Services

Telstra
PayPal
Sutter Health
BOEING
QATAR AIRWAYS القطرية
Allianz
Chevron
ERICSSON
NTT DATA

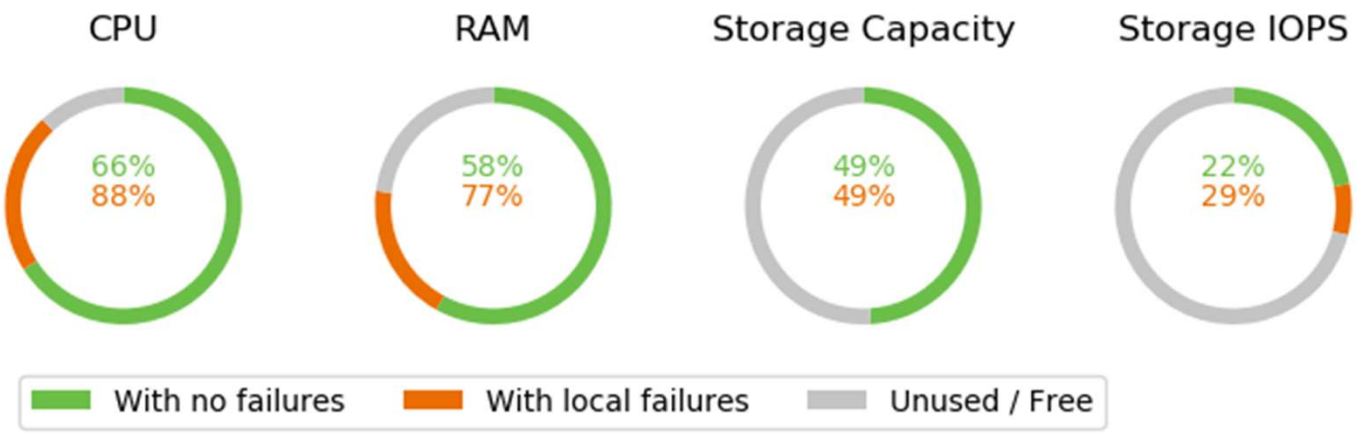
HyperFlex Sizer

Cluster 1 Summary

Sizing Summary

No. of Workloads	4
Number of Nodes	3 + 1 (FT)
Number of RU	4

Utilization - Cluster 1



Recommended Nodes (1/1)

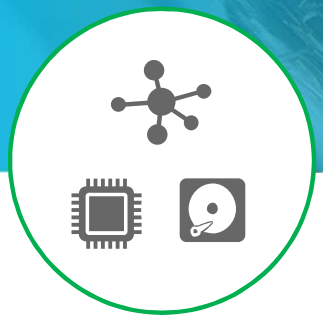
HXAF-220M5SX	
2xIntel Xeon Gold 6154 Processor, 18 cores, 3.00 GHz | 384 [12x32] GiB DDR4 RAM | 6x960GB, 2.5" SSD | 1x400GB SAS | 1 RU	4

https://hyperflexsizer.cloudapps.cisco.com/ui/index.html

Summary

Cisco UCS

Unified Computing System

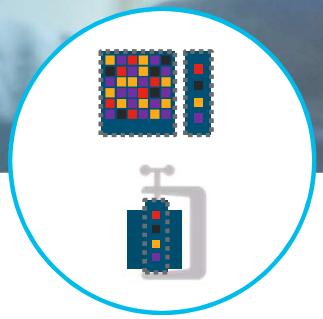


HyperFlex

Network, Compute & Storage Integrated

HyperFlex

Independent Scaling



Continuous Optimization

Always-on
Dedup & Compression

More information

https://www.cisco.com/c/dam/en/us/products/collateral/hyperconverged-infrastructure/hyperflex-hx-series/le-wp-hxt-op.pdf

CISCO