



# **HIGH AVAILABILITY CLUSTER**

## **Technologie White Paper**

Rev. 16-03-29

## Index

<b>Einleitung .....</b>	<b>4</b>
<b>Vorteile von HA Cluster .....</b>	<b>5</b>
<b>Hochverfügbare Architektur .....</b>	<b>6</b>
<b>Cluster Ereignisse.....</b>	<b>7</b>
<b>Hardware-Empfehlungen.....</b>	<b>8</b>
<b>Performance-Empfehlungen.....</b>	<b>9</b>
<b>Einschränkungen.....</b>	<b>11</b>
<b>Beispielkonfigurationen.....</b>	<b>12</b>

Copyright © 2015 euroNAS GmbH. All Rights Reserved.

euroNAS believes that the information in this publication is accurate of its publication date. The information in this publication is provided “as it is”. euroNAS GmbH makes no warranties of any kind with respect to the information in this publication. euroNAS GmbH specifically disclaims implied warranties or fitness for a particular purpose.

Microsoft, VMware, ESX, Xen and all other trademarks are registered trademarks and the property of their respective owners.

## Einleitung

Dieses Dokument erklärt die hochverfügbare Technologie hinter euroNAS HA Cluster und gibt eine kurze Übersicht über die Funktionalität.

In den letzten Jahren, Bedarf nach hochverfügbarem Speicher ist exponentiell gewachsen. Die Unternehmen werden zunehmend abhängig von der Verfügbarkeit deren Daten.

Von den Produktivsystemen werden eine ständige Erreichbarkeit und eine sehr schnelle Wiederherstellung in einem Notfall erwartet.

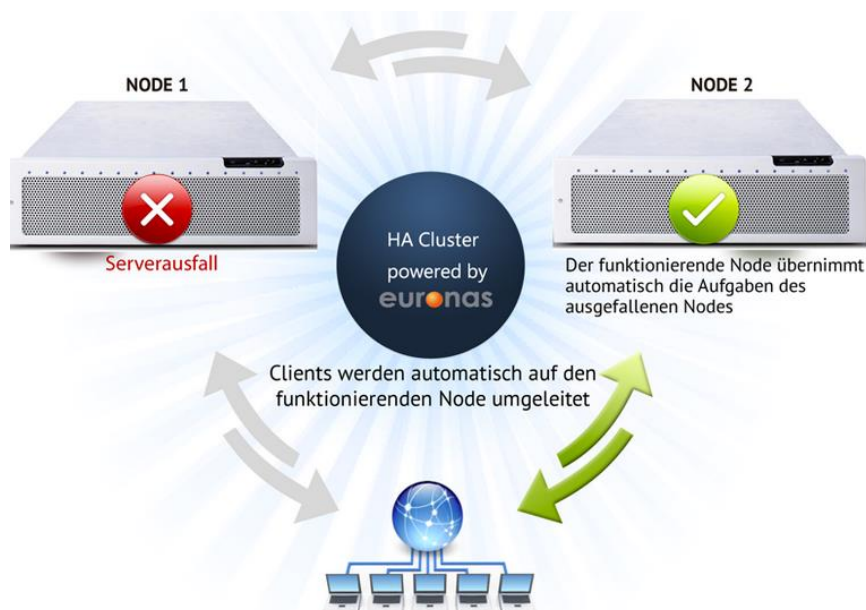
Mehr als jemals zuvor ist die Businesskontinuität selbst bei einem Totallausfall wichtig.

euroNAS HA Cluster erfüllt diese Anforderungen und bietet eine Storage-Spiegelung zwischen den Storage Servern in Echtzeit an.

Mit der Wahl von euroNAS Cluster bekommen Sie nicht nur die maximale Erreichbarkeit und Redundanz sondern auch ein flexibles Storage Betriebssystem welches nicht an einen bestimmten Hersteller gebunden ist.

Klassische Storage-Hersteller verwenden proprietäre Hardware und Software welche nur für deren Hardware bestimmt ist. Damit ist die Flexibilität für künftige Erweiterungen oder Hardwaretausch im Notfall sehr eingeschränkt. Bei sehr hohen Kosten ist man dem Hardwarehersteller ausgeliefert.

euroNAS ist sehr flexibel, nicht gebunden an bestimmte Hardware und jederzeit skalierbar. euroNAS kann mit Ihrem Unternehmen wachsen und damit die TCO Kosten enorm reduzieren. Die Wiederherstellung von dem Storage Server ist durch diese Freiheit wesentlich schneller, günstiger und einfacher.



## Vorteile von HA Cluster

### Hohe Verfügbarkeit

Standard NAS Server können Sie nicht gegen einen Totalausfall sichern – diese können RAID, redundante Netzteile oder mehrere Netzwerkkarten bieten, bleiben aber trotzdem ein “Single Point of Failure”.

Selbst mit ordentlichen Datensicherungen kann die Wiederherstellung mehrere Stunden dauern.

HA Cluster bietet alle Vorteile eines Standard NAS Servers mit zusätzlicher Hochverfügbarkeit. Selbst bei einem Totalausfall eines der gespiegelten Server wird der noch funktionierende Server die Netzwerkdienste nach einem kurzen Übergang weiterhin anbieten und die Clients können ganz normal weiterhin auf deren Daten zugreifen.

HA Clusterdienste werden automatisch Hardwarefehler entdecken und die Abfragen auf den noch funktionierenden Knoten umleiten.

### Datenspiegelung

Die Daten werden auf beiden Servern in Echtzeit gespiegelt - der euroNAS HA Cluster besitzt zusätzlich intelligente Dienste, die bei einem Serverausfall automatisch auf den anderen Server umschalten ohne dass ein manueller Eingriff notwendig ist.

### Redundante Dienste

euoNAS HA Cluster bietet alle relevanten Netzwerkprotokollen die maximale Flexibilität anbieten.

Alle diese Dienste können gleichzeitig benutzt werden. Folgende Dienste werden zurzeit unterstützt.

- SMB/CIFS/AFP
- Active Directory Authentifizierung (für CIFS und AFP)
- NFS
- iSCSI (incl. persistent reservation iSCSI Lun und Target Dienste)

### Höhere Performance durch die Ressourcenverteilung (active/active)

Cluster Management ermöglicht Ihnen zu entscheiden auf welchem Knoten eine Netzwerkfreigabe / iSCSI Target laufen sollte. Auf diesem Weg können beide Knoten effizienter benutzt werden und maximale Leistung anbieten. Bei einem Ausfall eines der Knoten, übernimmt der Funktionierende automatisch seine Dienste. Sobald der ausgefallene Knoten wieder online ist, werden die zu ihm zugewiesene Freigaben/iSCSI Targets automatisch wieder umgezogen.

### Einfache Verwaltung

HA Cluster Verwaltungsoberfläche basiert auf jahrzehntelanger Erfahrung und wurde auf maximale Einfachheit optimiert. Sie bekommen also nicht nur hochverfügbare Storage Lösung sondern eine Verwaltung die viel einfacher zu verwalten ist als die meisten Storage Lösungen auf dem Markt.

## Hochverfügbare Architektur

HA Cluster besteht aus folgenden Teilen

### Cluster Knoten

Cluster Knoten repräsentiert individuelles Mitglied im HA Cluster. Beide Knoten sind gleich – es ist egal über welchen der Cluster verwaltet oder überwacht wird.

### Cluster-Laufwerk

Clusterlaufwerk ist ein Spiegel von 2 Laufwerken auf den individuellen Servern.

Diese 2 Laufwerke können ein Laufwerk, Software oder Hardware Raid sein. Diese spiegeln sich zwischen den beiden Servern in Echtzeit. Auf diese Laufwerke werden Freigaben und iSCSI Targets konfiguriert und installiert.

Diese müssen nicht den vollen Speicherplatz der physikalischen Laufwerke belegen. Es ist möglich diese über beide Knoten zu verteilen.

Ein Beispiel :

Jeder Server hat einen RAID-Verbund von 16 TB. Sie könnten z.B. daraus 2 Clusterlaufwerke von 8 TB erstellen. Ein Clusterlaufwerk können Sie auf Knoten1 und den anderen auf Knoten2 laufen lassen. Dadurch kann die Last zwischen den Knoten ausgeglichen werden.

### Cluster-Ressource

Cluster-Ressourcen sind einzelne Netzwerkfreigaben oder iSCSI Targets welche Sie erstellt haben. Diese können SMB/CIFS, AFP oder NFS, bzw. iSCSI Target sein. Jede Ressource hat eigene Größe und ist über die IP Adresse des Clusterlaufwerks auf dem es sich befindet erreichbar

### Firmennetzwerk

Firmennetzwerk wird von Netzwerkbenutzern verwendet. Wenn der Cluster eingerichtet ist, ist die "Netzwerk-Zugriffs-IP-Adresse" für dieses Netzwerk auch eingerichtet.

### Replikationsnetzwerk

Replikationsnetzwerk wird für die Echtzeitreplikation, bzw. die Kommunikation zwischen den beiden Servern verwendet.

Replikationsnetzwerk sollte sich in einem anderen Subnetz als Firmennetzwerk befinden, wenn möglich, sollten diese auch physisch voneinander getrennt werden (z.B. auf einem anderen Switch).

### Netzwerk-Zugriffs-IP-Adresse

Cluster-Zugriffs-IP-Adresse ermöglicht Netzwerkdiensten und Benutzern auf die Freigaben und iSCSI Targets zuzugreifen. Der größte Vorteil ist, dass diese IP Adresse von beiden Servern geteilt wird - d.h. sollte ein Server ausfallen, übernimmt der andere Server automatisch diese Adresse ohne, dass die Benutzer einen Unterschied feststellen.

## ZU BEACHTEN

- Clusterfreigaben und iSCSI Targets sind nicht verfügbar bis eine Cluster-Zugriffs-IP-Adresse eingerichtet ist
- Sie können mehr als eine Cluster-Zugriffs-IP-Adresse pro Clusterlaufwerk definieren
- Cluster-Zugriffs-IP-Adresse verweist immer auf ein Clusterlaufwerk
- IP Adresse muss sich im Firmennetzwerk befinden

### Netzwerk Test IP Adresse

Während der Clustererstellung werden Sie nach einer "Netzwerk Test IP Adresse" gefragt.

Diese IP Adresse verwenden beide Server um zu sehen ob das Netzwerk noch verfügbar ist. Beide Server werden regelmäßig diese IP Adresse anpingen.

Es wird dringend empfohlen mehrere IP Adressen dort einzutragen. Das können die IP Adressen von Ihrem Router, Gateway, Mailserver, Netzwerkdrucker etc. ... sein.

## Gemeinsamer Speicher oder Spiegelung ?

HA Cluster unterstützt 2 verschiedene Arten von Speicheranbindung

### Storage-Spiegelung

Beide Knoten verwenden den eigenen Speicher der dann übers Netzwerk gespiegelt wird

Die Vorteile

- Datensicherheit – beide Knoten spiegeln die Daten – man kann z.B. die Daten in 2 verschiedenen Standorten oder Brandabschnitten gleichzeitig halten
- Kein „Single Point of Failure“ - Daten sind doppelt vorhanden

Die Nachteile

- Die Performance leidet etwas darunter, daß die Daten gespiegelt werden
- Kosten – man benötigt die doppelte Anzahl an Laufwerken

### Gemeinsamer Speicher

In diesen Konfigurationen verwenden beide Knoten einen gemeinsamen Speicher. Der Speicher wird via SAS Multipath Backplane, Fibre Channel oder iSCSI angebunden.

Die Vorteile

- Kostenersparnis (man benötigt nicht die doppelte Kapazität)
- Performance (man verliert keine Leistung wegen der Spiegelung)

Die Nachteile

- Die Daten werden nicht gespiegelt
- Wenn der gemeinsame Speicher ausfällt fallen beide Knoten aus



## Cluster-Ereignisse

Innerhalb des Clusters können verschiedene Ereignisse auftreten welche zu einem Umzug der Dienste auf den anderen Knoten führen können.

Abhängig von der Art kann der Umzug auch unterschiedlich lange dauern. Diese Zeiten sind auch von der Anzahl der Dienste ebenfalls abhängig.

In der Regel dauert es ca. 30 Sekunden bis der Cluster entdeckt hatte, daß der andere Knoten nicht mehr verfügbar ist, danach werden die Ressourcen schnell umgezogen.

Der Umzug von 10 Ressourcen auf einen anderen Server dauert in der Regel ca. 50-60 Sekunden. Bei 30 Ressourcen dauert es ca. 70-80 Sekunden.

Diese Zeiten können sich je nach Hardware und Speichergröße unterscheiden.

### Failover

Failover passiert wenn die Cluster-Dienste merken, daß ein Knoten ausgefallen ist. Die Dienste werden schneller als bei einem Switchover auf den noch funktionierenden Knoten umgezogen.

### Switchover

Switchover passiert, wenn die Ressourcen manuell von einem Server auf den Anderen umgezogen werden. Ein Beispiel wäre die Wartung eines der Knoten. Dieser Vorgang dauert in der Regel etwas länger als Failover.

## Hardware-Empfehlungen

### Installationslaufwerk

Für die optimalste Performance sollte euroNAS Cluster Software auf ein SSD Laufwerk von mind. 8 GB Größe installiert werden. Installationslaufwerk ist immer von den eigentlichen Daten getrennt.

Viele Hardware RAID Controller erlauben mit den gleichen Laufwerken 2 RAID-Verbünde zu erstellen, in diesem Fall sollte ein kleineres für das euroNAS Betriebssystem und ein großes für die Daten erstellt werden.

### Datenlaufwerke

Da Installationslaufwerke exklusive für das Betriebssystem verwendet werden, sind für die Daten zusätzliche Laufwerke notwendig.

Datenlaufwerk kann eine einfache Festplatte, Software oder Hardware RAID sein.

Sie können S-ATA, SAS, SSD und FC Laufwerke verwenden. Eine Liste der unterstützten Hardware RAID Controllern finden Sie auf unserer Webseite.

### Arbeitsspeicher

Mindestanforderung sind 8 GB RAM

### Netzwerk

Es müssen mehrere IP Adressen für das Firmen- und Replikationsnetzwerk konfiguriert werden.

Die beste Performance und Sicherheit erreicht man mit 4 individuellen Netzwerkkarten mit einer Mindestbandbreite von 1 Gigabit.

Diese sollten als 2 separate Netzwerkteams gebündelt werden.

Das eine Netzwerkbandel sollte mit dem Firmennetzwerk und das Andere mit dem Replikationsnetzwerk verbunden werden.

Sollten Dual-Port Netzwerkkarten verwendet werden, sollten die einzelnen Ports über die Netzwerke verteilt werden, in diesem Fall ist die Redundanz beim Ausfall einer Netzwerkkarte gesichert.

Ein Beispiel :

Karte 1 Port 1 -> bond0 (Firmennetzwerk)

Karte 1 Port 2 -> bond1 (Replikationsnetzwerk)

Karte 2 Port 1 -> bond0 (Firmennetzwerk)

Karte 2 Port 2 -> bond1 (Replikationsnetzwerk)

Die beste Performance erreicht man mit 10 Gigabit. Sollten die Netzwerkkarten für das Replikationsnetzwerk direkt, ohne Switch, angeschlossen werden empfehlen wir den Bonding Modus „Broadcast“.

DHCP darf nicht verwendet werden. Beim Wechseln der IP Adresse geht die Clusterkonfiguration verloren da diese an die Netzwerkeinstellungen gekoppelt ist.

## Performance-Empfehlung

Der Flaschenhals von HA Cluster ist meistens das Replikationsnetzwerk. Die Performance von dem gesamten HA Cluster hängt von der Verbindungsqualität zwischen den beiden Knoten. Für die optimalste Performance empfehlen wir 10 Gigabit Netzwerk.

Für die beste I/O Performance empfehlen wir Hardware RAID Controller oder SSD Laufwerke.

euoNAS HA Cluster ist Flexibel welche Hardware Sie einsetzen, trotzdem empfehlen wir für beide Knoten identische Hardware einzusetzen. Dadurch dass die Daten in Echtzeit auf beide Knoten gespiegelt werden wird die Performance von der langsamsten Komponente abhängen. Sollten z.B. langsamere Laufwerke auf dem 2. Knoten verwendet werden, werden die schnelleren Laufwerke auf dem 1. Knoten auf die langsamere Laufwerke warten müssen.

## Einschränkungen

HA Cluster ist in der Lage sehr viele Ausfallszenarien zu entdecken. Es gibt allerdings einige Szenarien welche nicht automatisch entdeckt werden können und ein manueller Eingriff notwendig ist.

### Replikationsnetzwerk ausgefallen, Firmennetzwerk über beide Knoten erreichbar

In diesem Fall denken beide Knoten, dass sie online sind und erreichbar. Um mögliche Datenverluste ("Split-Brain") zu vermeiden werden alle Dienste angehalten bis das Replikationsnetzwerk wieder online ist.

Solche Situationen lassen sich in der Regel mit Netzwerkportbündelung vermeiden.

### Masterknoten geht während der Synchronisation offline

Wenn beide Server sich nicht in Sync befinden und abgeglichen werden sollte der Server mit aktuelleren Daten nicht ausgeschaltet werden. In diesem Fall werden die Clusterdienste angehalten um mögliche Datenverluste zu vermeiden.

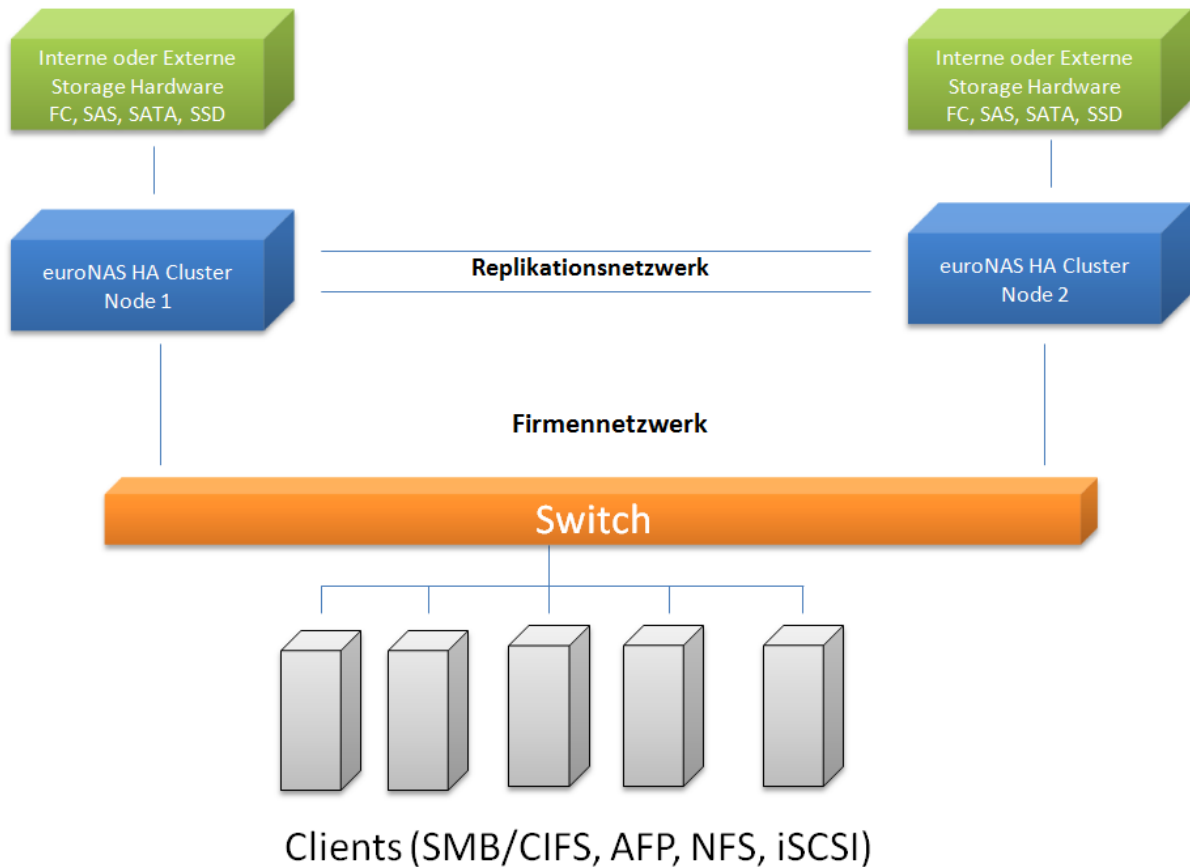
Nur der Server mit älteren Daten darf ausgeschaltet werden.

### Broadcast Modus für die Netzwerkportbündelung sollte für die Konfigurationen ohne Switch verwendet werden

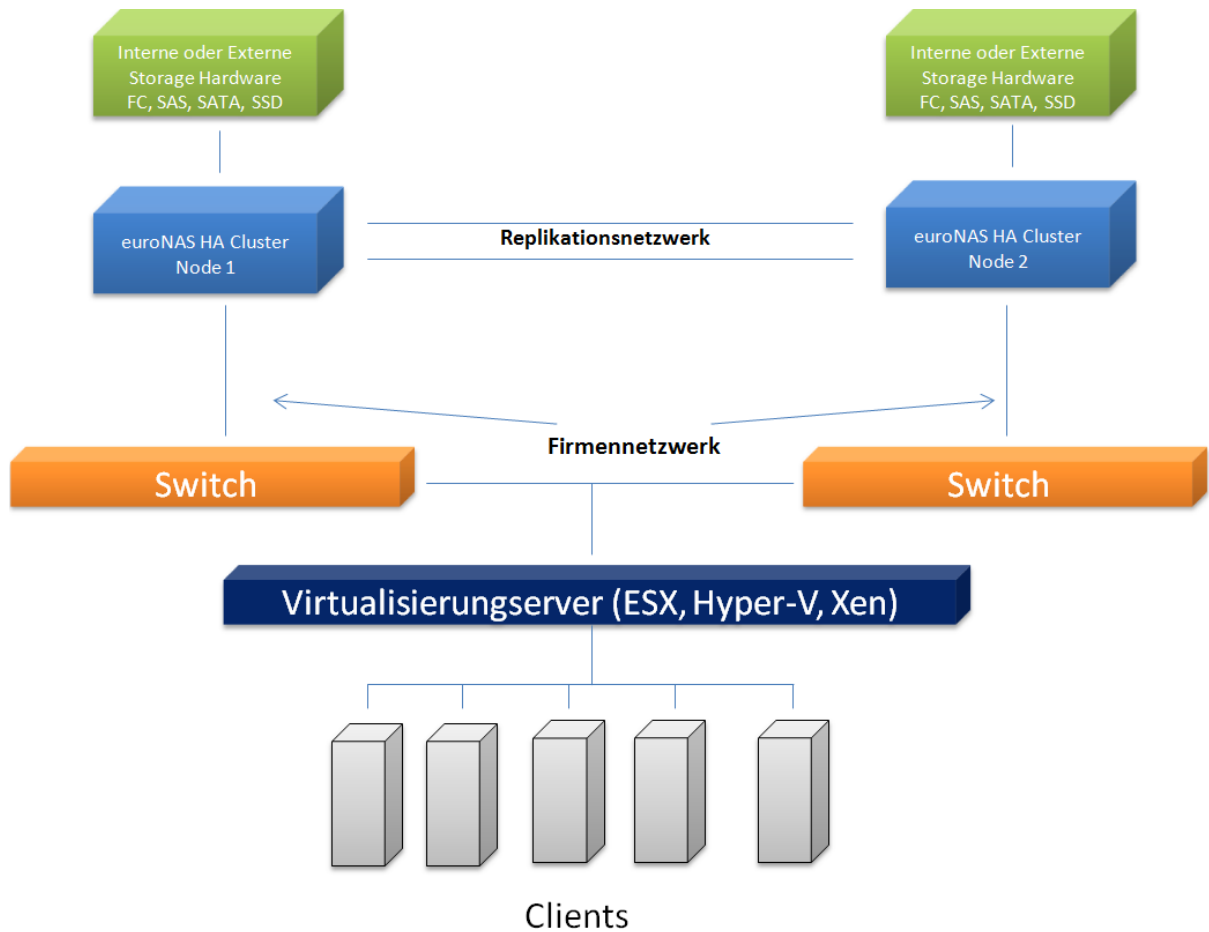
Andere Modi funktionieren ebenfalls aber in der Praxis hat sich diese Art der Verbindung als die stabilste Konfiguration in Umgebungen ohne Switch erwiesen.

## Beispielkonfigurationen (Storage-Spiegel)

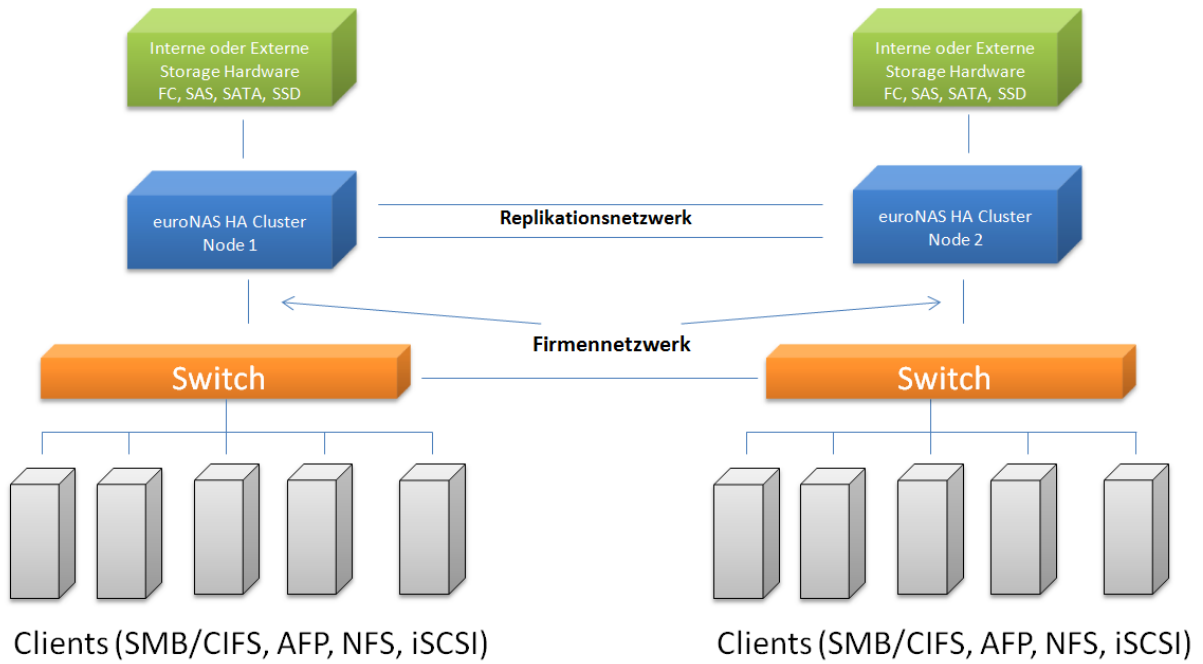
Aufgrund seiner großen Flexibilität gibt es sehr viele Möglichkeiten eine redundante Konfiguration mit euroNAS HA Cluster aufzubauen. Hier werden nur einige der Möglichkeiten angezeigt. Sie können auch komplett auf die Hardware verzichten und den HA Cluster als "Virtual Storage Appliance" innerhalb einer virtuellen Maschine betreiben (s. 4. und 5.)



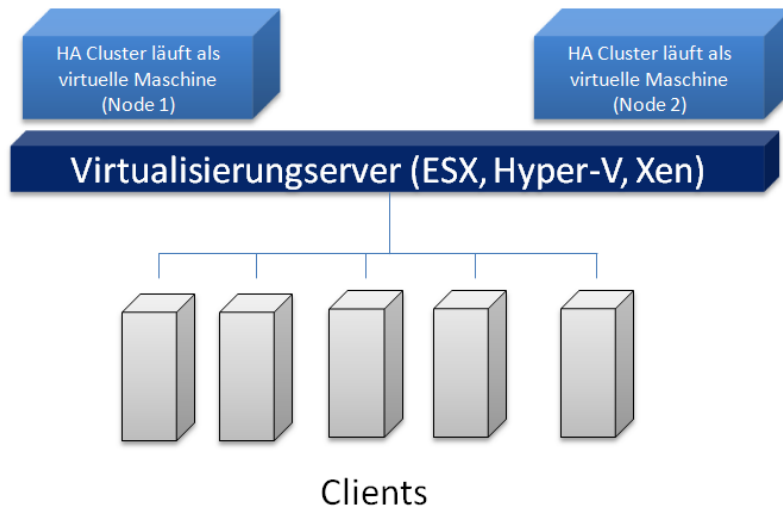
1. In dieser Konfiguration sind beide Cluster Knoten direkt via Switch verbunden. Die Clients können auf die Daten über das Firmennetzwerk zugreifen. Sollte einer der HA Cluster Knoten ausfallen, übernimmt der andere automatisch seine Dienste.



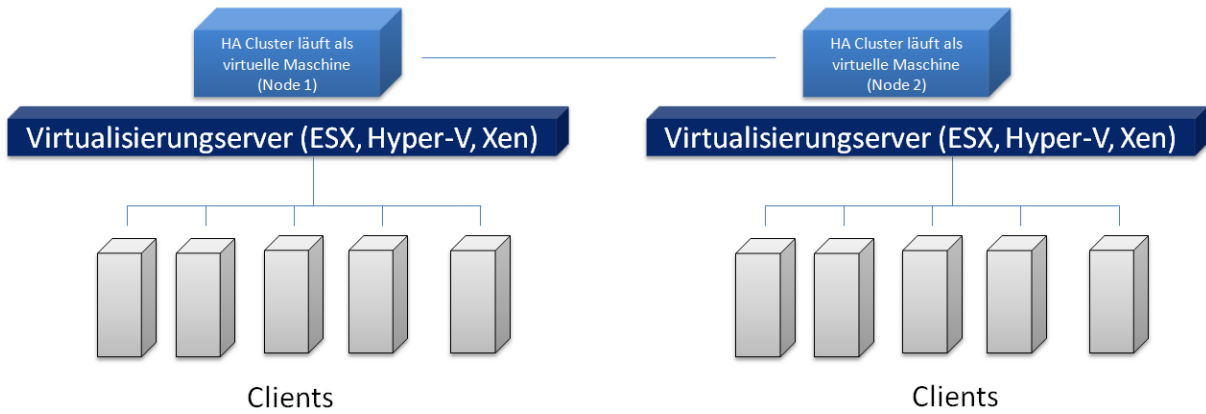
2. Diese Konfiguration zeigt eine redundante Verbindung zu einem oder mehreren Virtualisierung-Servern. Die Anbindung erfolgt via NFS oder iSCSI. Virtualisierung-Server verwenden dann den Speicher für die virtuelle Maschinen, bzw. An die angebundene Clients dieser VMs.



3. In dieser Konfiguration gibt es keine "Single-Point-of-Failure". Jede Komponente ist redundant ausgelegt – dabei können sich die Komponenten in verschiedenen Brandabschnitten oder Standorten befinden.

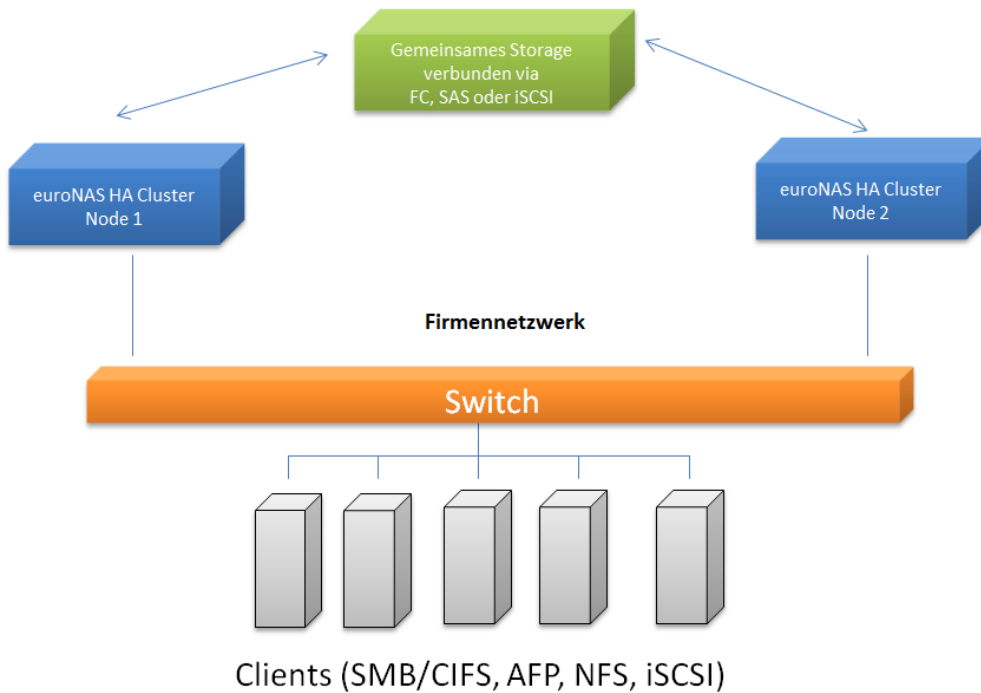


4. In dieser Konfiguration werden die beiden Knoten als Virtuelle Storage Appliance innerhalb des Virtualisierung-Servers innerhalb einer virtuellen Maschine betrieben.



5. In dieser Konfiguration werden die beiden Knoten als Virtuelle Storage Appliance innerhalb des Virtualisierung-Servers innerhalb einer virtuellen Maschine betrieben. Da sie sich auf 2 verschiedenen Virtualisierung-Servern befinden entfällt der "Single-Point-of-Failure".

### Beispielkonfiguration (Gemeinsamer Speicher)



6. In diesen Konfigurationen verwenden beide Knoten einen gemeinsamen Speicher. Der Speicher wird via SAS Multipath Backplane, Fibre Channel oder iSCSI angebunden.