Environmental Systems Research Institute, Inc., 380 New York St., Redlands, CA 92373-8100 USA • TEL 909-793-2853 • FAX 909-307-3014

Environmental Systems Research Institute, Inc.

ArcGIS Server 9.x High-Availability Configuration Testing Using Microsoft Network Load Balancing

Enterprise System Lab Test Report

Prepared by:

Ray Echevarria
ESRI Implementation Services
Redlands, California

Version History

Prepared By	ArcGIS Server Version	Revision Date
R. Echevarria	9.0/9.1	6/6/05
J. DeWeese	9.0/9.1	6/14/06

Test Objective

The objective was to test and document if ArcGIS Server 9.x (9.0 and 9.1) can be successfully integrated with Microsoft's Network Load Balancing (NLB) cluster solution. NLB functionality is included with Windows Server 2003 Standard Edition and is designed to load balance network traffic between multiple nodes. The report is not all inclusive and assumes reader is familiar with NLB concepts and terminology and certain ArcGIS Server installation and configuration tasks.

Test Hypothesis

If two ArcGIS Server Web applications are configured identically and connected via NLB, any incoming Web Server traffic using the virtual URL should be load balanced between the two Web Servers. A Web Server failure should force new requests to be processed by the remaining Web Server. When the failed Web Server is recovered, new client requests should once again be load balanced between the two Web Servers. Furthermore, configuring multiple Server Object Containers (SOC) that are cross-registered with Server Object Manager (SOM) machines should avoid the problem where a Web Server continues to send requests to a failed SOC. Combining NLB at the Web server level with multiple SOCs at the Map Server level should provide a high-availability load balanced ArcGIS Server Web application.

Test Hardware and Software Configuration

The test configuration involved four physical servers (see Figure 1); two identically configured Web Servers and two identically configured Map Servers, which will be hosting the SOCs. Two fully configured servers can be used as well, but four servers were used to demonstrate different failure recovery scenarios, etc. The Web Servers were configured with the following software:

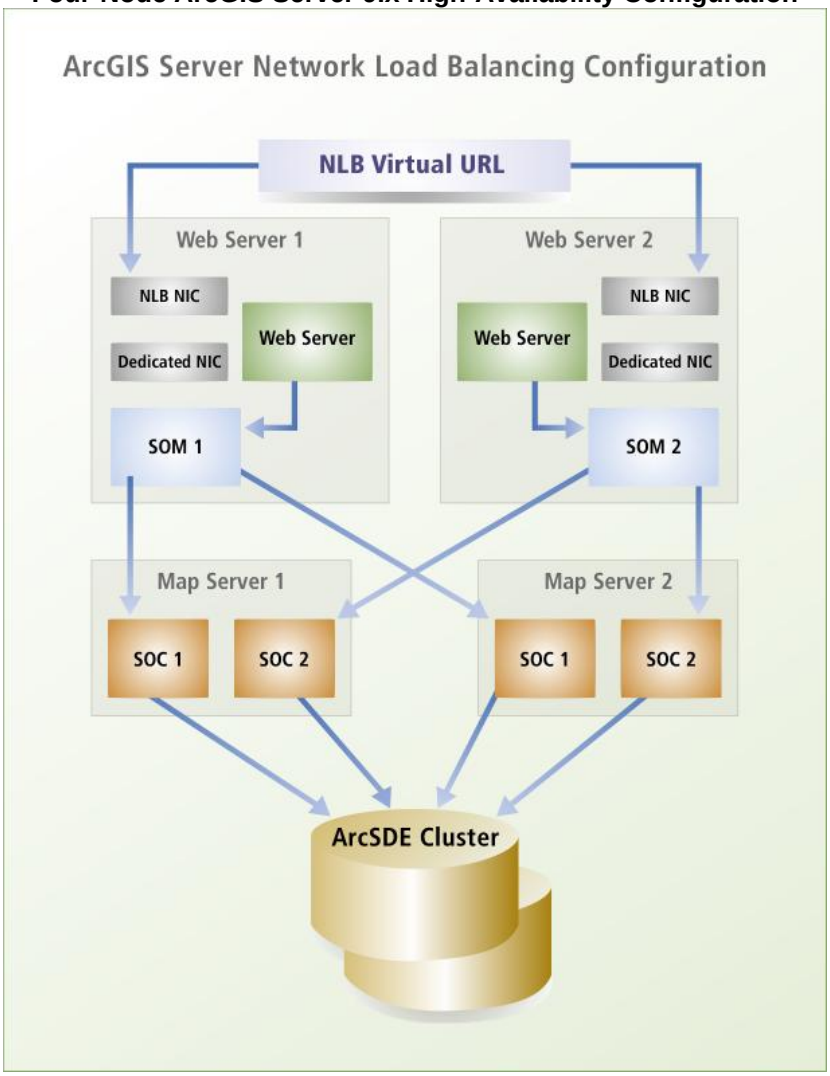
- Windows Server 2003 Standard Edition
- Network Load Balancing
- IIS 6.0 Web Server
- Microsoft .NET SDK Framework 1.1
- Microsoft Visual Studio .NET Enterprise Architect 2003
- ArcGIS Server 9.1 Server Object Manager
- ArcGIS Server 9.1 .NET Application Developer Framework (ADF)

The Map Servers were configured with the following software:

- Windows 2000 Advanced Server, SP4
- ArcGIS Server 9.1 Server Object Container

NLB provides load balancing and high-availability at the Web Server level. It only load balances based on network traffic and does not load balance based on the complexity of the Web request nor does it monitor system resources to make load balancing decisions. Furthermore, if a Web Server fails, the associated SOCs become detached from the configuration thus reducing site capacity. To avoid both load balancing and site capacity concerns, two SOCs should be configured for each SOM, one registered with Web Server 1 and other registered with Web Server 2 (Figure 1). This will ensure granular load balancing at the SOC level and will allow all the Map Servers to be utilized by the surviving Web Server in the event of a Web Server failure.

Figure 1
Four-Node ArcGIS Server 9.x High-Availability Configuration



Recommended NLB Software Configuration

NLB was configured with the Network Load Balancing Manager, which is the method recommended by Microsoft. NLB was configured on both nodes using a virtual IP address (10.15.11.77) and hostname (eslweb.esri.com) for a dual NIC unicast configuration. The first NIC (NLB NIC) was used for NLB communication. It was configured with the cluster IP address, which was reflected under the "Cluster IP Configuration" in the NLB Cluster Parameters Tab. Since this NIC was configured with only a cluster IP address, it was not necessary to configure a Dedicated IP Address in the NLB Host Parameters.

The second NIC (Dedicated NIC) served as the dedicated network interface and was used for any non-load balanced traffic. For more information on configuring dual NICs and NLB parameters in general refer to the NLB Manager "Help Topics" and look for a section called "Checklist: Enabling and Configuring Network Load Balancing."

In regards to affinity, single affinity is the default configuration and was used for the testing. With single affinity, NLB examines the client's IP address to determine which node to direct the request along with any subsequent requests. Therefore, the client becomes affined to a specific Web Server except in the case of a Web Server failure where the requests will be routed to a surviving node. With hundreds or thousands of potential users systems hitting the site, the statistical difference in their IP addresses will cause a statistical load balance to occur. This is important during test phases as it may take some effort to get test clients to load balance across the cluster.

An attempt was made to use "no affinity" to allow individual requests from the same client to be load balanced across multiple Web Servers, but this caused map rendering at the client to become erratic. The reason for this is that the request to create a map, for example, may go to Web Server 1, where the output image is ultimately placed for retrieval by the client. However, when the client application returns to the clustered Web site to retrieve the output image, NLB may send the request to Server 2, which does not have the image. An attempt was made to use image streaming (MIME) for output type, however this resulted in the same behavior on the client.

Recommended ArcGIS Server Configuration

Server Object Manager Installation (Web Server)

Install ArcGIS Server using standard installation procedures for both Web Server nodes. During initial installation unselect the Server Object Container component.

Server Object Container Installation (Map Server)

Install ArcGIS Server using standard installation procedure however just select the Server Object Container component, unselect all other components. Once installation is complete, open ArcCatalog and connect to both ArcGIS Server machines. Open the "Server Properties" for each ArcGIS Server connection and add the two Map Servers that will be hosting the server object. Once the connection to ArcGIS Server is established, add a server object. Start the server object and preview the service in ArcCatalog to verify the connection. Once both server objects have been verified go to one of the Web Servers and create a map viewer template in Visual Studio.

Test Tools

Internet Explorer 6.0 was used to access and test the ArcGIS Server Web application. ArcCatalog was used to connect to the ArcGIS Servers and administer the server objects.

Test Procedure

To ensure that ArcGIS Server functioned properly in an NLB configuration, the following tests were performed for an HTML client requesting map services from a pooled ArcGIS Server object. Non-pooled objects were not tested as they require custom ADF programming to ensure desired behavior, etc. The tests were performed on the clustered ArcGIS Server Web application using the virtual URL (http://eslweb.esri.com/Webapplication1). For the report, fail-over\fail-back is the time it takes client to successfully access ArcGIS Server Web application during either a Web Server or Map Server shutdown\reboot.

- Tested simulated failure of a Web Server to ensure that client requests were re-routed to the surviving Web Server via NLB control logic.
- Tested fail-back after rebooting a Web Server to ensure that client requests were once again load balanced between the two Web Servers via NLB control logic.
- Tested simulated failure of a Map Server to ensure that client requests were re-routed to the surviving SOC via SOM logic.
- Tested fail-back following a Map Server reboot to ensure that client requests were once again load balanced between the two SOCs via SOM logic.

Test Results and Analysis

Web Server fail-over and fail-back scenarios for pooled objects was successfully demonstrated. The Web Server fail-over time was approximately 5 seconds. That is, it took about 5 seconds before the client was able to successfully hit the ArcGIS Server Web application on the surviving node. Fail-back time took approximately 15 seconds for clients to sucessfully access the ArcGIS Server web application on rebooted Web Server once it was brought back online. Both fail-over and fail-back scenarios requires a new web browser instance to be opened. Performing any operations in existing web browser window will result in “Application error” web page.

Map Server fail-over and fail-back scenarios were also successfully tested. Following a Map Server shutdown, it took approximately 15 seconds for clients to start hitting the remaining SOCs on the second Map Server. For fail-back, the server object for both SOMs needs to be restarted to reinitialize the SOC processes (ArcSOC.exe) on the recovered Map Server. Once ArcSOC.exe processes are running, the SOMs will immediately start to load balance between Map Servers.

Recovery Procedure Following a Web Server reboot

Once a failed node is brough back online, clients will need to open a new web browser. If a client attempts to access the web site in an existing browser they may encounter a “Application error” web page.

Recovery Procedure Following Map Server reboot

Once the failed Map Server is brought back online, the server object will need to be restarted for both SOM configurations. After both server objects are restarted, clients will need to open a new web browser. Failure to open a new web browser can result in a “Application error” web page.

Conclusions

Testing demonstrated that ArcGIS Server 9.x could be integrated with a Microsoft NLB configuration, however the fail-over process is not completely transparent and manual intervention is required in certain scenarios. In the event of a Web Server failure, some clients were unable to access the ArcGIS Server Web application. For this scenario affected clients must open a new web browser instance to reconnect.

Configuration Support

It is important to realize that for enterprise configurations such as the one described in this paper, ESRI Support will typically be limited to supporting only the ESRI software components (ArcIMS, ArcSDE, etc.). ESRI Support strives to provide the best assistance possible, but problems or questions regarding third-party applications and components may require you to contact the support services provided by the respective vendor.

ESRI offers ongoing, dedicated assistance with the design and/or implementation of an enterprise configuration through our Professional Services Department. Professional Services staff has real-world experience with enterprise configurations, as well as expert knowledge of ESRI resources. To learn more about what ESRI Professional Services can offer your organization, please see the contact information posted at http://www.esri.com/consulting/contact.html.