White Paper

# Synology High Availability (SHA)

## Based on DSM 4.3

Synology Inc.

# Table of Contents

Chapter

1

# Introduction

Uninterrupted availability is a critical goal for all businesses; however, as many as 50% of SMBs worldwide remain unprepared for disaster[1]. Moreover, downtime costs a median of 12,500 USD daily. Assuming a median of six downtime events per year, the cost of unpreparedness begins to stack up.

The **Synology High Availability** solution helps users overcome this hurdle by ensuring non-stop storage services with maximized system availability to decrease the risk of unexpected interruptions and costly downtime.

[1] Symantec 2011 SMB Disaster Preparedness Survey, **http://www.symantec.com/about/news/resources/press_kits/detail.jsp?pkid=dpsurvey**

Chapter

2

# High-Availability Clustering

## 2.1 Synology High-Availability Cluster

The Synology High Availability solution is a server layout designed to reduce service interruptions caused by system malfunctions. It employs multiple servers to form a “**high-availability cluster**” (sometimes called “HA cluster) consisting of two compatible Synology servers. Once this high-availability cluster is formed, one server assumes the role of the active server, while the other acts as a stand-by passive server.

## 2.2 Service Continuity

Once the high-availability cluster is formed, data is continuously replicated from the active to the passive server. All files on the active server will exist in duplicate on the passive server. In the event of a critical malfunction, the passive server is ready to take over all services, equipped with a mirrored image of all data on the active server, allowing the high-availability cluster to continue functioning as normal, reducing downtime.

## 2.3 Data Replication Process

Within the high-availability cluster, all data stored in internal drives or expansion units will be replicated. Therefore when services are switched from the active to passive server, no data-loss will occur.

While data replication is a continual process, it has two distinct phases spanning the formation to the operation of a high-availability cluster:

- **Phase 1:** The initial data replication during cluster creation or the replication of differential data when connection to the passive server is resumed after a period of disconnection (such as when the passive server is switched off for maintenance). During this phase, the initial sync is not yet complete, and therefore switchover cannot be performed. Data changes made on the active server during this initial replication are also synced.
- **Phase 2:** Real-time data replication after the initial sync has been completed. After the initial sync, all data is replicated in real-time and treated as committed if successfully copied. In this phase, switchover can be performed at any time.

During both phases of data replication, all data syncing is performed at block-level. For example, when writing a 10 GB file, syncing and committing is broken down to block-level operations, and completed piecemeal to ensure that the active and passive servers contain identical data. As all data is maintained constantly up to date, switchover can be accomplished seamlessly.

Data or changes to be replicated include:

- **NAS Data Services:** All file services including CIFS/NFS/AFP are covered.
- **iSCSI Data Services:** High-availability clustering supports iSCSI, including iSCSI LUN and iSCSI Target services.
- **DSM and Other Services:** Management applications, including Synology’s DiskStation Manager (DSM) and its other services and Add-On Packages, e.g. Mail Server, Directory Server, will also be covered, including all settings and service statuses.

Chapter

3

# High-Availability Cluster Architecture

## 3.1 Physical Components

Synology's High Availability solution constructs a cluster composed of two individual storage systems, including an active and a passive server. Each server comes with attached storage volumes, and the two are linked by a "Heartbeat" connection which monitors server status and facilitates data replication between the two servers.

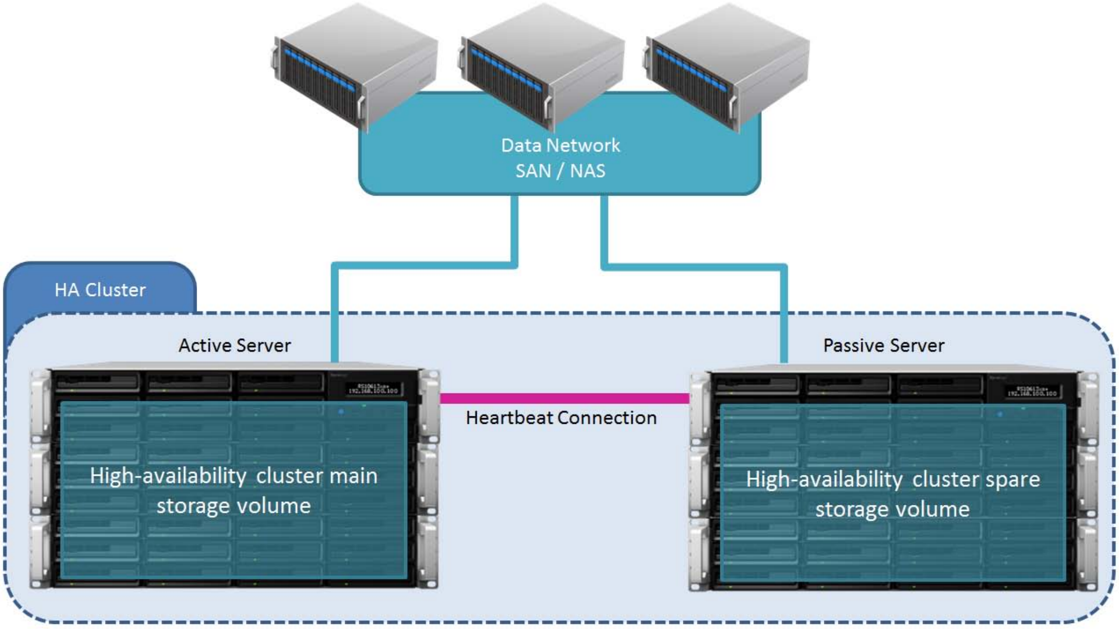
Data Network
SAN / NAS
HA Cluster
Active Server
Passive Server
Heartbeat Connection
High-availability cluster main storage volume
High-availability cluster spare storage volume

**Figure 1. Physical components of a typical Synology High Availability (SHA) deployment.**

- **Active Server**: Under normal conditions, all services are provided by the active server. In the event of a critical malfunction, the active server will be ready to pass service provisioning to the passive server, thereby circumventing downtime.
- **Passive Server**: Under normal conditions, the passive server remains in standby mode and receives a steady stream of data replicated from the active server.
- **Heartbeat Connection**: The active and passive servers of a high-availability cluster are connected by a dedicated, private network connection known as the "Heartbeat" connection. Once the cluster is formed, the Heartbeat facilitates data replication from the active server to the passive server. It also allows the passive server to constantly detect the active server's presence, allowing it to take over in the event of active server failure. The Heartbeat connection must be configured on the fastest network interface between the two servers. For instance, if servers are equipped with 10 GbE add-on network cards, the Heartbeat must be configured using 10GbE.

- **Note:** The passive server detects the presence of the active server via both the Heartbeat connection and data connection in order to prevent "split-brain" errors when the Heartbeat connection fails. A "split-brain" error occurs when both servers attempt to assume the role of active server resulting in service errors.

- **Main Storage:** The storage volume of the active server.
- **Spare Storage**: The storage volume of the passive server, which continually replicates data received from the main storage via the Heartbeat connection.

## 3.2 Virtual Interface

When the two servers are combined into a high-availability cluster, a virtual interface -- unique server name and IP address -- shall be configured. This virtual interface allows hosts to access the cluster resources using a single name space. Therefore, when a switchover is triggered and the provision of services is moved to the passive server, there will be no need to modify network configurations on hosts in the data network.

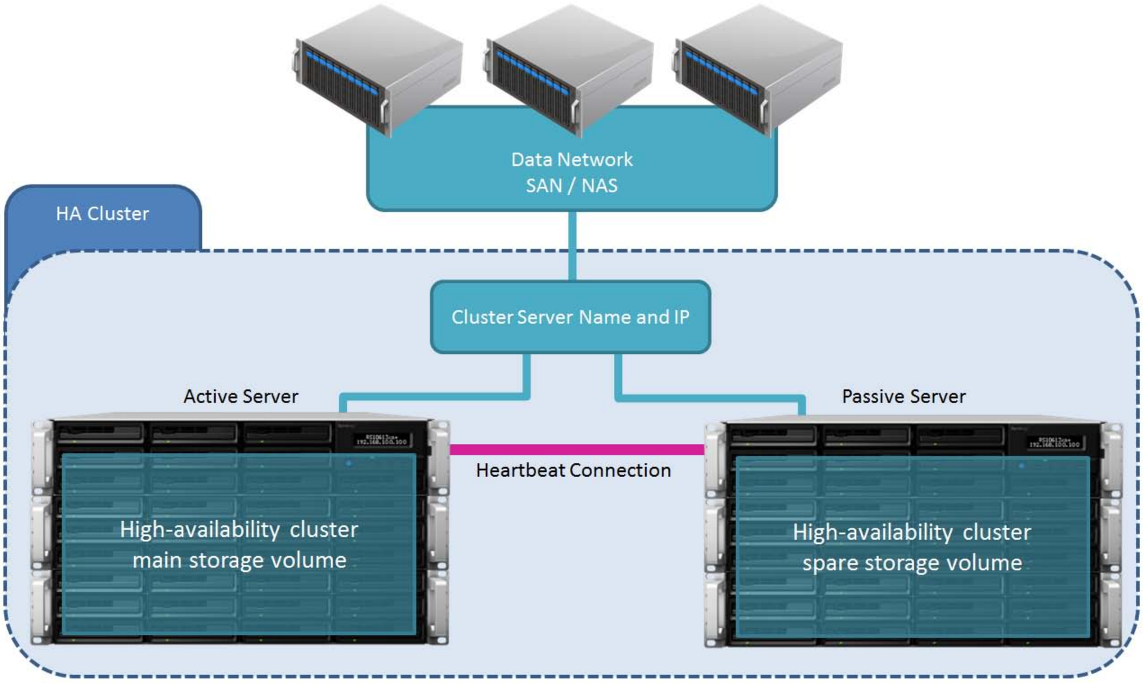
Data Network
SAN / NAS
HA Cluster
Cluster Server Name and IP
Active Server
Passive Server
Heartbeat Connection
High-availability cluster main storage volume
High-availability cluster spare storage volume

**Figure 2. Hosts and NAS clients access a Synology High Availability (SHA) cluster through a single virtual interface.**

- **Cluster Server Name and IP addresses:** Servers in the cluster will share IP addresses and a server name, which should be used in all instances instead of the original IP addresses and server name of individual servers.

## 3.3 Network Implementation

The physical network connections from the data network to the active server and passive server must be configured properly so that all hosts in the data network can seamlessly switch connection to the passive server in the event a switchover is triggered. The following section covers different configurations for various situations and Synology NAS models.

### Network Implementation for Synology NAS with two LAN ports

In situations where both servers have two network ports only, one network port on each server will be occupied by the heartbeat connection, so each server will have only one port available for the HA cluster to connect to the data network; therefore, there will not be sufficient network ports to accommodate redundant paths between the hosts in the data network and HA cluster. However, we still recommend using multiple paths to connect hosts to the data network, as well as more than one switch in your data network to provide redundancy.

Synology High Availability (SHA) provides an option to trigger a switchover when the active server detects network failure. When enabled, if connection failure occurs between the switch connected to the active server or the switch fails, services continuity will be maintained by switching over to the passive server (assuming the network connection of the passive server is healthy).

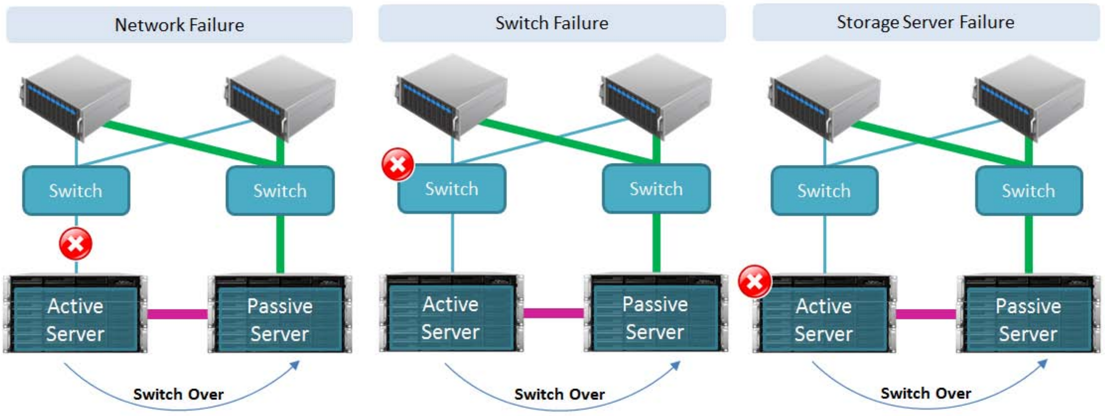

**Figure 3. High-availability cluster network configuration on 2 LAN models.**

### Network Implementation for Synology NAS with four or more LAN ports

The best way to create a high availability environment is to use Synology NAS with four network ports. In this case, you can connect multiple paths between the hosts and HA cluster, providing a redundant failover path in case the primary path fails. Moreover, I/O connections between the data network and each clustered server can be connected to more than one port, providing a load balancing capability when all the connections are healthy.

### Implementation for iSCSI storage

Connecting a host to more than one of the storage system's front end ports is called "multipathing." By implementing Multipath I/O (MPIO) or Multiple Connection per Session (MC/S) on the iSCSI connection, you may deliver a high quality and reliable storage service equipped with failover and load balancing capabilities, which is also one of the best practices for IT environments. When implementing an iSCSI multipath network, **network switches should be configured on separate subnets**.

The following diagram illustrates a full HA configuration that provides contingencies for path failure, switch failure, and storage server failure.

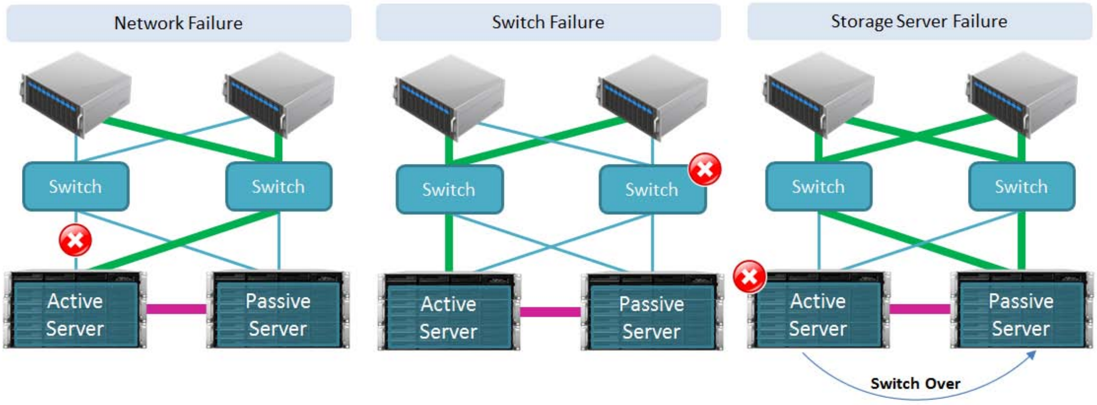

**Figure 4. High-availability cluster network configuration on models with four or more LAN ports (iSCSI).**

### Implementation for NAS storage

The link aggregation feature on Synology NAS can be leveraged to create a resilient HA network for file transfer services such as CIFS, NFS, AFT, and FTP. Link aggregation is a method of using two Ethernet ports in parallel to provide trunking and network fault tolerance. Link aggregation with the trunking feature enhances the connection speed beyond the limits that can be achieved with a single cable or port. Redundancy provides higher link availability and prevents possible disruption occurrences.

Please note that when creating link aggregation on two or more switches, stacked-switches are required for this configuration.

The following diagram demonstrates how link aggregation provides contingencies for path failover and failover that is triggered when a server fails.

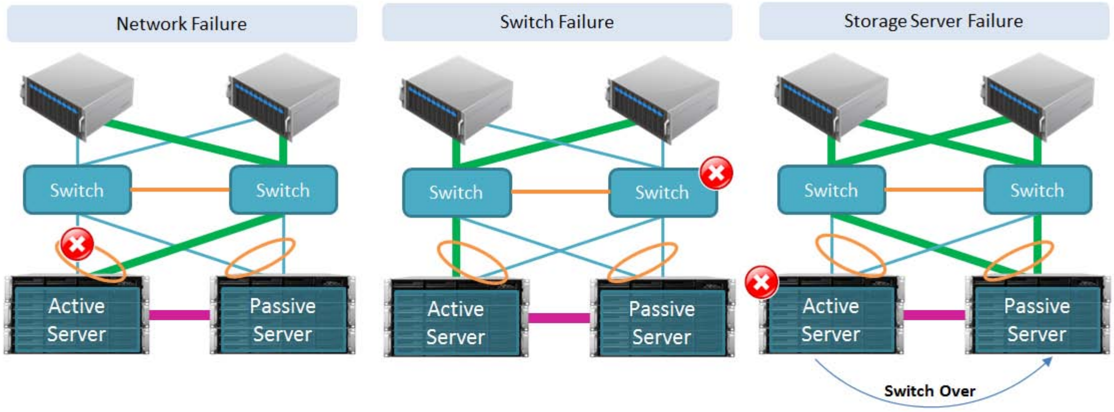

**Figure 5. High-availability cluster network configuration on models with 4 or more LAN ports (NAS).**

### Path redundancy for the heartbeat connection

For Synology NAS with four or more network ports, link aggregation may be implemented on the heartbeat connection to provide failover redundancy. This feature does not require a switch between the connections.

Chapter 4

# Ensuring Service Continuity

## 4.1 Switchover Mechanism

To ensure continuous availability, service provisioning can be switched from the active server to the passive server in a normally functioning high-availability cluster at any time. Switchover can be manually triggered for system maintenance, or automatically initiated in the event of the active server malfunctioning, which is known as "failover." After the servers exchange roles, the original active server assumes the role of the passive server and enters standby mode. As resources within the cluster are accessed using a single virtual interface, switchover does not affect the means of access.

- **Switchover:** The active and passive server can be manually made to exchange roles without interruption to service for occasions such as system maintenance.
- **Failover:** In the event of critical malfunction, the cluster will automatically initiate failover to maintain service availability.

The following situations can trigger system failover:

- **Crashed storage space:** If a storage space (e.g. volume, Disk Group, RAID Group, etc.) on the active server has crashed, while the corresponding storage space on the passive server is functioning normally, failover will be triggered unless there are no volumes or iSCSI LUNs (block-level) on the crashed storage space. Storage spaces are monitored every 30 seconds. Therefore, in the worst case, switchover will be triggered 30 seconds after a crash occurs.

> - **Note:** Switchover is not possible when the storage space on the passive server is busy with a Storage Manager related process (e.g. creating or deleting a volume).

- **Service Error:** If an error occurs on a monitored service, failover will be triggered. Services which can be monitored include CIFS, NFS, AFP, FTP, and iSCSI. Services are monitored every 30 seconds. Therefore, in the worst case, switchover will be triggered 30 seconds after an error occurs.
- **Power Interruption:** If the active server is shut-down or rebooted, both power units on the active server fail, or power is lost, failover will be triggered. Power status is monitored every 15 seconds. Therefore, in the worst case, switchover will be triggered 15 seconds after power interruption occurs.
- **Data Connection Lost:** If an error occurs on the data connection, and the passive server has more healthy data connections, failover will be triggered. For example, if the active server has three data connections and two of them are down, the active server will check whether the passive server has two or more available connections. If it does, failover will be triggered. Please note that for connections teamed with link aggregation, each teamed connection group is considered one connection.

After switchover has been performed, the faulty server may need to be replaced or repaired. If the unit is repaired, restarting the unit will bring the cluster back online and data-synchronization will automatically take place. If the unit is replaced, the cluster will need to be re-bound in order to recreate a functioning cluster. Any USB/eSATA devices attached to the active server will have to be manually attached onto the passive server once switchover is complete.

> **Note:** When a switchover occurs, all existing sessions are terminated. A graceful shutdown of the sessions is not possible, and some data loss may occur; however, retransmission attempts should be handled at a higher level to avoid loss. Please note that if the file system created on an iSCSI LUN by your application cannot handle unexpected session terminations, the application might not be able to mount the iSCSI LUN after a failover occurs.

## 4.2 Switchover Time-to-Completion

When switchover is triggered, the active server becomes the passive server, at which time the original passive server will take over. During the exchange, there will be a brief period where both servers are passive and services pause briefly. The time-to-completion varies depending on the number and size of volumes or iSCSI LUNs (block-level), and the number and total load of services on the cluster.

The following table provides estimated time-to-completion.

| Number of Volumes | Switchover | Failover |
|---|---|---|
| 10 | 60 seconds | 37 seconds |
| 32 | 115 seconds | 42 seconds |
| 64 | 185 seconds | 55 seconds |

*Tested on RS10613xs+ with CIFS, NFS, AFP, FTP and iSCSI enabled only, DSM version: DSM 4.1.

## 4.3 Switchover Limitations

Switchover cannot be initiated in the following situations:

- **Incomplete Data Replication:** When servers are initially combined to form a cluster, a period of time is required to replicate existing data from the active to passive server. Prior to the completion of this process, switchover may fail.
- **Passive Server Storage Space Crash:** Switchover may fail if a storage space (e.g. volume, Disk Group, RAID Group, etc.) on the passive server is crashed.
- **Power Interruption:** Switchover may fail if the passive server is shut down or rebooted, if both power units on the passive server malfunction, or if power is lost for any other reason.

# Deployment Requirements

Two identical Synology NAS servers which support the Synology High Availability (SHA) package are required for deployment. Before the two servers are combined to form a high-availability cluster, the Synology High Availability (SHA) Wizard will check for the following hardware and software limitations to ensure compatibility.

## 5.1 System Requirements and Limitations

- **Synology Servers:** Both active and passive servers must be identical models, or models specifically claimed as compatible by Synology, which support Synology High Availability (SHA).
- **DSM Version:** Identical DSM versions must be installed on both servers.
- **Package Version:** Identical Synology High Availability (SHA) package versions must be installed on both servers.

- **Note:** SSH functions will be enabled automatically once the high-availability cluster is formed.

## 5.2 Volume and Hard Disk Requirements and Limitations

- **Storage Volume:** In order to accommodate data replication, the storage capacity of the passive server must be equal to or larger than the capacity of the active server. It is strongly advised that the storage of capacity of both servers be identical to reduce chances of inconsistencies.
- **Quantity of Disks:** Both active and passive servers must have same quantity of disks. In addition, disk numbering and position must correspond.
- **Synology Hybrid Raid (SHR):** SHR format volumes are not supported.

## 5.3 Network Environment Requirements and Limitations

- **Network Settings:** Both servers must have static IP addresses belonging to the same subnet.
- **LAN Ports:** Both servers must have the same number of LAN ports, including the number of additional network card interfaces.

**Note:** Connecting via a proxy server, DHCP, IPv6, PPPoE, and Wi-Fi are not supported. Please ensure that these functions are disabled before attempting to form a high-availability cluster.

## 5.4 Storage Manager Limitations

Once a high-availability cluster has been formed, Storage Manager will no longer be able to change RAID types. However, the following actions will remain available after the formation of the high-availability cluster:

- Expand RAID Groups by adding or replacing hard disks (only for RAID Groups for multiple volumes or iSCSI LUNs).
- Expand volume or iSCSI LUN (block-level) size.
- Create, delete, or repair volumes and iSCSI LUNs.
- Change iSCSI LUN (file-level) size and location.
- Change iSCSI LUN target.

## 5.5 Expansion Units Requirements

Expansion units can be added to existing high-availability cluster configurations in order to increase storage capacity. As with other hardware requirements, identical expansion units are compulsory for both the active and passive servers.

Chapter

6

# Summary

Synology's High Availability solution provides a cost-effective and reliable means of insuring against service downtime. This white paper has outlined the basic principles and benefits of Synology High Availability (SHA). For more information and customized consultation on deployment, please contact Synology at **www.synology.com**.