Continuous Availability for Business-Critical Data

MetroCluster Technical Overview

Jean-Philippe Schwartz
Storage Systems Specialist - IBM / Gate Informatic

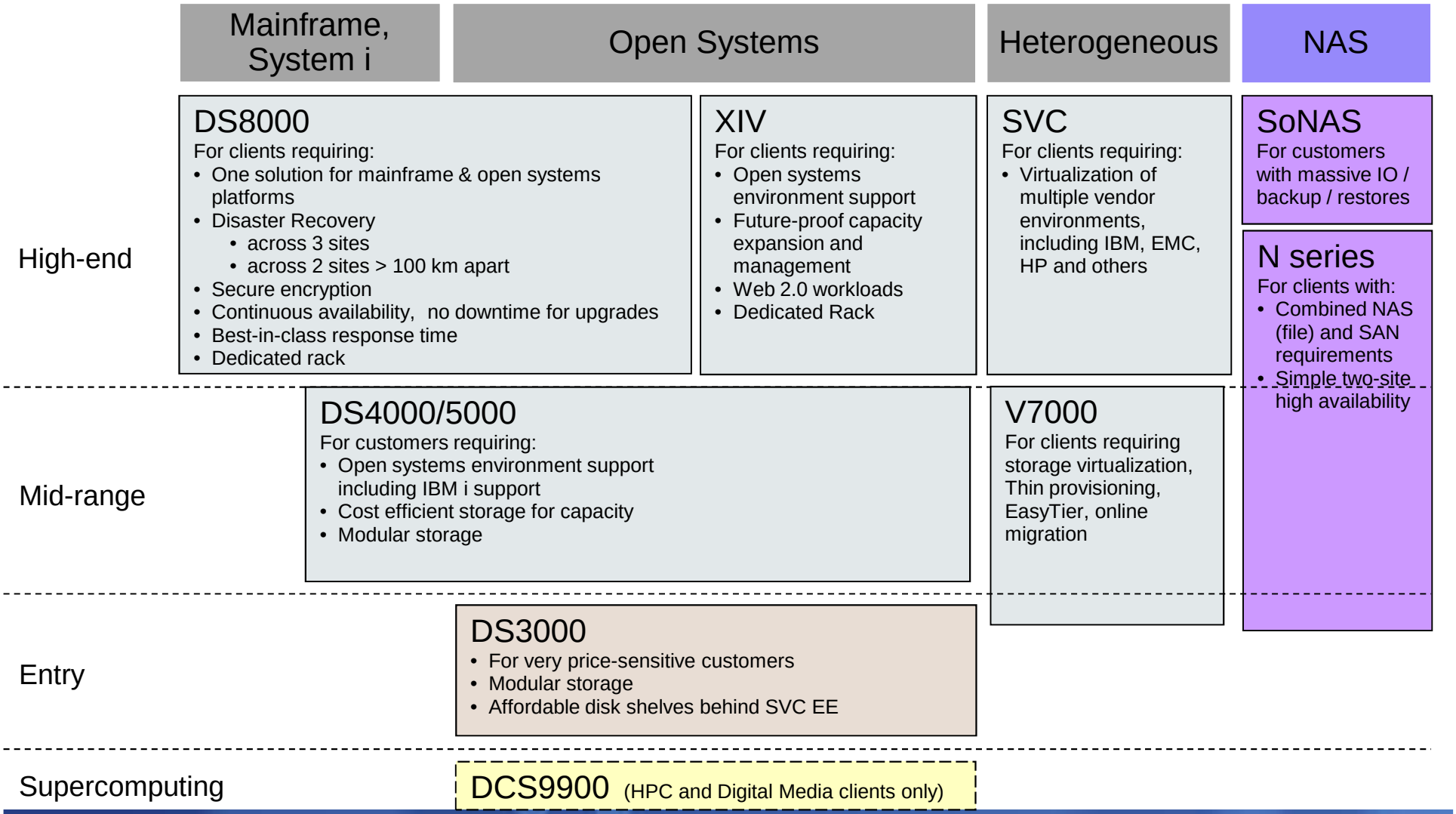
jesc@ch.ibm.com

Agenda

- NetApp / IBM Agreement
- Customer Challenges
- NetApp Business Continuity Solutions
- NetApp MetroCluster Overview
- MetroCluster Failure Scenarios
- When do I need MetroCluster
- IBM MetroCluster in SR
- One Pager

NetApp / IBM Agreement

Why NetApp ?



IBM and Network Appliance Join Forces

- Previous attempts by IBM to penetrate the NAS market in 2001-2004 have shown limited success (NAS100, NAS200, NAS300, NAS300G then later on NAS500G)
- In April, 2005, IBM and Network Appliance announced a strategic storage relationship to drive information on demand solutions and to expand IBM's portfolio of storage solutions:
  - IBM sells re-branded Network Appliance storage solutions as N series
  - Network Appliance extends its market presence with key partner
- IBM does additional testing on specific solutions to ensure the highest possible Availability, Reliability & Scalability

IBM N series Offer

For Customer NAS only or Unified Storage Requirements
Same HW and DOT as NetApp Platforms
Full range of NetApp Software Functions
Filer and Gateway Models

Highly-scalable storage systems for large enterprise data centers.

N7000/Gateway Series

Up to 1176TB

Excellent performance, flexibility and scalability at a lower TCO

N6000/Gateway Series

Up to 1.2PB

Entry level pricing, Enterprise Class Performance

N3000 Series

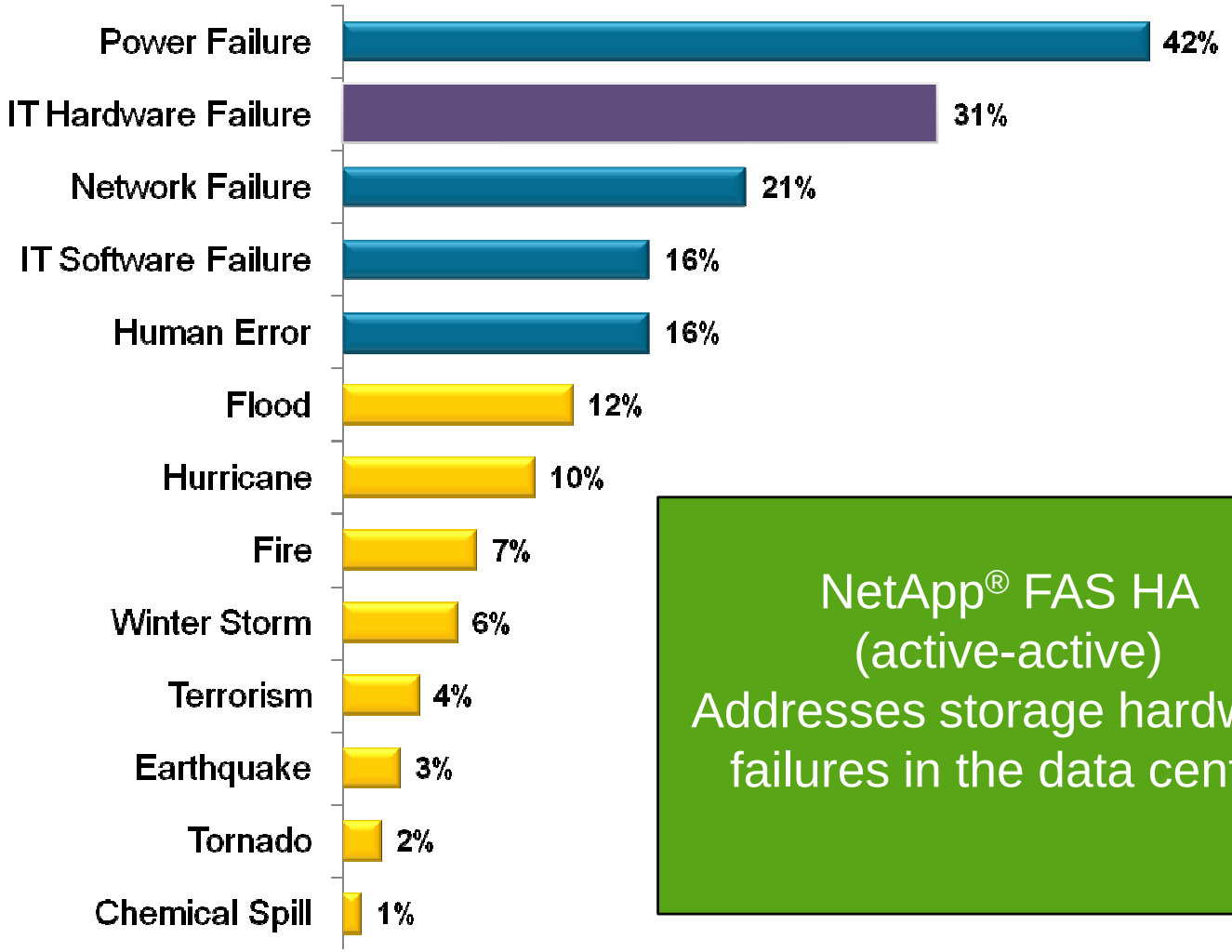
Up to 136TB

Customer Challenges

Customer Challenges

Internal Data Center Failures

External Data Center Failures

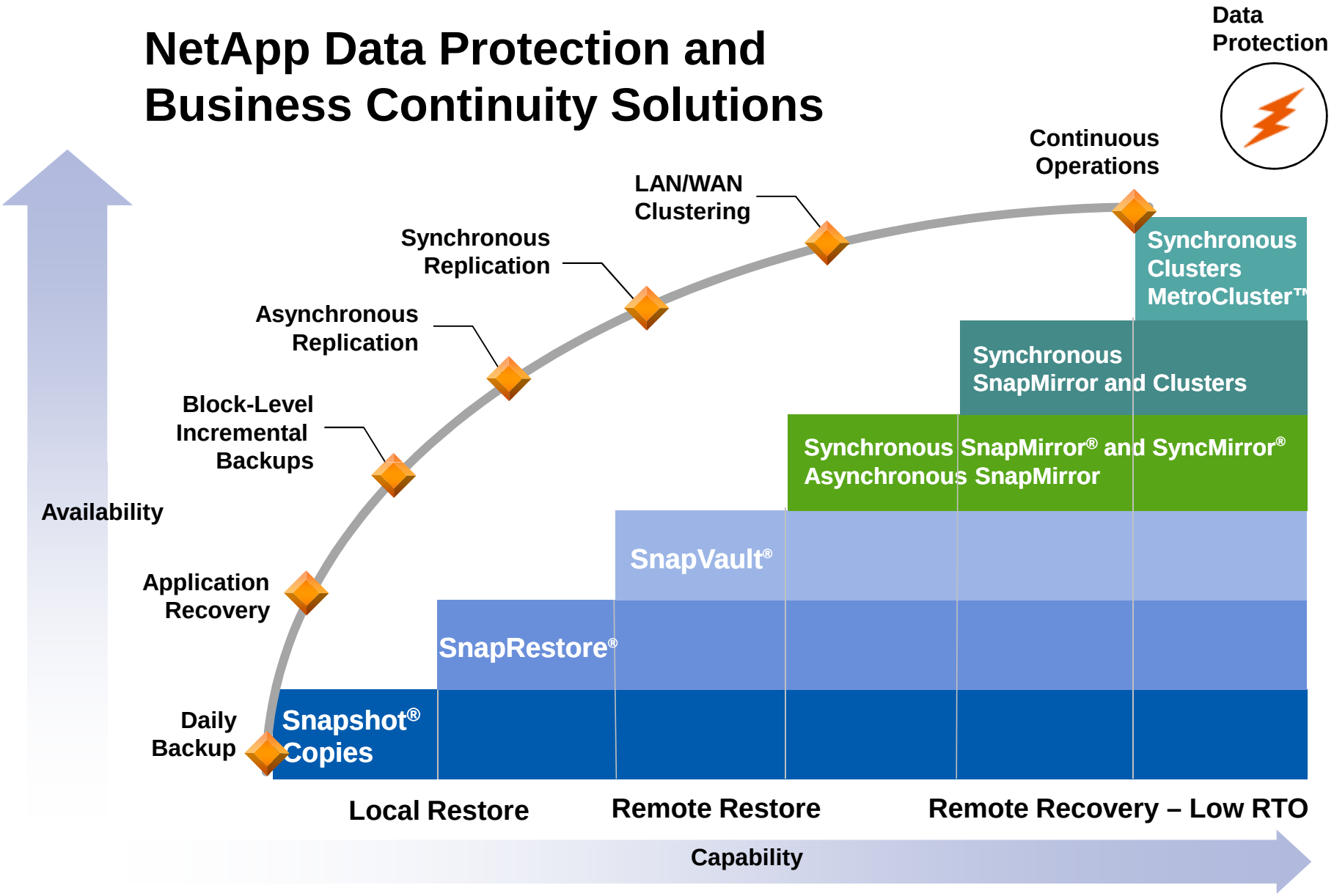


NetApp® FAS HA (active-active) Addresses storage hardware failures in the data center

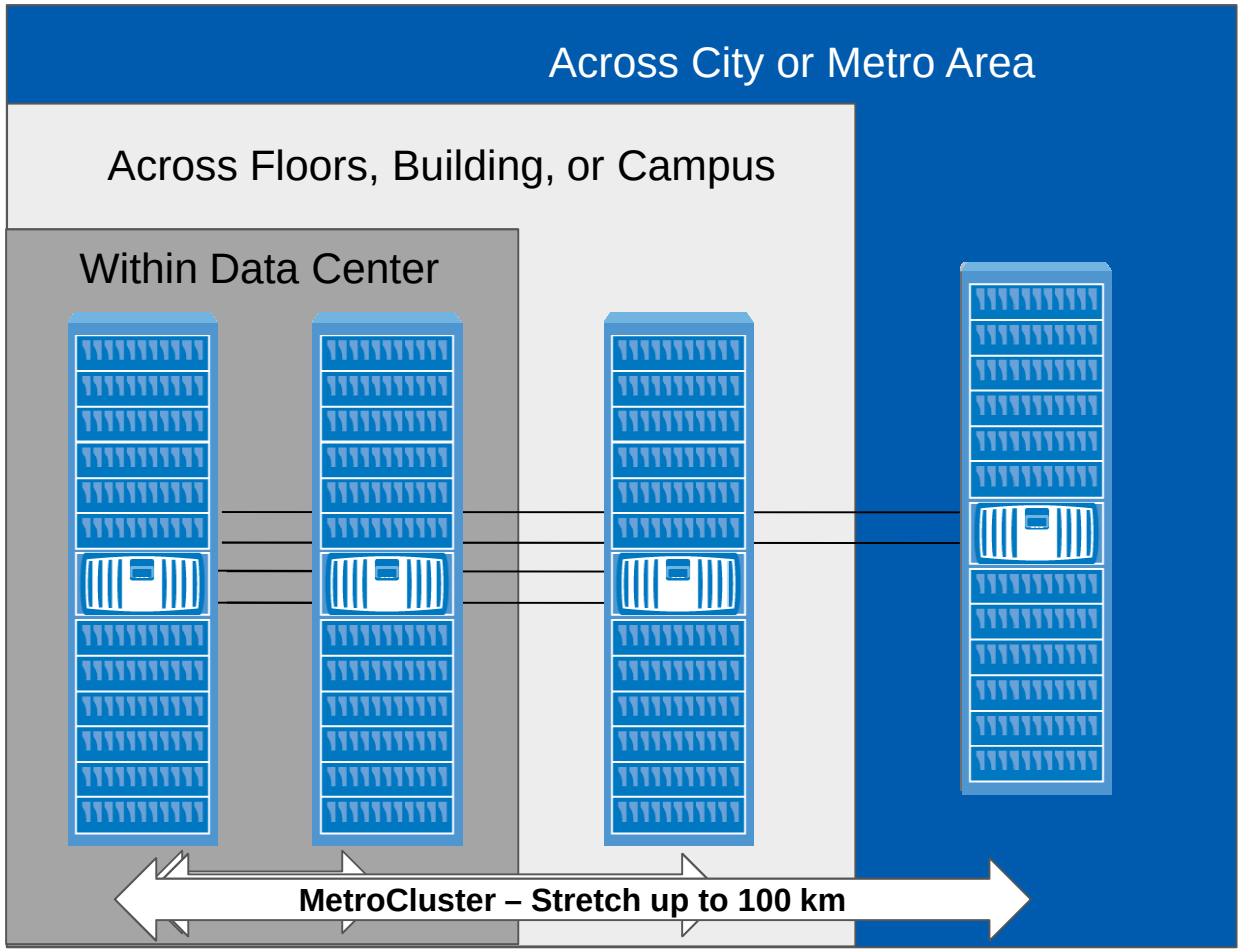
Source: Forester / Disaster Recovery Journal: Global Disaster Recovery Preparedness Online Survey, Oct, 2007

NetApp Business Continuity Solutions

NetApp Data Protection and Business Continuity Solutions

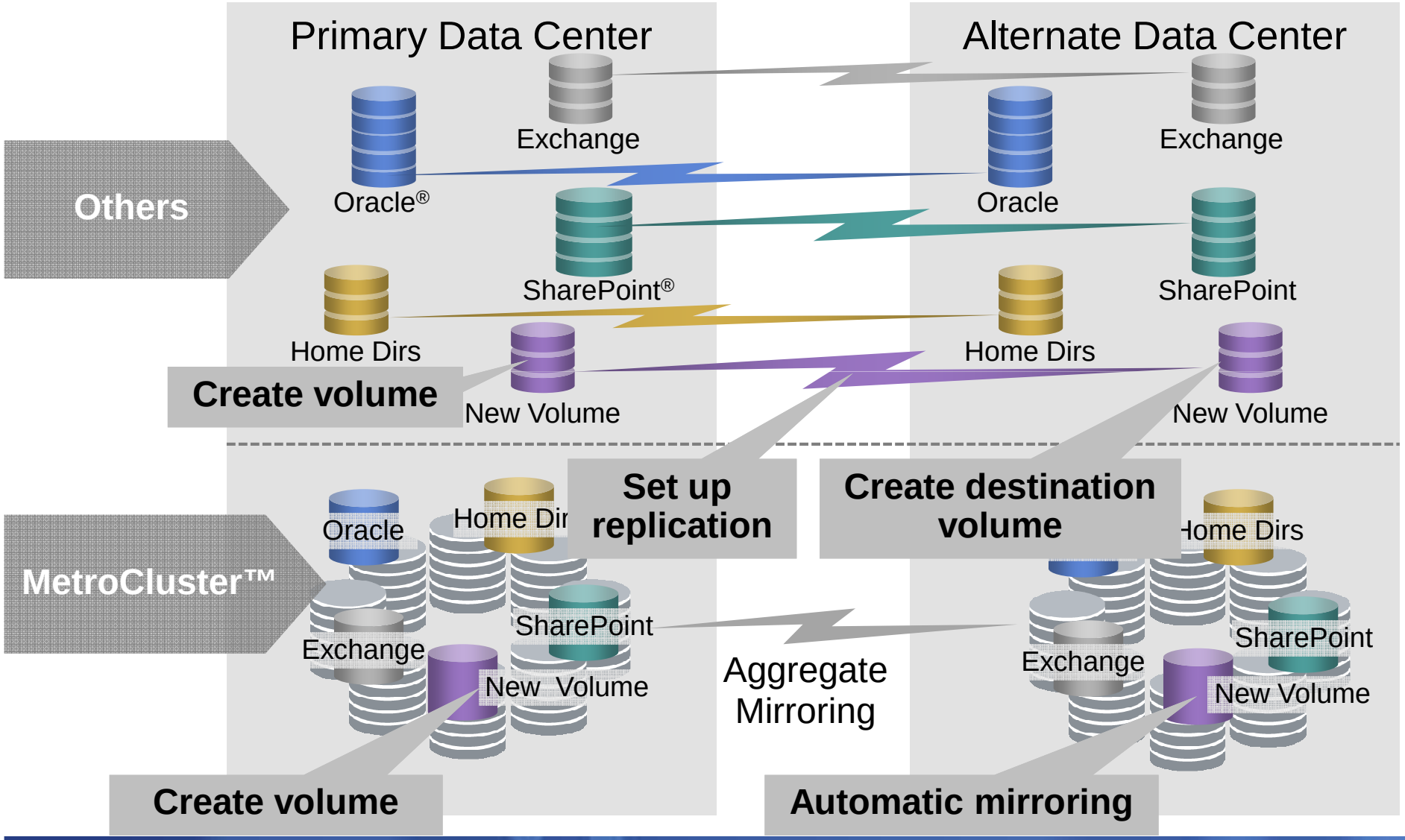


The MetroCluster Difference

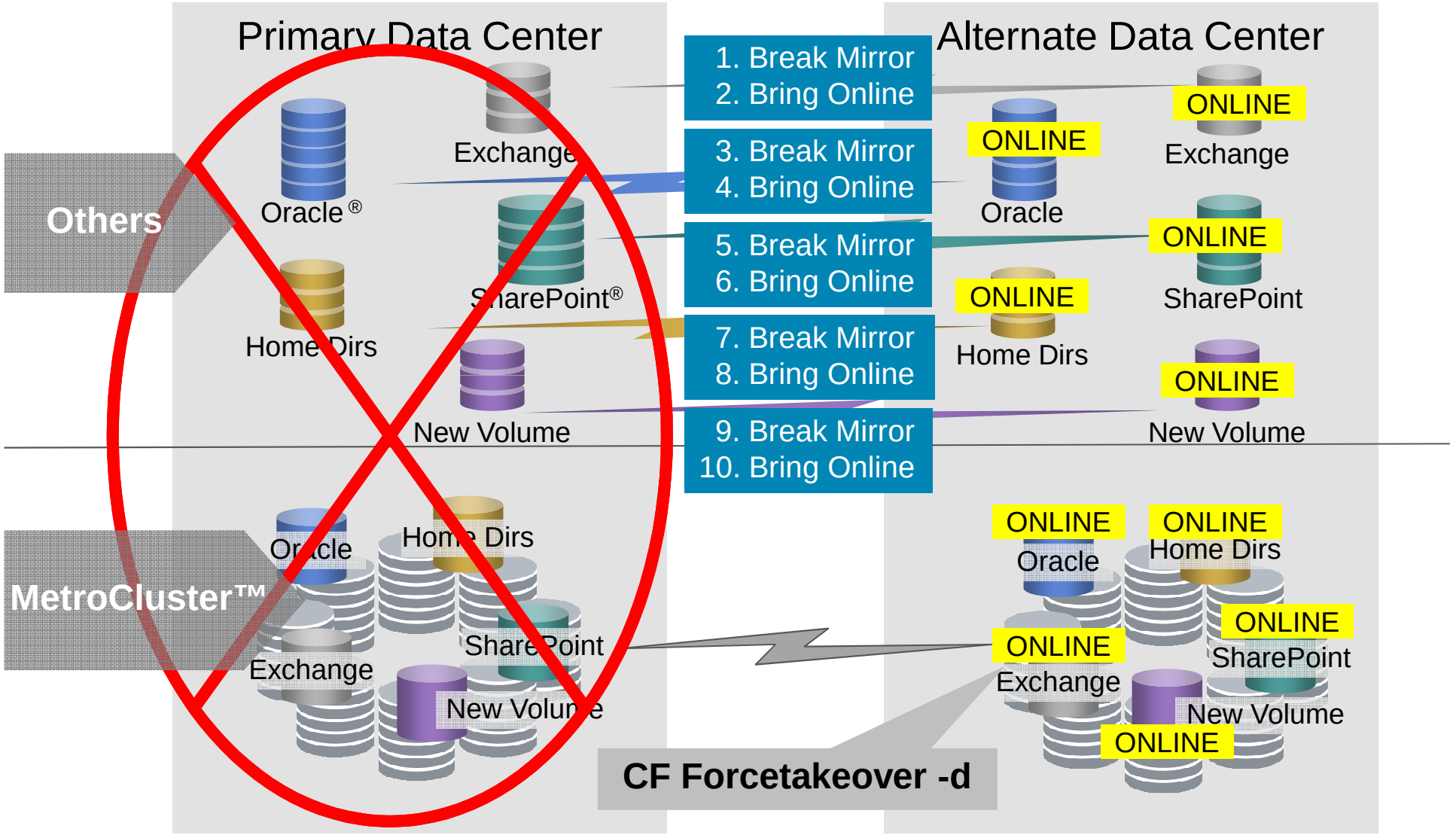


- NetApp® MetroCluster™ provides continuous availability within a single data center and across data centers in adjacent floors, buildings, and metro areas
- Support for N series N6000 & N7000

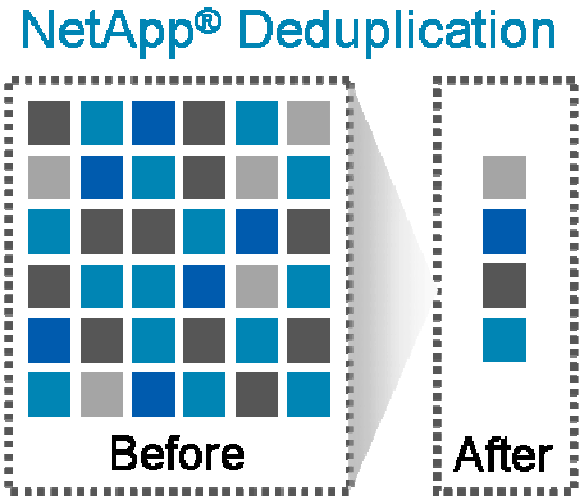
The MetroCluster Difference: Automatic Protection



The MetroCluster Difference: Quick Recovery Time



NetApp Efficiency



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

MetroCluster Overview

MetroCluster Components

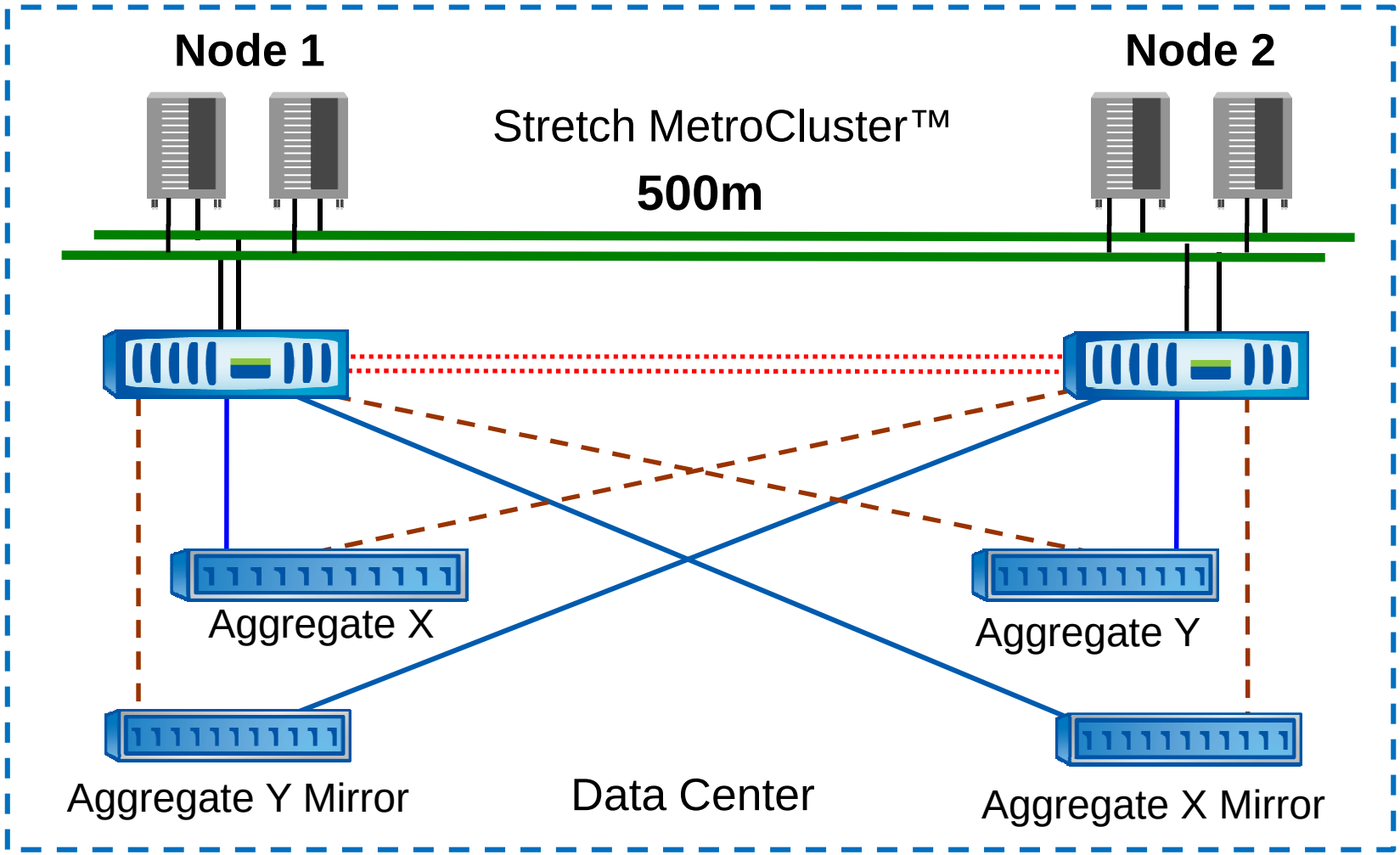
Cluster License Provides automatic failover capability between sites in case of most hardware failures.

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

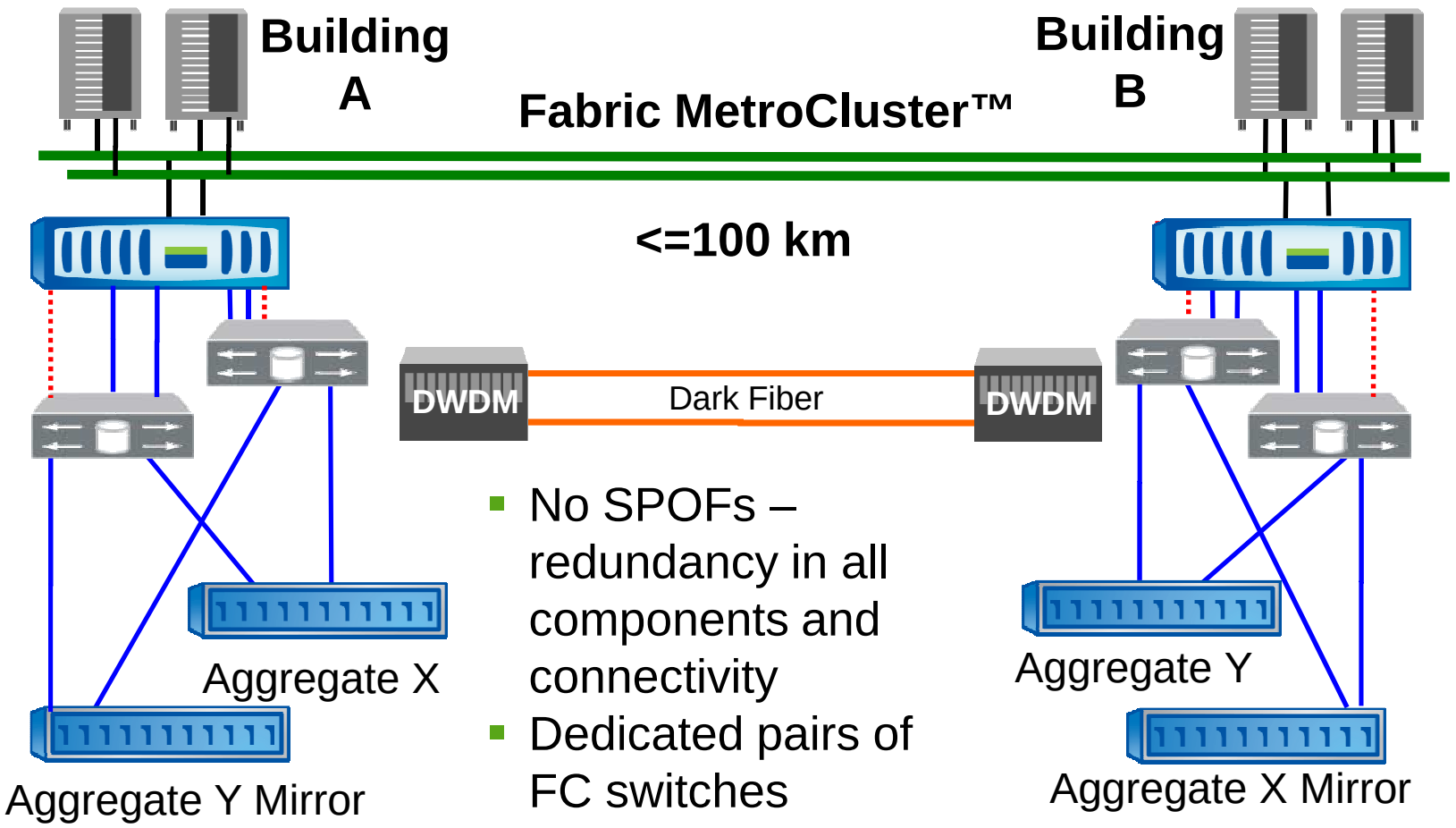
Cluster Remote Provides a mechanism for the administrator to declare a site disaster and initiate a site failover with a single command.

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

MetroCluster Deployment in a Single Data Center, Across Floors, and So On



MetroCluster Deployment Across City or Metro Areas

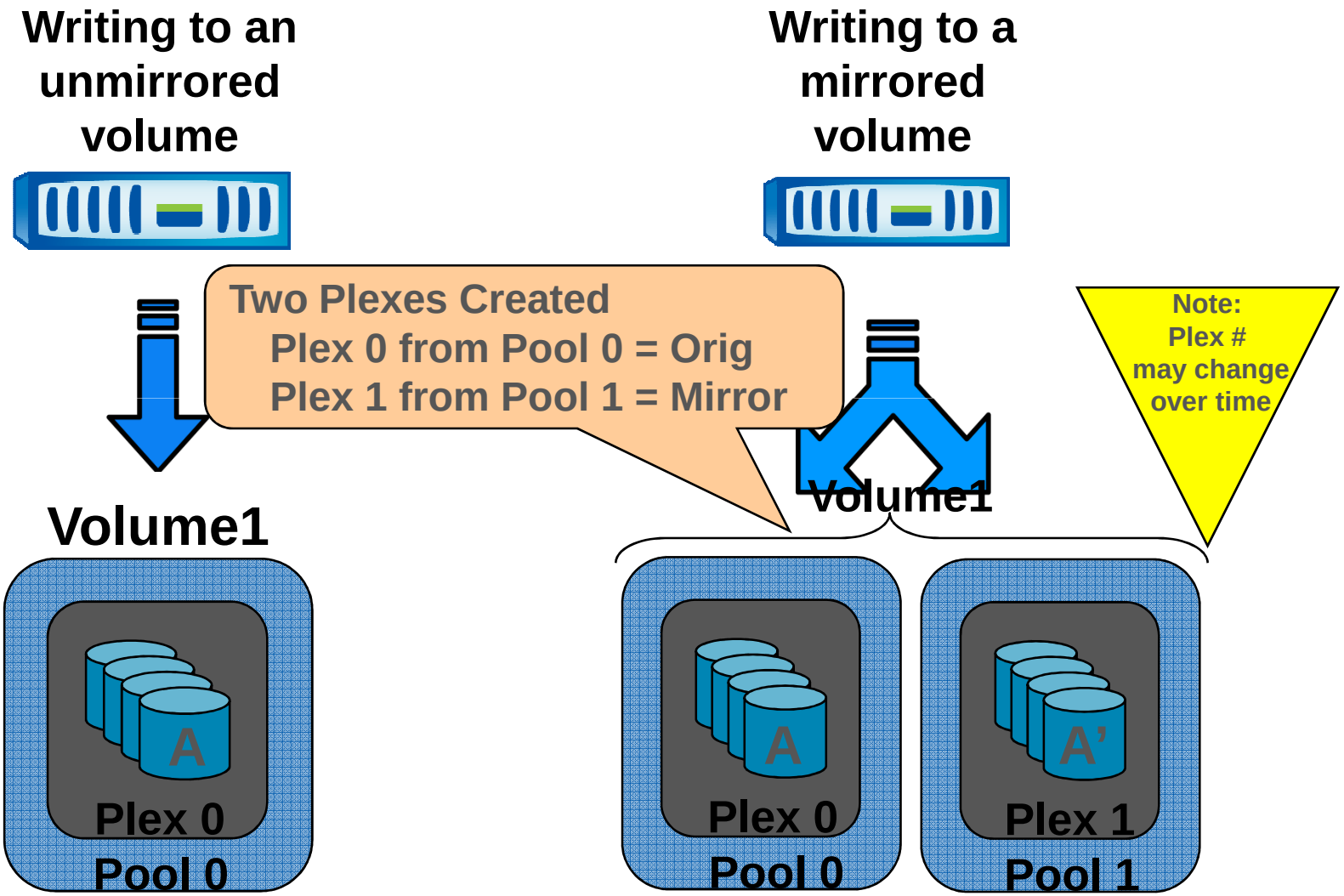


- No SPOFs – redundancy in all components and connectivity
- Dedicated pairs of FC switches

MetroCluster Component: Cluster Remote License

- Provides single-command failover.
- Provides the ability to have one node take over its partner's identity without a 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*)

MetroCluster Component: SyncMirror

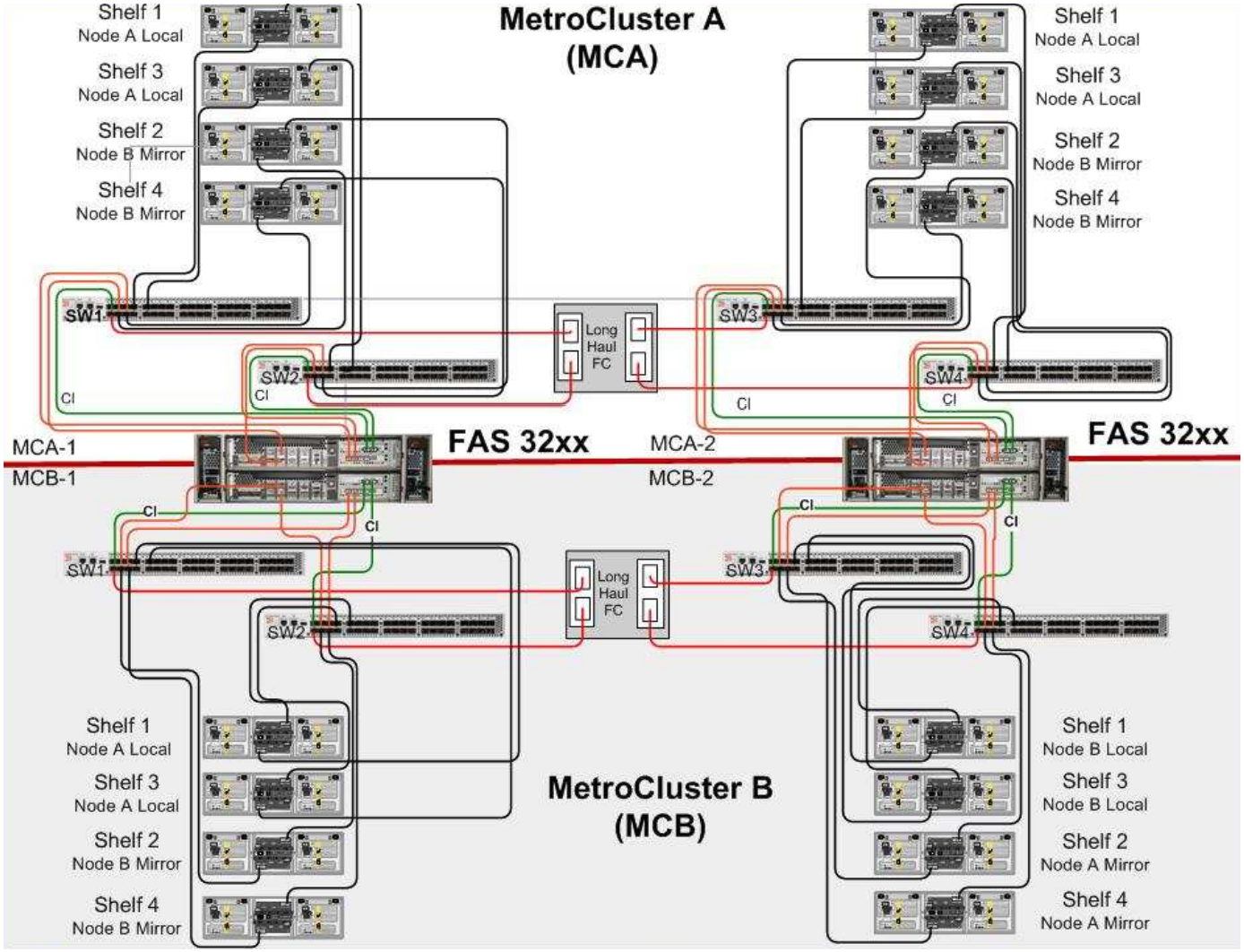


MetroCluster Support Matrix

	Stretch		Fabric	
FAS/V Model	Supported	Max. Disks	Supported	Max. Disks
3070	Yes	504	Yes	504[1]
3140	Yes	420	Yes	420
3170	Yes	840	Yes	840[2,3]
3210	Yes	480	Yes	480
3240	Yes	600	Yes	600
3270	Yes	960	Yes	840[2]
6030	Yes	840	Yes	840[2,3]
6040	Yes	840	Yes	840[2,3]
6070	Yes	1008	Yes	840[2,3]
6080	Yes	1166	Yes	840[2,3]
6210	Yes	1008	Yes	840[2,4]
6240	Yes	1176	Yes	840[2,4]
6280	Yes	1176	Yes	840[2,4]

[1] 504 spindle count requires Data ONTAP® 7.2.4 or later
[2] 672 spindle count requires Data ONTAP 7.3.2 to 7.3.4
[3] 840 spindle count requires Data ONTAP 7.3.5 or later
[4] 840 spindle count requires Data ONTAP 8.0.1 or later

“Twin” MetroCluster Configuration (DOT 7.3.3)



New in Data ONTAP 7.3.5/8.0.1

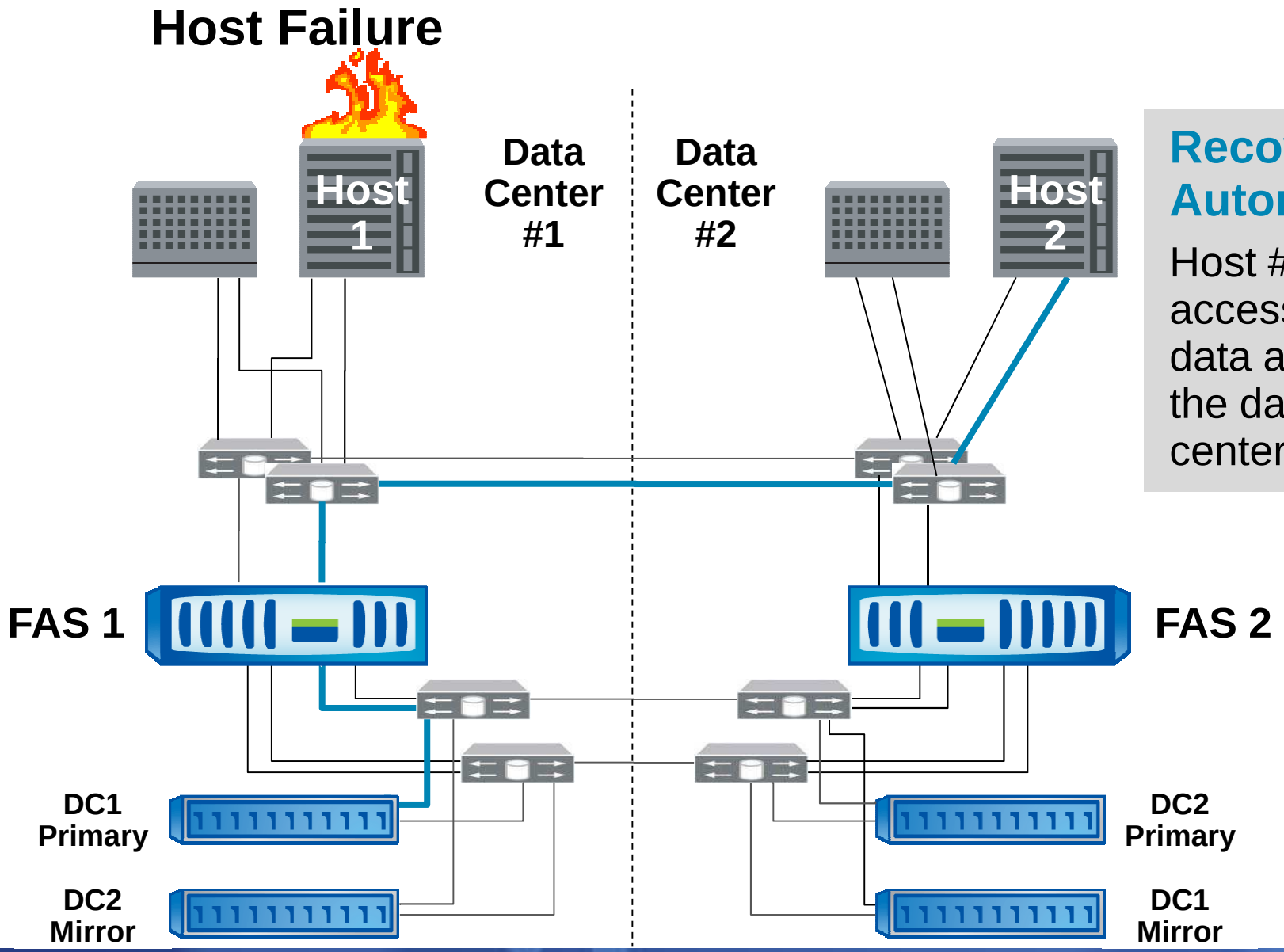
FAS 32xx

8Gb FCVI
X1927A-R6

- FAS32xx series support
- FAS62xx series Support
- Up to 840 spindles on fabric MetroCluster™
- 8Gb cluster interconnect

MetroCluster Failure Scenarios

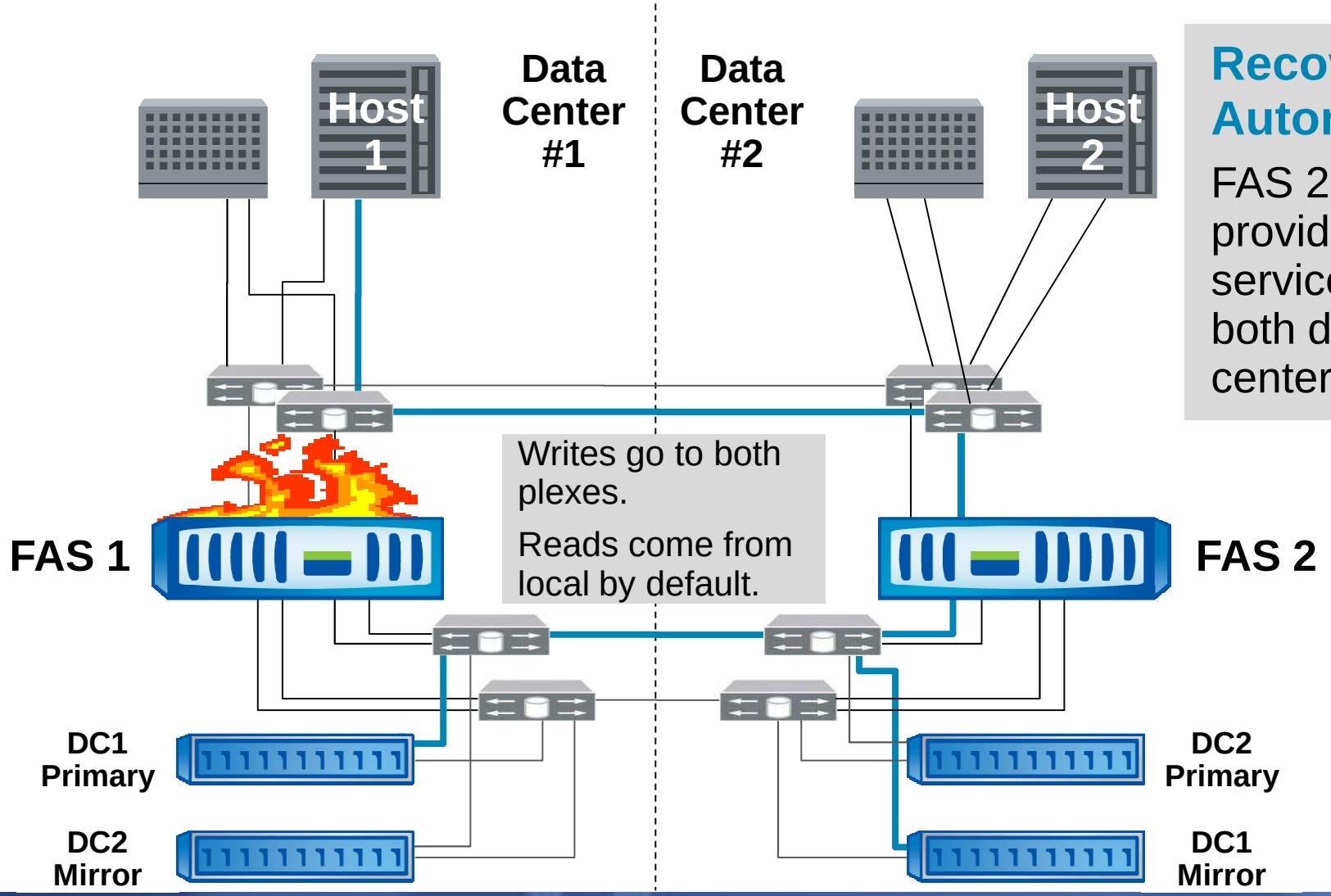
Host Failure



Recovery – Automatic

Host #2 accesses its data across the data centers

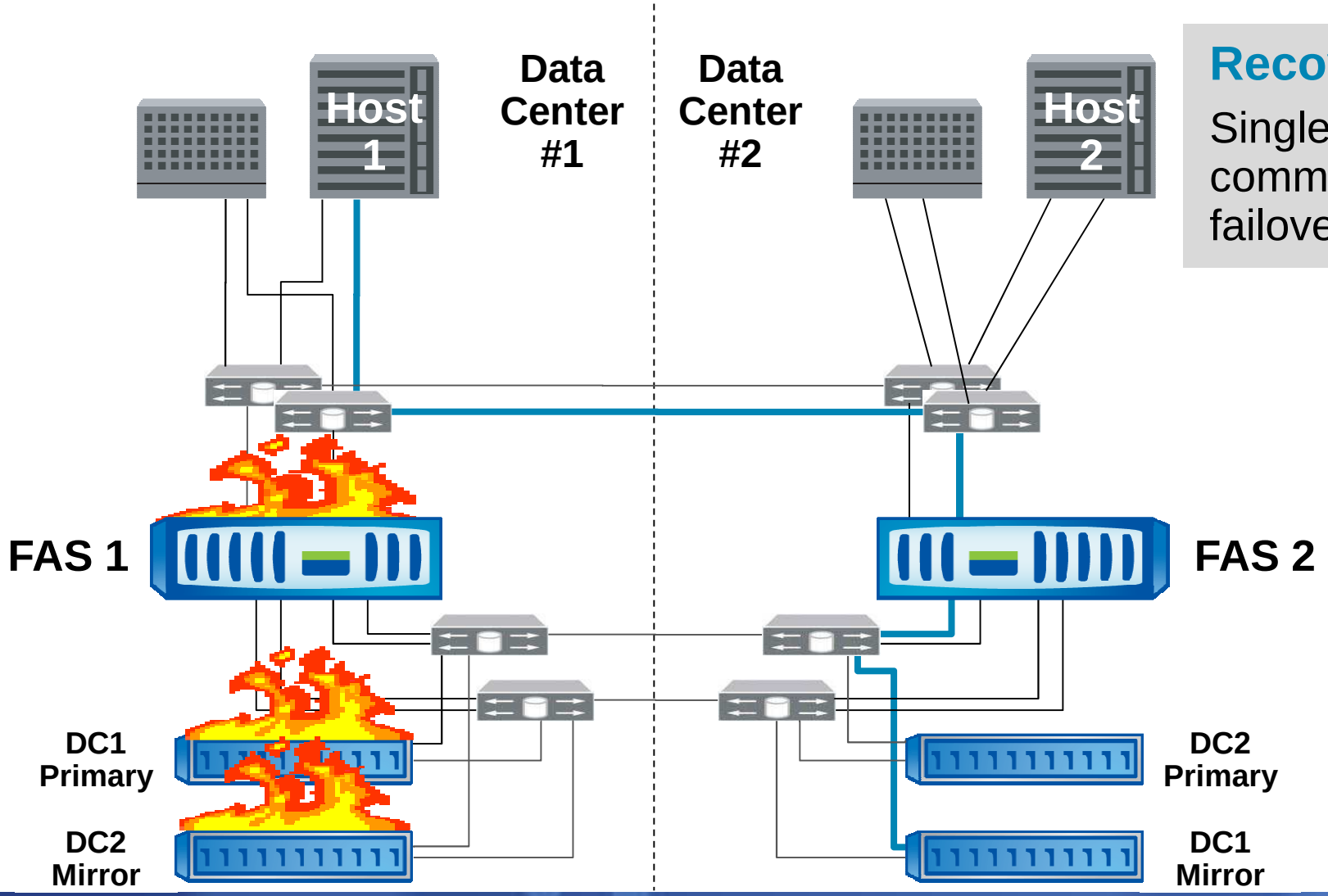
Storage Controller Failure



Recovery – Automatic

FAS 2 provides data services for both data centers.

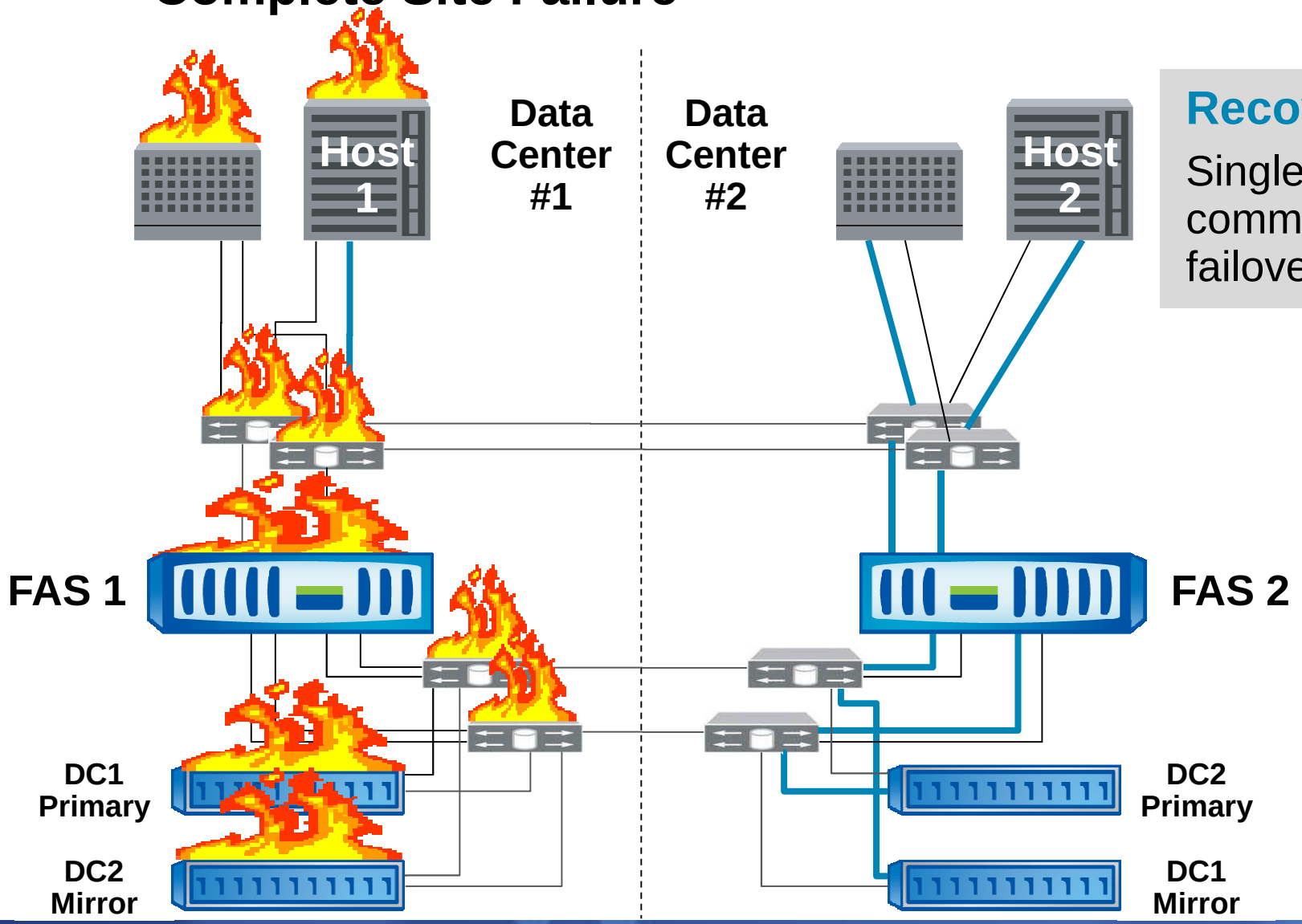
Storage Array Failure



Recovery

Single-command failover

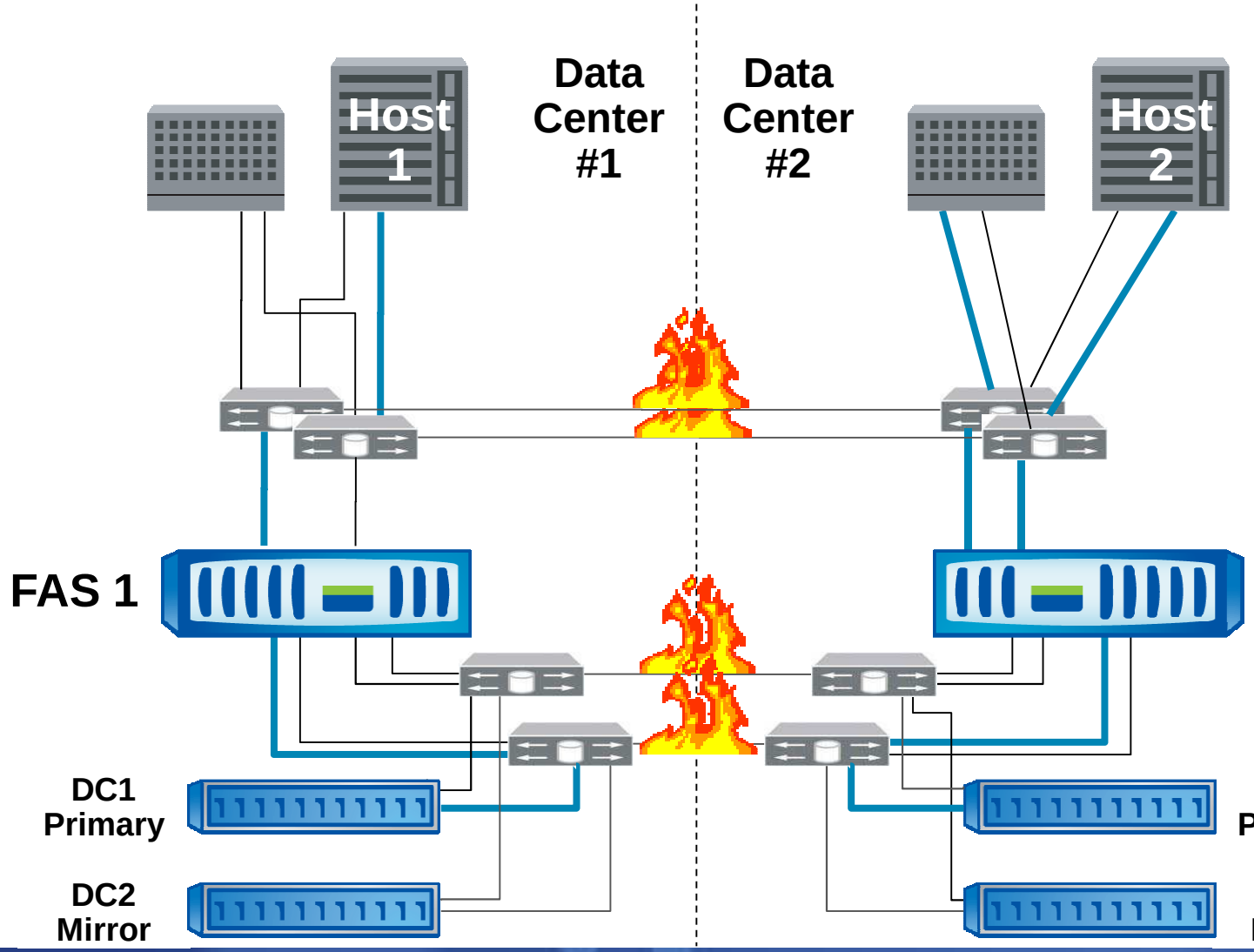
Complete Site Failure



Recovery

Single-command failover

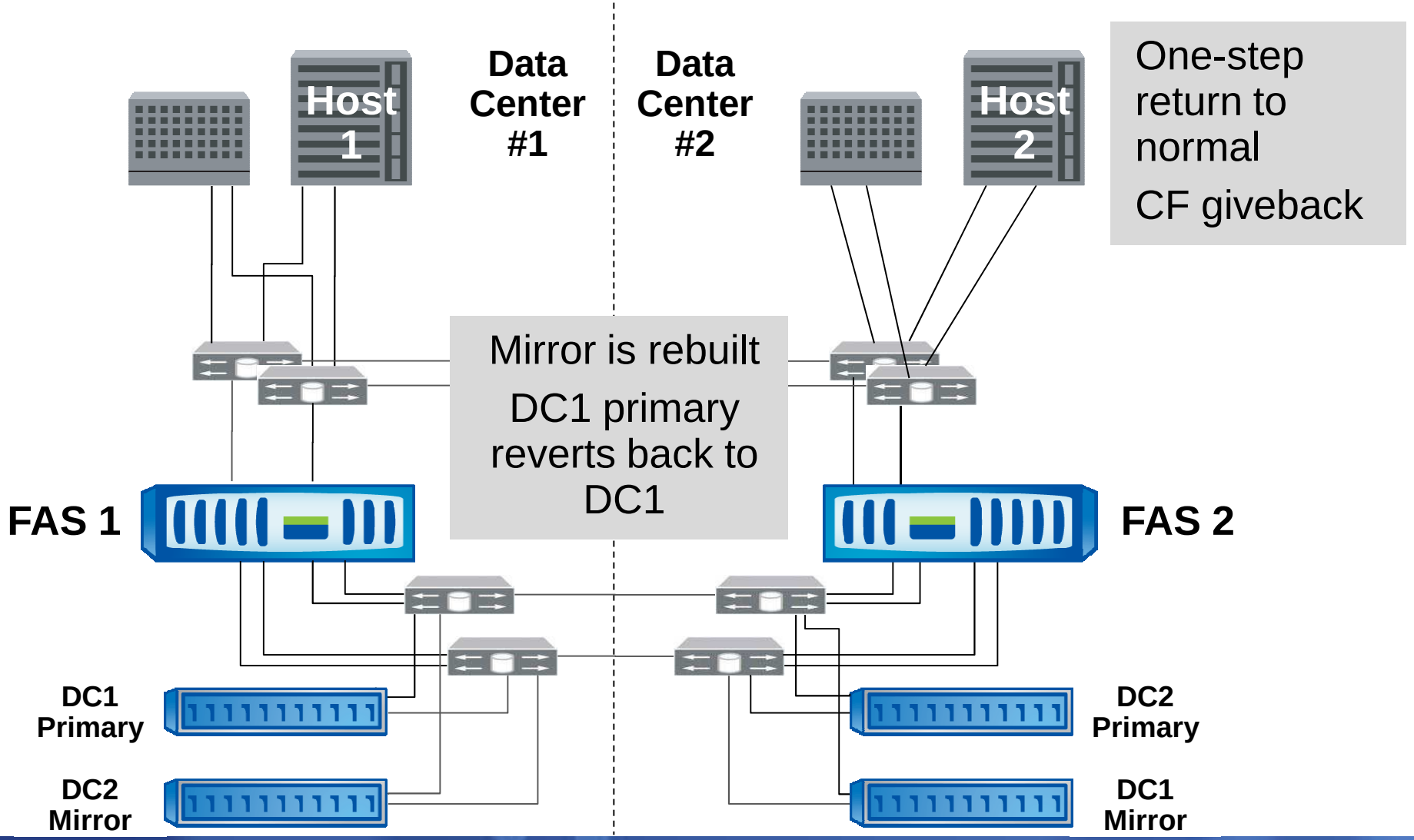
Interconnect Failure



Recovery

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

Return to Normal



When do I need MetroCluster

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 200km (latency driven)	100km
Replication between HA pairs	Yes	No
Manageability	FilerView®, CLI, Operations Manager	CLI
Multi-use replica	Yes (FlexClone®) No (dump)	Limited (app can read from both plexes)
Deduplication	Deduplicated 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

MetroCluster Best Practices

- 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 that mirroring requires 2x disks
- Remember MetroCluster™ spindle maximums

IBM MetroCluster in SR

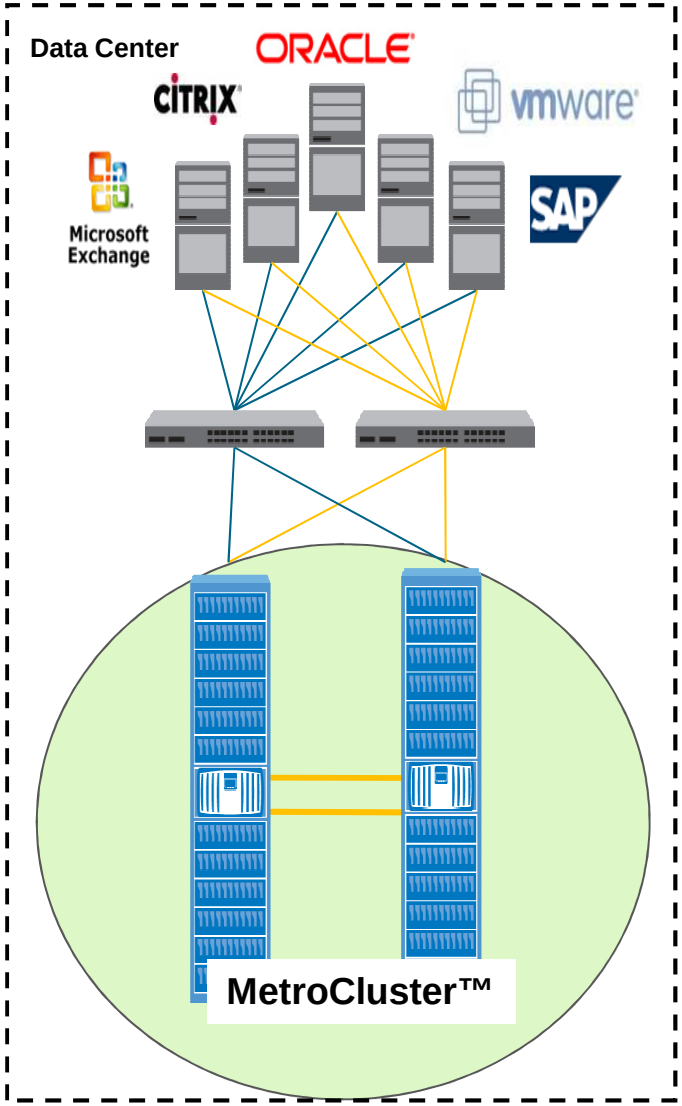
IBM MetroCluster Installations in Suisse Romande

- Total of 5 IBM MetroCluster Customers in Suisse Romande
- 3 Stretch Configurations
  - 1 x Across Rooms
  - 1 x Across Floors
  - 1 x Across Roads
- 2 Fabric Configurations
  - 1 x Owned LW Connections (1.2 km)
  - 1 x External Provider Service (30+ / 50+ km)
- Extremely High Customer Satisfaction

One Pager

Value of MetroCluster

- Simple
  - **Set it once:** Easy 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



Thank You