



This PDF is generated from authoritative online content, and is provided for convenience only. This PDF cannot be used for legal purposes. For authoritative understanding of what is and is not supported, always use the online content. To copy code samples, always use the online content.

SIP Server HA Deployment Guide

Windows NLB Cluster

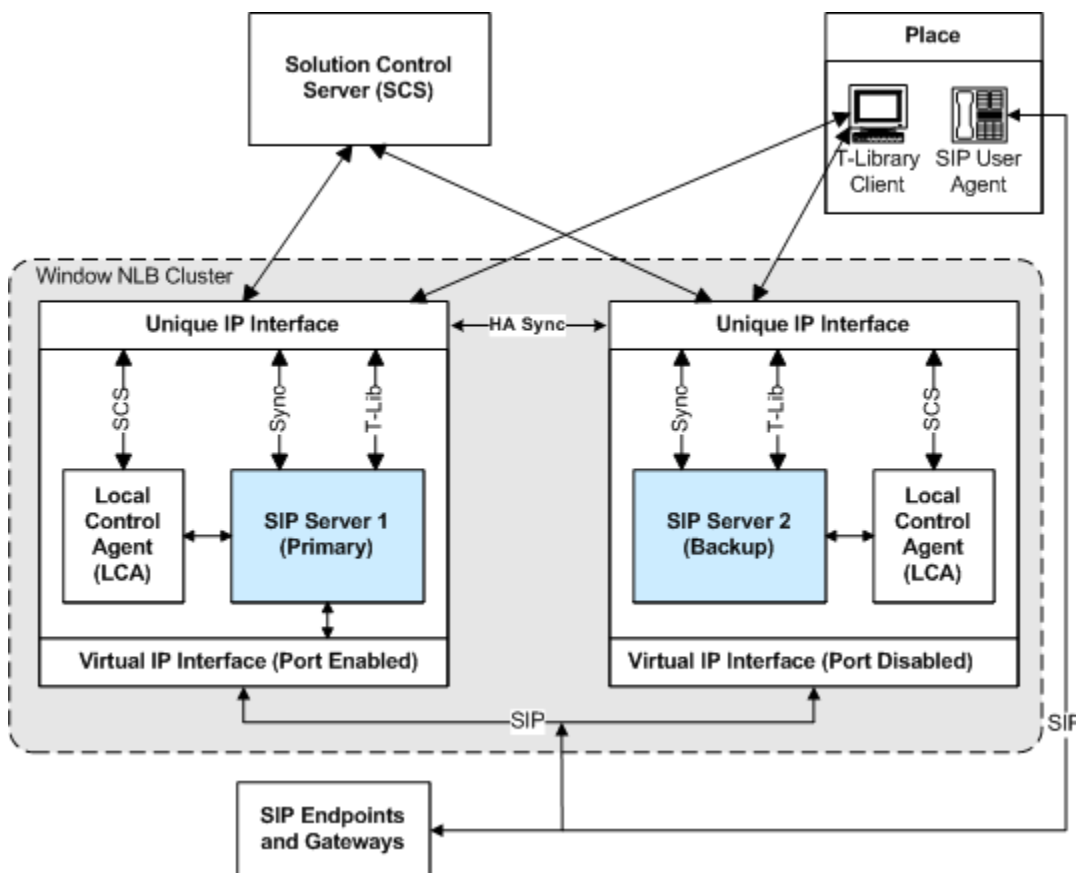
12/20/2025

Windows NLB Cluster

A SIP Server HA configuration using Windows Network Load Balancing (NLB) configuration is an alternative to a Windows IP Address Takeover configuration.

Microsoft's NLB cluster technology allows you to configure cluster hosts to receive requests at a single Virtual IP address. SIP endpoints and gateways are configured to send all requests to SIP Server by using this single Virtual IP address. The Windows NLB cluster technology delivers the requests to the SIP Server that is running in primary mode and reroutes traffic to the backup SIP Server when a failure is detected.

The **HA Windows NLB Cluster Configuration** figure shows a SIP Server HA configuration that uses Windows NLB. SIP endpoints and gateways are configured to communicate with SIP Server by using a single Virtual IP address, and the SIP Server port is enabled only at the SIP Server that is running in primary mode. When a switchover to the backup SIP Server occurs, the port at the backup SIP Server host is enabled, and traffic is directed to the active SIP Server.



HA Windows NLB Cluster Configuration

The Management Layer uses a Windows NLB utility (`wlbs.exe` or `nlb.exe`) to enable and disable ports that are occupied by SIP Server. The NLB utility is initiated by Cluster control scripts that are triggered by SIP Server Alarm Conditions that are configured for SIP Server log events that occur

when a SIP Server changes its mode from primary to backup or from backup to primary.

Windows NLB can be configured to distribute incoming requests by using either the Unicast or the Multicast method. When you deploy a SIP Server HA configuration, you must define the method that you want to use.

Unicast and Multicast methods are described in the following sections.

See [Windows NLB Cluster HA Workflows](#) for step-by-step descriptions of manual switchover, primary SIP Server failure, and primary SIP Server disconnect workflows. For deployment procedures, see [Windows NLB Cluster HA Deployment](#).

Unicast Method

In the Unicast method, all NLB cluster hosts share an identical unicast MAC address. NLB overwrites the original MAC address of the cluster adapter by using the unicast MAC address that is assigned to all of the cluster hosts. Unicast NLB nodes cannot communicate over an NLB-enabled network adapter. In the Unicast method, all switch ports are flooded with NLB traffic, including ports to which non-NLB servers are attached. A workaround for this issue is to place cluster hosts on separate VLANs.

Multicast Method

In a Multicast configuration, each NLB cluster host retains the original MAC address of the network adapter. In addition to the original MAC address of the adapter, the adapter is assigned a multicast MAC address that is shared by all cluster hosts. Client requests are sent to all cluster hosts at the multicast MAC address. Considerations for implementation of the Multicast distribution method include the following:

- Upstream routers might require a static Address Resolution Protocol (ARP) entry. Without an ARP entry, routers might not accept an ARP response that resolves unicast IP addresses to multicast MAC addresses.
- Without Internet Group Management Protocol (IGMP), switches might require additional configuration to define which ports the switch should use for multicast traffic.
- Upstream routers might not support mapping of a unicast IP address (the cluster IP address) to a multicast MAC address. In this case, you might be required to update or replace your router in order to use the Multicast method.