White Paper

Synology High Availability (SHA): An Introduction

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 6 downtime events per year, the cost of unpreparedness begins to stack up.

Synology's 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

Synology's 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" consisting of two compatible Synology servers, in which one assumes the role of the active serve while the other acts as the 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 replicate on the passive server. In the event of a critical malfunction, the passive server will be ready to take over all services, equipped with a mirrored image of all data on the active server, allowing the high-availability 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:

1st Phase: The initial 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.

2nd Phase: Real-time data replication after the initial synch has been completed. After the initial synch, all data is replicated in real-time and treated as committed if successfully copied. In this phase, switchover can be performed at any time.

Once the cluster has entered into the 2nd phase of data replication, all data synching is performed at block-level. For instance, when writing a 10 GB file, synching 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, the swap can be accomplished seamlessly.

Data or changes which will 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.

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 the "Heartbeat" connection which monitors server status and replicates data between the two servers.

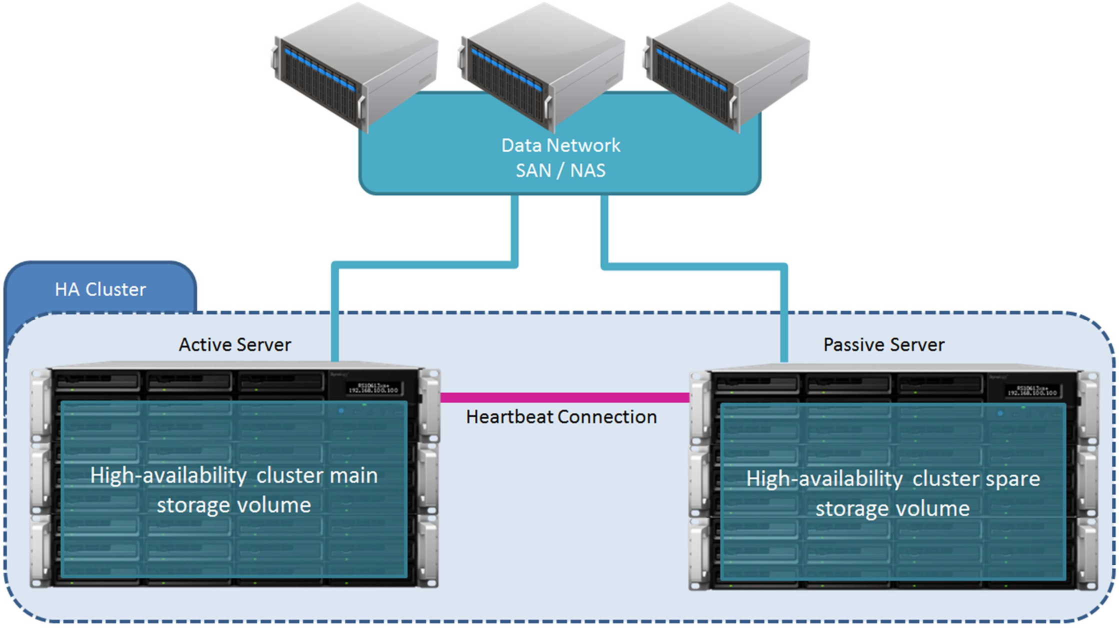
Data Network
SAN / NAS
HA Cluster
Active Server
Passive Server
Heartbeat Connection
High-availability cluster main storage volume
High-availability cluster spare storage volume

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 copied 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 synchronization for replicating data from active server to passive server. It also allows the passive server to constantly detect the active server's presence so as to allow the passive server to take over in the event of active server failure. The Heartbeat connection should 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 should 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 serve 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

Virtual interface allows servers from data network to access the HA cluster by a unified name space and prevent to change configuration on the servers with switchover event is triggered. Once a high-availability cluster is formed, hosts and clients must use its virtual interface to access cluster resources. IP addresses and a server name will be created on deployment for this purpose.

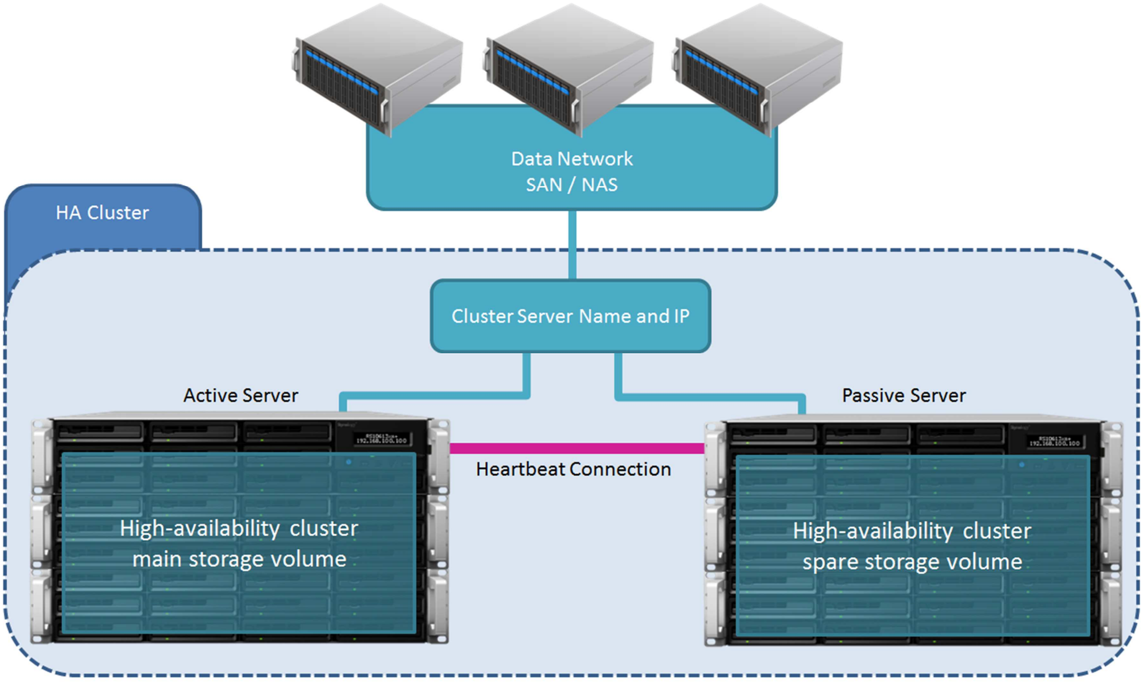
Data Network
SAN / NAS
HA Cluster
Cluster Server Name and IP
Active Server
Passive Server
Heartbeat Connection
High-availability cluster main storage volume
High-availability cluster spare storage volume

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

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

3.3 Network Configuration

As the cluster's virtual interface is mapped onto the active server, the physical network connections from the active server and passive server to the data network must be configured so that all hosts and clients can access both servers. The following image illustrates an example whereby all servers in the data network are able to connect to the cluster. Therefore in the event of a switchover, the data network will still be able to access data.

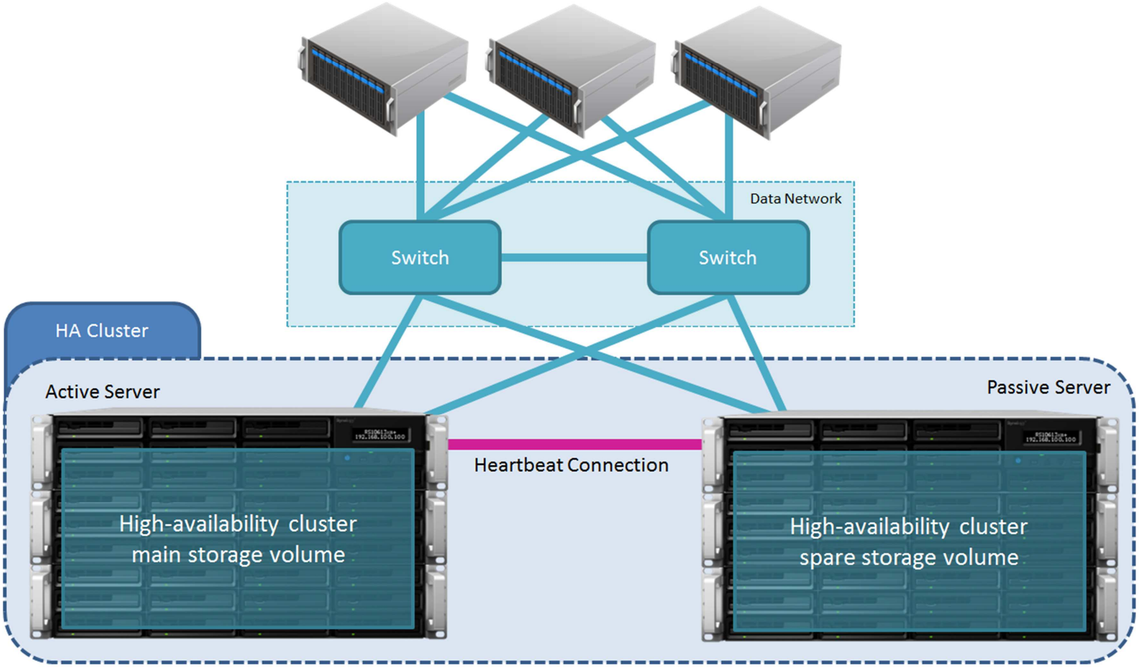
Data Network
Switch
Switch
HA Cluster
Active Server
Passive Server
Heartbeat Connection
High-availability cluster main storage volume
High-availability cluster spare storage volume

Figure 3. High-availability cluster network configuration.

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 active server malfunction, known as “failover.” After the servers exchange roles, the original active server will assume the role of the passive server and enter standby mode. As resources within the cluster are accessed using a single virtual interface, switchover will 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 RAID Group: If a RAID Group on the active server has crashed, while the corresponding RAID group on the passive server is functioning normally, failover will be triggered unless there are no volumes or iSCSI LUNs (block-level) on the crashed RAID group. The monitoring period is every 30 seconds. Therefore in the worst case, switchover will be triggered 30 seconds after the crash.
- 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. The monitoring period is every 30 seconds. Therefore in the worst case, switchover will be triggered 30 seconds after the error.
- 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. The monitoring period is every 15 seconds. In the worst case, switchover will be triggered 15 seconds after the interruption.

Post-switch over, 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, after which the originally 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 Failure to Complete Switchover

Switchover may fail 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 RAID Group Crash: Switchover may fail if any RAID groups on the passive server crashes.
- 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.

- ***Note:*** In the event of a manually triggered switchover failure, the system will automatically attempt to switch servers back to the active server.

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 compliance.

System Requirements and Limitations

- Synology Servers: Both active and passive servers must be identical models 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.

5.1 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.

5.2 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:*** PPPoE, link aggregation, Wi-Fi, and VLAN are not supported. Please ensure that these functions are disabled before attempting to form a high-availability cluster.

5.3 Storage Manager Limitations

- Once a high-availability cluster has been formed, Storage Manager will no longer be able to perform certain actions: Edit or expand volume and iSCSI LUN (block-level) size.
- Change RAID types.

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).
- Create, delete, or repair volumes and iSCSI LUNs.
- Change iSCSI LUN (file-level) size and location.
- Change iSCSI LUN target.

5.4 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 5

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**.