White Paper

High Availability Cluster for RC18015xs+

Shared Storage Architecture

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 in the case of disaster[1]. Moreover, downtime costs a median of 12,500 USD per day. Assuming a median of six downtime events per year, the cost of unpreparedness begins to stack up.

The **High Availability** solution for RC18015xs+ helps users overcome this hurdle by ensuring non-stop storage services with maximized system availability to mitigate the risk and impact 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 with shared storage solution is a server layout designed to reduce service interruptions caused by system malfunctions. It employs two computing servers (RC18015xs+ units) and a shared storage (RXD1215sas) to form a “**high-availability cluster**” (also called “HA cluster”). Once this high-availability cluster is formed, one computing server assumes the role of the active server, while the other acts as a standby passive server.

2.2 Service Continuity

Once the high-availability cluster is formed, the health of active server is monitored continuously. In the event of a critical malfunction, the passive server is ready to take over all services. The passive server will enable the high-availability cluster to continue functioning as normal, reducing downtime.

Chapter 3

High-Availability Cluster Architecture

3.1 Physical Components

Synology's High Availability solution constructs a cluster composed of two **RC18015xs+** units (computing servers), an active and a passive server. Both servers have been directly attached to a set of shared SAS enclosures (**RXD1215sas**), and the two are linked by a "**Heartbeat**" connection that monitors server status. To ensure successful boot, please make sure the Heartbeat, data network connection and shared storage are connected to active and passive server correctly.

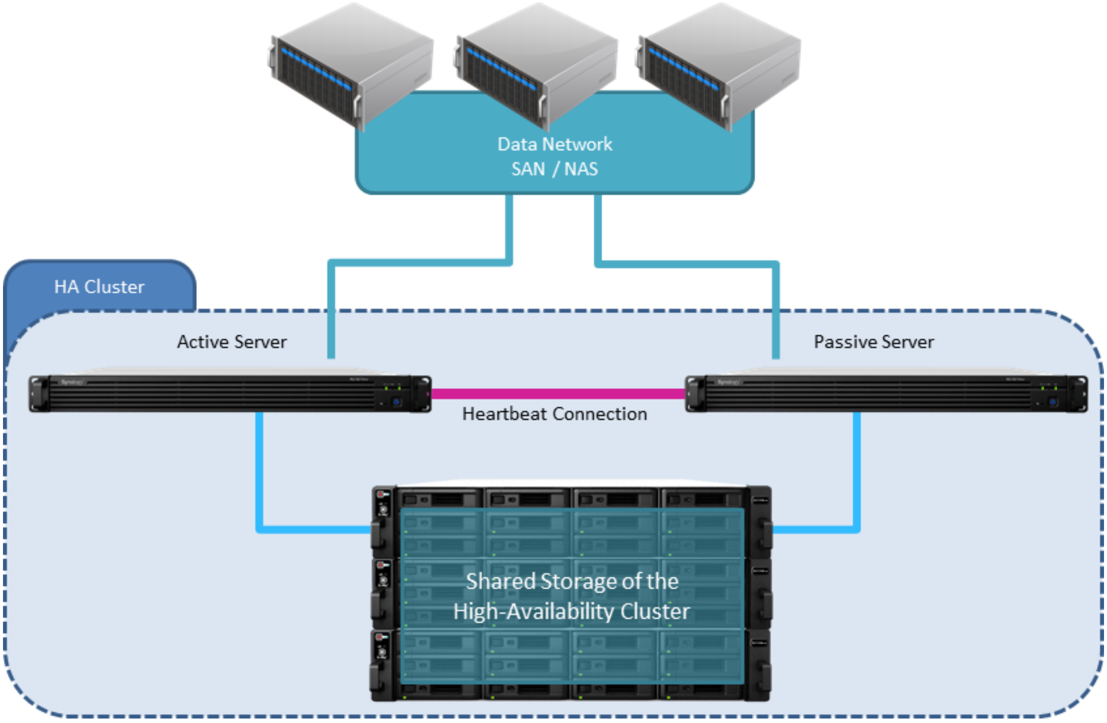

Figure 1: Physical components of a typical High Availability with Shared Storage deployment

- **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 Heartbeat signal from the active server.
- **Shared Storage Path Connectivity**: Under normal conditions, the number of drives connected to active and passive server should be the same. If some connectivity issues occurred in storage connectivity, the active server will pass service provisioning to the passive 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. 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.

- **Note:** The passive server detects the presence of the active server via the Heartbeat connection and data connection, as well as the hardware components in the SAS enclosure in order to prevent "split-brain" errors when the Heartbeat connection fails.

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 connections 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 storage

We recommend using multiple paths to connect hosts to the data network, as well as more than one switch in your data network to provide 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.

High Availability Manager provides an option to trigger a switchover when the active server detects network failure. When enabled, if connection failure occurs in the switch connected to the active server or the switch fails, service continuity will be maintained by switching over to the passive server (assuming the network connection of the passive server is 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 can 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.

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

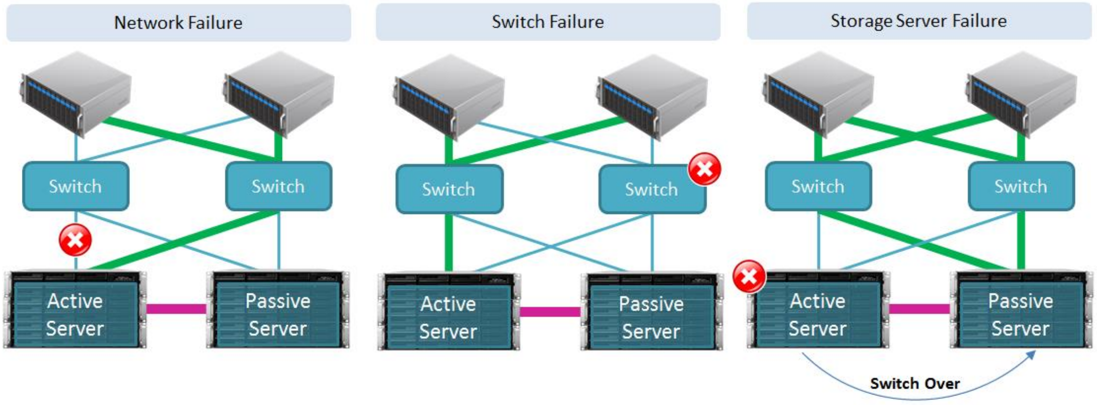

Figure 4: High-availability cluster network configuration (iSCSI)

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, AFP, and FTP. Link aggregation is a method of using multiple Ethernet ports in parallel to provide trunking and network fault tolerance. Link aggregation with trunking 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 disruptions.

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

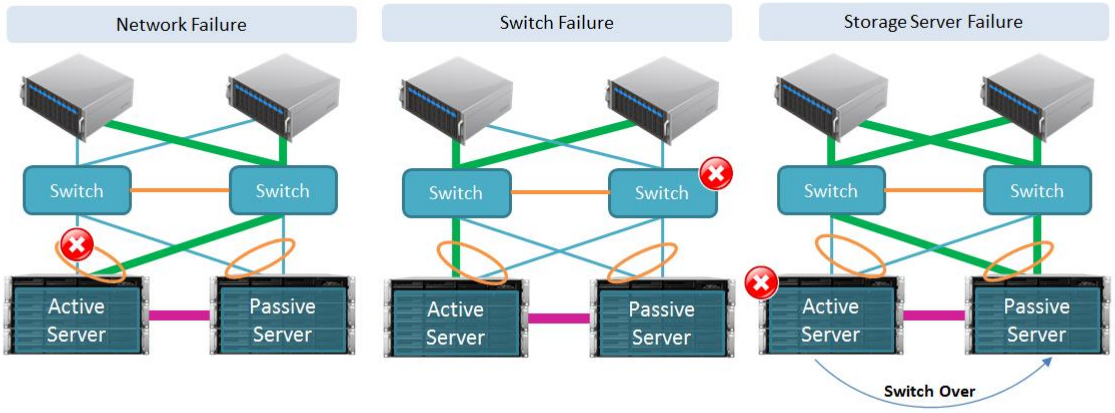

Figure 5: High-availability cluster network configuration (NAS)

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 triggered 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 switchover to maintain service availability.

System failover can be triggered by the following situations:

- **Service Error:** If an error occurs in a monitored service, failover will be triggered. Services that can be monitored include CIFS, NFS, AFP, FTP, and iSCSI. Services are monitored every 15 seconds. Therefore, in the worst case, failover will be triggered 15 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 cut off, failover will be triggered. Power status is monitored every 15 seconds. Therefore, in the worst case, failover will be triggered 15 seconds after power interruption occurs.
- **Data Connection Lost:** If a monitored network interface on the active server is disconnected, and the passive server has healthy data connections for all monitored interface, failover will be triggered. Failover will be triggered when all monitored connections on the active server are disconnected and the passive server have at least one healthy connection for any of the monitored interfaces.

After failover has occurred, the faulty server may need to be replaced or repaired. If the unit is repaired, restarting the unit will bring the cluster back online. If the unit is replaced, the cluster will need to be re-bound in order to recreate a functioning cluster. Any USB devices attached to the active server will have to be manually attached onto the passive server once the failover is complete.

Note: When a failover 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 by the application server. 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 are paused. 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.

The Sum of Volume Size	Switchover	Failover
60 TB	60 seconds	56 seconds
450 TB (200 TB + 200 TB + 50 TB)	156 seconds	132 seconds

*Tested on RC18015xs+ with DSM 5.1

4.3 Switchover Limitations

Switchover cannot be initiated in the following situations:

- **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 cut off for any other reasons.
- **Hardware Misconfiguration:** Switchover cannot be performed if two servers do not have identical hardware, including memory size, network adapters or storage connections.
- **No Heartbeat connection:** If the Heartbeat connection between the active and passive servers is disconnected, failover cannot be performed.
- **Unknown Server**: A new computing server is connected to the existing HA cluster and a repair operation in web user interface is required. Switchover will be possible after the new unknown server is initialized and joined in the HA cluster.
- **No Static IP address**: If a monitored network interface lacks a static IP address, failover cannot be performed.
- **Volume Crash**: If a volume is crashed and cannot be repaired by performing a failover operation, failover cannot be performed.
- **DSM Update:** When installing DSM updates, all services will be stopped and then come back online after the completion of DSM update process.

Chapter 5

Deployment Requirements & Best Practices

To create HA cluster with shared storage architecture, please look over the following hardware and software limitations to ensure compatibility.

5.1 System Requirements and Limitations

- **Synology Servers:** The hardware of active and passive servers must be identical, including the memory sizes, network adapters, and connections to external storage of both servers.

5.2 Network Environment Requirements and Limitations

- **Network Settings:** Static IP addresses must be assigned for the HA cluster.
- **LAN Ports:** Both servers must have the same number of LAN ports, including the same number of additional network card interfaces. Wi-Fi is not supported.
- **External Devices:** Only USB storage is supported.

5.3 Storage Requirements

- **Hard Drives:** Only SAS drives are supported, and at least 3 drives are required for normal operation.

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 high availability clustering. For more information and customized consultation on deployment, please contact Synology at **www.synology.com**.