euronas

euroNAS SAN Cluster

Technology White Paper
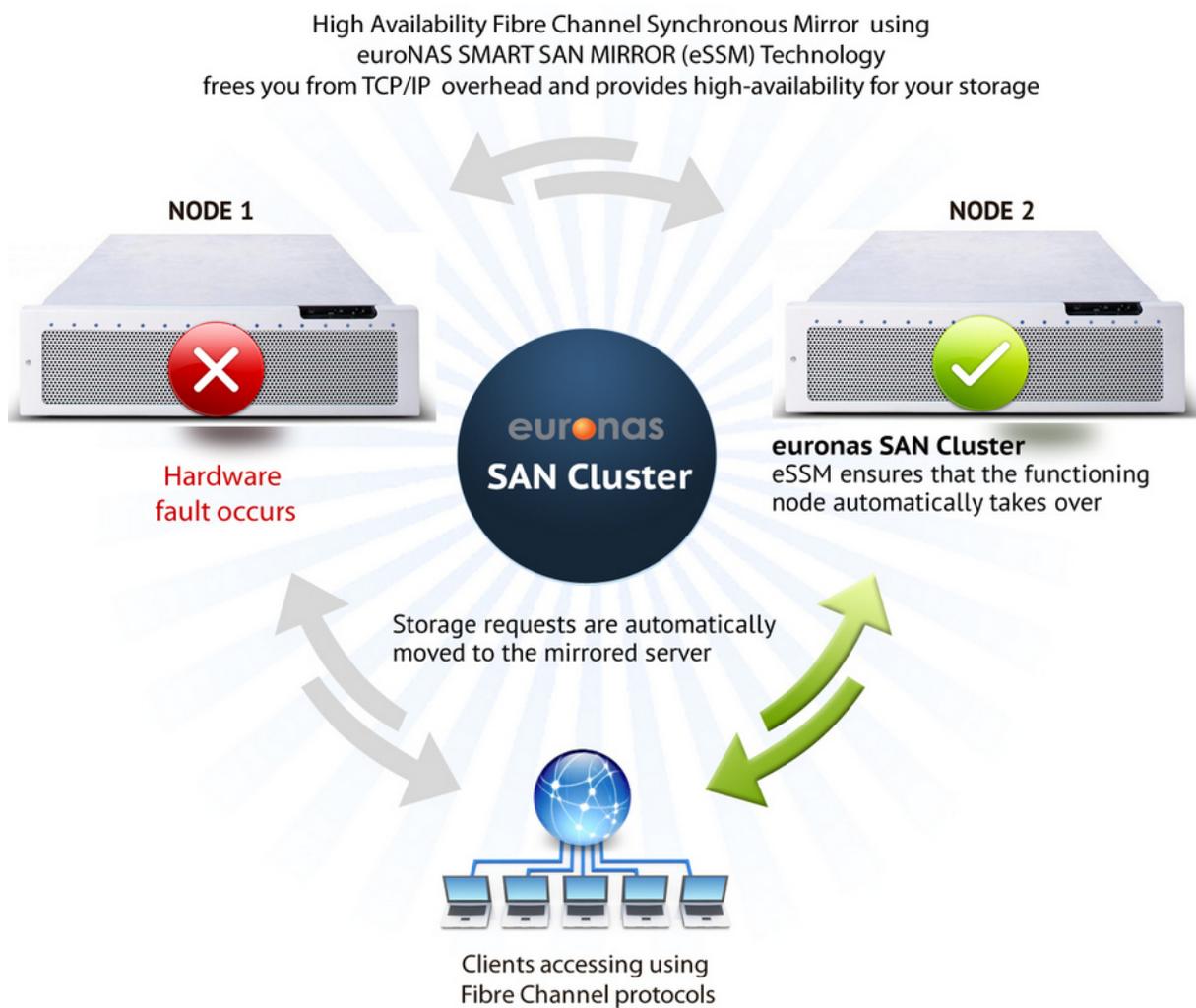
Rev. 14-09-19

Index

Introduction

This document describes how to use and build a highly-available SAN storage that ensures business continuity even in the event of total storage server failure.

Until recently, building a highly available, performance fibre channel storage that provides high availability was very expensive. euroNAS changes this by providing a powerful, scalable SAN solution that is both affordable and easy to deploy. It brings to small and midsize companies the enterprise SAN functionality, usually affordable only to large companies.

High Availability Fibre Channel Synchronous Mirror using euroNAS SMART SAN MIRROR (eSSM) Technology frees you from TCP/IP overhead and provides high-availability for your storage

NODE 1

NODE 2

Hardware fault occurs

euronas SAN Cluster eSSM ensures that the functioning node automatically takes over

euronas SAN Cluster

Storage requests are automatically moved to the mirrored server

Clients accessing using Fibre Channel protocols

Benefits of Fibre Channel

Our choice to build a product based on fibre channel high-end storage connectivity was because of its proven reliability, high bandwidth and its guaranteed in-order delivery of data packets.

What makes euroNAS SAN Cluster different from other SAN solutions?

On the market there are many SAN solutions providing high speeds and redundancy. They will protect you from disk failures or failure of one RAID controller.

euroNAS goes a step further by protecting you from a total failure of one of the units and also enabling you to simultaneously have the same data on 2 different places.

This will protect your data in case of some catastrophic events like fire, earthquake or theft and provide you business continuity at the same time.

Advantage of euroNAS SAN software not being locked to specific hardware

euroNAS runs on standard x64 servers. This enables you total freedom in choosing the hardware that fits to your performance, scalability and budget.

Legacy storage manufacturers often use proprietary hardware and an operating system which will only run on the hardware in question. This makes it difficult to expand in the future and forces you to buy new units. euroNAS will grow with the needs of your company. It is not locked to proprietary hardware – this makes it a highly flexible and scalable solution.

By selecting euroNAS storage software you are not only getting highest possible redundancy in case of server failure but also hardware tolerant operating system that runs on any hardware.

Why is Fibre Channel better than Ethernet?

Ethernet is one of the most popular network protocols on the market. It is low cost and provides relatively easy setup. However, Ethernet is not designed for transferring block data in a networked storage environment.

One of the greatest problems of the Ethernet is its way of handling data collisions – when more than one computer tries to transmit data simultaneously. Under heavy load condition, too many packet collisions will greatly reduce the whole network efficiency due to retransmissions.

Fibre channel on the other hand with its asynchronous protocol design ensures that even under heaviest load, collisions are handled efficiently and maintains maximum throughput.

What makes Fibre Channel better than iSCSI?

iSCSI has been developed as cost efficient alternative to fibre channel. It is built on an underlying TCP/IP protocol and is usually implemented only as a software initiator that incurs significant processing overhead.

Times have changed since iSCSI introduction - today prices for Fibre Channel are very close to enterprise NICs. In contrast to iSCSI, the fibre channel commands have been directly implemented into HBAs, making it much more efficient and less CPU consuming. It is more suitable for the data transfers.

One of the biggest advantages of Fibre Channel is constant data throughput. While fibre channel constantly performs very close to its max speed (8 Gb = 800 MB/s or 16 GB = 1600 MB/s), collision management on Ethernet makes it very difficult to achieve same consistent performance.

Network length is another argument for Fibre Channel. Currently maximum theoretical distance for long-haul Fibre Channel networks using fiber-optic links is 10 KM.

euroNAS SAN Cluster is a native fibre channel product. It is designed to provide critical shared data on SAN workloads

Fibre Channel is the safest and most reliable solution for data storage providing consistent high performance at the same time.

Benefits of euroNAS SAN Cluster

Higher availability

When using standard FC Storage your network applications depend on this single point of failure.

Whereas many fibre channel appliances will provide you with redundant controllers and components – most of them will not protect you from a total unit failure.

When using FC Cluster your network applications will continue to work - no matter which server fails, clients can still access their data. Even with proper backup policy it takes a while to recover and manual intervention to get server and data back online.

With FC Cluster, should a hardware or software fault occur, it is detected through intelligent features (eSSM) and the software automatically and immediately moves storage requests to the other server mirrored in the cluster.

euroNAS SMART SAN MIRROR (eSSM)

eSSM is a set of advanced replication and monitoring services built into euroNAS Fibre Channel Cluster. eSSM manages real-time server replication by providing controlled I/O optimization for the highest consistent performance.

It will provide you the

- Highest redundancy for your mission-critical applications

- Highest I/O performance thanks fibre channel technology

- Advanced Data Integrity protection against data corruption (T10)*

- Multipath I/O technology for even higher redundancy

- Synchronous Data Mirroring using native fibre channel protocol: real-time data mirroring across a 2 node storage cluster

- Automatic failover: services will automatically move on to working node

- Failback with fast synchronization: failback to an original system after an automatic failover

- FC port management – allocate FC Targets to single or multiple ports

VMware vSphere® Storage APIs – Array Integration (VAAI)

euroNAS SAN Cluster supports VMware vSphere® Storage APIs – Array Integration (VAAI), also referred to as hardware acceleration or hardware offload APIs. SAN Cluster will communicate with VMware ESXi and enable it to offload certain storage operations to the Cluster, which reduces resource overhead on the ESXi hosts and can significantly improve performance for storage-intensive operations. Please keep in mind that this functionality is not supported during failover.

Great distance between the servers

Fibre channel is perfect for the environments where data must be stored in different part of the building or on a different site. Depending on speed and cable type even distances of up to 10 km are possible.

Speed	OM1 Link Distance	OM2 Link Distance	OM3 Link Distance	OM4 Link Distance	OS1 Link Distance
	62.5 um core / 200 MHZ*km	50 um core/500 Mhz*km	50um core/2000 Mhz*km	50um core/4700 Mhz*km	9um core/~infinite Mhz*km
1 Gb	300 m	500 m	860 m	not defined	10 km
2 Gb	150 m	300 m	500 m	not defined	10 km
4 Gb	50 m	150 m	380 m	400 m	10 km
8 Gb	21 m	50 m	150 m	190 m	10 km
16 Gb	33 m	82 m	300 m	not defined	10 km
	15 m	35 m	100 m	125 m	10 km

Low Latency

SAN Cluster will help you speed up response and throughput providing you the lowest latency.

Simple Management interface

SAN Cluster is perfect for IT Professionals. You are not only gaining a high available storage solution with continuous replication but also an efficient and simple to use management interface.

SAN Cluster components and features

In order to achieve high-availability SAN Cluster consists of following parts

Cluster node

Cluster node represents an individual member server of the cluster.

Minimum requirements are 4 GB RAM and at least 2 QLogic Fibre Channel HBA ports. One is used for internal replication, other for providing FC targets to the clients.

Active node

Active node is the server that is providing storage to the clients. It replicates data to the passive node and makes sure that data is consistent on both nodes

Passive node

Passive node contains the same data as the active node and stays in standby mode as long as the other node is active. In case of failure of an active mode, it will automatically take over and continue to provide storage to the clients

Cluster drive

Cluster drive represents a mirror of 2 drives on each individual server.

This 2 drives can be a single disk or raid array. They are mirrored in real-time. On this drives Fibre Channel Targets are installed.

You can have up to 50 cluster drives however each additional drive will have impact on rebuild and failover time.

Corporate connection

Corporate connection is the FC connection used by clients for accessing fibre channel cluster targets. You can use more than one FC port for this connection.

Replication connection

Internal fibre channel connection used for data replication between the nodes.

You can use more than one FC port for this connection.

MPIO/ALUA (Asymmetric Logical Unit Access)

MPIO/ALUA support delivers high availability by establishing multiple sessions from a client to each node of SAN Cluster through Fibre Channel. In the event that a device in the path fails, I/O requests will be automatically redirected to an alternate path for continued application availability.

On euroNAS SAN Cluster the active node will be automatically recognized as active and the passive node will be automatically recognized as standby node by the client.

Cluster Events

Within cluster there are several types of events that can occur. Depending on type transfer of services can take different time. Initial time for servers to realize that other server has failed is around 30-40 seconds. After this resources are moved pretty quickly.

Estimated time to move around 10 resources on another server is around 10-20 seconds. To move around 30 resources usually takes about 50 seconds.

Failover

Failover occurs when server detects server fault and automatically moves all resource on another server. On failover services are moved faster on another server. One example would be total crash of one of the servers.

Switchover

Switchover happens when access to a resource is manually moved from one server to another. This is usually done if you wish to perform system maintenance of one of the nodes. Moving services takes little bit longer than on failover.

Hardware recommendations

Installation Disk

For best performance the euroNAS cluster software should be installed on an SSD with a minimum size of 8 GB. Installation disk is separate from data.

Data Disk

Since installation disk is exclusively used for the OS, for data at least one additional disk is needed. Data disk can be simple disk or disks managed by the hardware RAID controller.

System RAM

The server should have at least 8 GB of RAM

Fibre Chanel HBA

Per node at least 2 fibre channel ports are required. One port will be used for replication, other for presenting FC targets to the initiators.

Currently for use with SAN Cluster only QLogic HBAs are supported. On clients side it can be any controller.

Networking

Networking is not used for replication but it is necessary for management and internal communication between the nodes.

There must be one or more reliable IP Addresses on the Corporate Network that both servers can reach for proper operation of the High-Availability/Custer features

Since it is only used for internal communication and management – 1 Gigabit network is sufficient.

DHCP should not be used unless static leases are provided

Using DHCP without static leases will cause the cluster to lose communication to other node and fail on IP address change

Performance recommendations

The bottleneck of SAN Cluster is the disk I/O and the speed of the FC controller. We recommend the use of SAS or SSD disks and hardware RAID controller.

euroNAS HA Cluster is hardware independent however we still recommend using identical hardware. Due to server replication in real time, the performance will depend on the weakest link. Using slower disks on one server will slow the server with faster disks as well.

Limitations

SAN Cluster detects automatically many possible fail scenarios and moves resources automatically. There are some scenarios where SAN cluster is unable to detect the problem.

Powering down master node during sync process

During the sync process the node with more current data must not be powered off. In this case server replication service will stop the drives in order to prevent data corruption.

Only server with less current data can be powered off. This does apply if both nodes are in sync. In this case either of the servers can be powered off or rebooted.