Achieving Continuous Availability

Configuring Server Blades for Fault and Disaster Tolerance

Jerry Melnick
VP Engineering
Marathon Technologies Corporation

Simple, Affordable Uptime

Why is Availability Important?

Downtime is Costly

- Lost Revenue
- Reduced Employee Productivity
- Disrupted Operations
- Damaged Public Reputation
- Drained IT Resources
- Damaged Customer Satisfaction
- Lowered Financial Performance
- Regulatory Fines

Cost of Downtime by Industry

Business Operation	$Rev./hour
Manufacturing	$1,610,000
Finance	$1,495,000
Retail	$1,107,000
Pharmaceuticals	$1,082,000
Healthcare	$ 636,000
Utilities	$ 643,000

Source: META Group

Cost of Downtime by Industry

Business Operation	$Rev./hour
Utilities	$ 643,000

Source: META Group

"In a recent survey, 30% of companies did not have a disaster recovery program in place and 2 in 3 companies felt that their data backup and disaster recovery plans had significant vulnerabilities."

Source: Imation Data Protection Survey (August, 2003)

Cost of Downtime by Industry

Business Operation	$Rev./hour
Manufacturing	$1,610,000
Finance	$1,495,000
Retail	$1,107,000
Pharmaceuticals	$1,082,000
Healthcare	$ 636,000
Utilities	$ 643,000

"30 to 40 percent of businesses never reopen after a disaster"

Source: Dept. of Homeland Security, Federal Emergency Management Agency

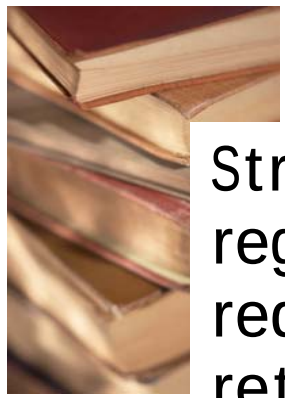
Source: META Group

Trend Toward Continuous Availability

Dependence on existing automation and increasing number of systems automated

More Threats to Security and Physical Systems

Server Density & Consolidation (Blades)



Stringent regulations requiring data retention and availability

Inhibitors to Adopting Availability Solutions

- Increased complexity
- Unclear benefit
- Technology limitations
- Poor cost justification
- Lack of knowledge about solutions

Recovery Technology Tradeoffs

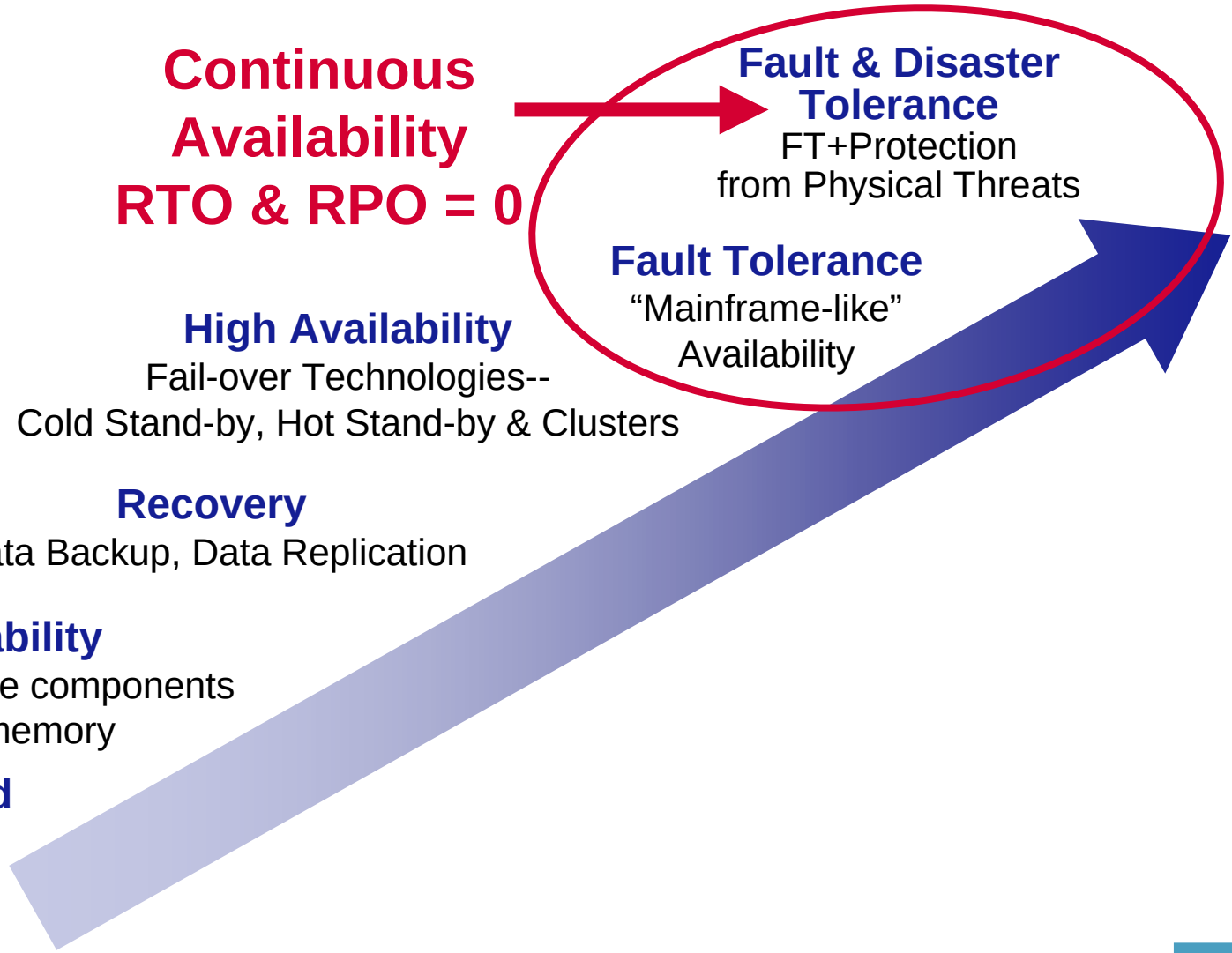
Recovery Time Objective (RTO)

- How long can you afford your system to be down?

Recovery Point Objective (RPO)

- How out-of-date can you afford the data to be when the system comes back up?

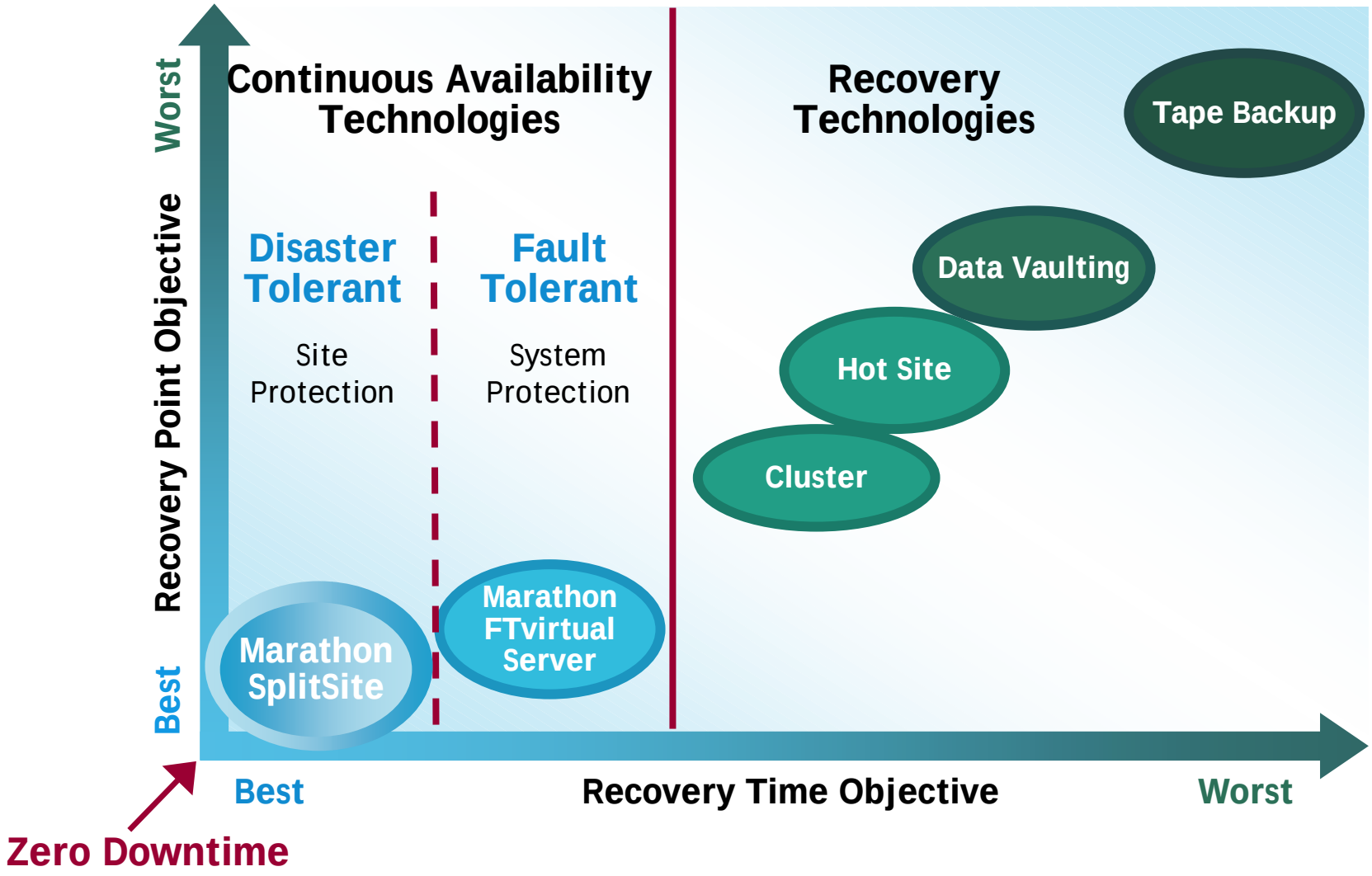
Availability Spectrum



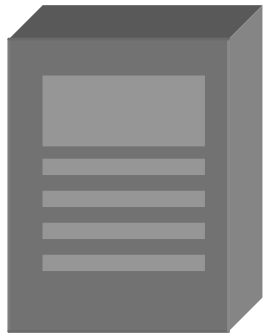
Continuous Availability Through Fault and Disaster Tolerance

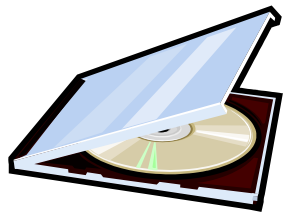
✓ **Continuous Operation**	■ Fault tolerance: no downtime during and after a system failure ■ Disaster tolerance: no downtime during and after a data center disaster ■ Zero recovery time (RTO) ■ Zero lost transactions, current data (RPO)
✓ **Highest Level of Protection**	■ ***No loss*** of system, transaction, or data state in the event of a failure
✓ **Simplest Technology**	■ Deployment and management the same as standard server

Disaster Tolerance vs. Disaster Recovery



Achieving Continuous Availability

Replication/backup solutions preserve data =  **Storage**

Clusters preserve application operations =  **Application**

Reliable application recovery must preserve data, application, and OS = 

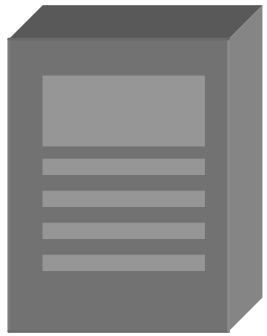
Application

OS

Data

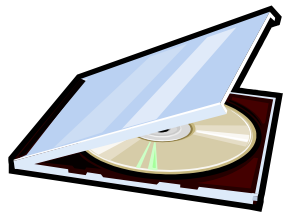
Achieving Continuous Availability

Replication/backup solutions preserve data =



Storage

Clusters preserve application operations =



Application

Continuous availability preserves application, OS, and data – redundant & synchronized =

Application

OS

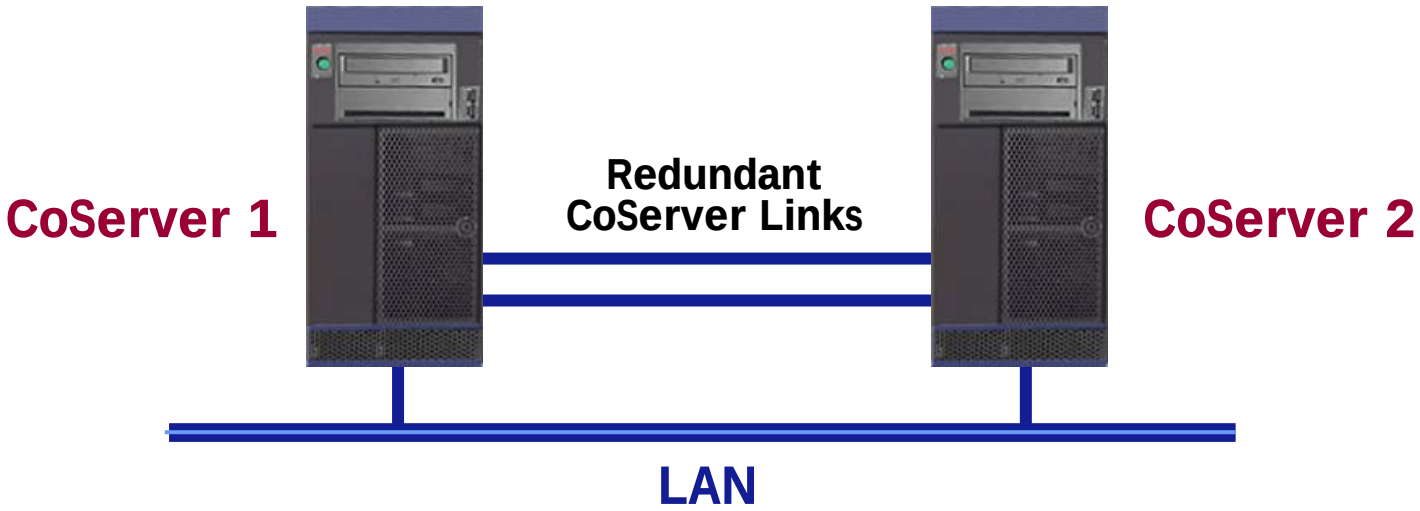
Data

Marathon Fault & Disaster Tolerance

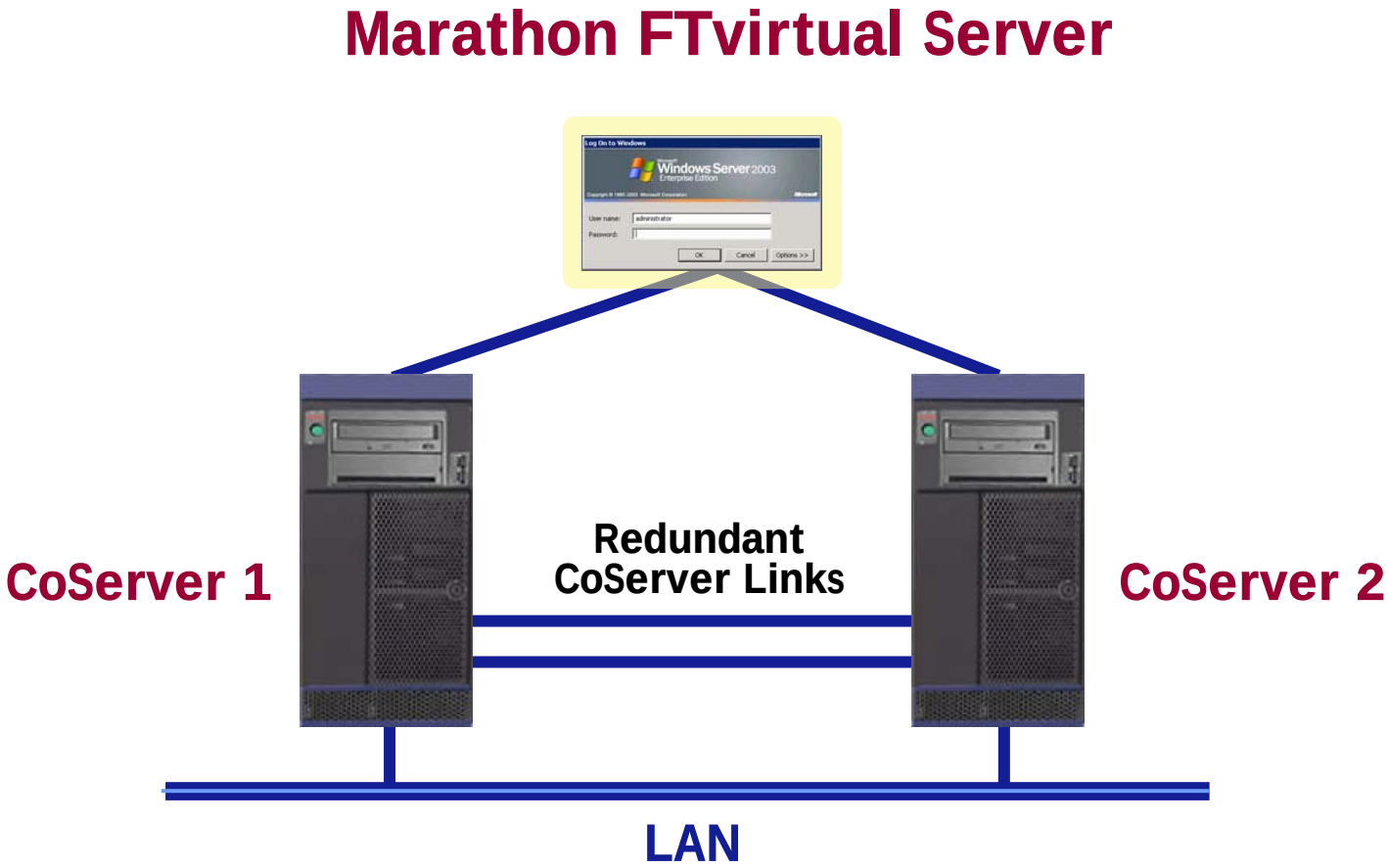
- Fault and disaster *tolerance* — not just recovery
  - Continuous application and data access through failures
  - No loss of state or application context
  - No failover — zero downtime
- In-flight data transactions are preserved
- Failures do not affect performance
- Repairs while system is operating
- Integrated mirrored storage, delta copy recovery
- Compatible with external storage
- Appropriate for blades and conventional servers

How Does it Work?

- Two standard Intel-based server or blades
- Two standard gigabit Ethernet links
- Marathon software synchronizes redundant components:
  - CPU, Memory,Networks, Disk

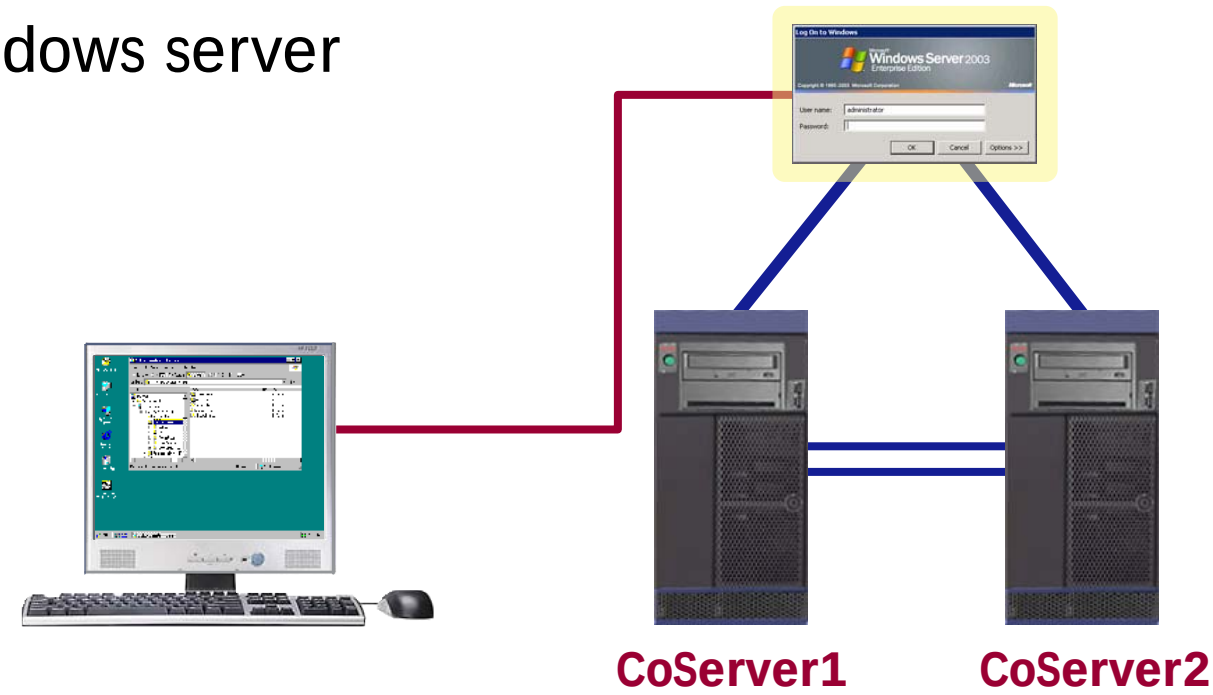


Fault & Disaster Tolerant Virtual Server

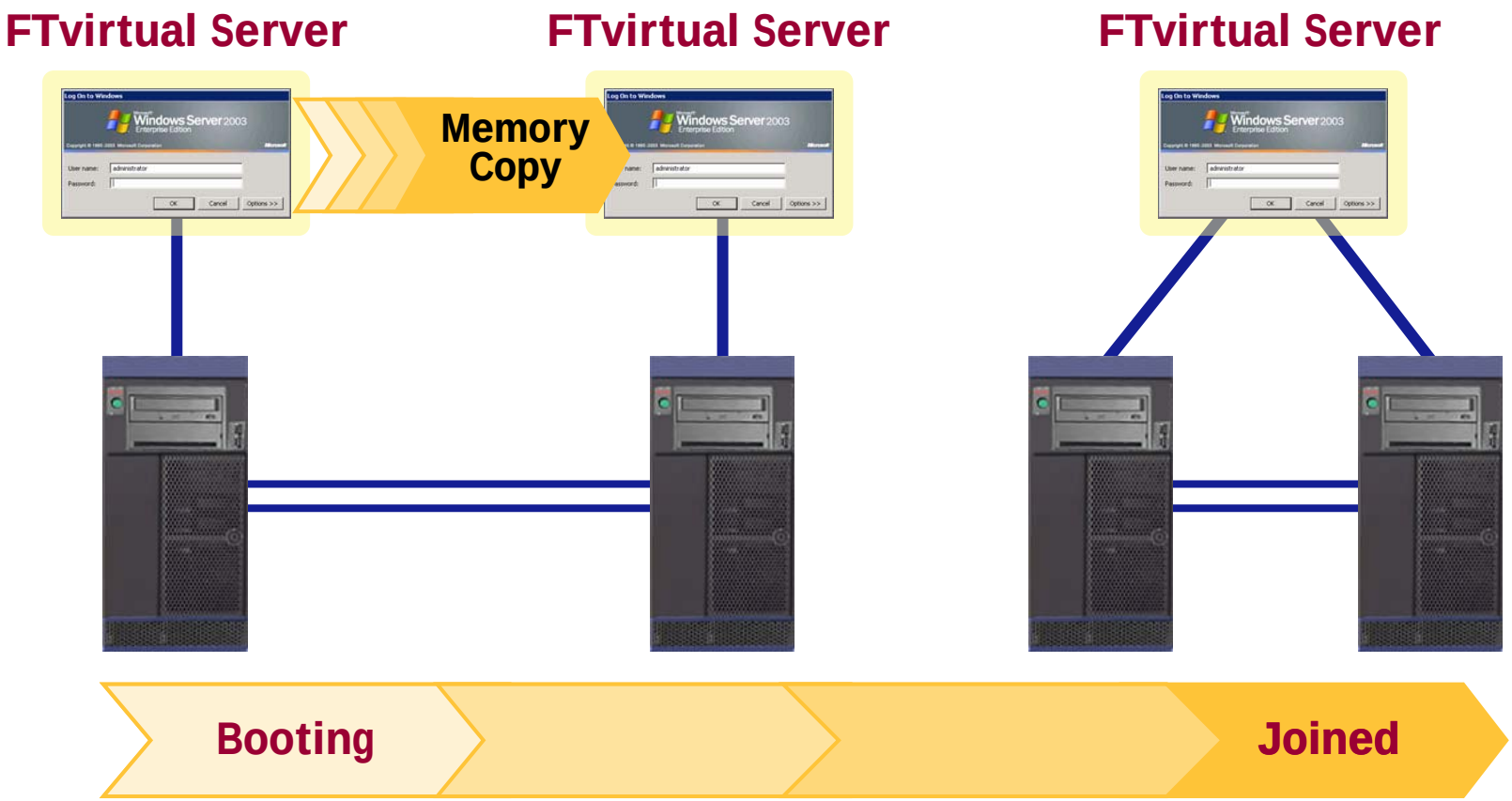


Administrator/User Perspective

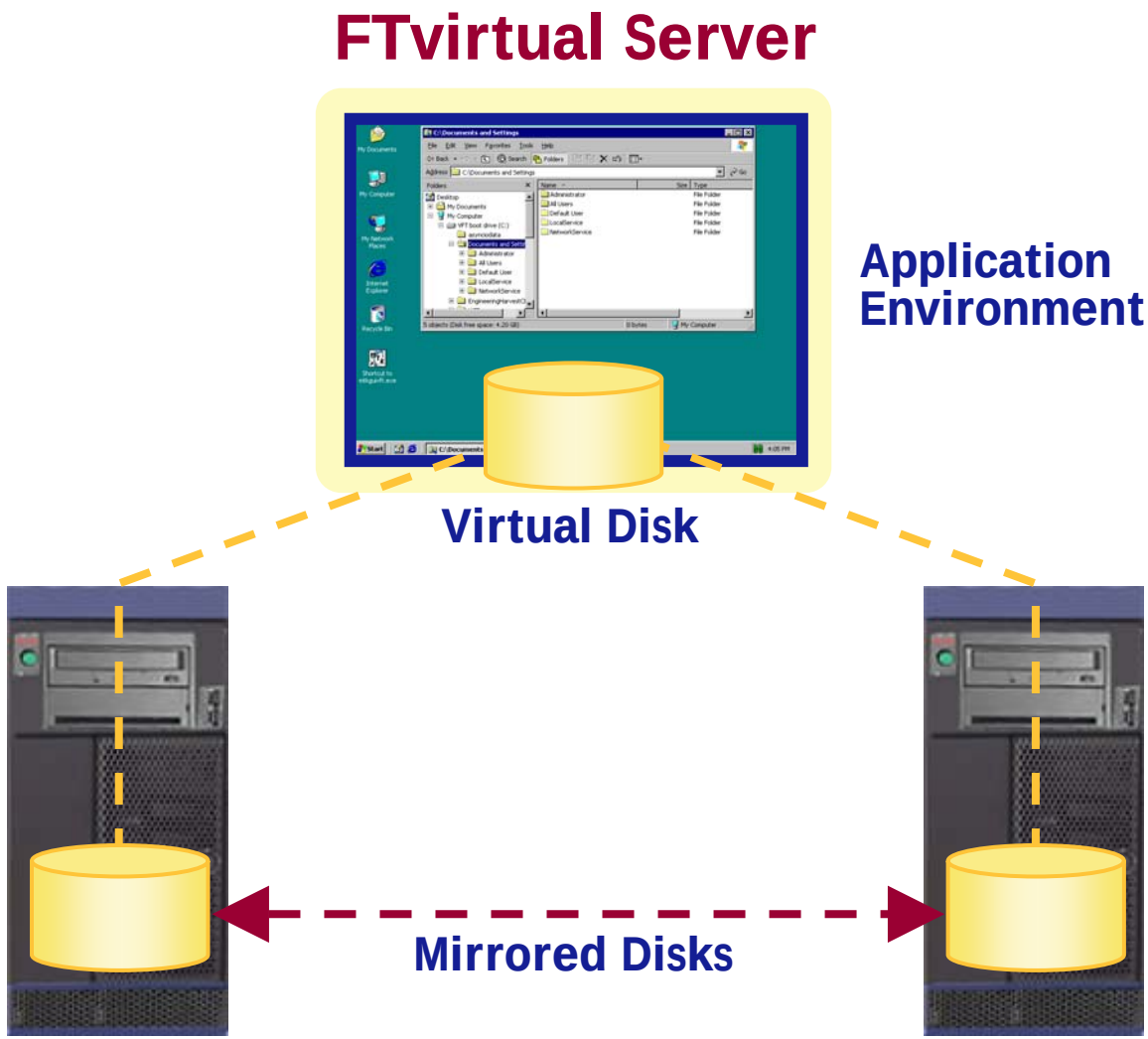
- Application management is performed in the FTvirtual OS environment
- Identical to a standard Windows server



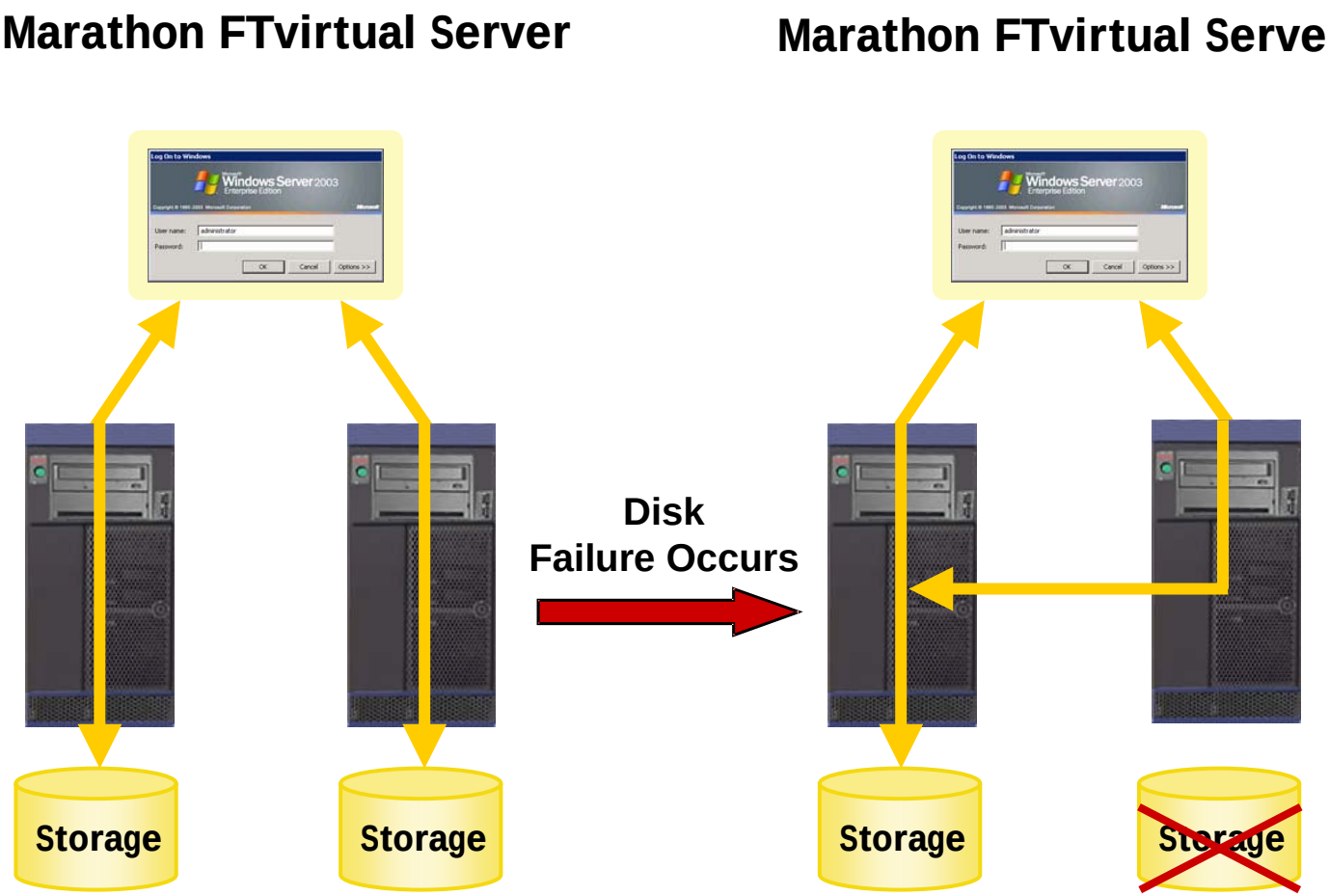
Marathon FTvirtual Server™ Joining and Synchronization



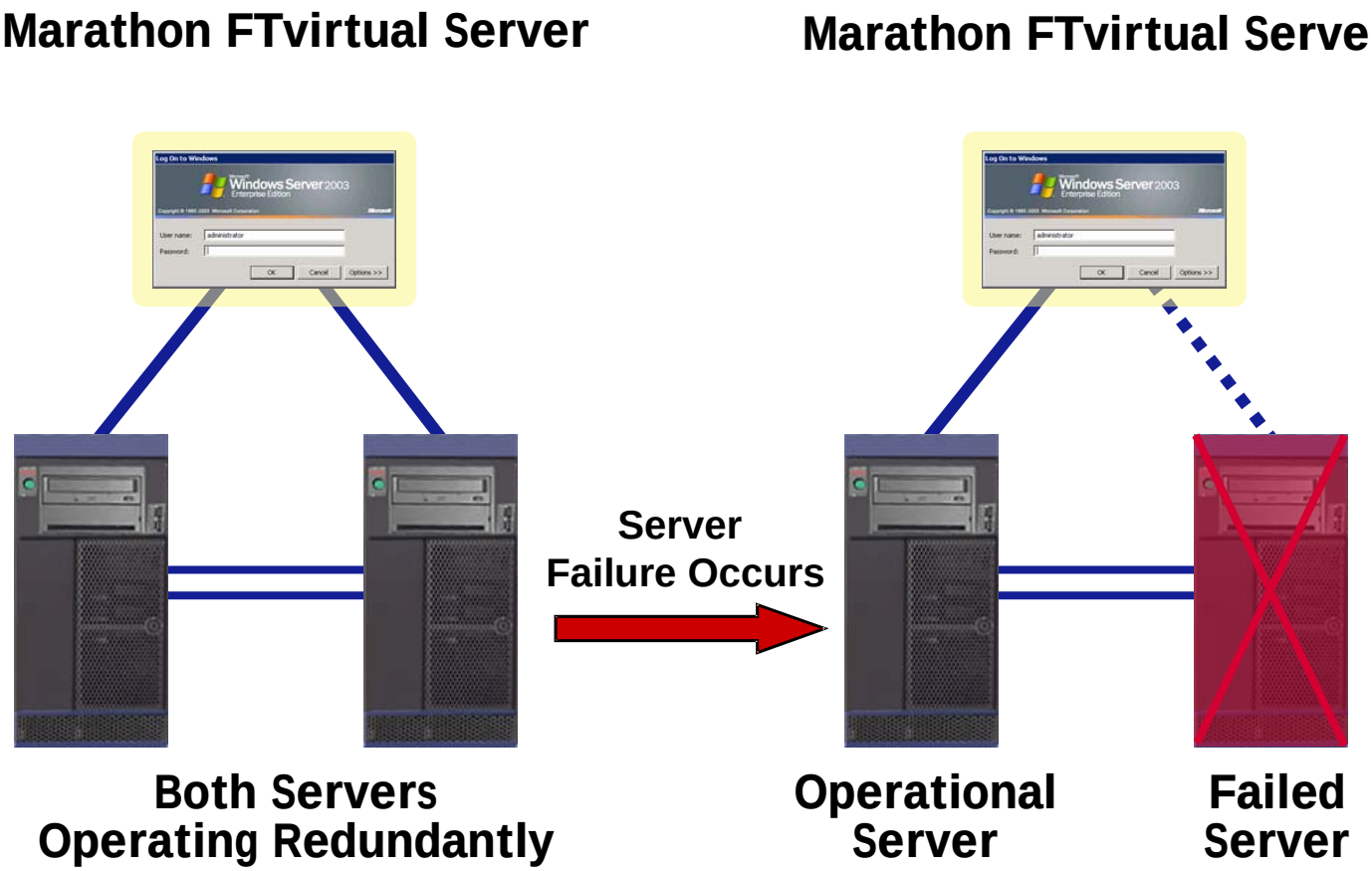
Device Virtualization
Disk Mirroring



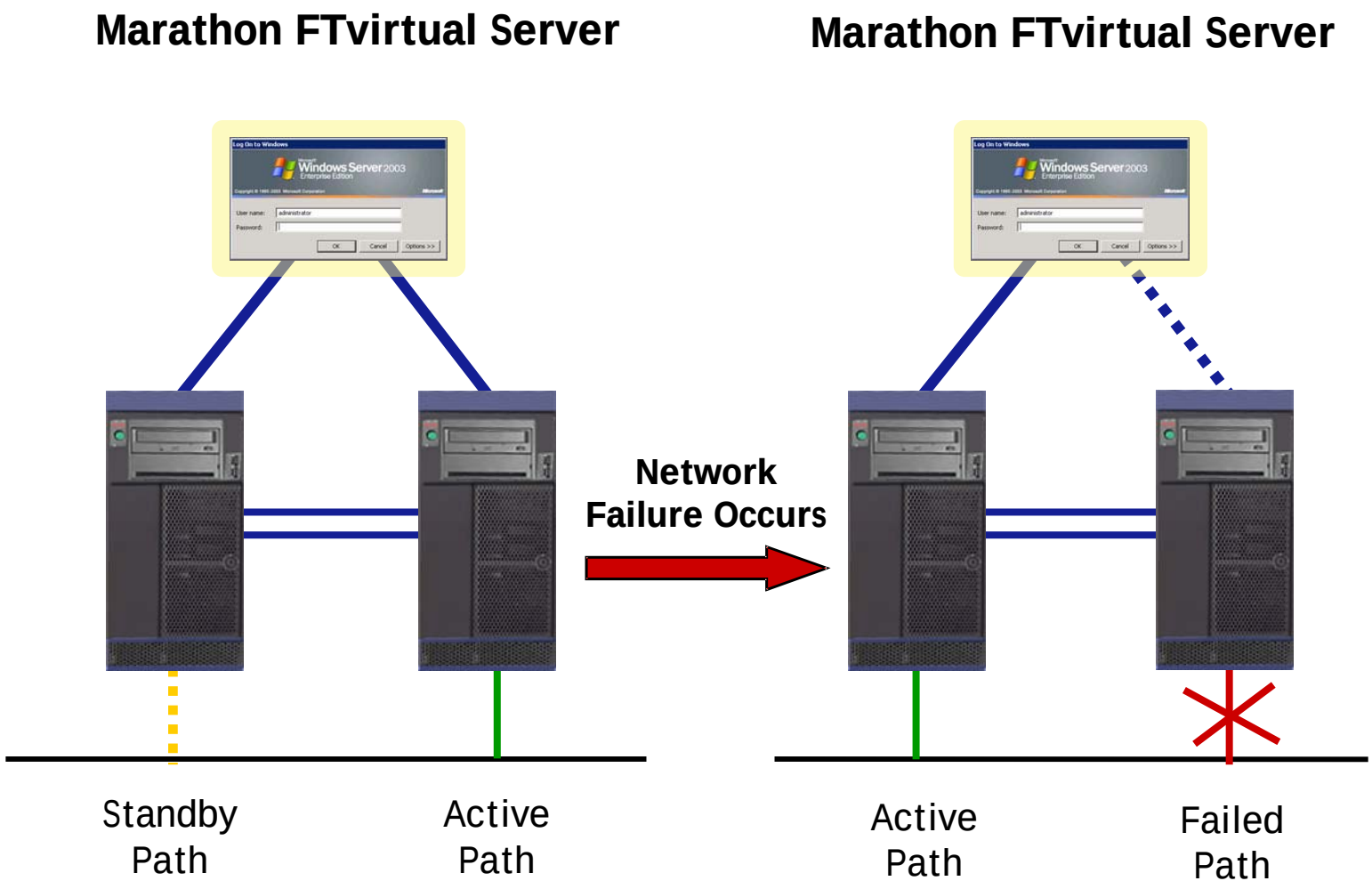
Fault Management
Disk Failure



Fault Management
Server Failure



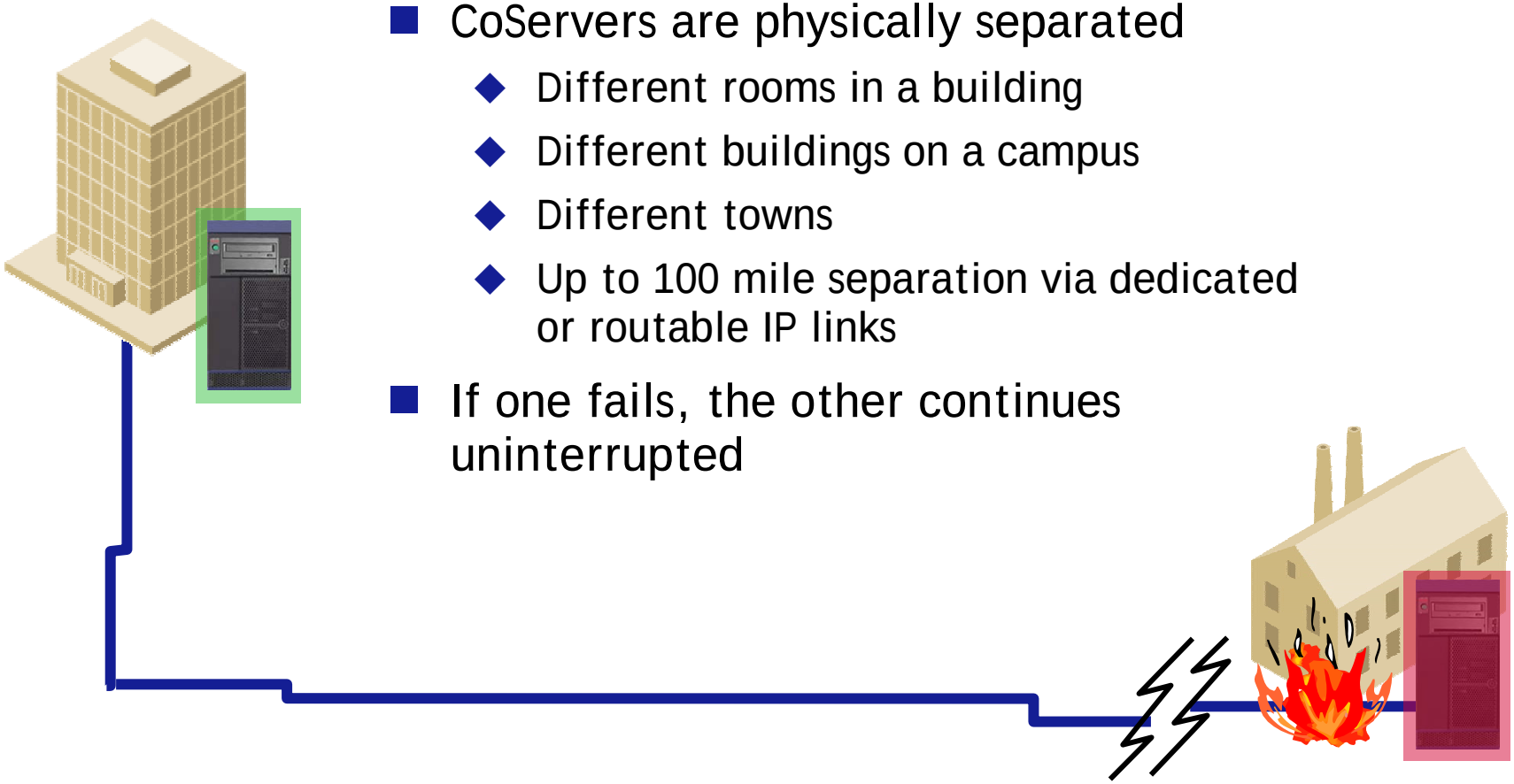
Fault Management
Network Failure



SplitSite® for Disaster Tolerance

Surviving a Disaster Rather than Recovering from One

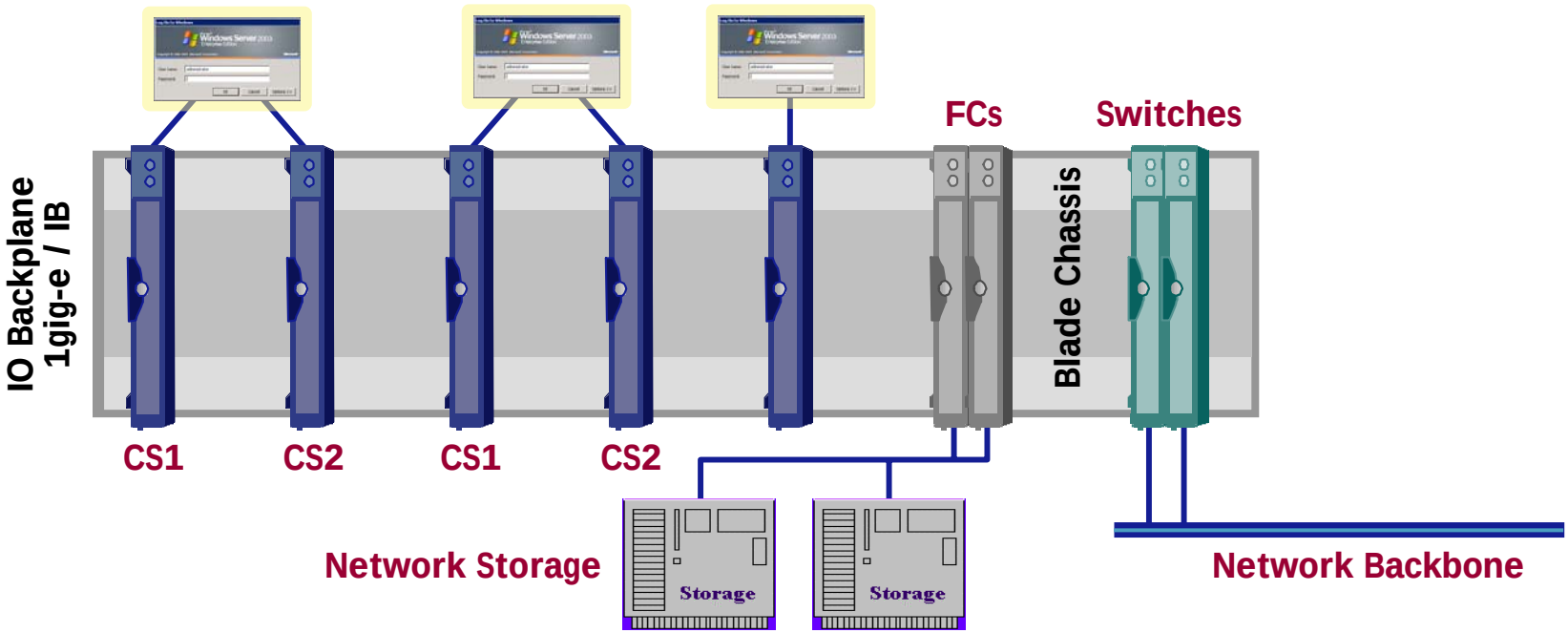
- CoServers are physically separated
  - Different rooms in a building
  - Different buildings on a campus
  - Different towns
  - Up to 100 mile separation via dedicated or routable IP links
- If one fails, the other continues uninterrupted



Fault Tolerance for Server Blades

- Marathon FTvirtual Server complements blades' high availability features
  - Eliminates Single Points of Failure
  - Simple Repair
- Blade = CoServer
- Blades and on-board storage become fault tolerant for critical applications

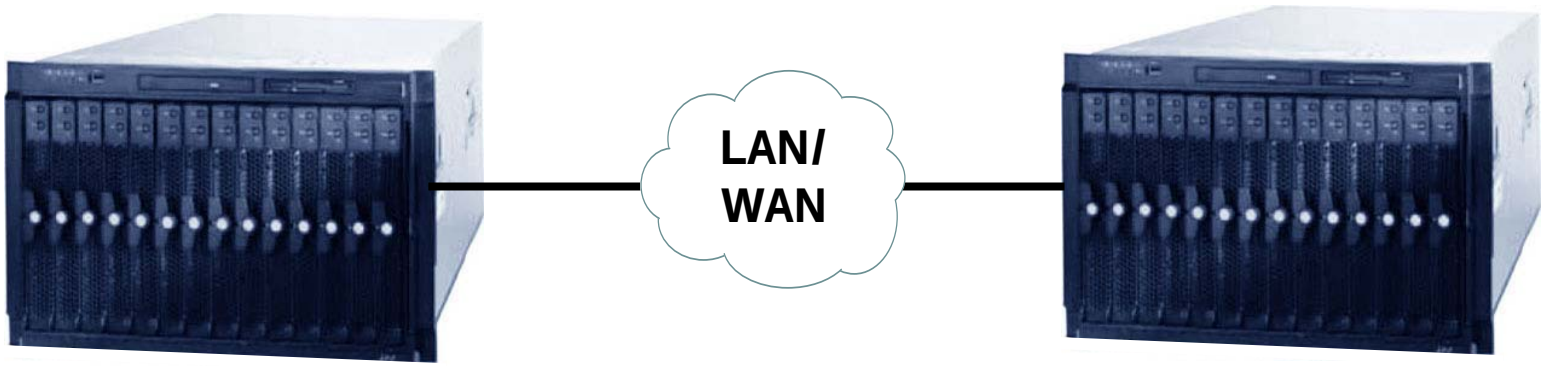
Typical Fault Tolerant Server Blade Configuration



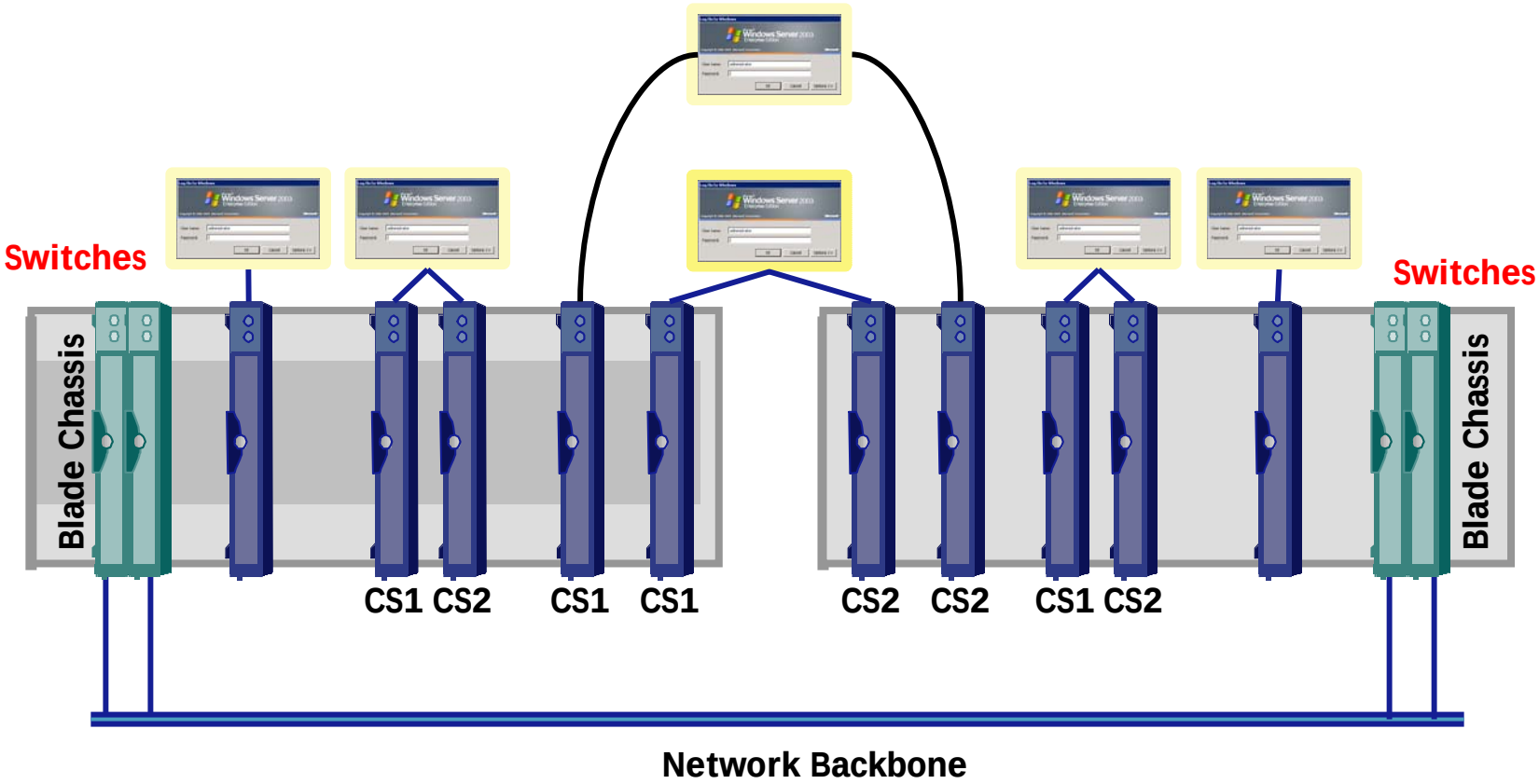
- Blade = CoServer
- CoServer links via blade chassis backplane
- On-blade or network storage
- Multiple FTvirtual Servers per chassis
- FT and non-FT servers in same chassis

Disaster Tolerance for Server Blades

- Two chassis – 1 in each data center
- CoServer blades paired across chassis
- Chassis separated across LAN / WAN



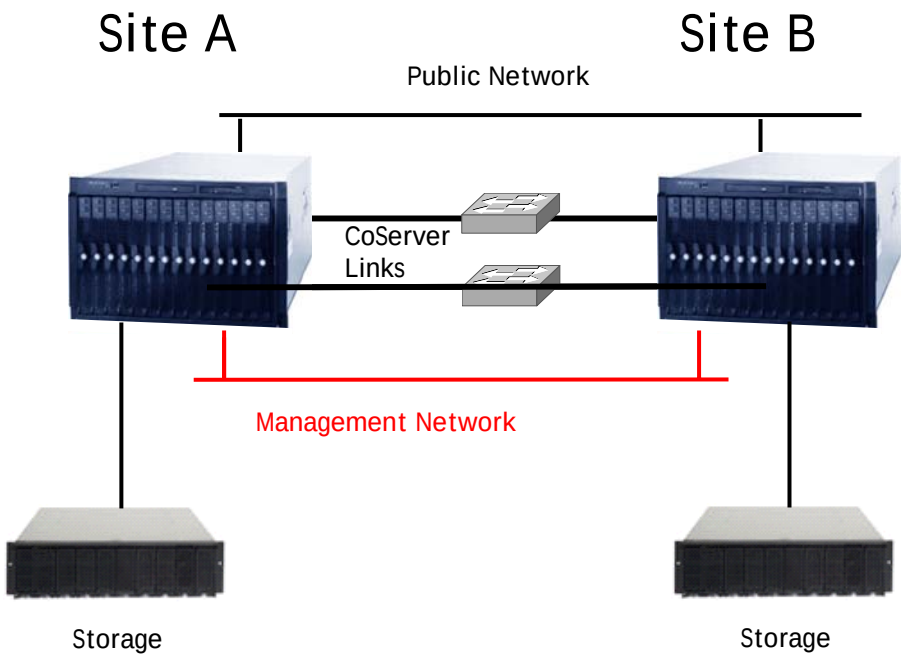
Typical Disaster-Tolerant Server Blade Configuration



Blade Repair While Online

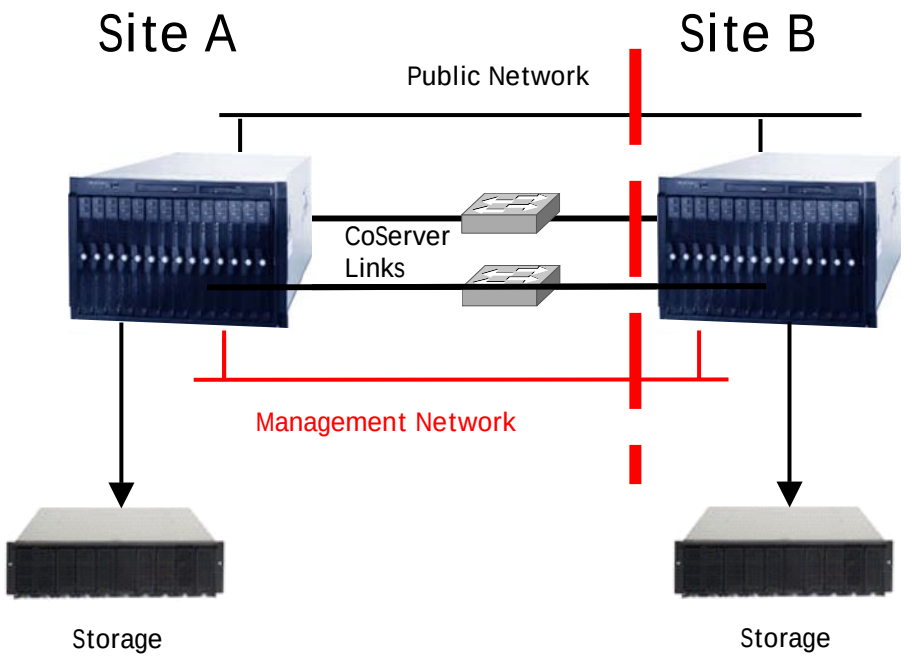
1. Remove, repair and replace the blade
2. Provision onboard disk with CoServer boot image
3. Insert the repaired blade
4. Blade automatically rejoins and data is remirrored to on-board disks

Managing Network Failures Across a Wide Area



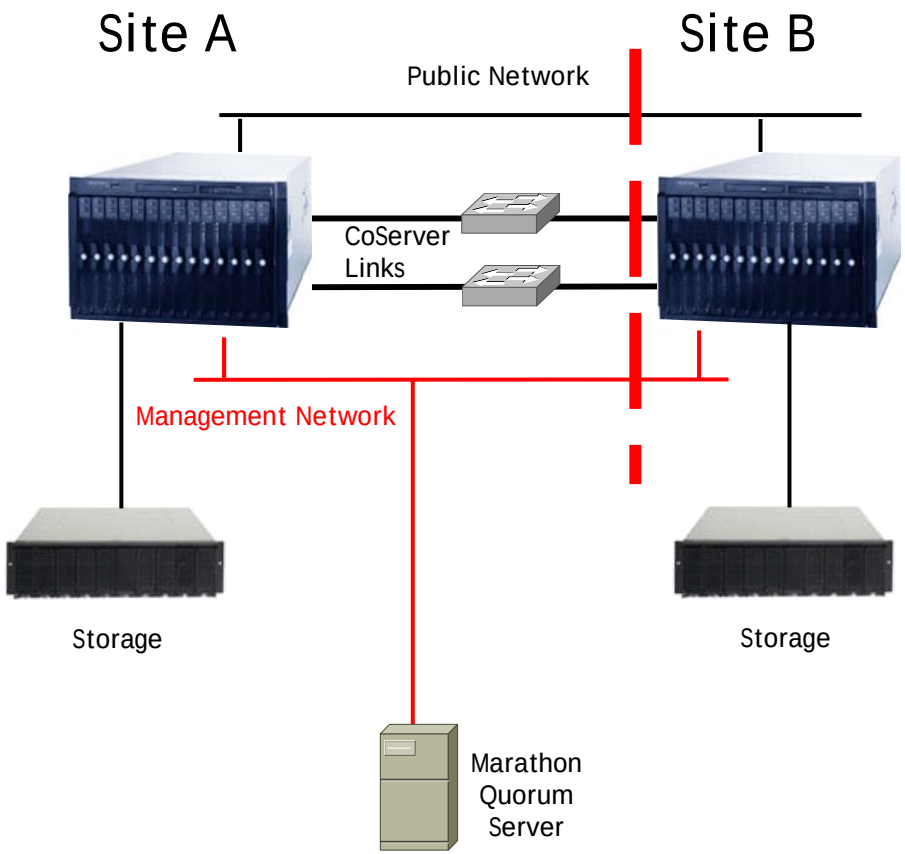
- "Heart-beat" datagrams continuously communicate CoServer health

Managing Network Failures Across a Wide Area



- "Heart-beat" datagrams continuously communicate CoServer health
- Simultaneous break in all communications causes "Split-brain":
  - Both halves runaway and independently update data

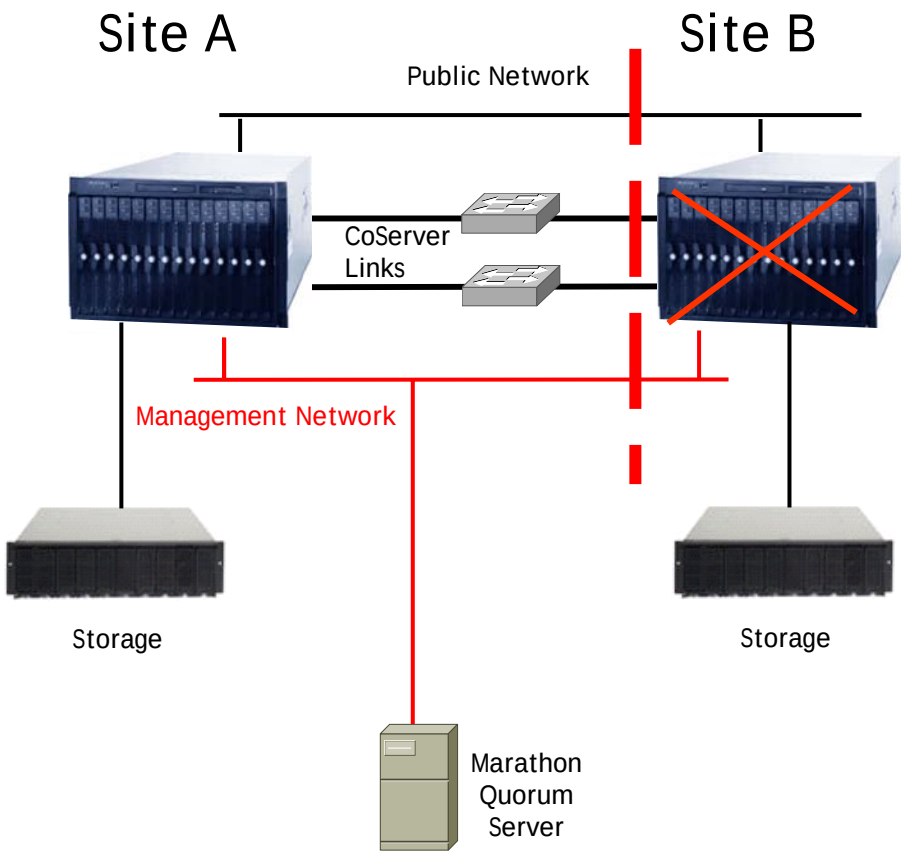
Managing Network Failures Across a Wide Area



- "Heart-beat" datagrams continuously communicate CoServer health
- Simultaneous break in all communications causes "Split-brain":

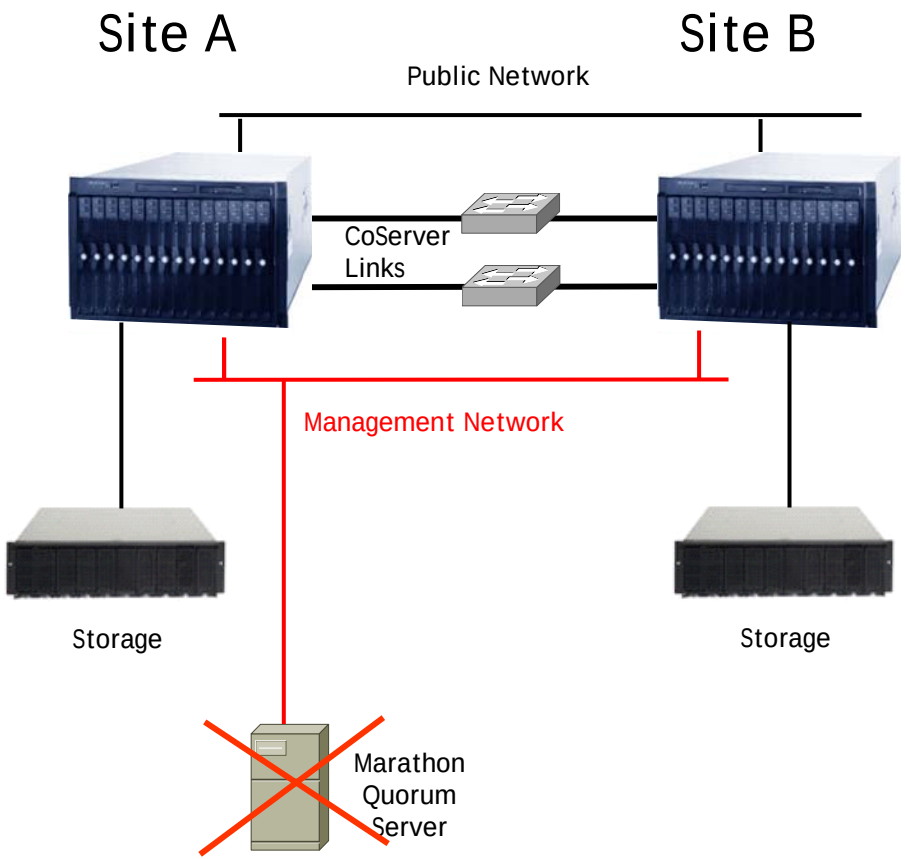
  Both halves running and independently update data
- Quorum server negotiates split-brain scenario

Managing Network Failures Across a Wide Area



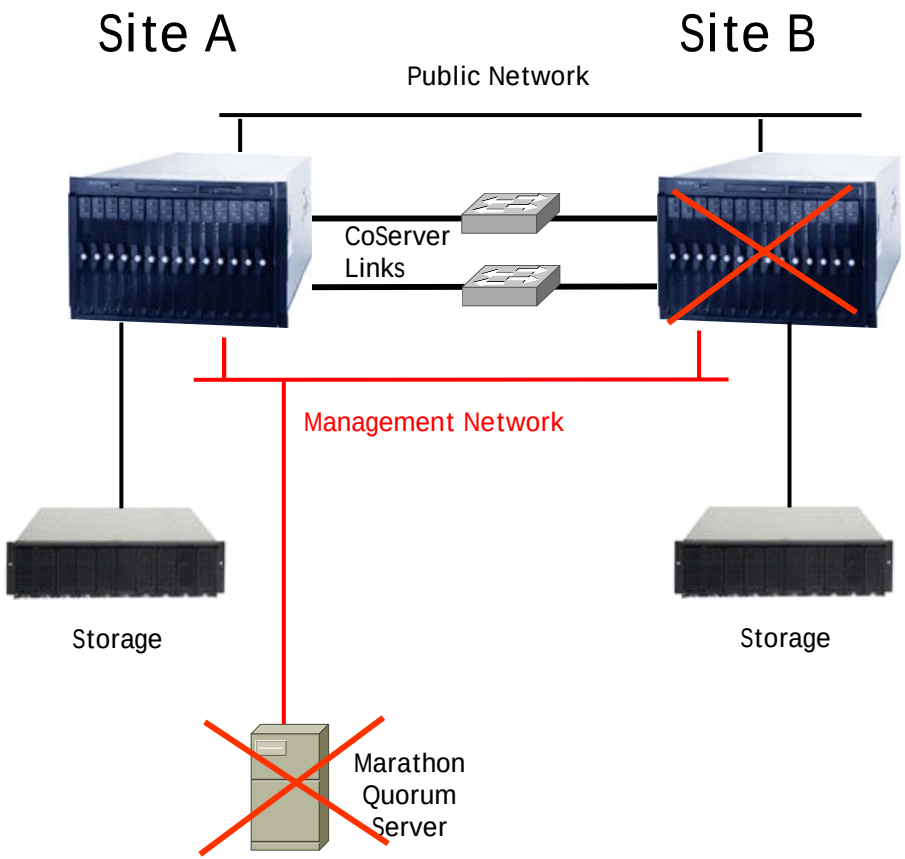
- "Heart-beat" datagrams continuously communicate CoServer health
- Simultaneous break in all communications causes "split-brain":
  - Both halves run away and independently update data
- Quorum server negotiates split-brain scenario
- CoServer losing quorum server communication goes offline
- Continuous availability provided through failure

Quorum Server Failures



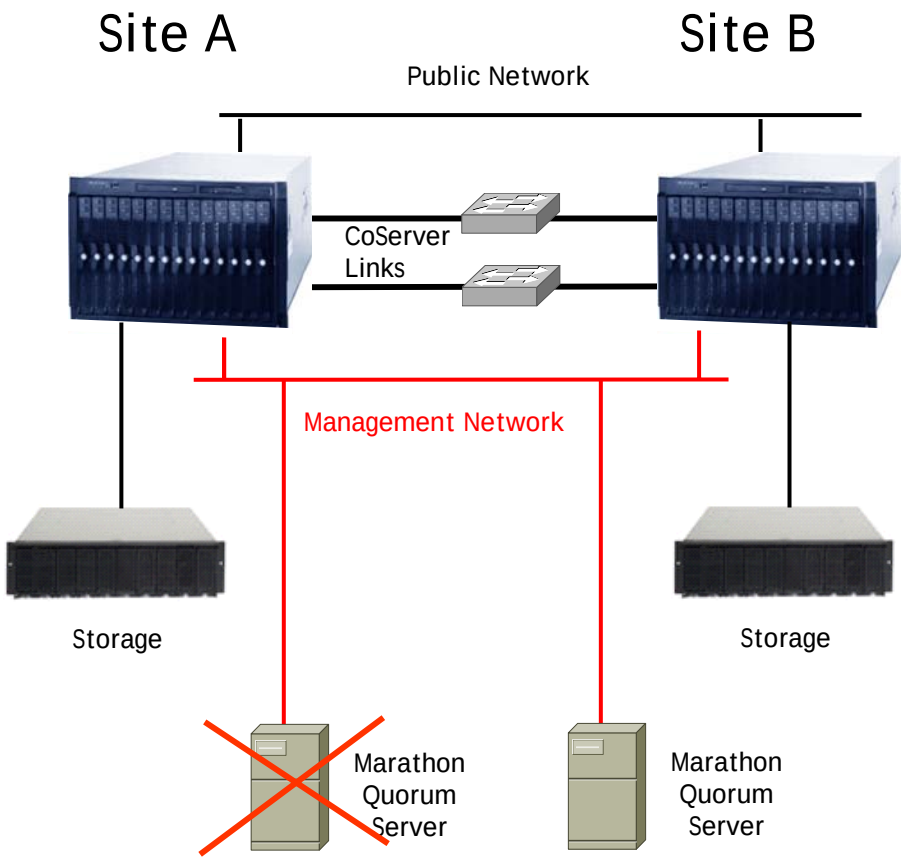
- Quorum Server failure leaves blades susceptible to "split-brain"

Quorum Server Failures



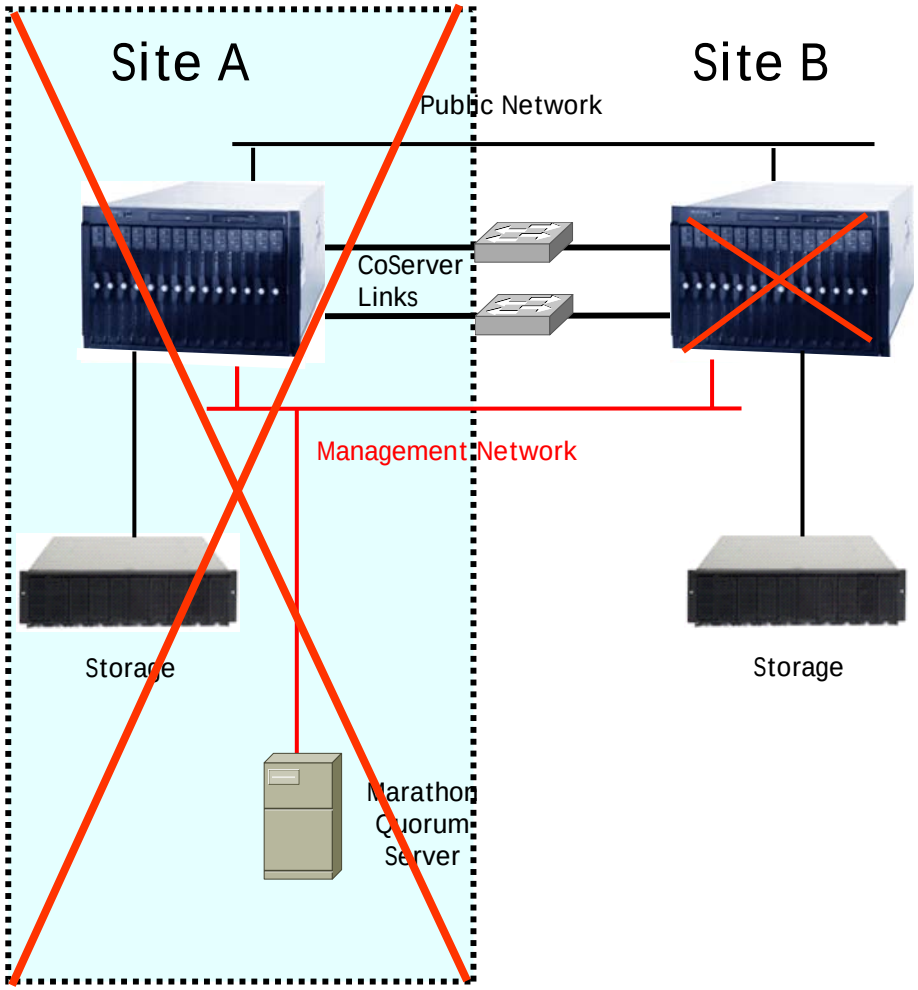
- Quorum Server failure leaves CoServer blades susceptible to “split-brain”
- Requiring one CoServer blades to shutdown
  - Application availability continues but redundancy is reduced

Quorum Server Failures



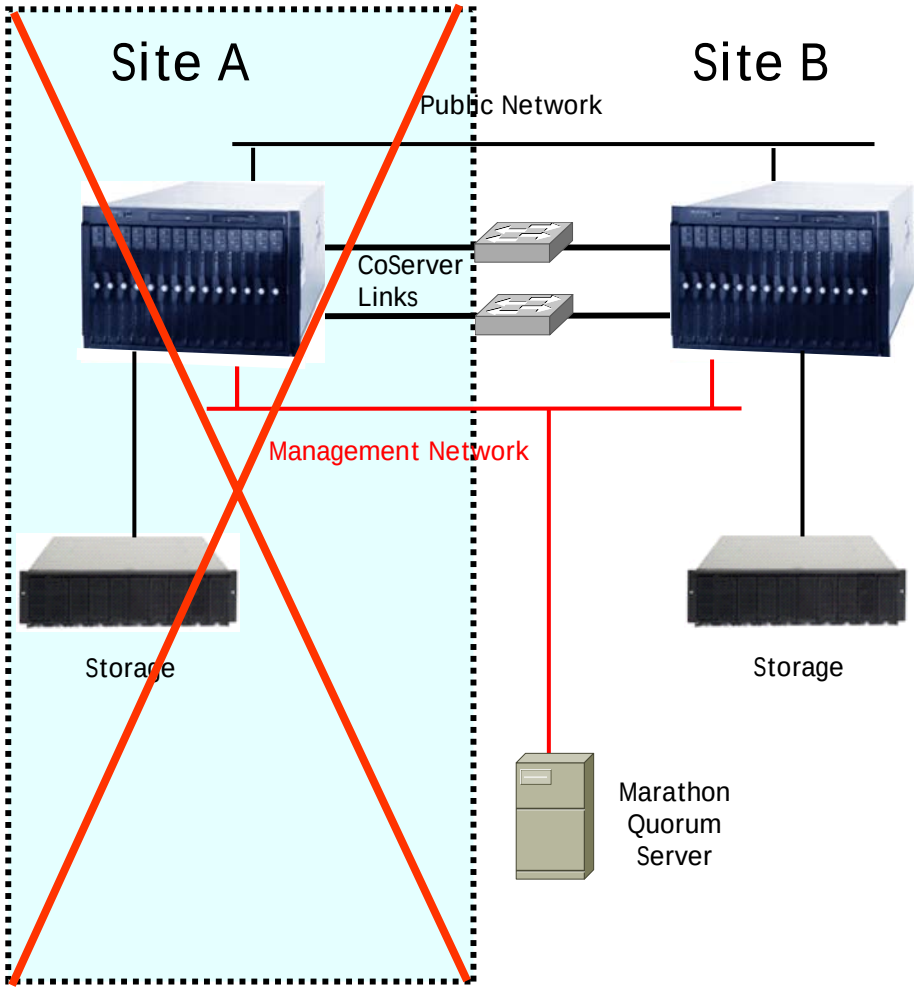
- Quorum Server failure leaves blades susceptible to “split-brain”
- Requiring effected CoServer blades to shutdown
  - Application availability continues but redundancy is reduced
- A second Quorum Server provides needed redundancy

Quorum Server Site Selection



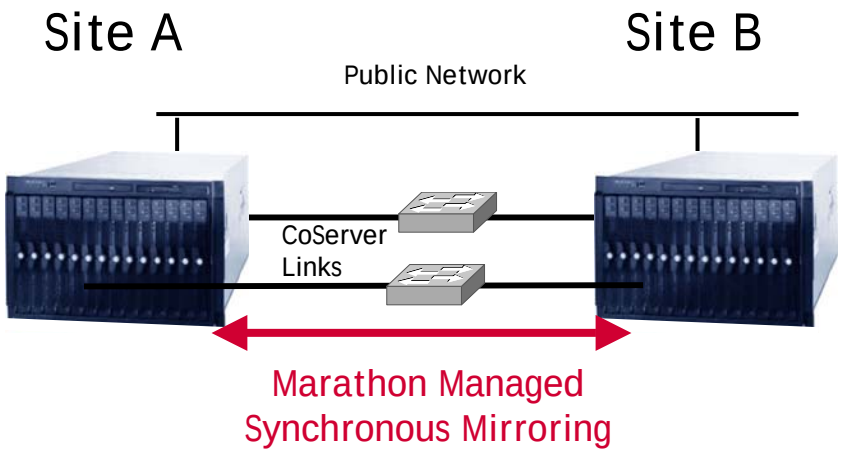
- Co-locating Quorum Server and blades leaves blades susceptible to “split-brain” on site failure

Quorum Server Site Selection



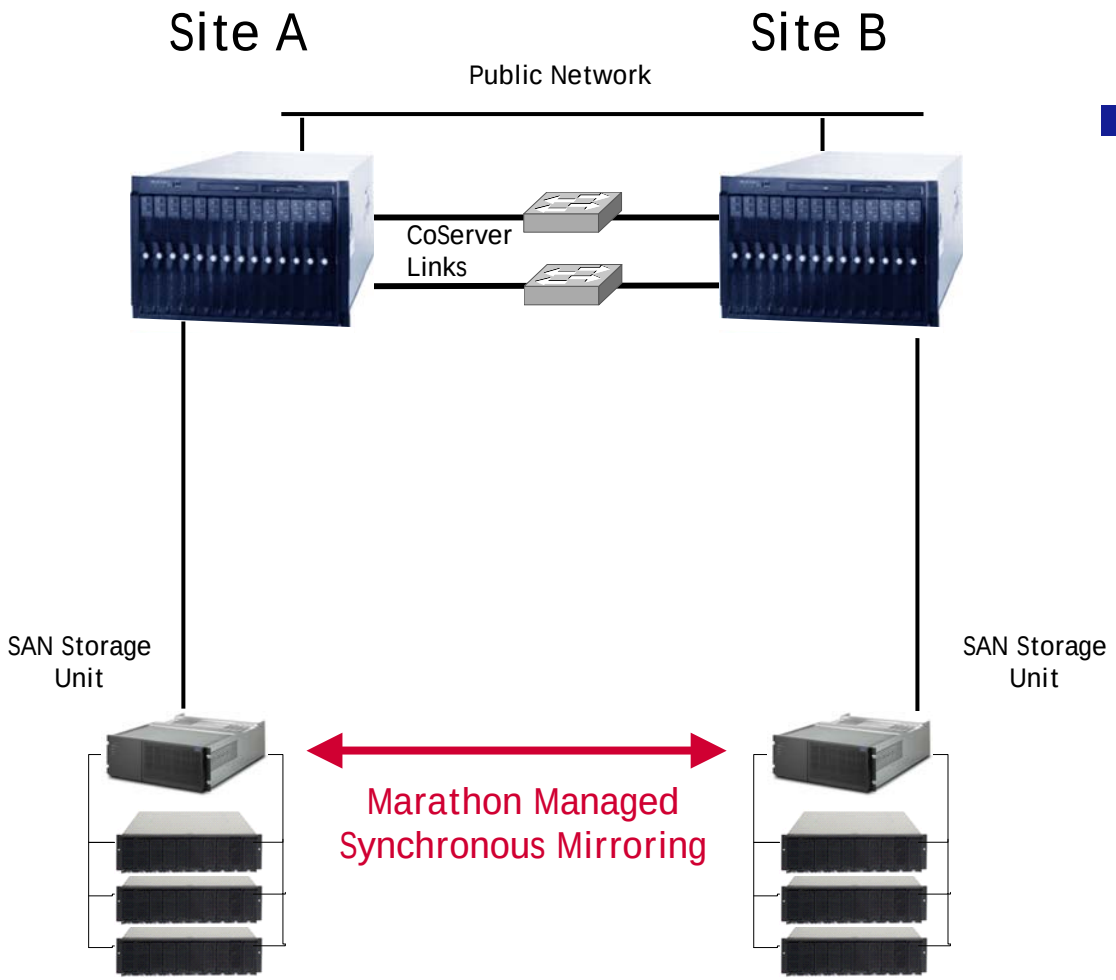
- Co-locating Quorum Server and blades leaves blades susceptible to “split-brain” on site failure
- Locating Quorum Server at independent site preserves continuous availability

Storage Configurations For Disaster Tolerance



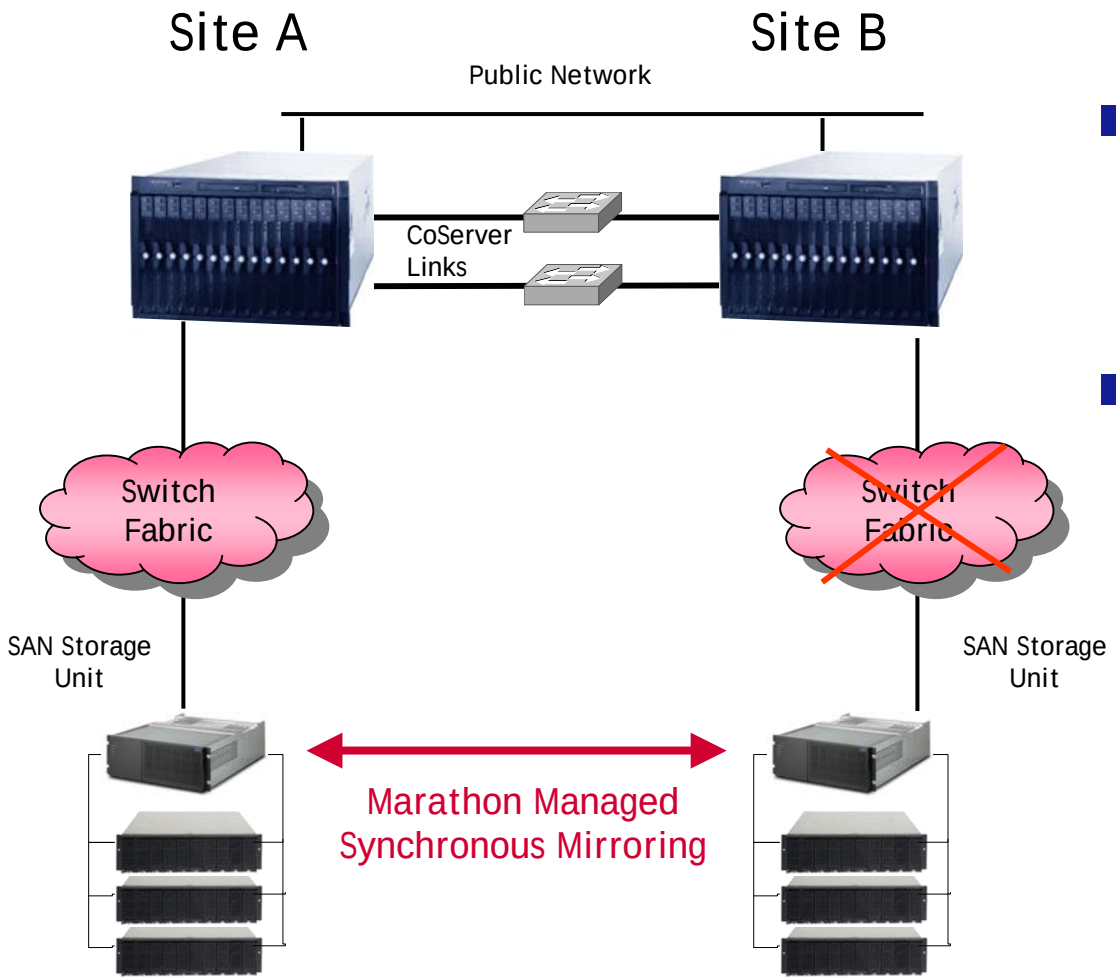
- Blade attached internal storage mirrored by Marathon software

Storage Configurations For Disaster Tolerance



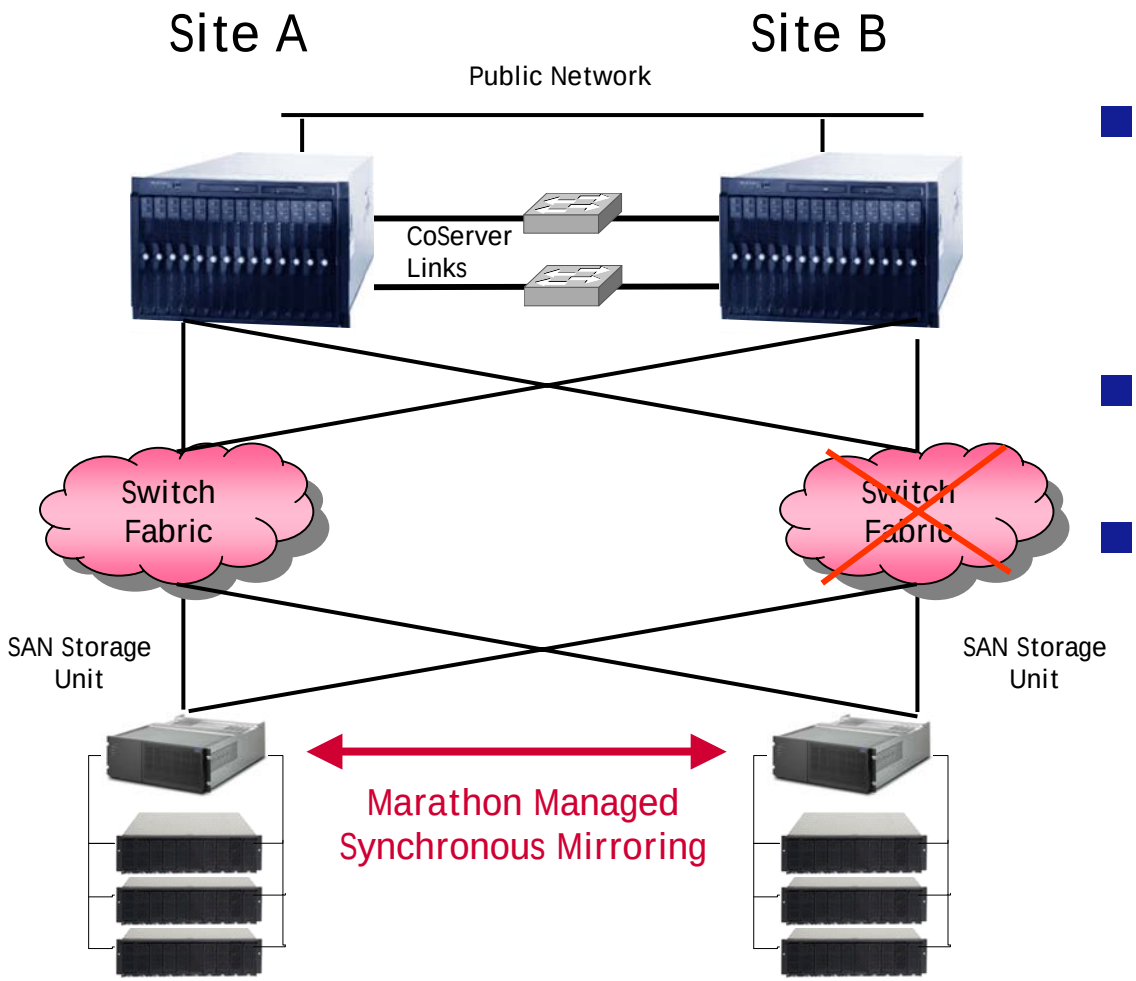
- SAN Storage device at each location
- Marathon software manages synchronous data updates at each site

Storage Configurations For Disaster Tolerance



- SAN Storage device at each location
- Marathon software manages synchronous data updates at each site
- Fabric failures cause loss of redundancy

Storage Configurations For Disaster Tolerance



- SAN Storage device at each location
- Marathon software manages synchronous data updates at each site
- Fabric failures cause loss of redundancy
- Multipath IO solutions provide higher availability
  - No loss of redundancy on fabric or HBA failure
  - Mirror copy avoided on service restoration
  - Multipath load balancing

Conclusion

- Availability is critical to many IT systems
- Emerging business and technology trends such as server blades increase the need for availability
- Continuous availability is the most effective HA technology
- Marathon software supports cost-effective, flexible blade configurations that provide continuous availability in the event of a site disaster

Thank You

Contact Information:

Jerry Melnick
VP Engineering
Marathon Technologies
295 Foster Street
Littleton, MA 01460
Tel: 978-489-1100

Email: Jmelnick@marathontechnologies.com
Website: www.marathontechnologies.com
Download Presentation:
http://www.marathon1.com/BladeSummit/BladeSummit.htm