

# White Paper FUJITSU Storage ETERNUS DX Storage Cluster

Introduction to the concept, technology and setup of the ETERNUS DX transparent failover solution for optimal business continuity



## Content

Management Summary	2
Business Continuity, Server Virtualization and Storage	3
Basic Architecture	4
Storage Cluster Features and Benefits	5
Storage Cluster in Detail	6
Primary and Secondary Storage	6
SAN Infrastructure	6
Storage Cluster Controller	7
ETERNUS SF Management Server	7
Failover Mechanism	8
Failover and Failback	9
Storage Cluster in Virtualized Server Environments	10
Conclusion	11
Related Documents	12

---

## Management Summary

As data is one of the most important assets in any company, storage systems play a vital role in the IT infrastructures of all enterprises. IT administrators need to ensure that mission-critical data is continuously accessible, ensuring round-the-clock business continuity despite any system or site failures, as well as unplanned or planned outages.

Fujitsu's ETERNUS DX family, with its Business-centric Storage approach, addresses this issue with a unique transparent application and server failover solution – Storage Cluster.

Storage Cluster is the high-availability feature of the ETERNUS DX S3 family of storage arrays. Data is synchronously mirrored between two interlinked storage systems. If the primary (active) system should fail, all primary host connections are switched instantly to the secondary (standby) system. This failover is transparent for both servers and applications, and it ensures uninterrupted operations.

This advanced feature helps fulfill service levels and delivers predictable operation for business-critical applications – particularly in virtualized server environments.

Storage Cluster is integrated and controlled by the ETERNUS SF Management Suite and features simple setup and administration functionalities.



# Business Continuity, Sever Virtualization and Storage

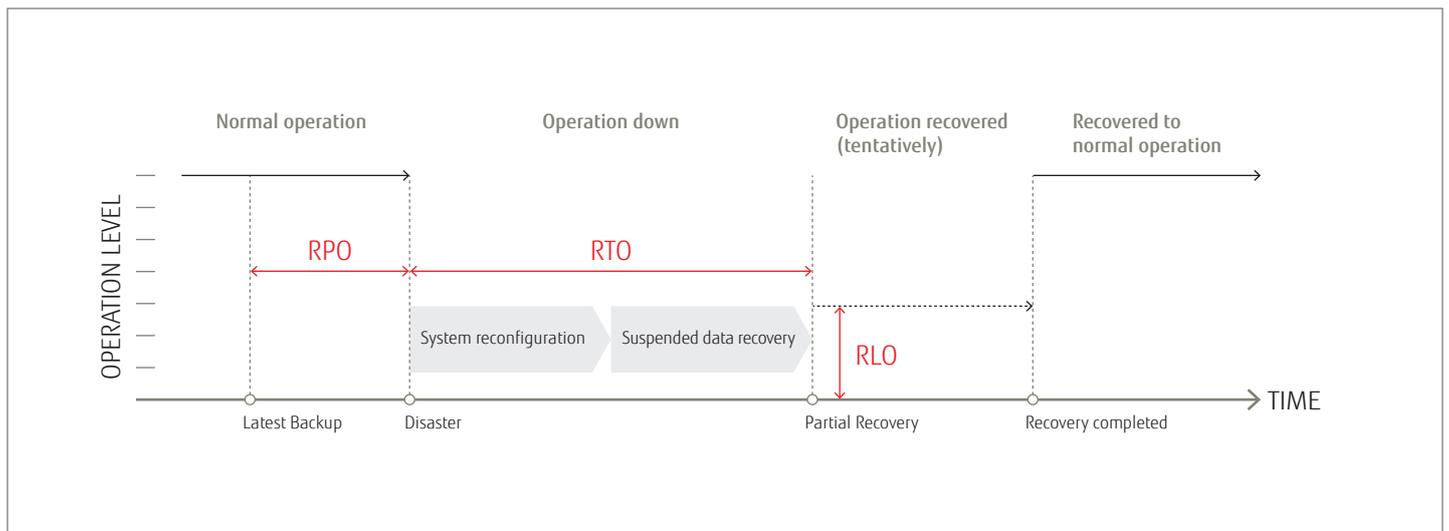
A business continuity plan is built on the 3R fundamental objectives: RLO, RTO and RPO.

- The Recovery Level Objective (RLO) is defined according to the priority of the business application – it defines the granularity with which the recovered data must be enabled – i.e., whether it is necessary to recover a whole database, a web application, a site collection, a site, a list or library, or just an item.
- Based on the RLO, acceptable values are defined for both Recovery Point Objective (RPO) and Recovery Time Objective (RTO)

All objectives should be set for each business application based on

- the expected loss for the business with the objective
- the cost of achieving the objective

## Recovery flow from disaster, tentative recovery to complete recovery



### RTO (Recovery Time Objective)

Metric for the amount of time it takes to recover from a data loss event, and how long it takes to return to service.

### RPO (Recovery Point Objective)

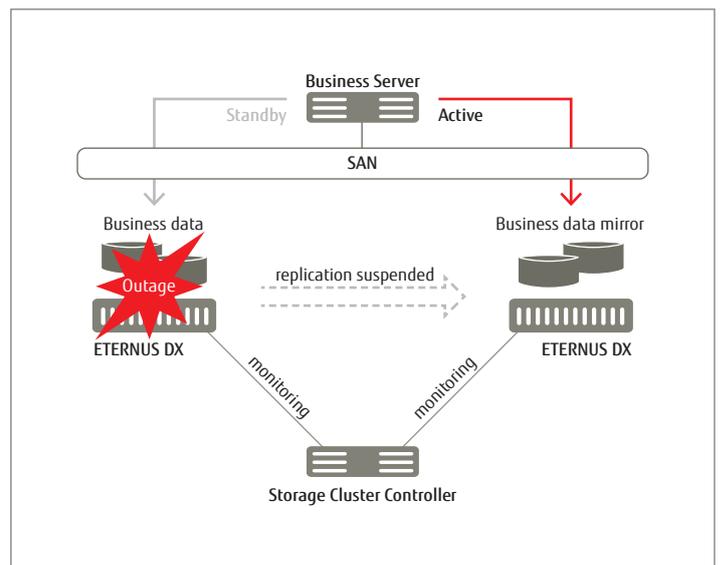
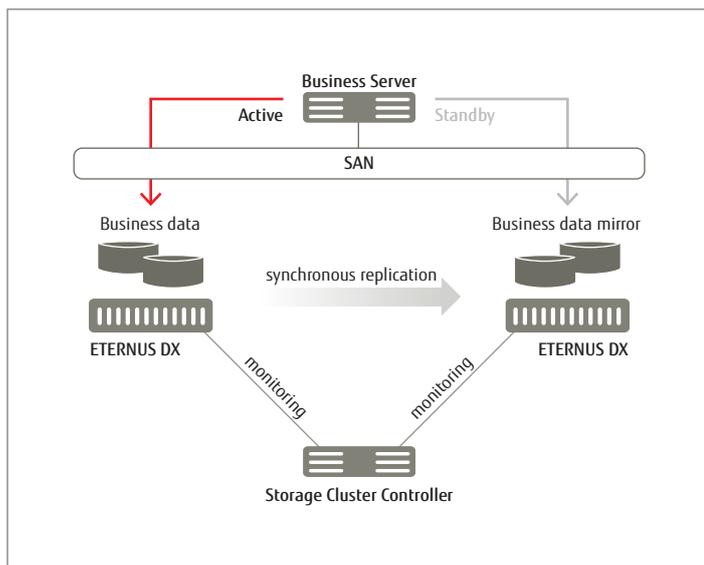
Metric for the amount of data at risk of being lost during a recovery.

### RLO (Recovery Level Objective)

Granularity level of the recovered data.

# Basic Architecture

The concept of Storage Cluster is based on the deployment of a secondary storage system and a supervising instance – the Storage Cluster Controller.

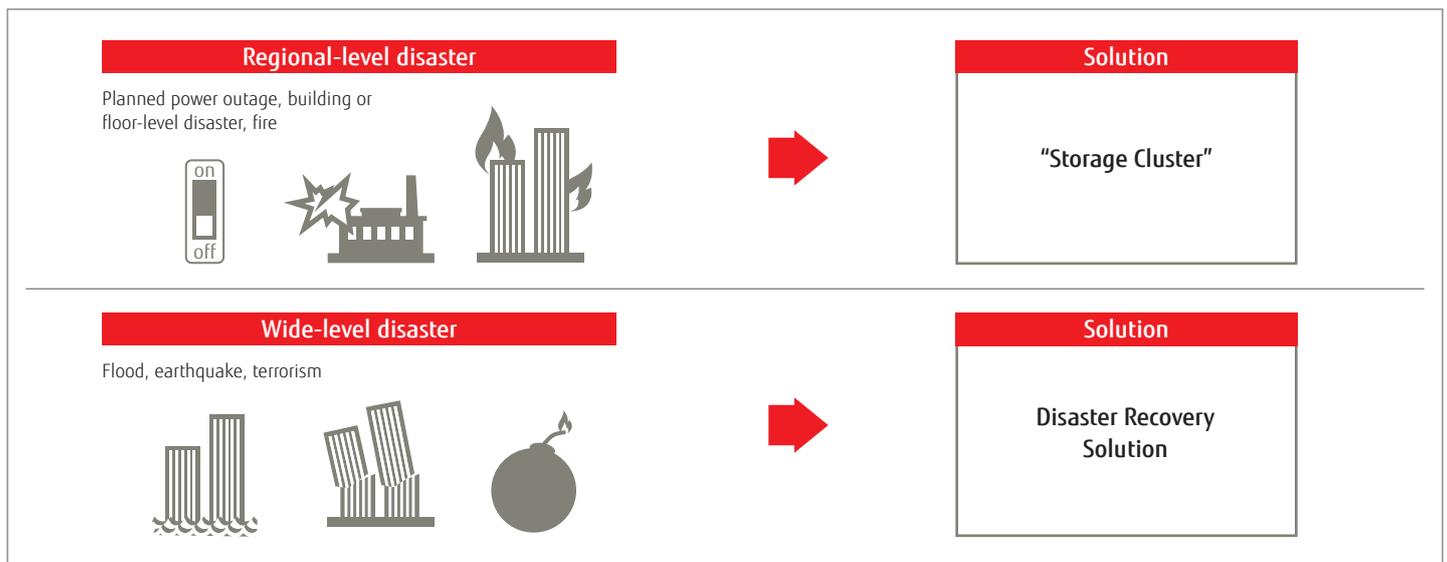


As long as the primary storage system is running, data is transferred from it to the secondary system via a synchronous replication function. The Storage Cluster Controller continuously checks the status of the primary storage. If a failure is detected, it runs the failover logic, and the primary storage information (e.g., LUN ID/WWN) is shifted over to the secondary storage in order to transparently identify the volume via the server I/O. Hence, operations run smoothly and ensure business continuity

Due to the restrictions on synchronous mirroring between the storage systems, Storage Cluster can be deployed in building, campus and metro environments. For regional or global business continuity and disaster recovery scenarios, ETERNUS DX can be deployed with asynchronous remote mirroring functionality resulting in increased RPO and RTO, and without automated and transparent failover.

# Storage Cluster Features and Benefits

Storage Cluster technology is the foundation for a non-disruptive business continuity solution. By transferring data in real time and switching storage resources automatically and transparently for servers and applications, data is protected 100 percent and the business process is safe from any interruption.



- Data consistency is managed by the storage systems without the need for specific volume management software. This frees up server resources while the secondary storage is securely synced with consistent data sets.
- Data is replicated along with the automated storage tiering layout and auto QoS settings.
- Failover to the standby resources is triggered automatically – there is no need for any human intervention.
- For manual failover, failover test or restoration after failover (failback), the failover operation can also be triggered from the user interface – it is simple and secure.
- Seamless server affinity – no change in access path, drive letter or mount point information after the failover. Volumes continue to be accessed transparently from the business application. No workload for volume reallocation or remount is needed or generated.
- After the recovery of the failed system, Storage Cluster automatically executes the failback to the initial state. System operation is always stable and secure.

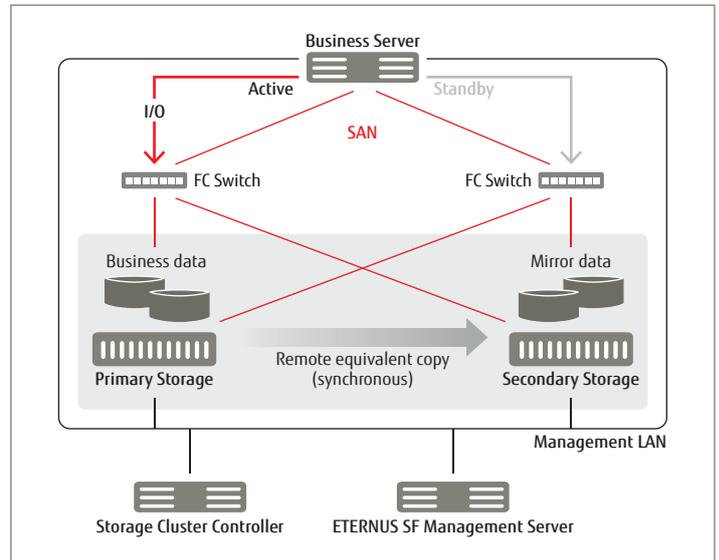
## Limitations and hints

- Maximum distance between the sites is limited by the roundtrip time for the synchronous replication which must not exceed 10 milliseconds.
- To avoid misbehavior in case of complete site outages it is recommended to deploy the Storage Cluster Controller and the ETERNUS SF Management Server on an independent third site.
- A redundant multipath configuration with at least one path to each storage controller is recommended.
- Storage Cluster Controller can be run as virtual machine without needing dedicated hardware.

# Storage Cluster in Detail

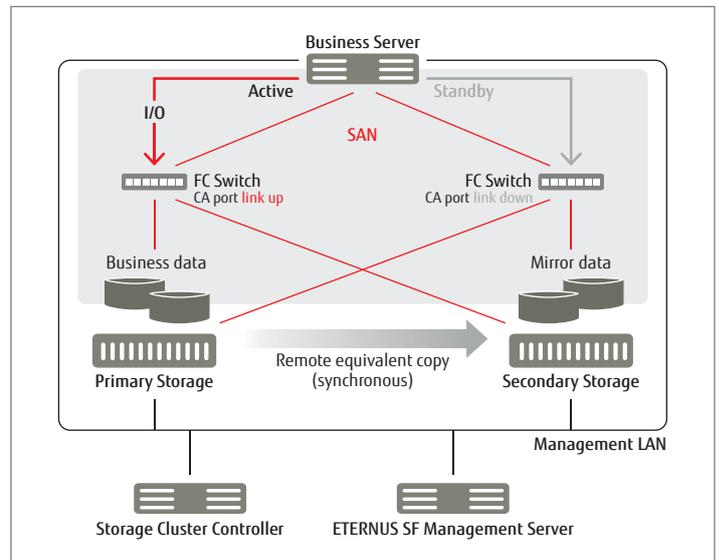
## Primary and Secondary Storage

The primary and secondary ETERNUS DX S3 storage array should preferably be located in different fire compartments – even better in different buildings or in metropolitan dispersion. Storage Cluster is set up using Transparent Failover Volumes (TFOV) which are part of a special copy group – the TFO group. Layout of the TFOVs and TFO groups are identical in the primary and secondary storage, even including the configuration settings for automated storage tiering, snapshots, etc. TFOVs are synchronously replicated from the primary to the secondary array.



## SAN Infrastructure

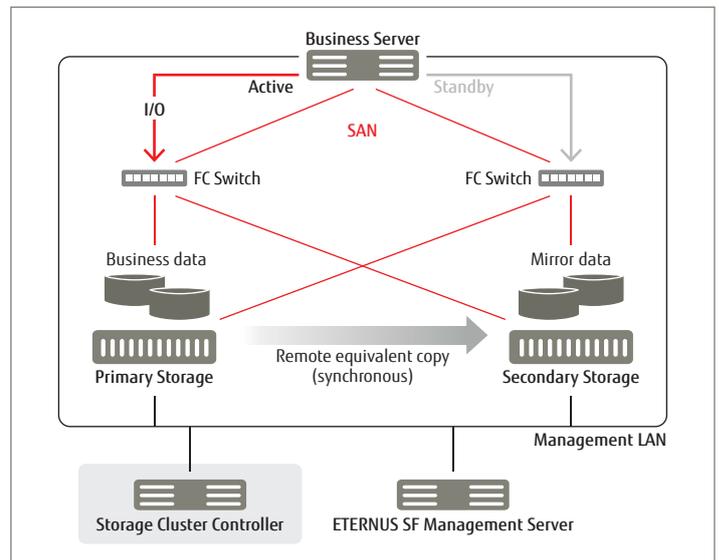
Primary and secondary storage have paired channel adapter (CA) ports. Paired means that the Fibre Channel ports on both sites have the same identity regarding WWN/WWPN. Under normal conditions the CA port on the primary site is in "link up" status, and the port on the secondary site is in "link down" status, so all server I/O is processed to the primary storage. The CA port states, as well as the REC state, are controlled by the Storage Cluster feature.



### Storage Cluster Controller

The Storage Cluster Controller is a server or virtual machine including an agent to connect with the ETERNUS SF management software. It monitors the health of primary and secondary storage in order to detect outages of the active system.

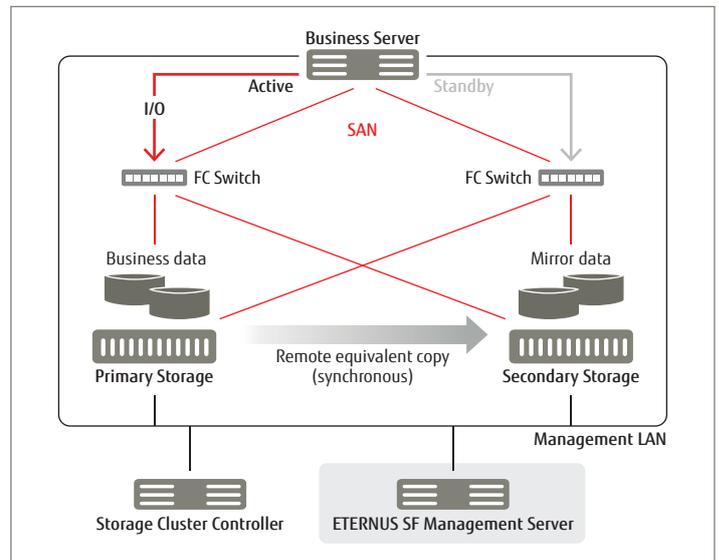
The Storage Cluster Controller triggers the automatic failover in this scenario only and is not involved in cases of administrator-triggered manual failover or in cases of automated failover caused by RAID failures.



### ETERNUS SF Management Server

ETERNUS SF management is the prerequisite for setting up the Storage Cluster configuration with regard to TFO groups, TFOVs, copy groups and REC pairs. It also executes the failover and failback operations, either triggered by the Storage Cluster Controller in cases of automatic failover, or manually by an operator. It also executes the automated failover in cases of RAID failures on the primary array.

ETERNUS SF and the Storage Cluster Controller can be installed on the same physical or virtual server.



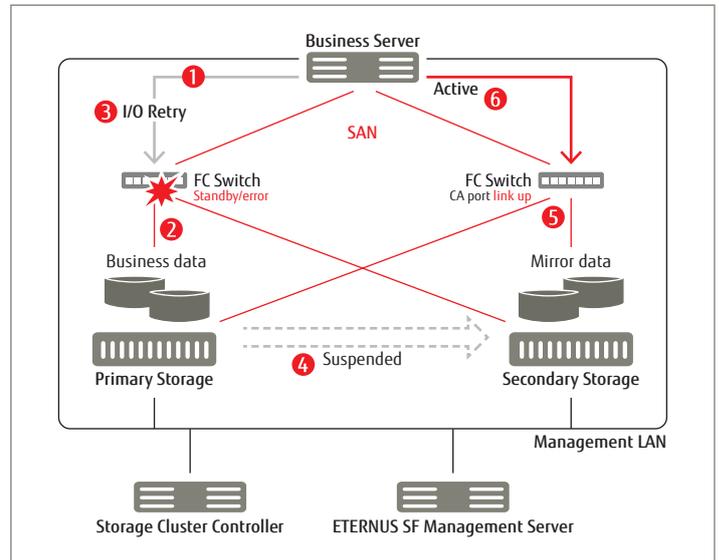
### Failover Mechanism

Storage Cluster reroutes I/O access from one array to the other as seen in this figure.

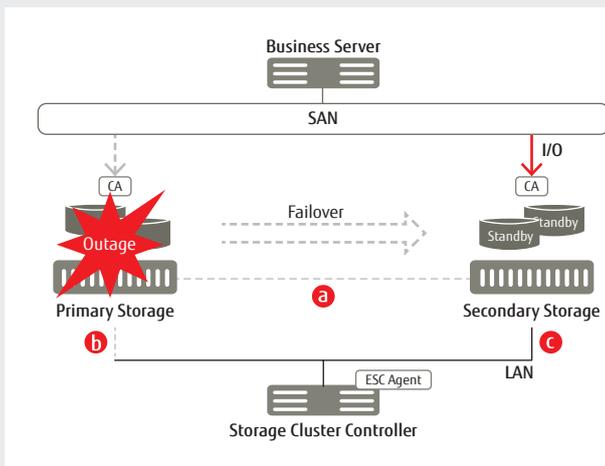
If an outage occurs the failover sequence is executed as follows:

- ① The server sends I/O requests to the primary storage.
- ② The primary CA port does not respond, the Storage Cluster Controller detects the primary ETERNUS is unreachable and reports it to ETERNUS SF.
- ③ The server retries the I/O after a preset time-out
- ④ ETERNUS SF suspends the remote mirroring (REC) session, the replicated data becomes the actual business data.
- ⑤ The CA port on the secondary array is activated (link up) with the same identity (WWN/WWPN) as the primary CA port.
- ⑥ The server I/O is processed by the secondary storage before the retry time-out is exceeded. The application continues running without any restrictions.

Such an automatic failover is typically completed within less than three seconds, which is sufficient for most applications to keep on running smoothly.



### Conditions for Automatic Failover



	Link status			Array Status	
	(a)	(b)	(c)	Primary	Secondary
1	OK	OK	OK	Normal	Normal
2	X	OK	OK	Normal	Normal
3	OK	X	OK		
4	OK	OK	X		
5	X	X	OK	Down	Normal
6	X	OK	X	Normal	Down
7	OK	X	X	Normal	Normal
8	X	X	X	Down	Down

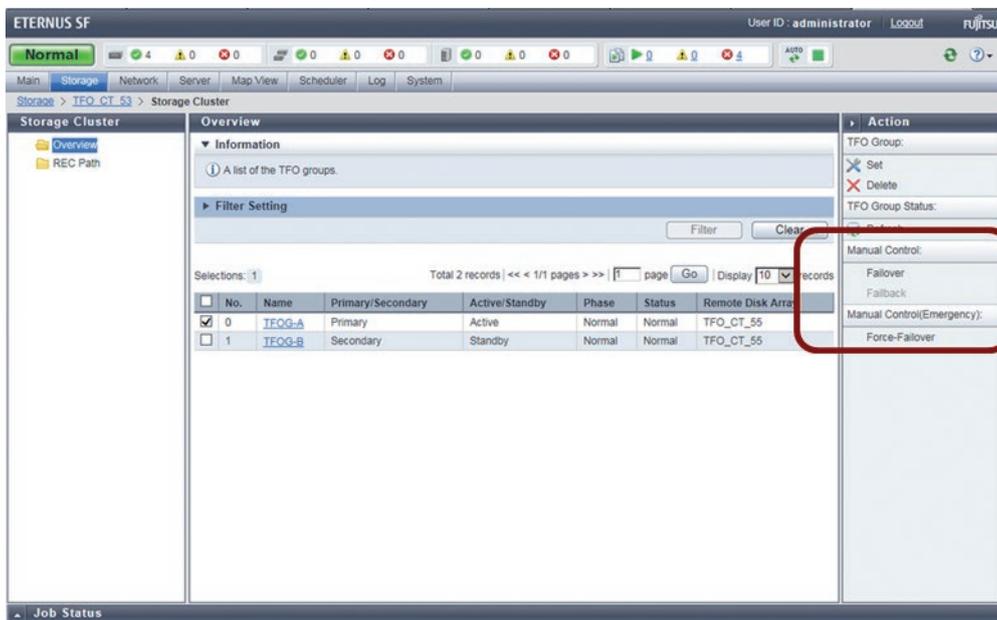
Automatic failover is only triggered if both the primary storage health supervision from Storage Cluster Controller (Link b) and the

remote replication (Link a) are down (5). In all other cases a failover is either not necessary or not possible (8)

## Failover and Failback

Storage Cluster can handle different failover types and scenarios:

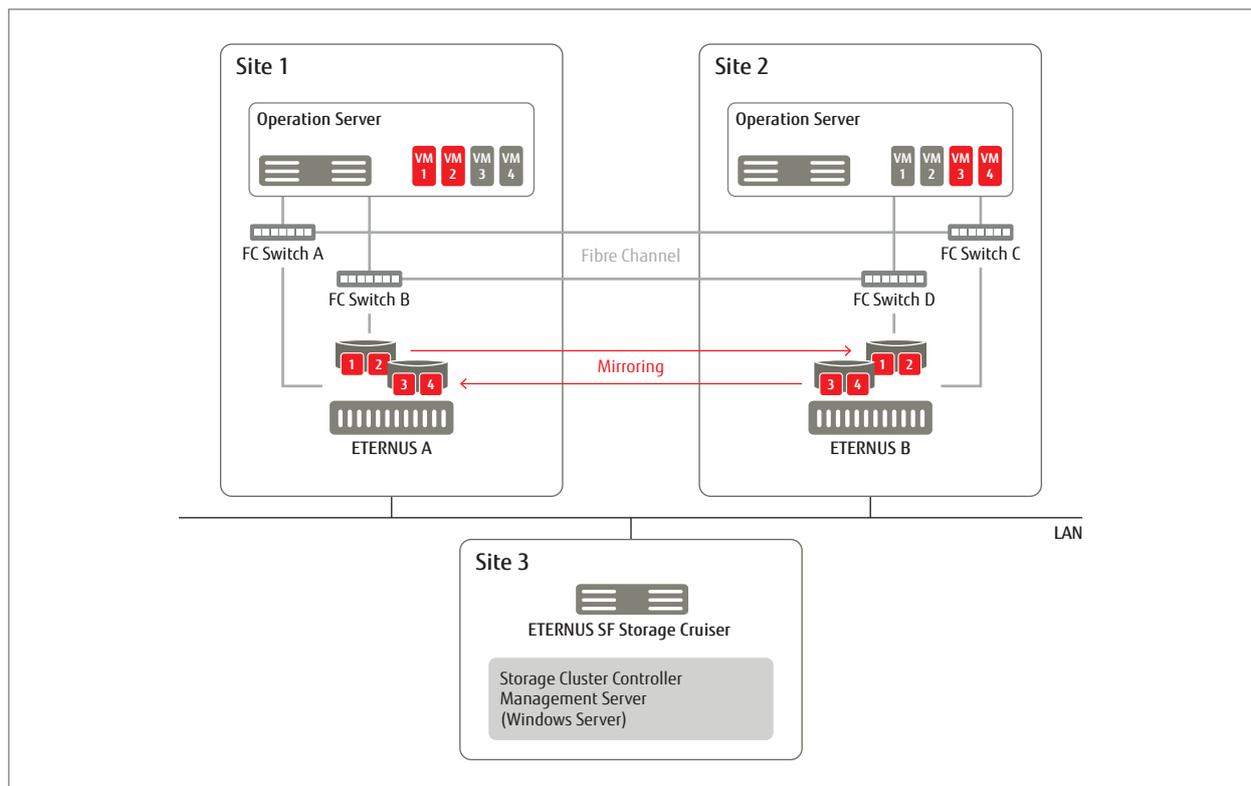
- **Auto failover:** Failover is triggered automatically when the primary storage becomes unreachable or a RAID group becomes unavailable. This mode ensures business continuity in cases of unpredictable failures or a disaster at the primary site.
- **Manual failover:** Failover is triggered manually from the ETERNUS SF user interface by stopping access on the primary storage and activating the secondary storage. This mode ensures business continuity when planned downtime is required on the primary site, e.g., for maintenance, disruptive upgrades or planned power shutdowns. It can also be used for general testing of the failover mechanism.
- **Force failover:** Failover is triggered manually from the ETERNUS SF user interface by activating the secondary storage regardless of the status of the primary storage. This mode ensures business continuity in cases of emergency when the primary storage is unreachable, and for any reason the auto failover cannot be executed.
- **Auto failback:** The failover back from the secondary site to the primary site is triggered automatically under these conditions: The Storage Cluster Controller confirms that all systems are operative, the REC session is established, and the business data and mirror data are consistent.
- **Manual failback:** The failover back from the secondary site to the primary site is triggered manually via the ETERNUS SF user interface. This mode resets normal operation manually. The conditions of auto failback apply as well.



# Storage Cluster in Virtualized Server Environments

In combination with VMware's HA/FT functionalities, Storage Cluster helps overcome even complete site outages in virtualized multisite server configurations. It provides instantaneous, non-disruptive failover in the event of server or site failures, delivering protection from even the slightest lapse, disruption or data loss.

Storage Cluster configuration is based on remotely replicated TFOVs which can be freely configured and paired. Thus on one site both primary and secondary TFO groups and TFOVs can be configured – connected via linked-up (active) or passive (linked-down) CA ports, respectively.



Under normal conditions, VM1 and VM2 at site 1 and VM1 and VM2 at site 2 are connected via Fibre Channel to the respective ETERNUS DX arrays located at the same site – this is called the active (primary) site. All data is synchronously replicated to the standby (secondary) site, while in this example ETERNUS DX 2 is secondary for VM1 and 2 and ETERNUS DX 1 is secondary for VM 3 and 4, respectively.

Both Fibre Channel ports on both sites have the same identity regarding WWPN. As the Fibre Channel port on the primary site is in "link-up" status and the port on the secondary site is in "link-down" status, all server inputs/outputs are processed to the primary storage.

# Conclusion

ETERNUS DX Storage Cluster provides zero-downtime, zero-data-loss continuous availability for any application, without the cost or complexity of traditional solutions like additional storage virtualization appliances or instances.

- Eliminates expensive downtime or data loss due to storage or site failures.
- Copes with both unplanned and planned downtime.
- Provides continuous service to any application, regardless of operating system.
- Provides easy setup through an intuitive administrative interface.

## Advantages – at a glance

Feature	Advantages
Transparent Failover	Ensures business continuity for any kind of outages
Automated control	Protection against hardware, network or environmental failures to prevent unplanned downtime
Manual control	Administrator-initiated failover supports business continuity in cases of planned downtime like disruptive upgrades
Symmetrical configuration	End-to-end continuous data availability for virtualized server environments
Ease of use	Simple system administration with wizard-based setup of transparent failover volumes and relationships
Flexible infrastructure options	Can be deployed for in-house fire compartments, campus-wide or metro-wide infrastructures

**Related documents:**

Global ETERNUS DX Storage Cluster Website

Flyer: ETERNUS Storage Cluster

Whitepaper: ETERNUS DX Feature Set

Whitepaper: ETERNUS SF

---

Published by  
Fujitsu Limited  
Copyright © 2015 Fujitsu Limited  
[www.fujitsu.com/eternus](http://www.fujitsu.com/eternus)

All rights reserved, including intellectual property rights. Technical data subject to modifications and delivery subject to availability. Any liability that the data and illustrations are complete, actual or correct is excluded. Designations may be trademarks and/or copyrights of the respective manufacturer, the use of which by third parties for their own purposes may infringe the rights of such owner.  
For further information see [www.fujitsu.com/eternus](http://www.fujitsu.com/eternus)