



NetApp®

Go further, faster®

Continuous Availability for Business-Critical Data

MetroCluster Technical Overview





Agenda

- Causes of Downtime
- MetroCluster Overview
- MetroCluster Deployment Scenarios
- MetroCluster Failure Scenarios
- New Features
- Summary

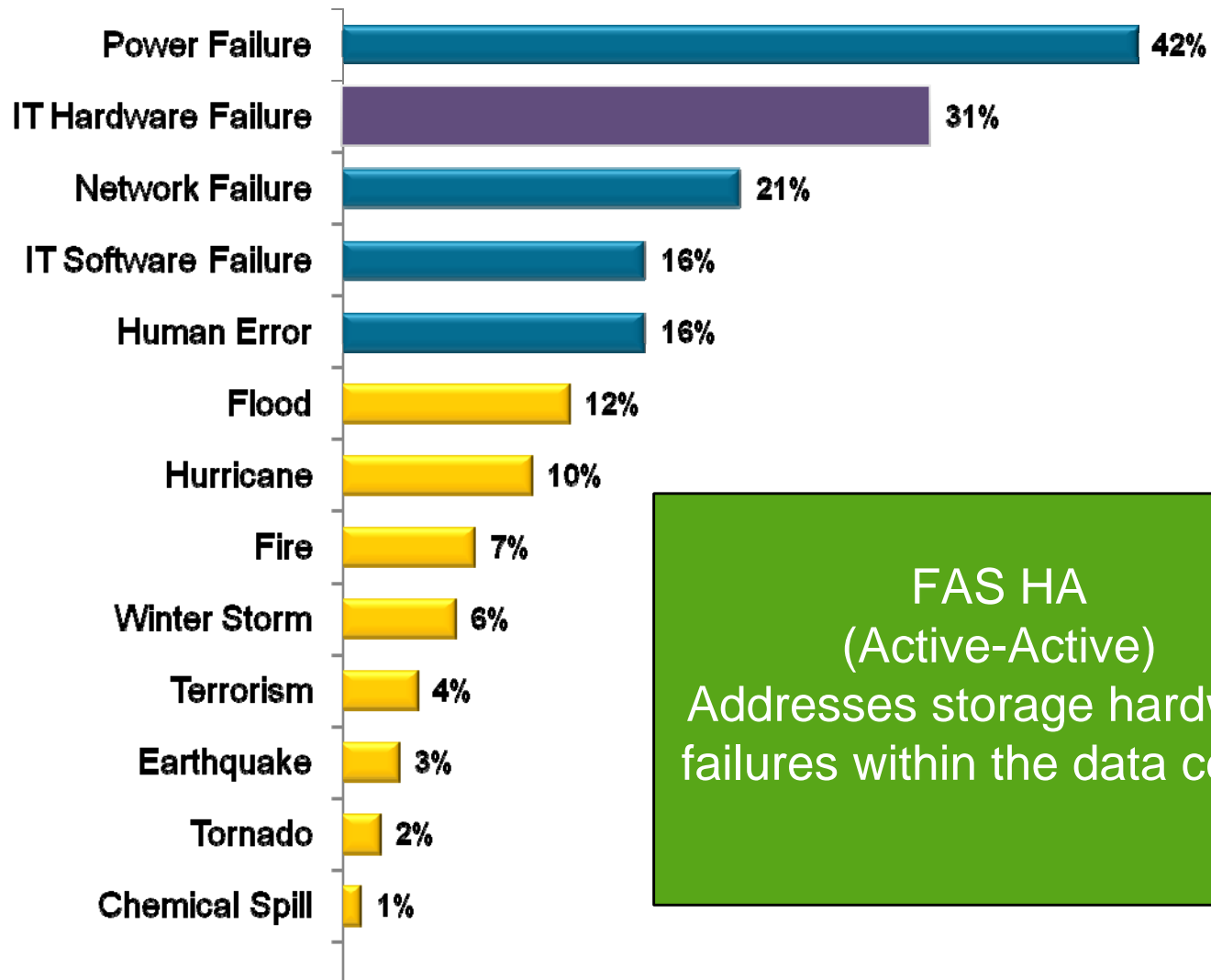




Customer Challenges

**Internal
Data Center
Failures**

**External
Data Center
Failures**



**FAS HA
(Active-Active)**
Addresses storage hardware failures within the data center

Source: Forester / Disaster Recovery Journal Global Disaster Recovery Preparedness Online Survey, Oct, 2007



Defining Availability and Disaster Recovery

- High Availability (HA) refers to eliminating the business impact and downtime due to a hardware failure
 - RAID-DP®
 - Active-Active controllers

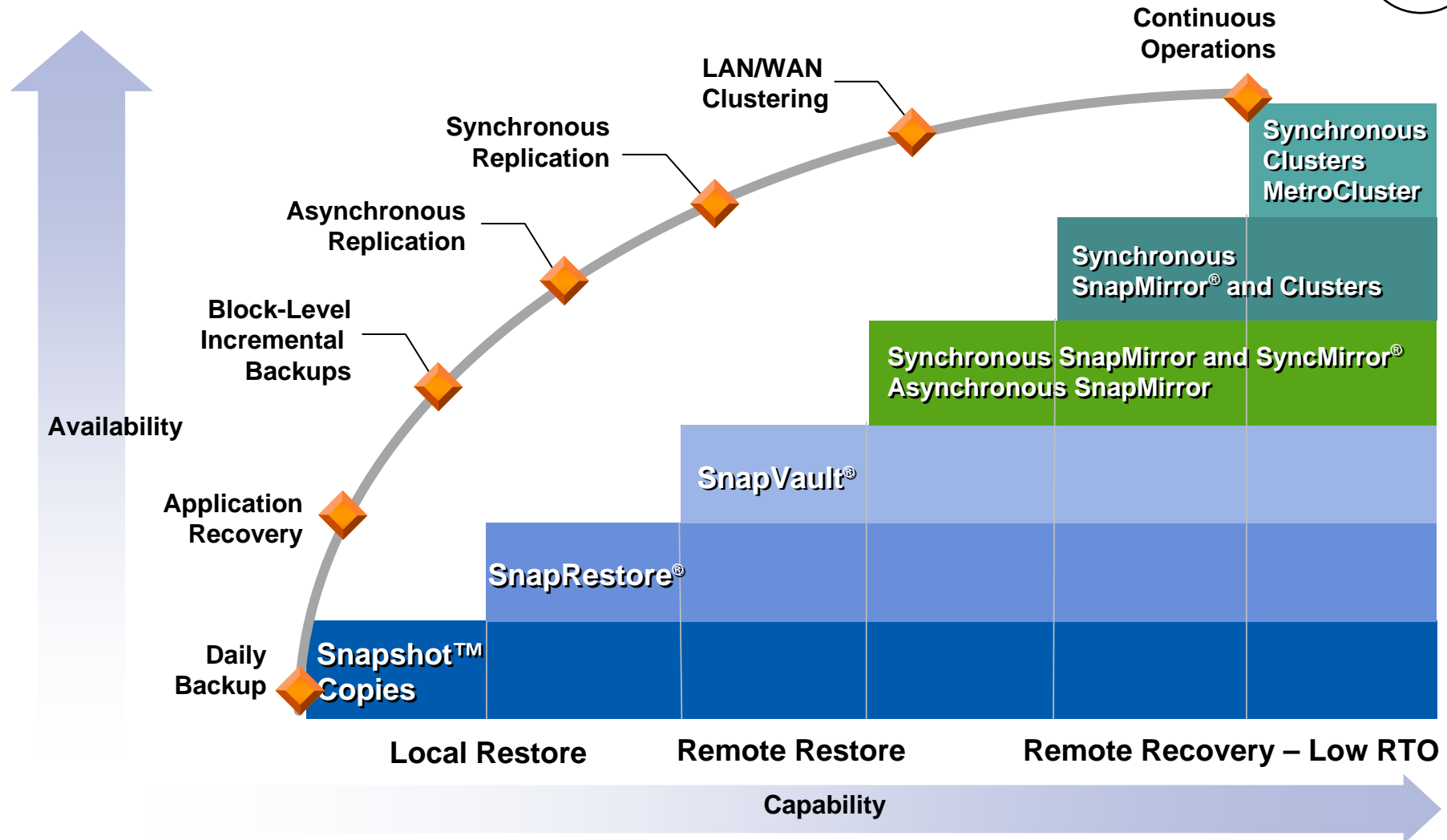
- Continuous Availability extends HA by providing protection from external events such as power, cooling and network failures, eliminating planned and unplanned downtime and data loss
 - MetroCluster

- Disaster Recovery provides site resiliency protecting against failures that affect an entire data center
 - SnapMirror®



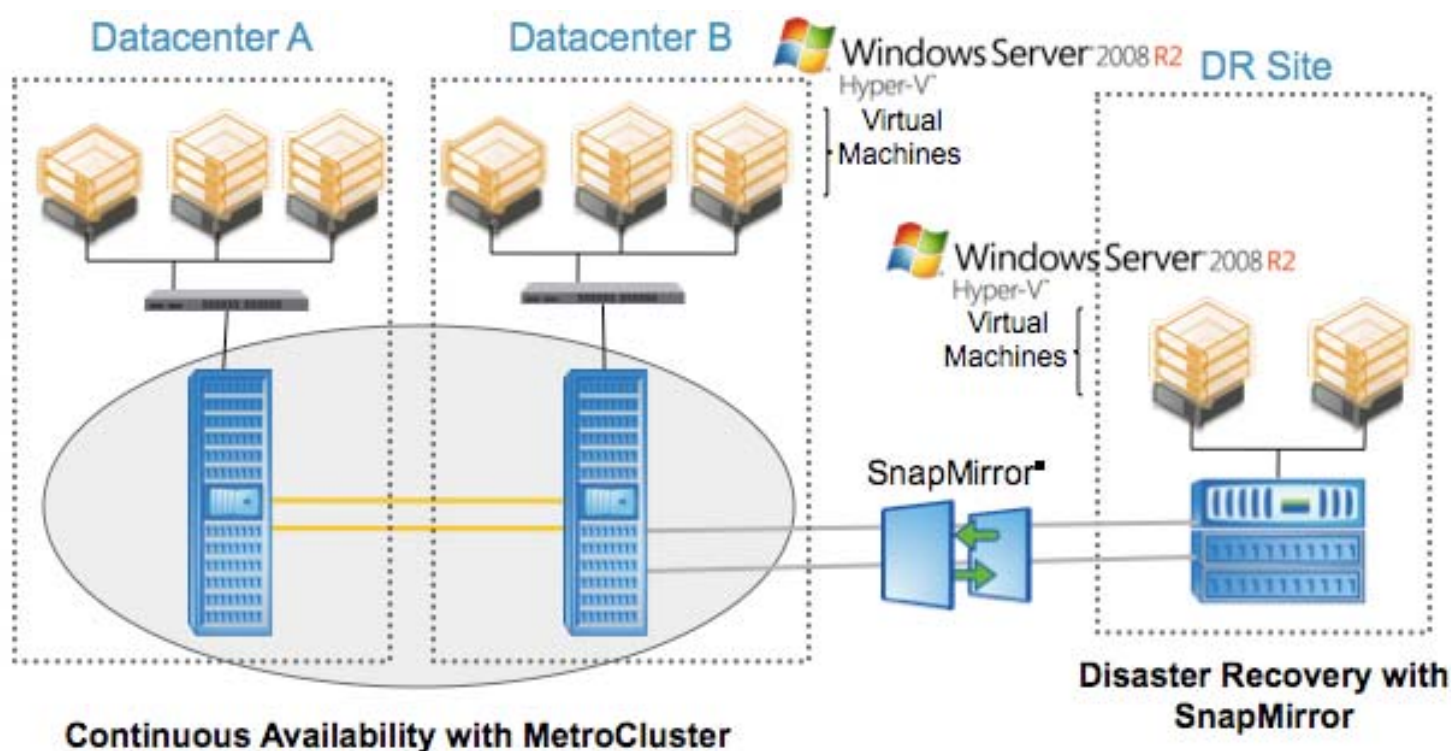
NetApp Data Protection/ Business Continuity Solutions

Data Protection





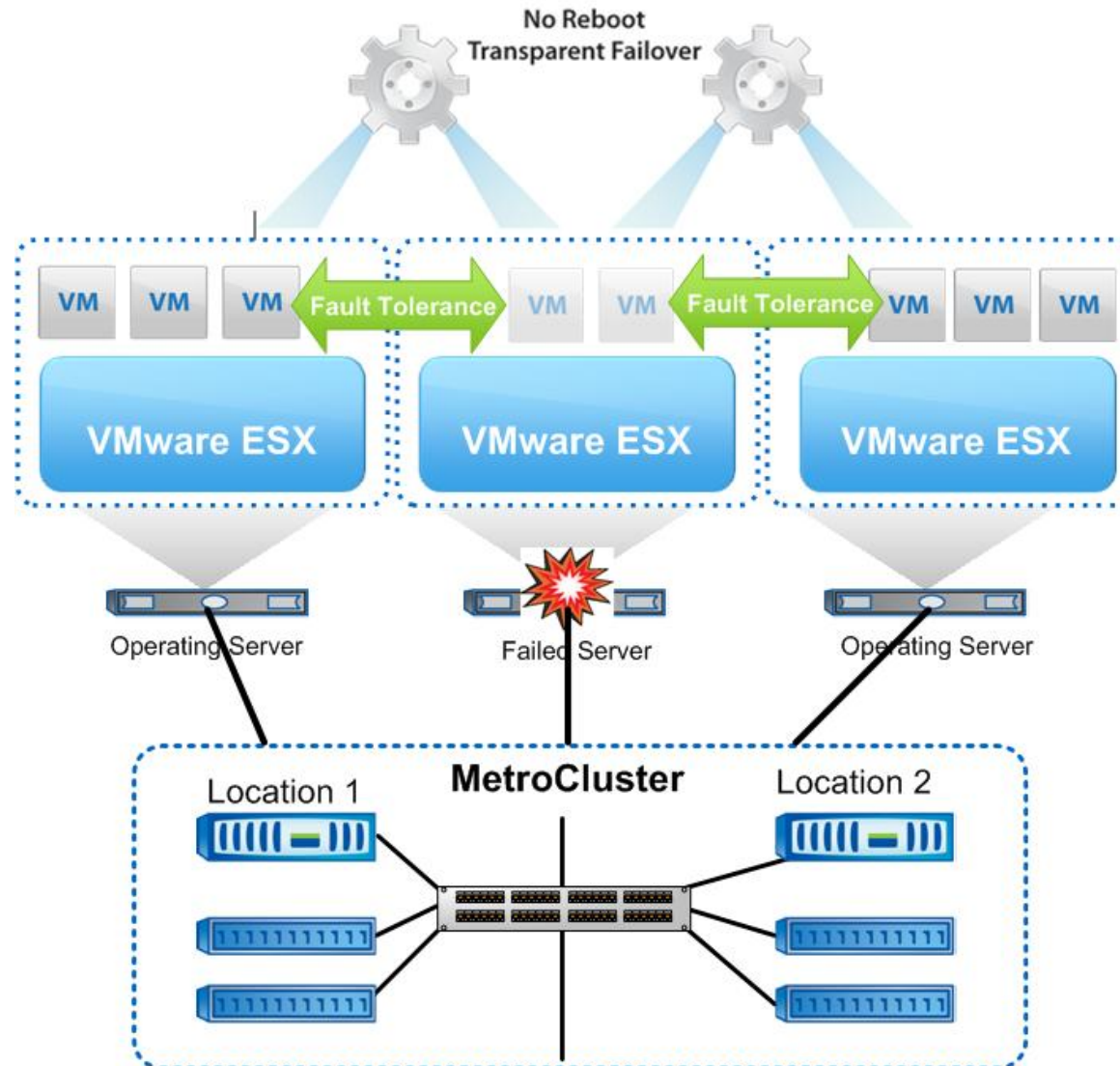
NetApp Solution for Hyper-V



- Automatic failover within data centers or regional area
- Practical continuous data availability (CA) and DR for all apps
- Replication of VMs and data
- Hyper-V VM and application consistency
- Asynchronous replication for fast performance
- Synchronous replication for up-to-the-minute data



VMware FT and MetroCluster Top-down Continuous availability



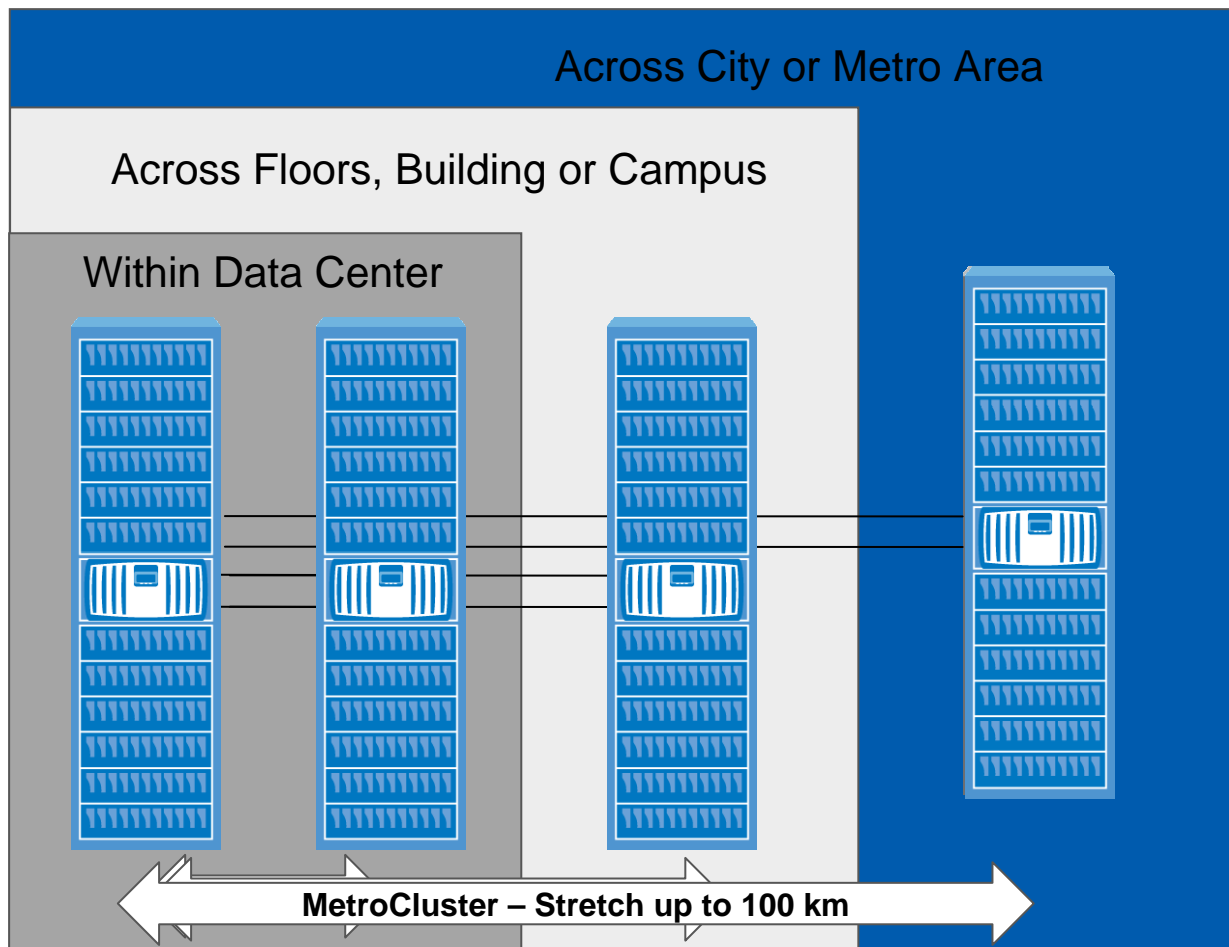


MetroCluster Overview





The MetroCluster Difference

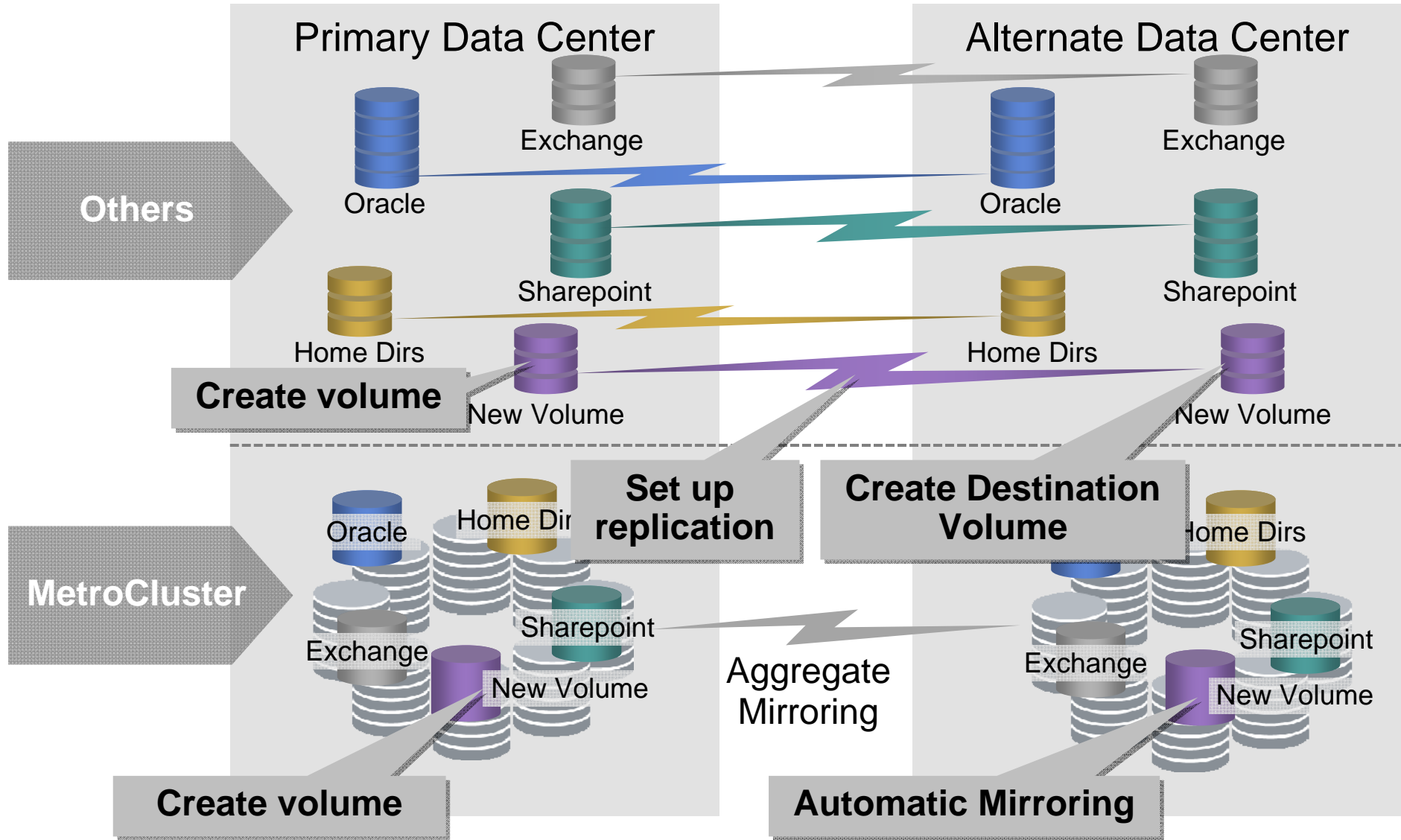


- MetroCluster provides continuous availability within a single data center and across data centers in adjacent floors, buildings and metro areas
- Support for FAS3000, FAS3100, FAS6000 and V-Series



The MetroCluster Difference

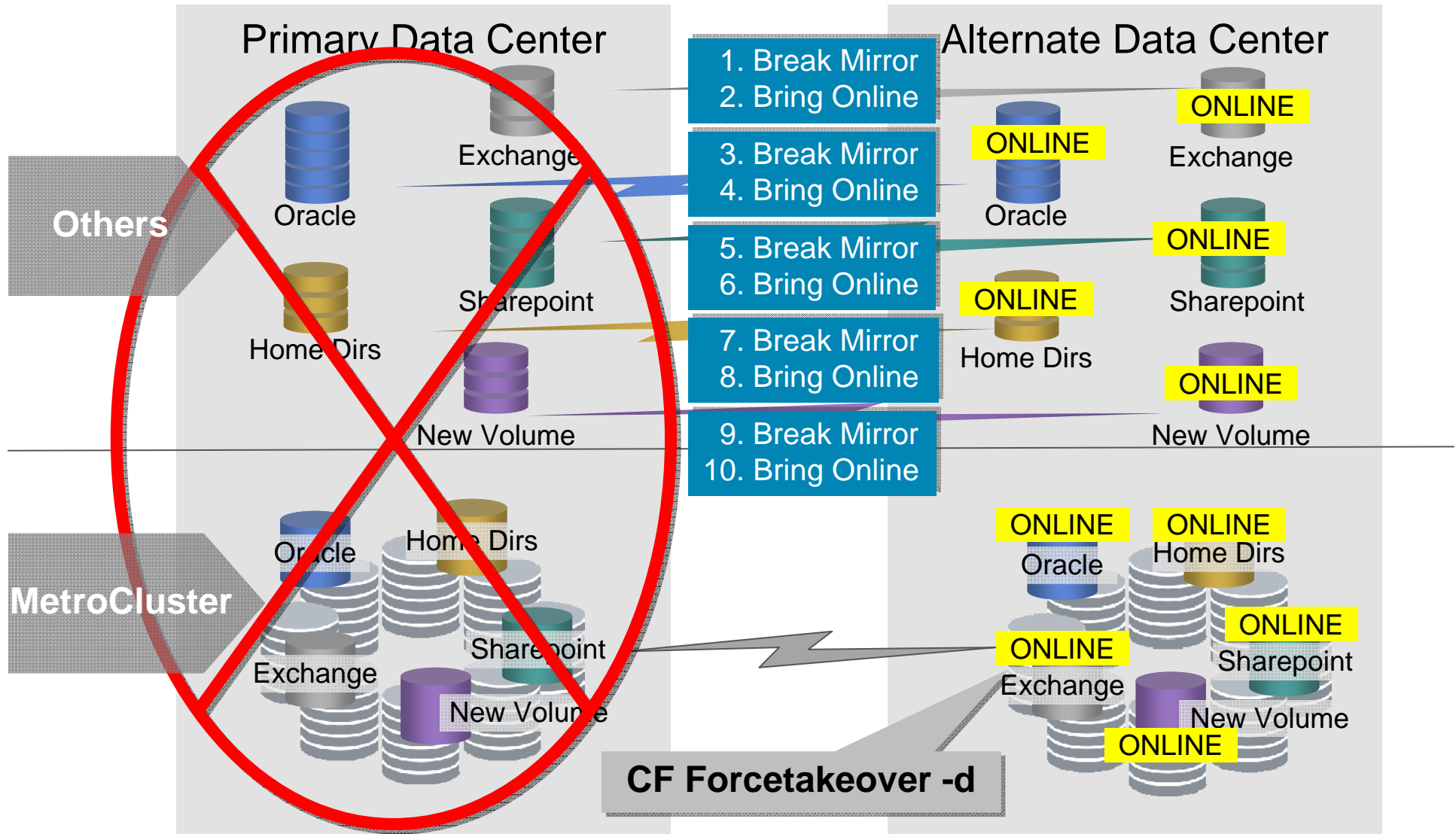
Automatic Protection





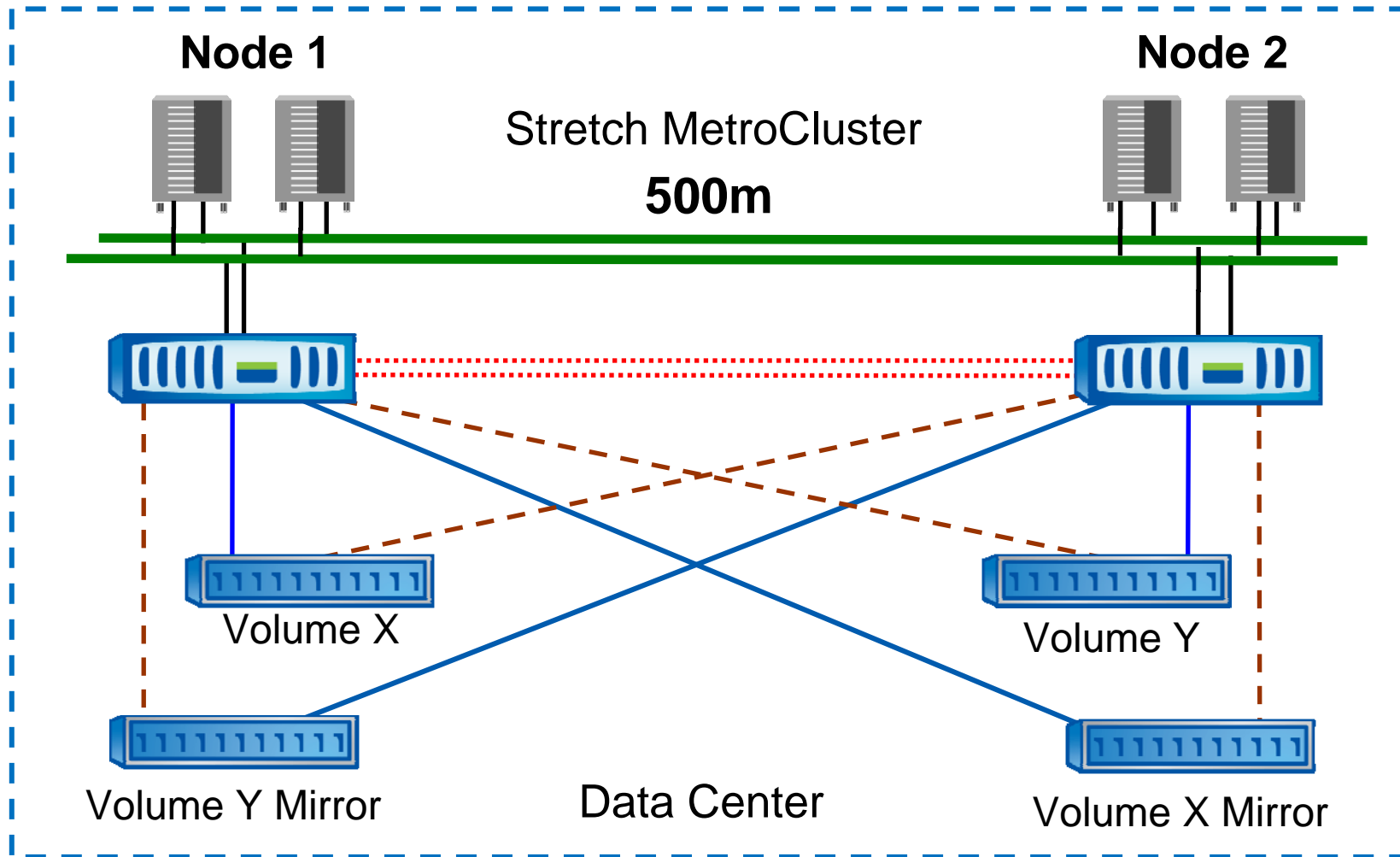
The MetroCluster Difference

Quick Recovery time



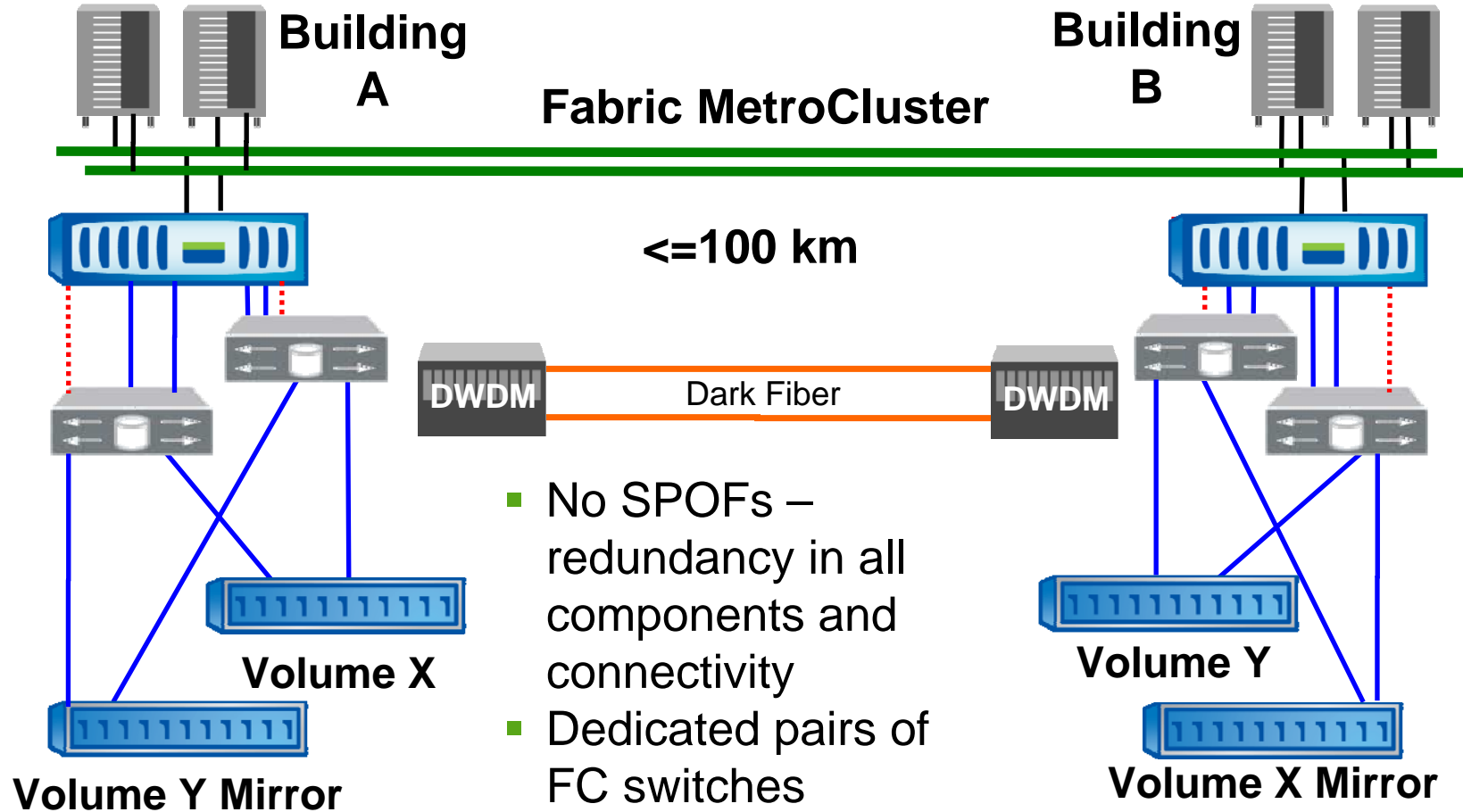


MetroCluster Deployment Within a Single Data Center, across floors, etc.





MetroCluster Deployment Across City or Metro Areas





MetroCluster Components

Cluster License Provides automatic failover capability between sites in case of most hardware failures

SyncMirror[®] Provides an up-to-date copy of data at the remote site; data is ready for access after failover without administrator intervention

Cluster Remote Provides a mechanism for administrator to declare site disaster and initiate a site failover via a single command for ease of use

FC Interconnect Provides appliance connectivity between sites that are greater than 500 meters apart; enables controllers to be located a safe distance away from each other



MC Component: Cluster Remote License

- Provides single command failover
- Provides the ability to have one node take over its partner's identity without quorum of disks
- Root volumes of both appliances **MUST** be synchronously mirrored
- By default all volumes are synchronously mirrored and are available during a site disaster
- Requires administrator intervention as a safety precaution against a “split brain” scenario (*cf forcetakeover -d*)
 - *IMPORTANT: Site-specific disasters are not the same as a normal cluster failover!*



MC Component: SyncMirror

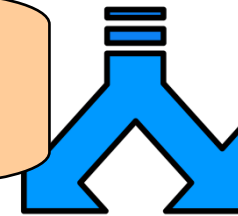
Writing to an unmirrored volume



Writing to a mirrored volume

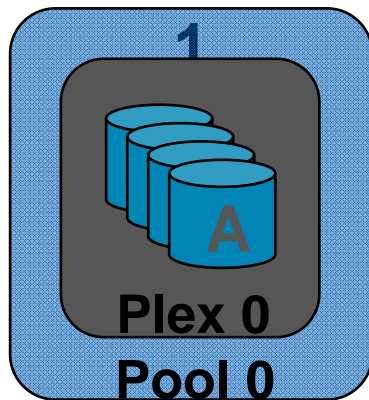


Two Plexes Created
Plex 0 from Pool 0 = Orig
Plex 1 from Pool 1 = Mirror

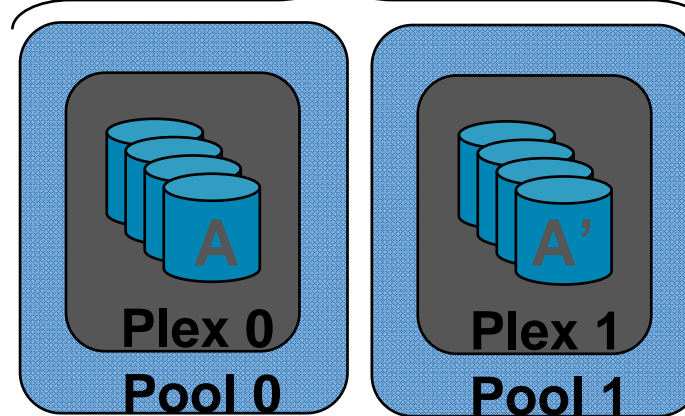


Note:
Plex #
may change
over time

Volume



Volume1



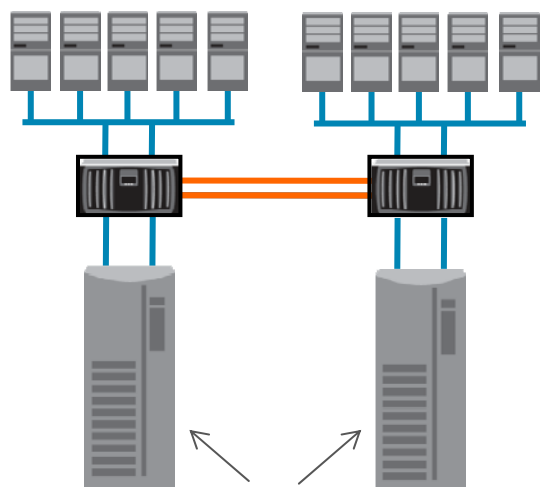


Disk Ownership

- Software
 - Disks assigned to pools by administrator
 - FAS3040, FAS3070, FAS31xx, FAS6000
 - FAS3020, FAS3050 (Stretch MC only)
- Hardware
 - Determined by physical connection
 - 900 series, FAS3020, FAS3050



MetroCluster for V-Series



NetApp
V-Series
Open Storage
Controller

Heterogeneous Storage Arrays

V-Series MetroCluster	FAS MetroCluster
Up to 30km distance	Up to 100km
Dedicated or shared SAN	Dedicated
Brocade or Cisco Switches	Brocade
LUN limit = Data ONTAP Spindle limit	504 Max spindles on Fabric
Fabric or Stretch Configurations	Fabric or Stretch Configurations
3 rd Party Storage only	NetApp Storage

Refer to V-Series Compatibility Guide at:

http://now.netapp.com/NOW/knowledge/docs/V-Series/supportmatrix/V-Series_SupportMatrix.pdf

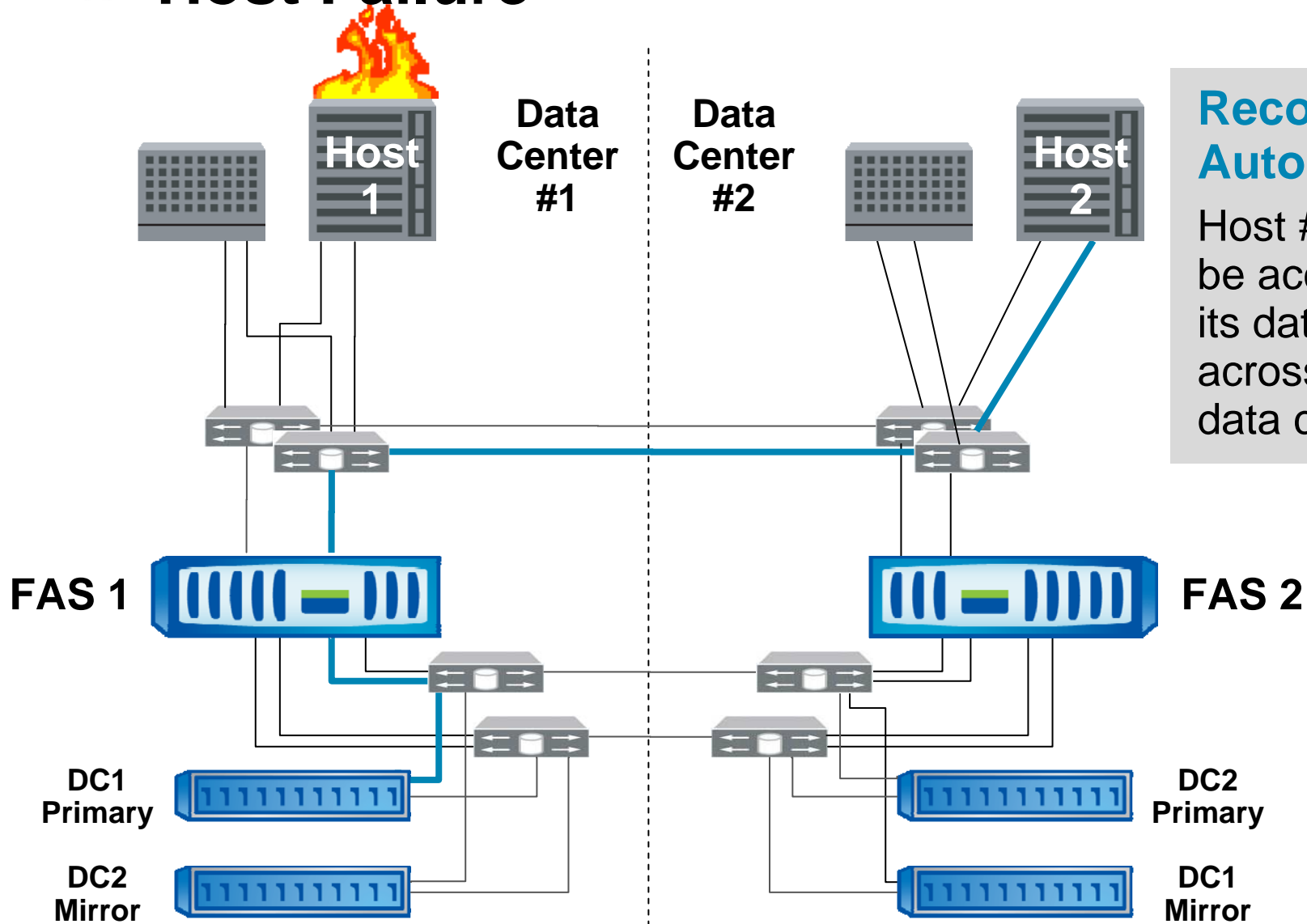


MetroCluster Failure Scenarios





MetroCluster Failure Scenarios: Host Failure

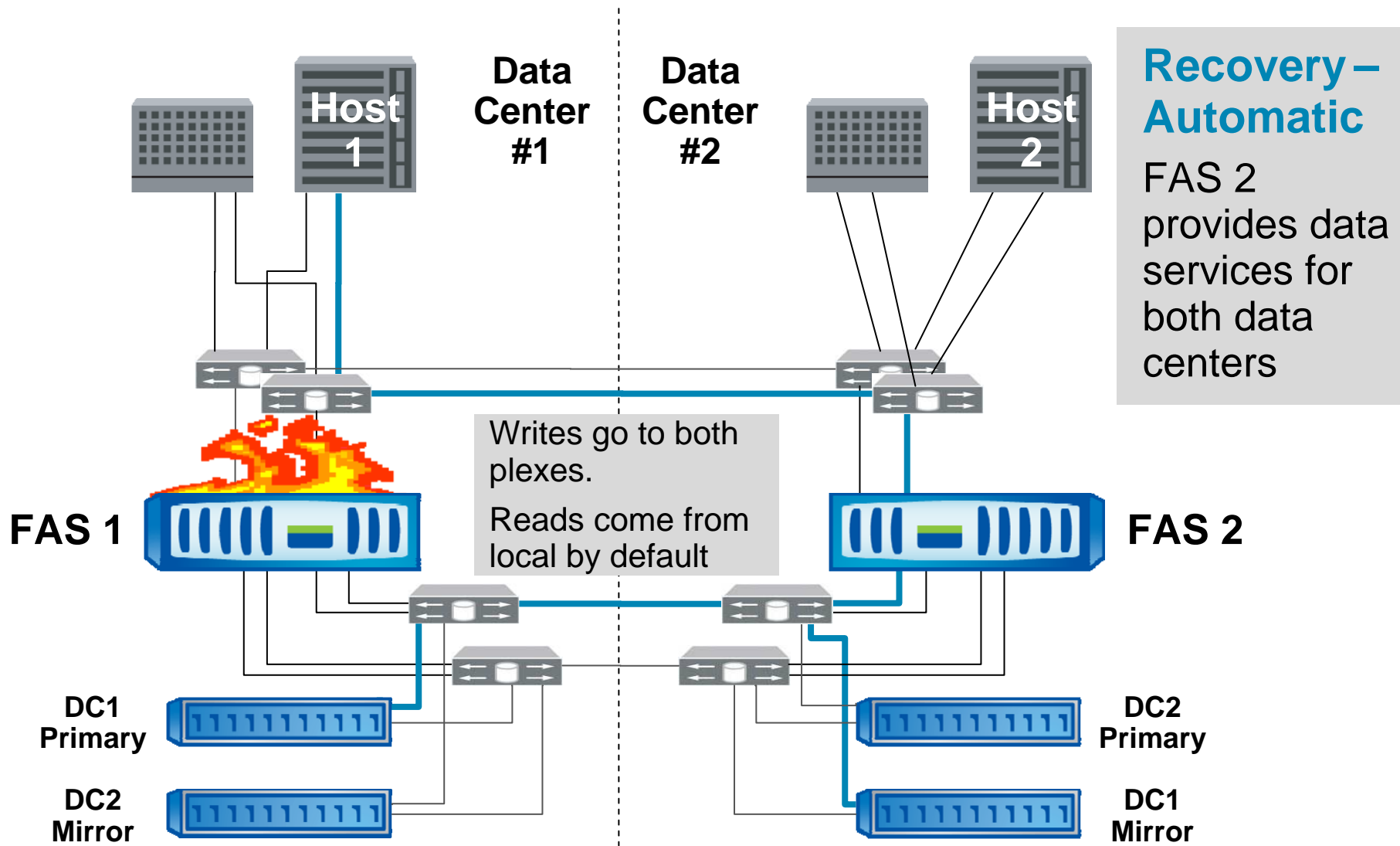


Recovery – Automatic

Host #2 will be accessing its data across the data centers

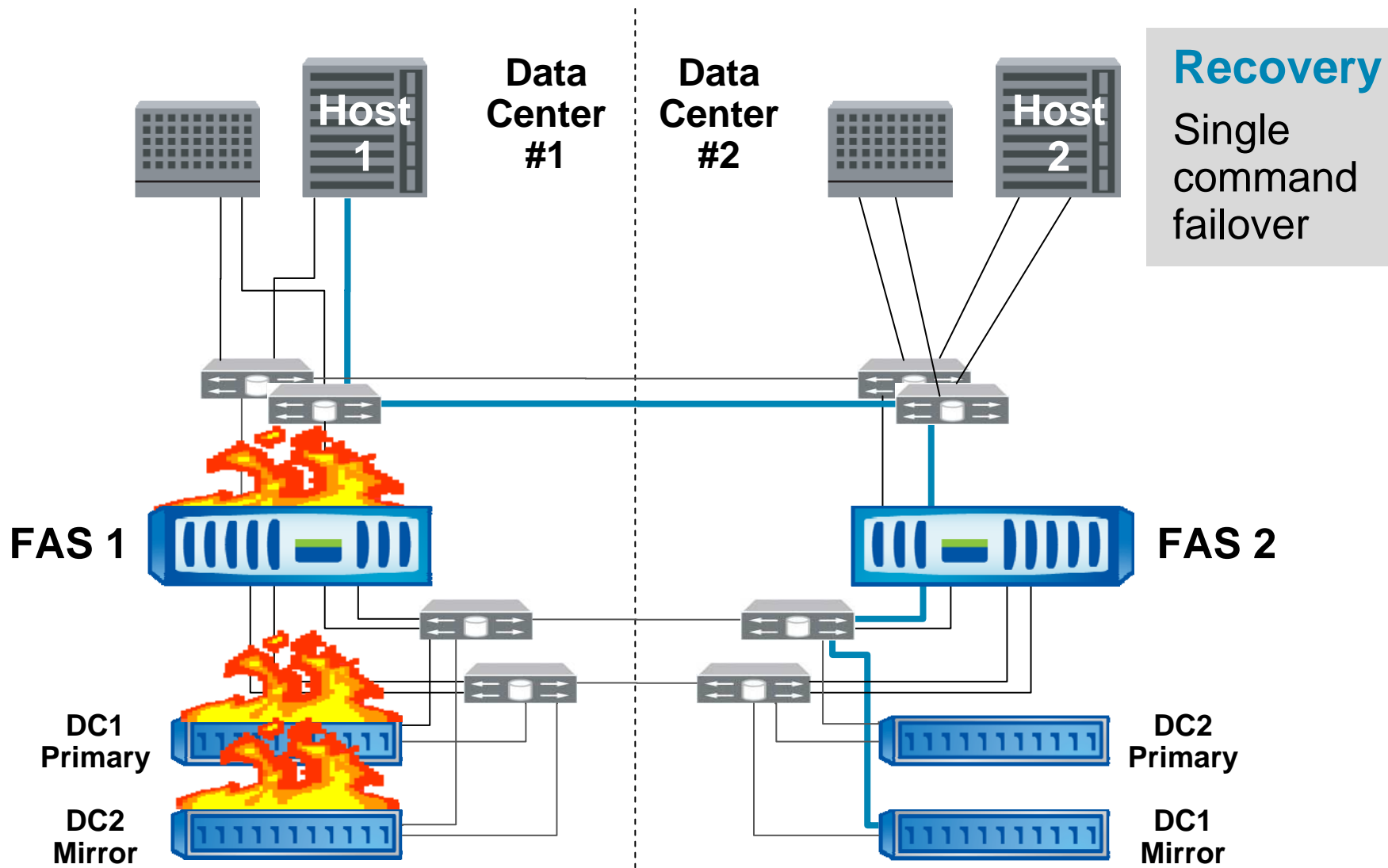


MetroCluster Failure Scenarios: Storage Controller Failure



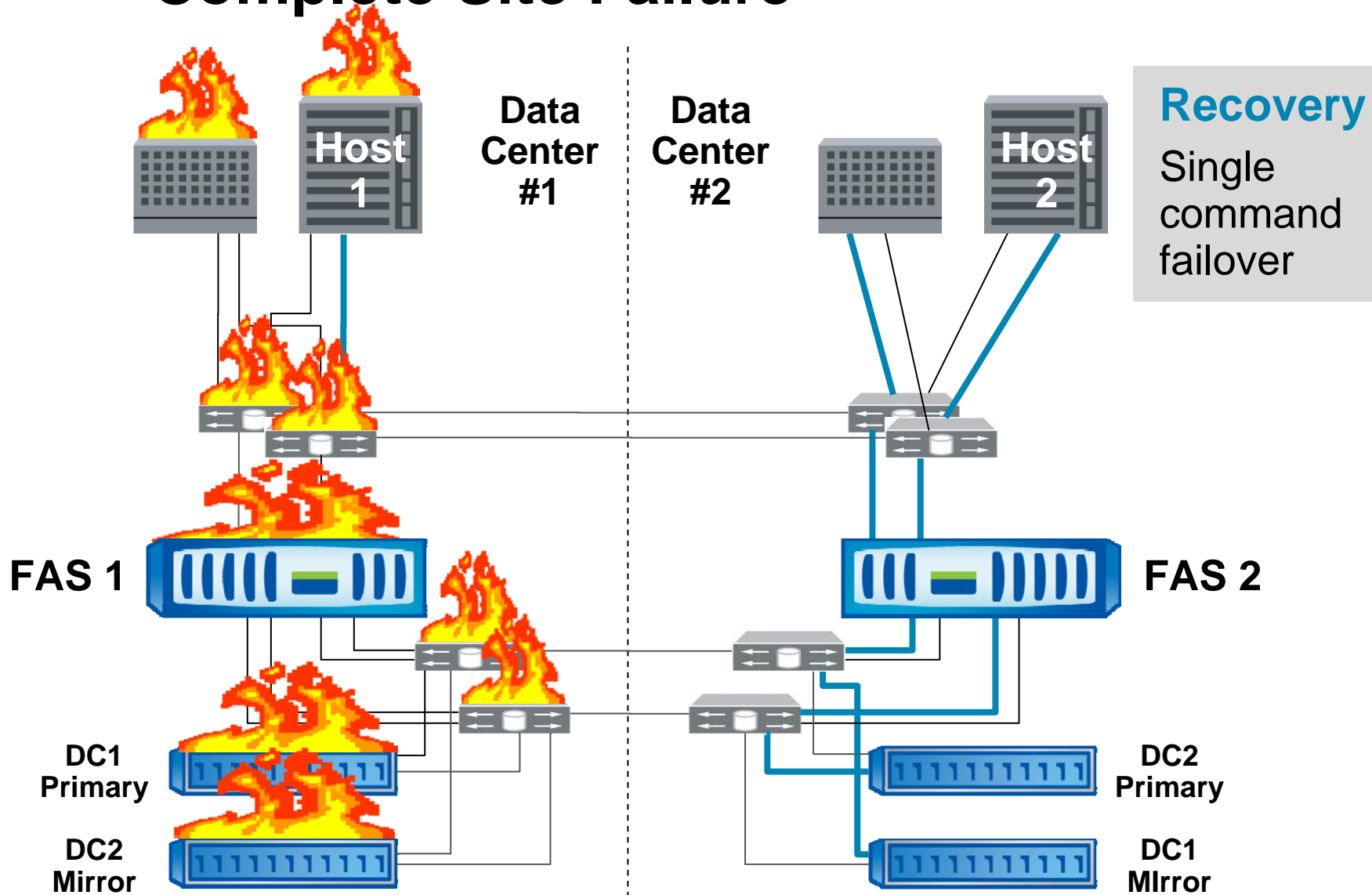


MetroCluster Failure Scenarios: Storage Array Failure



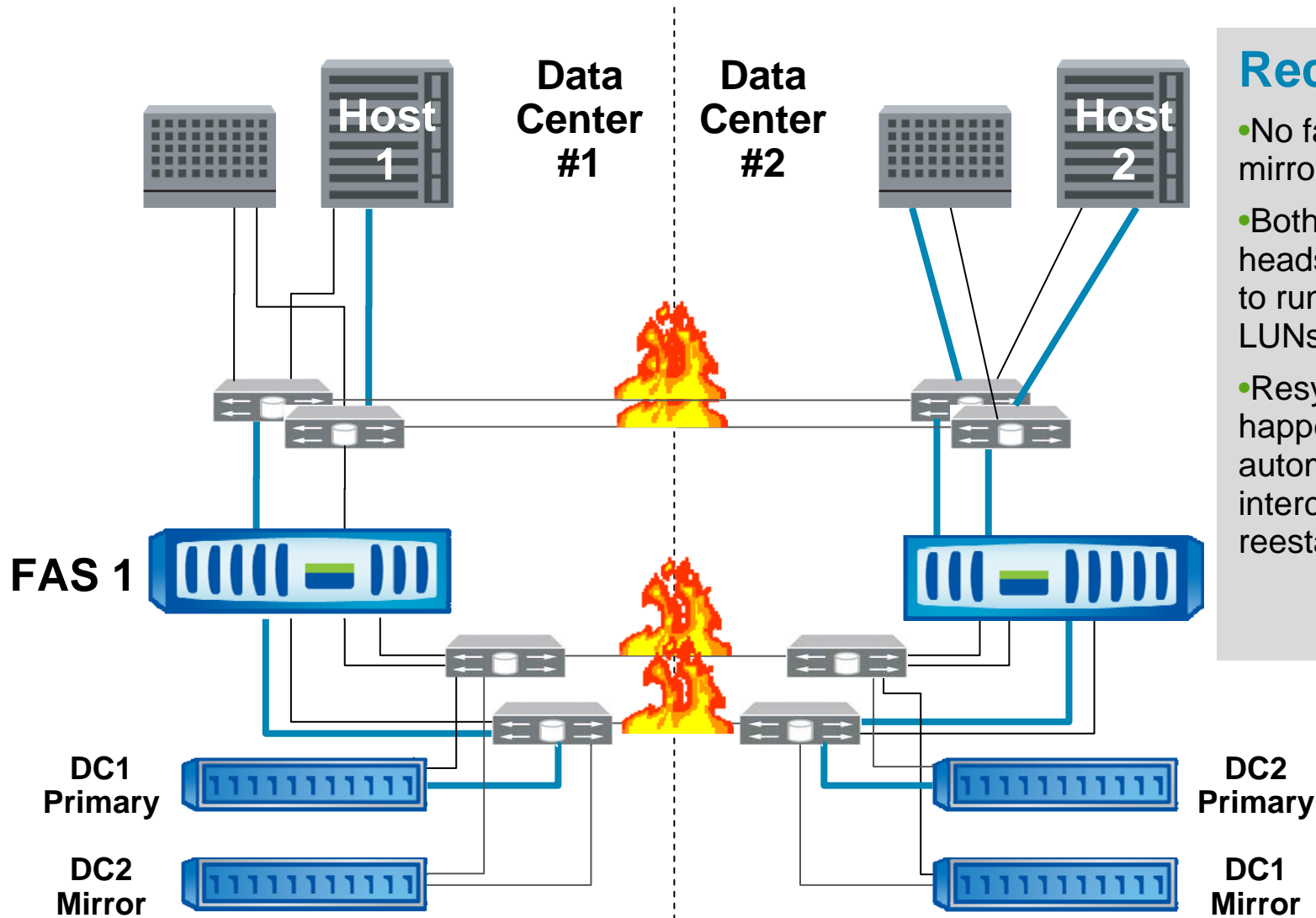


MetroCluster Failure Scenarios: Complete Site Failure





MetroCluster Failure Scenarios: Interconnect Failure

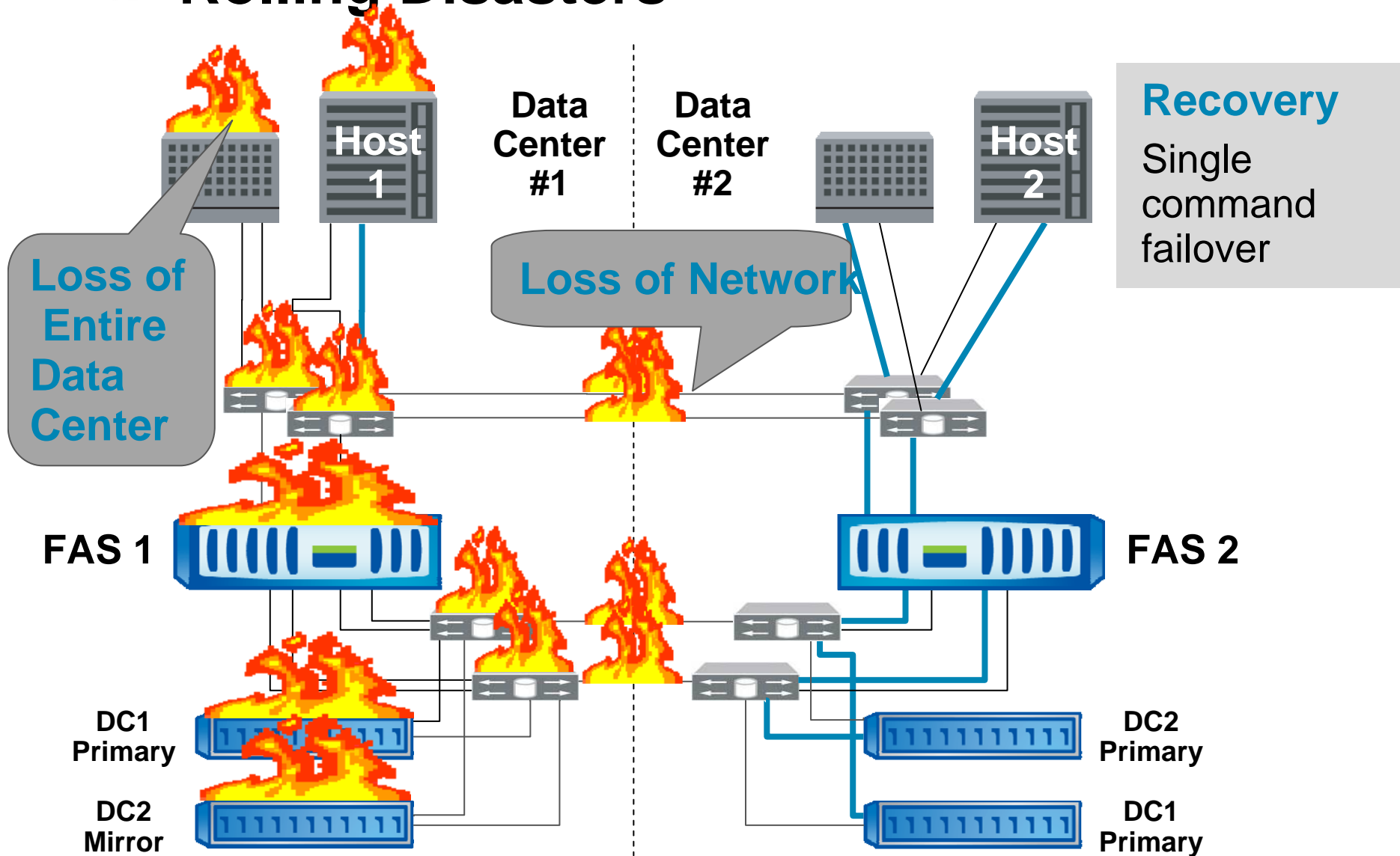


Recovery

- No failover; mirroring disabled
- Both appliance heads will continue to run serving their LUNs/volumes
- Resyncing happens automatically after interconnect is reestablished

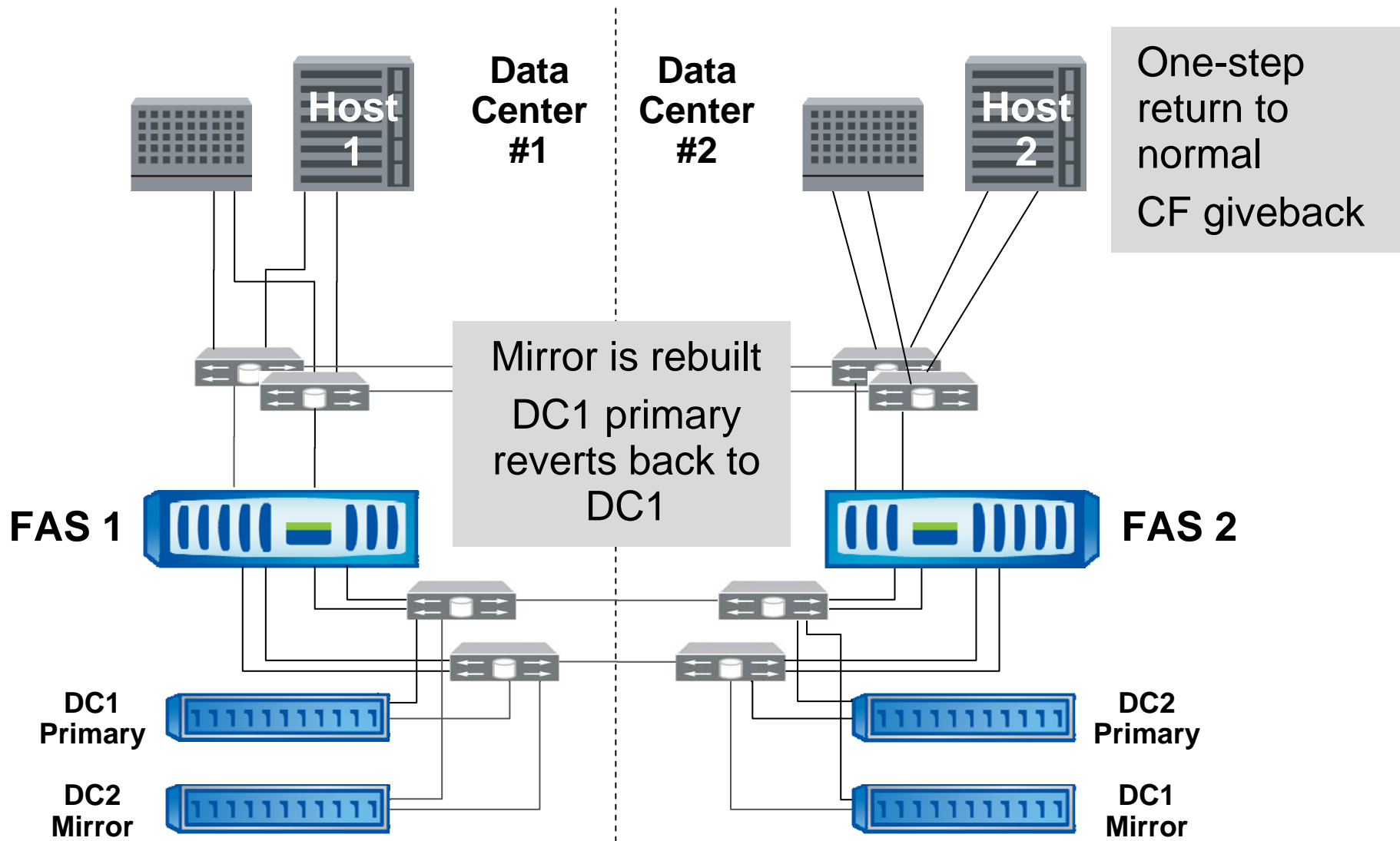


MetroCluster Failure Scenarios: Rolling Disasters





MetroCluster Failure Scenarios: Return to normal





Support Matrix

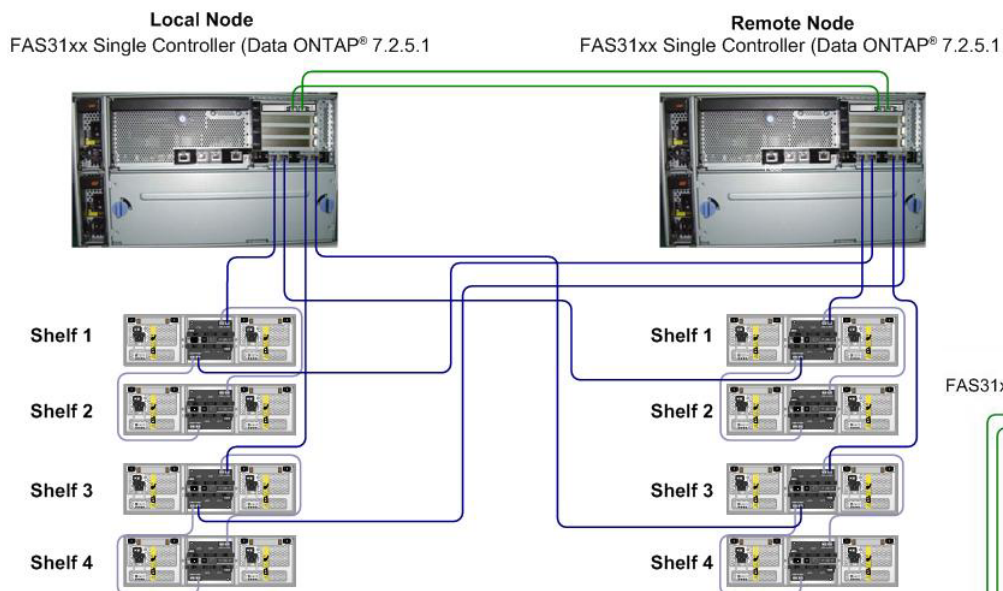
	Stretch		Fabric	
Model	Supported	Max. Disks	Supported	Max. Disks
3020	YES	168	YES	168
3040	YES	336	YES	336
3050	YES	336	YES	336
3070	YES	504	YES	504*
3140	YES	420	YES	420
3170	YES	840	YES	672*
6030	YES	840	YES	672*
6040	YES	840	YES	672*
6070	YES	1008	YES	672*
6080	YES	1166	YES	672*

* 504 spindle count requires Data ONTAP® 7.2.4 or later

* **672 spindle count requires Data ONTAP 7.3.2 or later**

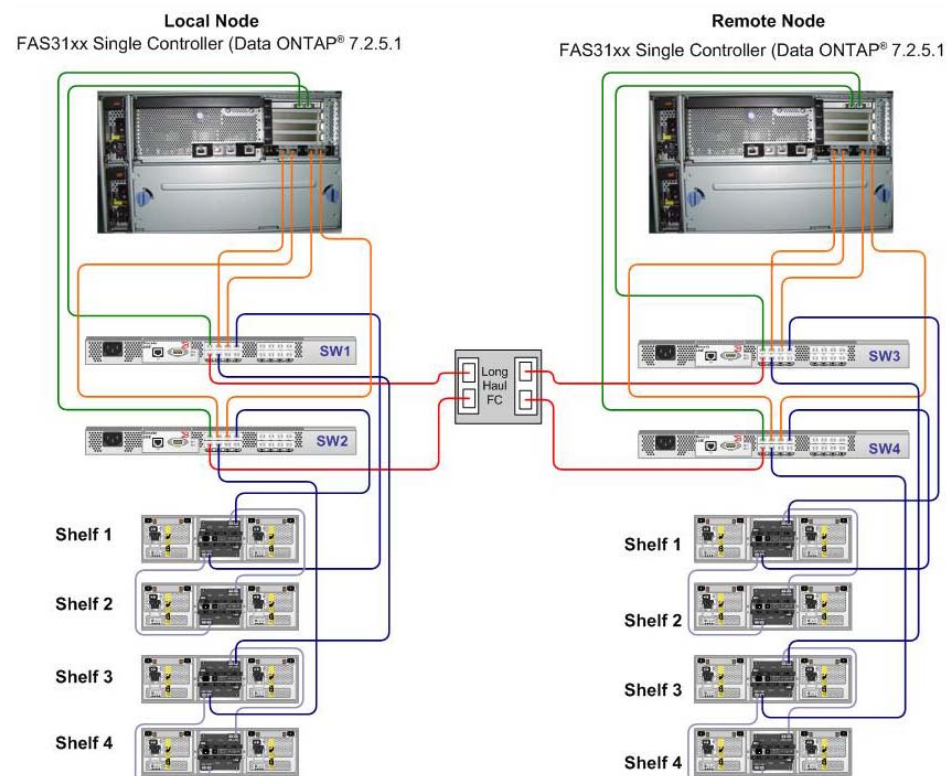


31xx Configurations



NOTE:

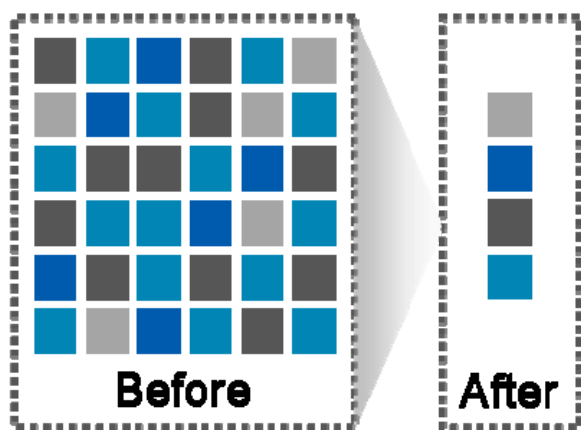
1. Single controller chassis only
2. FC/MI card required in both Stretch and Fabric configuration





Deduplication for MetroCluster

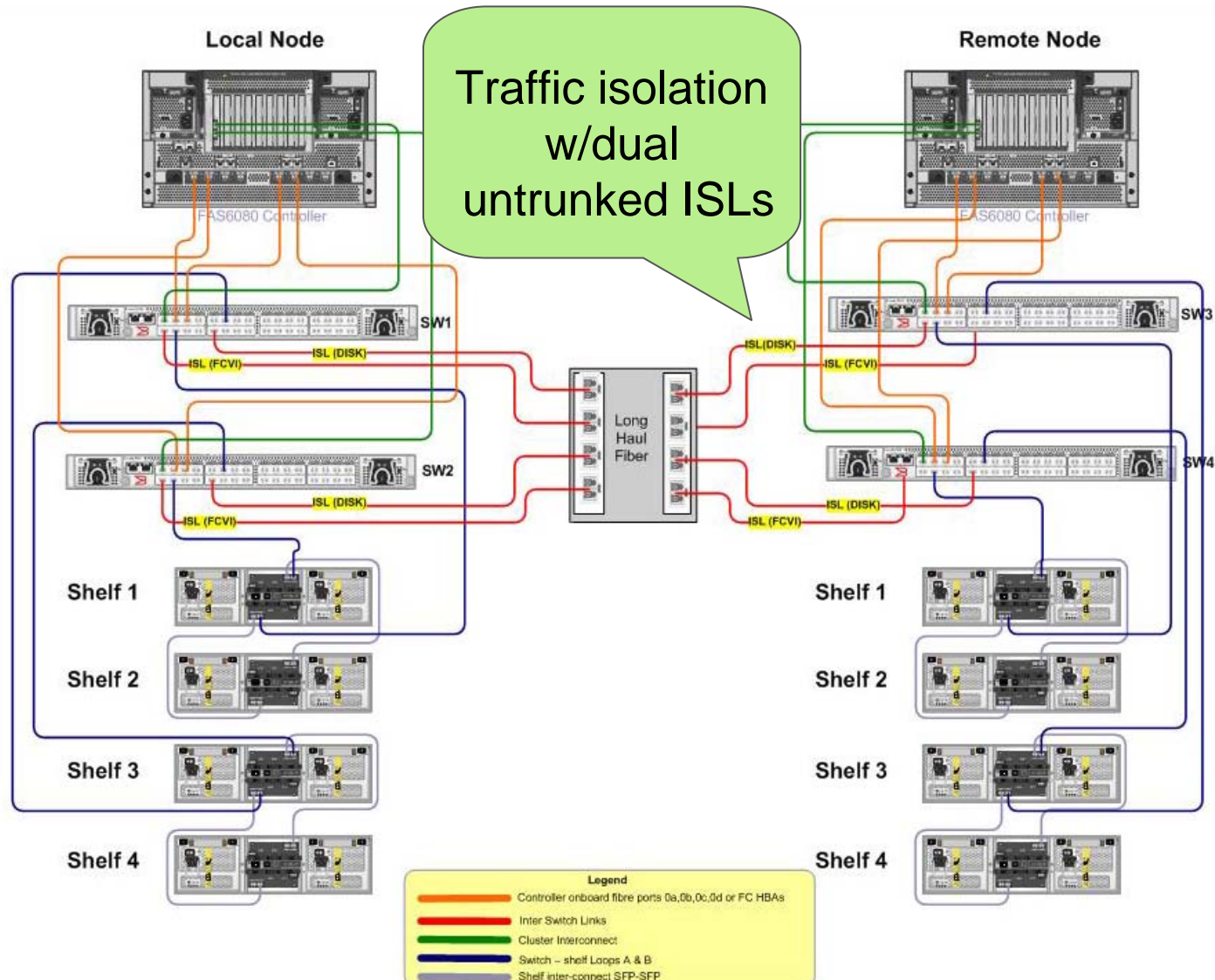
NetApp® Deduplication



- Deduplication provides cost savings for MetroCluster
 - Deduplication efficiencies for both the source and mirrored arrays
 - Investment protection and storage efficiency for other vendor storage with V-Series
- PVR no longer required
 - Must be licensed on both nodes
 - Must be running Data ONTAP® 7.2.5.1 or later, 7.3.1 or later
 - Subject to same considerations as non-MetroCluster systems



Traffic Isolation: Available since 7.2.6



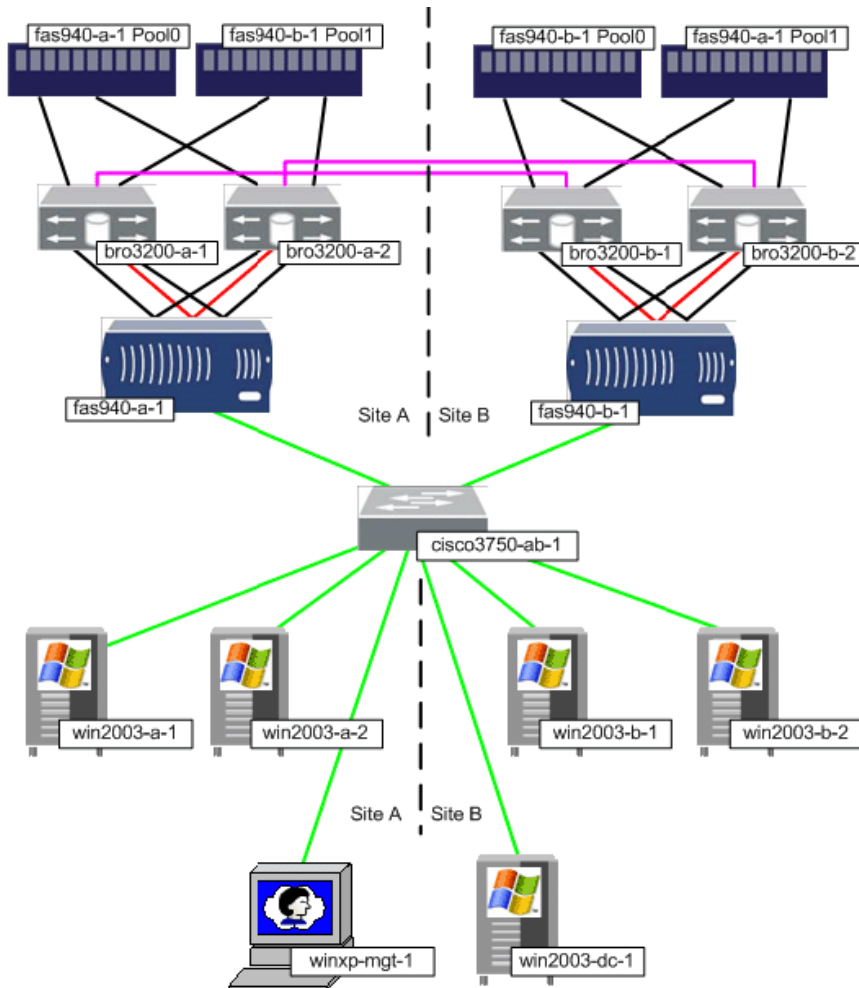


Best Practices: MetroCluster

- Choose the right controller
- Connections are important
 - Calculate distances properly
 - Use the right fiber for the job
 - Verify correct SFPs, cables required, and patch panels
- Remember mirroring requires 2x disks
- Remember MetroCluster spindle maximums
- Remember speed restrictions



MetroCluster Complements Host Clustering



■ Solution Highlights

- No SPOF
- High failure resiliency—no failover required for any single component failures:
 - Disks, disk shelves, disk loop, MetroCluster switches, ISLs, cables, interconnects, single Exchange node, appliance head
- No storage failover required, even in case of failure of both Exchange nodes
- No/minimal disruption to service availability for any of above failure conditions*
- Failover to the DR site within minutes after site-wide failure
 - No loss of data
 - Easy failover operation aided by server cluster and storage cluster
- Seamless backup and recovery across cluster resources with SME
- Support of active-active configurations—fully utilized HW at both sites
- Minimal performance degradation during and after recovery

MetroCluster works with other host clustering software such as Microsoft® Cluster Server, Veritas™ Cluster Server, and SunCluster.

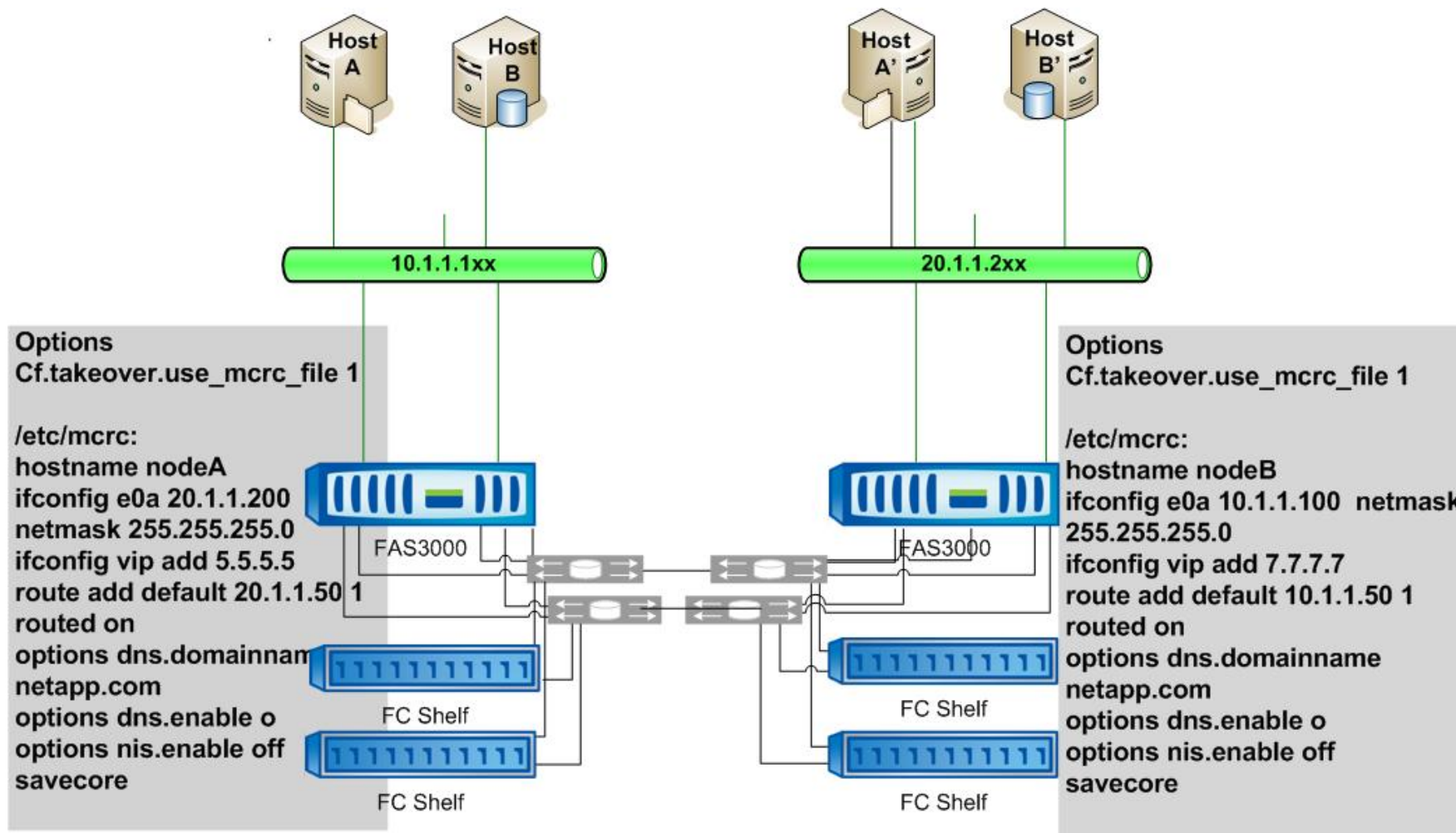


Recent Additions (7.3.2)

- Up to 672 spindles on Fabric MC
- Multiple Subnet Support
- Nondisruptive disk shelf replacement
- Nondisruptive upgrades



Multiple Subnet Support





What's New?

- MetroCluster TieBreaker Solution
- Vmware Certification with HA
- Vmware Certification with FT

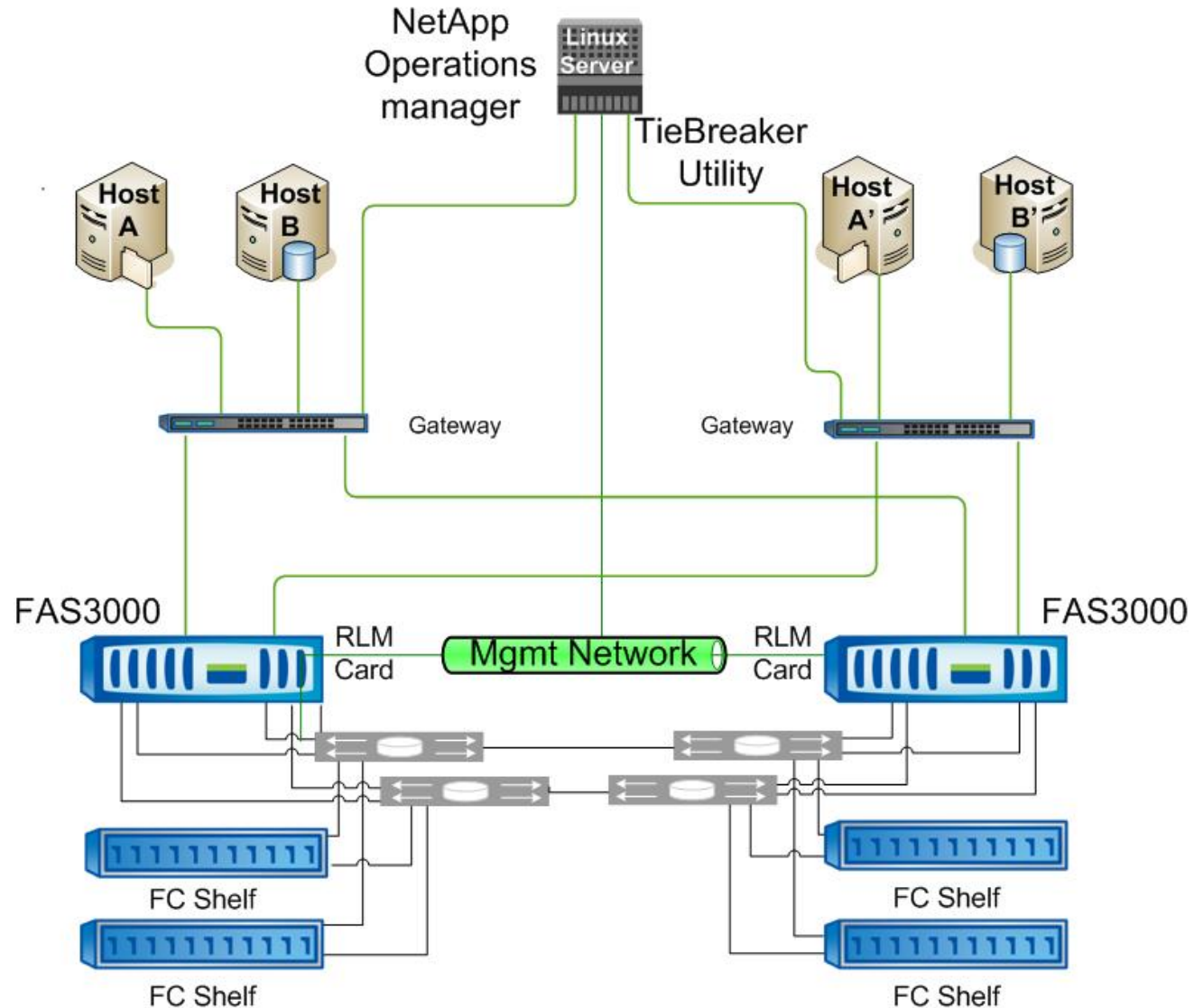


MetroCluster Tiebreaker Solution

- Automates site failover (cfod)
- Developed by Professional Services and RRE
- Supported via RRE
- Available via PVR only
- Not required (or necessary) in most instances
- Requirements
 - Operations Manager
 - Perl installed on Ops Mgr Server
 - RLM (optional for fencing)

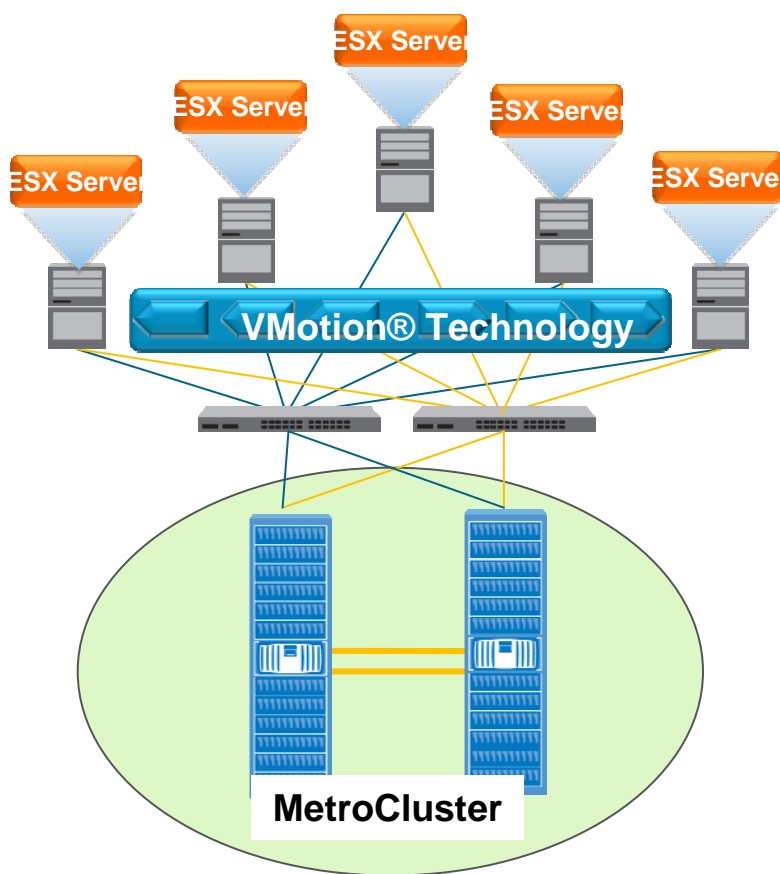


TieBreaker Diagram





MetroCluster for VMware Environments



- Delivers continuous data availability
 - Complements VMware® host environment with complete server to storage availability
- Improves read performance
 - Interleave reads from mirrored arrays
- Simplifies operations
 - Zero interdependencies
- Deploys with confidence
 - Tested interoperability
- Leverages “Virtualization Guarantee Program”



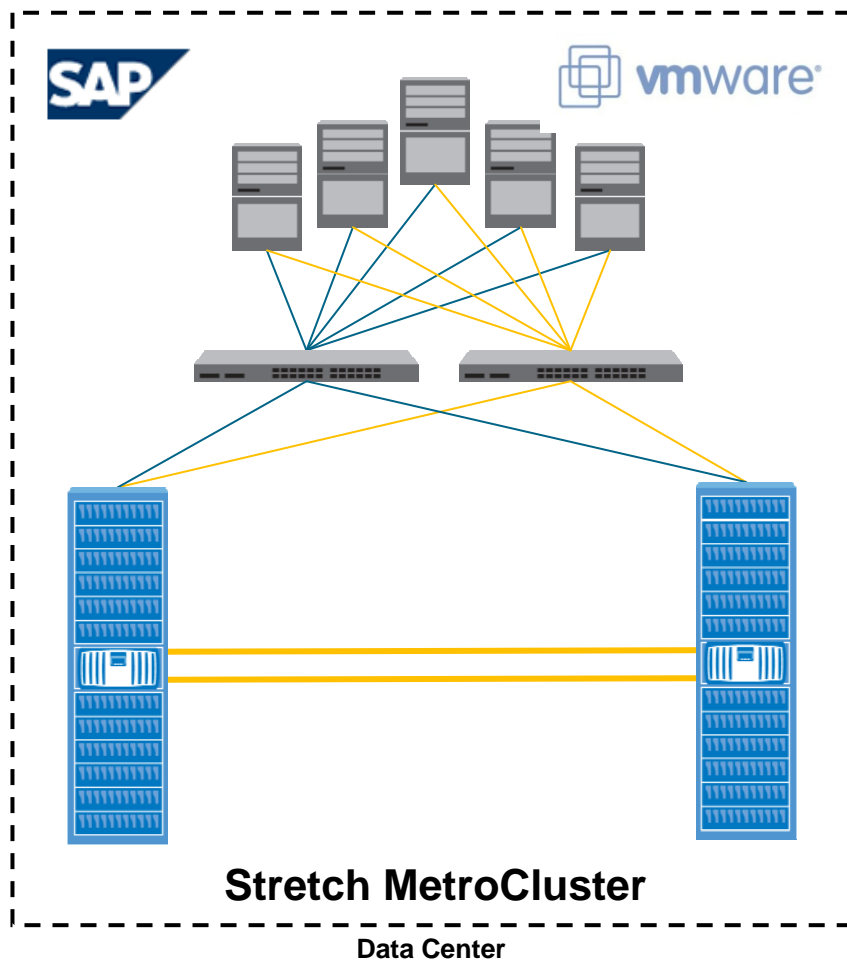
MetroCluster – Reference Stories

- Lee County Court of Clerks – Public Sector, N.A.
 - Full Story: <http://media.netapp.com/documents/LeeCounty.pdf>

- United Health Network – Healthcare, N.A.
 - Full Story: <http://media.netapp.com/documents/UHN.pdf>

- Bitburger Brewery Group – Retail, Germany
 - Full story: <http://www.netapp.com/us/library/customer-stories/bitburger.html>

- New customer stories coming soon



Challenges

- Business requiring 24x7 operations
- Risk of downtime was unacceptable
- Replace its ERP system based on Oracle® with an ERP infrastructure based on SAP®
- Implement a VMware® server virtualization solution

Solution

- FAS6080 and FAS3140 storage systems
- Located at opposite ends of the data center and configured as a stretch MetroCluster
- Provides Lexmark with a local campus disaster recovery and high availability solution
- Deduplicate data from Oracle
- Initial tests have shown a 60% space savings



Proposing a MetroCluster

- Information to be gathered
 - Distance between locations
 - Number of disks – consider mirroring, growth
- New system or upgrade?
- Upgrade considerations
 - Hardware or software ownership
 - Existing data
 - 31xx issues



MetroCluster Positioning

- Qualification questions
 - Is continuous data availability important to one or more of your product applications?
 - Would providing geographic protection be of value?
 - What distance between controllers would provide the best protection (building, floor campus, metro, or regional)?
 - If you are employing an HA solution for a virtualized environment, why not have the same resilience at the storage level?



SnapMirror Sync versus MetroCluster

Feature	SnapMirror Sync	MetroCluster
Replication Network	IP or FC	FC only
Limit on concurrent transfers	Yes	No limit
Distance limitation	Up to 200 KM (latency driven)	100 KM
Replication between HA pairs	Yes	No
Manageability	FilerView, CLI, Ops Mgr	CLI
Multi-use replica	Yes (FlexClone) No (dump)	Limited (app can read from both plexes)
Deduplication	Dedupe volume and sync volume cannot be in the same aggregate	No restrictions
Use of secondary node for async mirroring	Yes	No. Async replication has to occur from primary plex



Best Practices: MetroCluster

- Choose the right controller
 - Based on normal sizing tools and methods
- Connections are important
 - Calculate distances properly
 - Use the right fiber for the job
 - Verify correct SFPs, cables required, and patch panels
- Remember mirroring requires 2x disks
- Remember MetroCluster spindle maximums
- Remember speed restrictions



References

- **TR3548 - MetroCluster Design and Implementation Guide**
 - <http://media.netapp.com/documents/tr3548.pdf>
- **TR3517 – MetroCluster Upgrade Planning Guide**
 - <http://media.netapp.com/documents/tr-3517.pdf>
- **TR3552 – Optical Network Installation Guide**
 - <http://media.netapp.com/documents/tr-3552.pdf>
- **TR3606 - High Availability and Disaster Recovery for VMware using NetApp SnapMirror and MetroCluster**
 - <http://media.netapp.com/documents/tr-3606.pdf>
- **Data ONTAP Active/Active Configuration Guide**
 - <http://Now.netapp.com>
- **MetroCluster Compatibility Matrix**
 - http://now.netapp.com/NOW/knowledge/docs/olio/guides/metrocluster_compatibility/



Value of MetroCluster

■ Simple

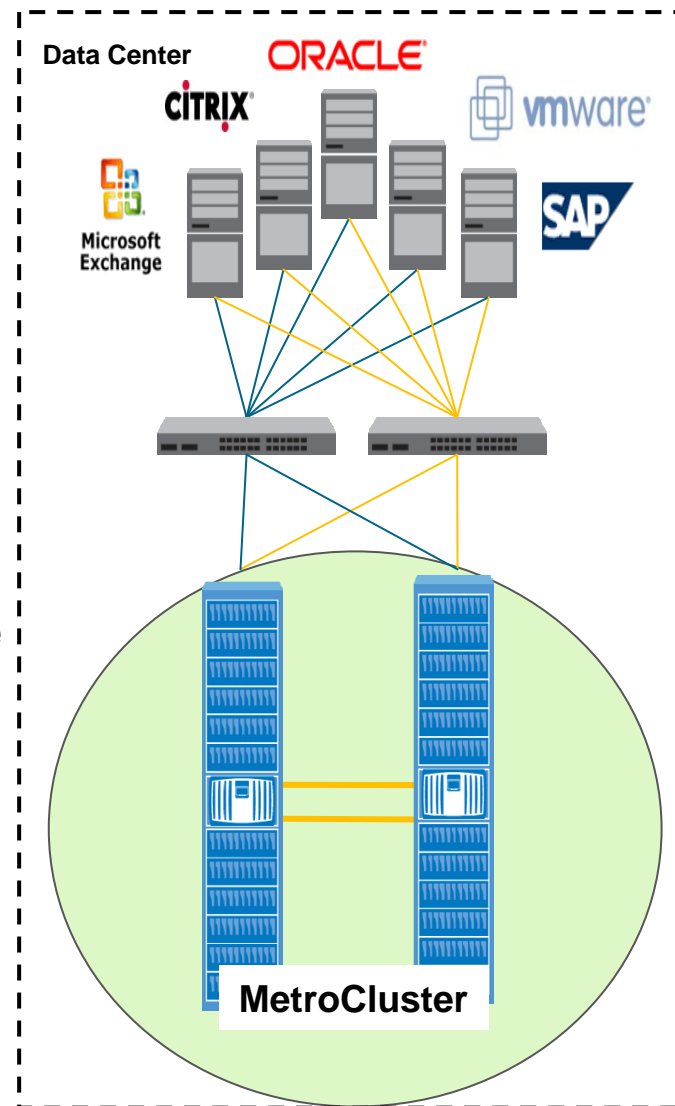
- **Set it Once:** Simple to deploy, with no scripts or dependencies on the application or operating system
- **Zero change management** through automatic mirroring of changes to user, configuration, and application data

■ Continuous Availability

- **Zero unplanned downtime** through transparent failover with protection for hardware plus network and environmental faults
- **Zero planned downtime** through nondisruptive upgrades for storage hardware and software

■ Cost Effective

- **50% lower cost**, no host-based clustering required, added efficiency through data deduplication and server virtualization





NetApp®

Go further, faster®

Thank You

© Copyright 2009 NetApp, Inc. All rights reserved. No portions of this document may be reproduced without prior written consent of NetApp, Inc. NetApp, the NetApp logo, Go further, faster, Data ONTAP, NOW, RAID-DP, SnapMirror, and SyncMirror are trademarks or registered trademarks of NetApp, Inc. in the United States and/or other countries. Microsoft is a registered trademark of Microsoft Corporation. Oracle is a registered trademark of Oracle Corporation. Veritas is a trademark of Symantec Corporation or its affiliates in the U.S. and other countries. SAP is a registered trademark of SAP AG. VMware and VMotion are registered trademarks of VMware, Inc. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.

