Microsoft Network Load Balancing on UCS−B Series Servers Deployment Configuration Example

Document ID: 118262

Contributed by Charles Stizza, Vishal Mehta, and Vincent La Bua,
Cisco TAC Engineers.
Aug 14, 2014

Contents

Introduction

This document describes the implementation of Microsoft Network Load Balancing (NLB) mode on Cisco Unified Computing System−B (UCS−B) series with Fabric Interconnect (FI) in End−Host mode. There are also a number of requirements on the upstream devices to facilitate the correct forwarding of NLB traffic which are described in this document. The configuration sample focuses on multicast Internet Group Management Protocol (IGMP) mode.

Prerequisites

Requirements

Cisco recommends that you have knowledge of these topics:

- Microsoft Network Load Balancing
- Cisco UCS B−Series Servers
- Cisco Catalyst and/or Nexus Switches

Components Used

This document is not restricted to specific software and hardware versions.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, make sure

that you understand the potential impact of any command.

Background Information

Microsoft NLB functions in three different operational modes: unicast, multicast, and multicast IGMP. A group of NLB nodes is collectively known as an NLB cluster. An NLB cluster services one or more virtual IP (VIP) addresses. Nodes in the NLB cluster use their load balancing algorithm in order to agree on which individual node will service the particular traffic flow destined for the NLB VIP.

This document does not make specific deployment recommendations for Microsoft NLB on UCS. As described in this document, NLB relies on unconventional methods for delivery of cluster bound traffic. It has been observed that both multicast and multicast IGMP modes appear to have stable and consistent operation on UCS−B Series servers. While NLB sizing guidelines are beyond the scope of this document, it is a solution generally recommended for smaller deployments.

Configuration

Microsoft NLB Modes

Unicast Mode

The NLB default setting is unicast mode. In unicast mode, NLB replaces the actual MAC address of each server in the cluster to a common NLB MAC address. Typically, something in the 02bf:xxxx:xxxx range. All the nodes in the NLB cluster understand what the NLB VIP and MAC address is. Traffic, which includes Address Resolution Protocol (ARP) replies from NLB nodes, is *never* sourced from the NLB MAC or IP address. Instead NLB nodes use an assigned MAC address based on the host ID of the member. The MAC address is usually in the 0201:xxxx:xxxx, 0202, 0203, and so on range, one for each node in the cluster. This is the source address in the Layer 2 (L2) header when an ARP request is answered. However, the ARP header information contains the NLB MAC address. Thus, hosts that wish to correspond to the NLB VIP address send traffic towards the NLB MAC address.

IEEE compliant switches (L2 devices) build their MAC address table based on the L2 source header and not the information contained in the ARP payload. This means that traffic forwarded to the NLB cluster is sent to the NLB MAC address, which is never the source of traffic. Therefore traffic destined for the NLB MAC address is flooded as unknown unicast.

Caution: NLB in unicast mode relies on unknown unicast flooding for delivery of cluster bound packets. *Unicast mode will not work on UCS B−Series servers when the FI is in End−Host Mode since unknown unicast frames are not flooded as required by this mode.* For more details on the L2 forwarding behavior of UCS in End−Host mode, see Cisco Unified Computing System Ethernet Switching Modes.

Mulitcast/Multicast IGMP Mode

Multicast mode assigns the cluster unicast virtual IP address to a non−Internet Assigned Numbers Authority (IANA) multicast MAC address (03xx.xxxx.xxxx). IGMP snooping does not dynamically register this address, which results in flooding of the NLB traffic in the VLAN as unknown multicast.

Multicast IGMP mode assigns the cluster virtual IP address and a multicast MAC address within the IANA range (01:00:5E:XX:XX:XX). The clustered nodes send IGMP membership reports for the configured multicast group and thus the FI dynamically populates its IGMP snooping table to point towards the clustered servers.

There is a slight operational advantage to the use of multicast IGMP mode since state information (via IGMP membership reports and IGMP snooping) about the interested L2 ports can be maintained both upstream and downstream. Without the optimization of IGMP snooping, NLB relies on unknown multicast flooding into the NLB VLAN for delivery to the cluster via the UCS designated broadcast/multicast receiver. In releases later than UCS Release 2.0, the designated broadcast/multicast receiver is chosen on a per−VLAN basis.

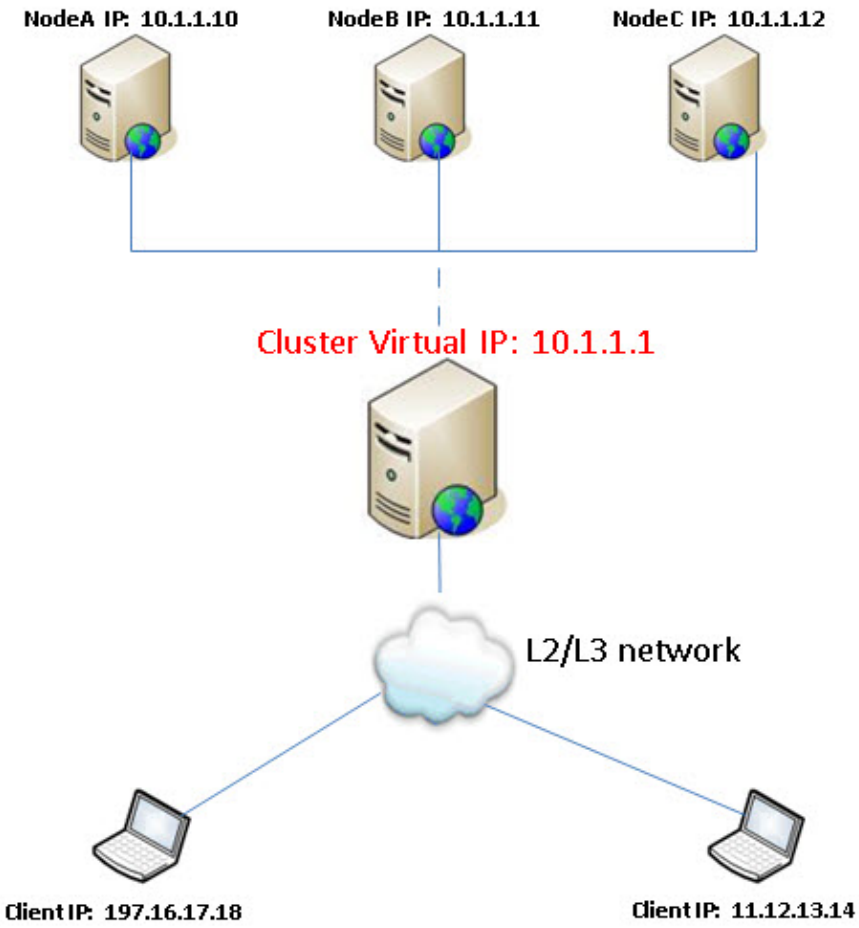
Caution: Regardless of the version of multicast mode chosen, the NLB VIP address requires a static ARP entry on the upstream device, which is typically the Switched Virtual Interface (SVI) for the VLAN. This is a workaround since the ARP replies from the NLB nodes contain a multicast MAC address. Per RFC 1812, ARP replies that contain a multicast MAC address should be ignored. Therefore the VIP MAC address cannot be dynamically learned on RFC 1812 compliant devices.

A summary of the steps required to support NLB in multicast IGMP mode is shown here:

1. Static ARP entries for Virtual NLB IP addresses are typically on the VLAN SVI. If you use Hot Standby Router Protocol (HSRP) or First Hop Redundancy Protocol (FHRP), be sure that both devices have the static ARP entry.
2. An IGMP snooping querier in the NLB VLAN. In releases later than UCS Release 2.1, snooping querier functionality is supported in UCS Manager.
3. IGMP snooping needs to be enabled on all switches, which includes UCS. Note that most platforms that include UCS have IGMP snooping enabled by default.

Tip: These configuration guides are for Cisco switches. They include details on how to implement different modes of Microsoft NLB.

A basic setup of NLB, the nodes can be virtual machines (VMs) or bare−metal installation of Windows Server OS, is shown in this diagram.



NLB VLAN 10 that has IP subnet 10.1.1.0 /24. The MAC address is truncated for brevity.

NLB VIP (MAC = 01, IP = 10.1.1.1)

NODE−A (MAC = AA, IP = 10.1.1.10)

NODE−B (MAC = BB, IP = 10.1.1.11)

NODE−C (MAC = CC, IP = 10.1.1.12)

Microsoft NLB Data Flow

Static ARP entry on the upstream switch SVI points to VIP address 10.1.1.1 to MAC 01.

Microsoft NLB nodes send the IGMP membership report. Note that the IGMP snooping table can take 30−60 seconds to populate.

With IGMP snooping and the VLAN querier, the snooping table is populated with the NLB MAC address and groups that point to the correct L2 ports.

1. Off−subnet clients send traffic to NLB VIP address 10.1.1.1.
2. This traffic is routed into the VLAN 10 interface which uses a static ARP entry in order to resolve the MAC address (01) of the NLB VIP.
3. Since this is a multicast frame destination, it is forwarded per the IGMP snooping table.
4. The frame arrives at all NLB nodes (Node A,B,C).

5. The NLB cluster uses its load balancing algorithm in order to determine which node will service the flow. Only one node responds.

See these documents for more information:

- Catalyst Switches for Microsoft Network Load Balancing Configuration Example
- Microsoft Network Load Balancing on Nexus 7000 Configuration Example

Special Consideration for Nexus 1000v

Nexus 1000v only supports unicast Microsoft NLB mode. So on deployments of Nexus 1000v with UCS, multicast IGMP mode will only work after you disable snooping on Nexus 1000v. When this is done, Microsoft NLB packets on that VLAN are flooded as unknown multicast.

In order to minimize the impact of flooding:

1. Disable snooping only on that VLAN in the Nexus 1000v.
2. Use a dedicated VLAN for Microsoft NLB traffic.

Verify

The verification procedures for the configuration examples described in this document are provided in the respective sections.

Troubleshoot

There is currently no specific troubleshooting information available for this configuration.

Related Information

- ***Network Load Balancing Technical Overview***
- ***Cisco Support Community Discussion***
- ***Cisco Unified Computing System Ethernet Switching Modes*** (search for Microsoft Network Load Balancing)
- ***Technical Support & Documentation − Cisco Systems***

Updated: Aug 14, 2014 Document ID: 118262