Implementing Network Load Balancing

Eli Chang

MSSA Cohort #2

Lab Summary 10

1/28/2020

## Implementing Network Load Balancing

As a need for scalable web applications has increased, the admin needs to develop a program to test NLB on Windows Server 2016 hosts. NLB is a scalable, high-availability technology to balance traffic based on node utilization. Also, it is failure-aware, which means one resource is aware of other resource states. It detects a server failure, rather than application failure. When a web application fails while the server remains operational, NLB continues to forward traffic to the node. It seems like a failover clustering but uses for specific applications like web applications.

Before implementing the NLB cluster, there are two standalone servers: LON-SVR1 and LON-SVR2. These servers should be verified that both are working independently. NLB cluster creates a networking environment once it is deployed. Both servers have an image file of the Internet Information Service logo, and one of both has some modification of the picture because it would verify there are working independently when the Domain Controller displays each server.

During the installation of the NLB cluster, the operation mode can be either multicast or unicast. The unicast mode is suitable for the clusters with multiple network adapters, whereas the multicast mode is suitable for a single network adapter NLB cluster. In the lab, since the NLB is installed in the virtual machine environment, both modes are set up for practice purposes. In real-time, the modes should be chosen depending on the number of network adapters. Once it is installed, each server should add to the NLB cluster to DNS. Both servers would have two IP addresses, such as a virtual IP address and their IP address. The virtual address will fall into NLB so that when other networks try to send and receive data, they could go through the virtual switch.

As the installation and addition of the cluster are completed, the validation for the NLB cluster is necessary. In the Network Load Balancing Manager, each node should display on the screen with the status of "converged." Then, each node configures NLB port rules. The port rules determine how traffic goes to the cluster nodes, depending on ports of TCP or UDP. Also, it does not have to be in specific ports. In this case, the ports 80 and 5678 are used because the admin wants to have all incoming requests to come into these ports. The port can be set with the default port rule.

Again, the filtering mode will have three options, but for the practice purpose, multiple and singles hosts would be chosen. Since the NLB cluster is specifically for web applications, the cluster can be validated in the web applications. For the port rule validation, each port number should be typed right next to the web address.

NLB cluster should have a static IP address with the same subnet. Also, it should be configured with unicast or multicast. The NLB cluster works with the heartbeats, which transmit every second between the nodes in the cluster. When the node misses five times of heartbeats, the NLB cluster will remove the node from the cluster. These options should be validated before deploying NLB clusters. Then, each node should be converged into the cluster in the NLB Manager.

## IMPLEMENTING NETWORK LOAD BALANCING

