Load Balancing
Insignia Medical Systems

Deployment Guide
v1.0.0
1. About this Guide

This guide details the steps required to configure a load balanced Insignia Medical System environment utilizing Loadbalancer.org appliances. It includes details on load balancing DICOM & HL7.

For more information about initial appliance deployment, network configuration and using the Web User Interface (WebUI), please also refer to the relevant Administration Manual:

- v7 Administration Manual
- v8 Administration Manual

2. Loadbalancer.org Appliances Supported

All our products can be used with Medical Imaging and Information Systems. The complete list of models is shown below:

<table>
<thead>
<tr>
<th>Discontinued Models</th>
<th>Current Models *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterprise R16</td>
<td>Enterprise R20</td>
</tr>
<tr>
<td>Enterprise VA R16</td>
<td>Enterprise MAX</td>
</tr>
<tr>
<td>Enterprise VA</td>
<td>Enterprise 10G</td>
</tr>
<tr>
<td></td>
<td>Enterprise 40G</td>
</tr>
<tr>
<td></td>
<td>Enterprise Ultra</td>
</tr>
<tr>
<td></td>
<td>Enterprise VA R20</td>
</tr>
<tr>
<td></td>
<td>Enterprise VA MAX</td>
</tr>
</tbody>
</table>

* For full specifications of these models please refer to: [http://www.loadbalancer.org/products/hardware](http://www.loadbalancer.org/products/hardware)

3. Loadbalancer.org Software Versions Supported

- V8.4 & later

4. Medical Systems Supported

- Any systems that utilizes medical system standards & protocols such as DICOM and HL7
5. Medical Information System Standards & Protocols

DICOM
The Digital Imaging and Communications in Medicine (DICOM) Standard describes the means of formatting, storing and exchanging medical images and image related information to facilitate the connectivity of medical devices and systems. The DICOM Standard endorsed by the National Electrical Manufacturers Association (NEMA) is a result of joint efforts of users and manufacturers of medical imaging and health-care information technology.

Today, virtually all imaging devices (Modalities) that are used in radiology, such as CT, MRI, Ultrasound, RF, and other digital rooms, supports the DICOM standard for the exchange of images and related information.

HL7
Health Level Seven (HL7) is an American National Standards Institute accredited Standards Developing Organization (SDO) operating in the health-care arena. Since its inception, HL7 has specified standards for a large number of application areas. HL7 standards cover generic application fields such as patient administration, patient care, order entry, results reporting, document and financial management. In addition to that, HL7 addresses the departmental information system communication needs of clinical specialties like laboratory medicine and diagnostic imaging. HL7 is the language used for communication between health-care IT systems.

6. Load Balancing Overview

Basic Concepts
To provide resilience and high availability, multiple Virtual Services (VIPs) are configured for the various protocols and systems. Clients and systems then connect to these VIPs rather than directly to the application servers. Each VIP can be configured in one of the following ways:

- **Load balanced mode**
  
  *Load is distributed across all configured servers/ endpoints*

- **Failover mode**
  
  *The second server is used only when the first server/endpoint fails*

Load Balancer Deployment
The following diagram shows a simplified view of Insignia Medical System in load balancing mode:
The following diagram shows a simplified view of Insignia Medical System in failover mode:

**Notes:**

- **VIP (Virtual IP)** – This is IP address presented by the load balancer. Clients and other systems connect to this rather than directly to the back end servers/endpoints.
- A single load balancer appliance can be used to load balance all services. More that one load balancer appliance may be required depending on throughput and physical network topology.
- All Loadbalancer.org models support unlimited VIPs except the Enterprise R20 which supports up to 5 VIPs, each with up to 4 load balanced servers.
Load Balancing Deployment Modes

The load balancer supports the following deployment modes:

**Layer 4 DR Mode** – this mode offers the best performance and requires limited physical Real Server changes. The load balanced application must be able to bind to the Real Servers own IP address and the VIP at the same time. This mode requires the “ARP Problem” to be solved as described on page 18. This mode is transparent, i.e. the Real Servers will see the source IP address of the client.

**Layer 4 NAT Mode** – this mode is also a high performance solution but not as fast as DR mode. It requires the implementation of a two-arm infrastructure with an internal and external subnet to carry out the translation (the same way a firewall works). Also each Real Server must use the load balancer as the default gateway. This mode is transparent, i.e. the Real Servers will see the source IP address of the client.

**Layer 4 SNAT Mode** – this mode is also a high performance solution but not as fast as the other layer 4 modes. It does not require any changes to the Real Servers and can be deployed in one-arm or two-arm mode. This mode is ideal for example when you want to load balance both TCP and UDP but you’re unable to use DR mode or NAT mode due to network topology or Real Server related reasons. This mode is non-transparent, i.e. the Real Servers will see the source IP address of the load balancer.

**Layer 7 SNAT Mode** – this mode offers greater flexibility but at lower performance levels. It supports HTTP cookie insertion, RDP cookies, Connection Broker integration and works very well with either Pound or STunnel when SSL termination is required. It does not require any changes to the Real Servers and can be deployed in one-arm or two-arm mode and. HAProxy is a high performance solution, but since it operates as a full proxy, it cannot perform as fast as the layer 4 solutions. This mode is non-transparent, i.e. the Real Servers will see the source IP address of the load balancer.

**Our Recommendation**

Where possible we recommend that Layer 4 Direct Routing (DR) mode is used. This mode offers the best possible performance since replies go directly from the Real Servers to the client, not via the load balancer. It’s also relatively simple to implement. Ultimately, the final choice does depend on your specific requirements and infrastructure.

Note: If you are using Microsoft Windows Real Servers (i.e. the backend servers) make sure that Windows NLB (Network Load Balancing) is **completely disabled** to ensure that this does not interfere with the operation of the load balancer.

**Load Balanced Ports & Services**

The following tables shows the typical ports/services that are load balanced.

<table>
<thead>
<tr>
<th>Port</th>
<th>Protocols</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>104</td>
<td>TCP/DICOM</td>
<td>Exchange of images and related information</td>
</tr>
<tr>
<td>2575</td>
<td>TCP/HL7</td>
<td>Communication between health-care IT systems</td>
</tr>
</tbody>
</table>
Persistence (Server Affinity)
Source IP address persistence is used for all protocols. This ensures that a particular client will connect to the same load balanced server/endpoint for the duration of the session.

Server Health Checking
The default health-check used for new VIPs is a TCP port connect. This verifies that the port is open and accepting connections. However, it does not necessarily guarantee that the associated service is fully operational. Also, repeated ongoing connections to the service port may cause multiple log entries reporting incomplete connections or other issues.

More robust service oriented health-checks can be configured for both layer 4 and layer 7 services using the negotiate option. This effectively tests and verifies the running service.

For example, the load balancer can be configured to look for specific content on an HTTP web page on the load balanced Real Server. If the page can be opened and the content can be found, the check will have passed. If not, the check will fail and the server/endpoint will be marked as down.

If the service running is not HTTP based, a custom page could be setup on the load balanced servers that simply indicates service status. The load balancer can then use this for health checking.

The page to check and the content to be verified can easily be configured for layer 4 and layer 7 VIPs using the WebUI. Select the required negotiate option and configure the required settings. For more details on configuring health-checks please refer to Chapter 8 in the Administration Manual.

Note: The configuration examples in this guide use a TCP port connect (the default) to check the health of load balanced servers.
7. Loadbalancer.org Appliance – the Basics

Virtual Appliance Download & Deployment
A fully featured, fully supported 30 day trial is available if you are conducting a PoC (Proof of Concept) deployment. The VA is currently available for VMware, Virtual Box, Hyper-V, KVM and XEN and has been optimized for each Hypervisor. By default, the VA is allocated 1 CPU, 2GB of RAM and has an 8GB virtual disk. The Virtual Appliance can be downloaded here.

Note: The same download is used for the licensed product, the only difference is that a license key file (supplied by our sales team when the product is purchased) must be applied using the appliance’s WebUI.

Note: Please refer to the Administration Manual and the ReadMe.txt text file included in the VA download for more detailed information on deploying the VA using various Hypervisors.

Initial Network Configuration
The IP address, subnet mask, default gateway and DNS settings can be configured in several ways as detailed below:

**Method 1 - Using the Network Setup Wizard at the console**
After boot up, follow the instructions on the console to configure the IP address, subnet mask, default gateway and DNS settings.

**Method 2 - Using the WebUI**
Using a browser, connect to the WebUI on the default IP address/port: https://192.168.2.21:9443
To set the IP address & subnet mask, use: Local Configuration > Network Interface Configuration
To set the default gateway, use: Local Configuration > Routing
To configure DNS settings, use: Local Configuration > Hostname & DNS

Accessing the Web User Interface (WebUI)

1. Browse to the following URL: https://192.168.2.21:9443/lbadmin/
   (replace with your IP address if it's been changed)
   * Note the port number → 9443
2. Login to the WebUI:

   **Username:** loadbalancer  
   **Password:** loadbalancer

   Note: To change the password, use the WebUI menu option: Maintenance > Passwords.

Once logged in, the WebUI will be displayed as shown below:
Clustered Pair Configuration

Loadbalancer.org recommend that load balancer appliances are deployed in pairs for high availability. In this guide a single unit is deployed first, adding a secondary slave unit is covered in section 1 of the Appendix on page 16.
8. Appliance & Server Configuration

Load Balancing Mode

As mentioned on page 6, Virtual Services can be configured in one of four fundamental ways, i.e. Layer 4 DR mode, Layer 4 NAT mode, Layer 4 SNAT mode or Layer 7 SNAT mode. The following sections illustrate how to configure the Virtual Services using various modes. If a different mode is required for a particular VIP, please refer to one of the other sections that uses that mode for guidance. Please also don’t hesitate to contact our support team: support@loadbalancer.org.

Health-Check Configuration

As mentioned on page 7, health checks can be configured in several different ways. The sections below all use a TCP port connect on the service port.

Load Balancing DICOM
(Using Layer 4 DR Mode)

Setting up the Virtual Service (VIP)

1. Using the WebUI, navigate to: Cluster Configuration > Layer 4 – Virtual Services and click Add a New Virtual Service
2. Enter the following details:

3. Enter an appropriate name (Label) for the Virtual Service, e.g. DS_DICOM
4. Set the Virtual Service IP address field to the required IP address, e.g. 172.26.11.70
5. Set the Virtual Service Ports field to the required port(s), e.g. 104
6. Set Protocol to TCP
7. Set Forwarding Method to Direct Routing
8. Click Update
9. Now click Modify next to the newly created Virtual Service
10. Set Persistent Timeout to 300, i.e. 5 minutes
11. Set the Check Type to External Script
12. Set the External script to DICOM-C-ECHO
13. Set the Check Port to the required port - by default this is set to the first port (104)
14. Click Update

Setting up the Real Servers (RIPs)
1. Using the WebUI, navigate to: Cluster Configuration > Layer 4 – Real Servers and click Add a new Real Server next to the newly created Virtual Service
2. Enter the following details:

   ![Layer 4 Add a new Real Server - DS_DICOM](image)

3. Enter an appropriate name (Label) for the first DICOM server, e.g. DS1
4. Change the Real Server IP Address field to the required IP address, e.g. 172.26.11.100
5. Click Update
6. Now repeat for your remaining DICOM server(s)

Configuring the load balanced DICOM servers
1. As mentioned on page 6, the ARP problem must be solved for all load balanced servers. Please refer to section 2 in the appendix on page 18 for more details.
Load Balancing HL7
(Using Layer 7 SNAT Mode)

Setting up the Virtual Service (VIP)

1. Using the WebUI, navigate to: Cluster Configuration > Layer 7 – Virtual Services and click Add a New Virtual Service
2. Enter the following details:

3. Enter an appropriate name (Label) for the Virtual Service, e.g. HL7
4. Set the Virtual Service IP address field to the required IP address, e.g. 172.26.11.71
5. Set the Virtual Service Ports field to the required port, e.g. 2575
6. Set the Layer 7 Protocol to TCP Mode
7. Click Update
8. Now click Modify next to the newly created Virtual Service
9. Ensure Persistence Mode is set to None
10. Set the Health Checks to External Script
11. Set the Check script to ping.sh
12. Set the Fallback Server IP address field to that of the Standby node e.g. 172.26.11.103
13. Set the Port field to 2575
14. Click Update

Setting up the Real Servers (RIPs)

1. Using the WebUI, navigate to: Cluster Configuration > Layer 7 – Real Servers and click Add a new Real Server next to the newly created Virtual Service
2. Enter the following details:
3. Enter an appropriate name (Label) for the first HL7 server, e.g. IMS1
4. Change the Real Server IP Address field to the required IP address, e.g. 172.26.11.101
5. Set the Real Server Port field to 2575
6. Click Update

Reload HAProxy
1. To apply the new settings restart HAProxy using the WebUI option: Maintenance > Restart Services and clicking Reload HAProxy

Note: If you will be configuring additional layer 7 services, you can reload HAProxy once all layer 7 Virtual Services and Real Servers have been defined.
9. Testing & Verification

Using the System Overview
Verify that all VIPs & associated RIPs are reported as up (green) as shown below:

![System Overview](image1)

If certain servers are down, i.e. failing their health check, they will be highlighted red as shown below:

![System Overview](image2)

System Logs & Reports
Various system logs & reports can be used to help diagnose problems and help solve appliance issues. Logs can be accessed using the WebUI options: Logs & Reports.

10. Technical Support
If you have any questions regarding the appliance or would like assistance designing your deployment, please don't hesitate to contact our support team: support@loadbalancer.org.

11. Further Documentation

12. Conclusion
Loadbalancer.org appliances provide a very cost effective and flexible solution for highly available load balanced Medical Imaging Systems environments.
13. Appendix

1 - Clustered Pair Configuration – Adding a Slave Unit
If you initially configured just the master unit and now need to add a slave - our recommended procedure, please refer to the relevant section below for more details:

Note: A number of settings are not replicated as part of the master/slave pairing process and therefore must be manually configured on the slave appliance. These are listed below:

- Hostname & DNS settings
- Network settings including IP addresses, bonding configuration and VLANs
- Routing configuration including default gateways and static routes
- Date & time settings
- Physical – Advanced Configuration settings including Internet Proxy IP address & port, Firewall table size, SMTP relay and Syslog server
- SNMP settings
- Graphing settings
- Firewall Script & Firewall Lockdown Script settings
- Software updates

Version 7:
Please refer to Chapter 8 – Appliance Clustering for HA in the v7 Administration Manual.

Version 8:
To add a slave node – i.e. create a highly available clustered pair:
1. Deploy a second appliance that will be the slave and configure initial network settings
2. Using the WebUI, navigate to: Cluster Configuration > High-Availability Configuration
   Specify the IP address and the loadbalancer users password (the default is 'loadbalancer') for the slave (peer) appliance as shown above

3. Click Add new node

4. The pairing process now commences as shown below:

   ![CREATE A CLUSTERED PAIR](image)

5. Once complete, the following will be displayed:

   ![HIGH AVAILABILITY CONFIGURATION - MASTER](image)

6. To finalize the configuration, restart heartbeat and any other services as prompted in the blue message box at the top of the screen

   Note: Clicking the Restart Heartbeat button on the master appliance will also automatically restart heartbeat on the slave appliance.

   Note: Please refer to chapter 9 – Appliance Clustering for HA in the Administration Manual for more detailed information on configuring HA with 2 appliances.
2 – Solving the ARP Problem

Layer 4 DR mode works by changing the MAC address of the inbound packets to match the Real Server selected by the load balancing algorithm. To enable DR mode to operate:

- Each Real Server must be configured to accept packets destined for both the VIP address and the Real Servers IP address (RIP). This is because in DR mode the destination address of load balanced packets is the VIP address, whilst for other traffic such as health-checks, administration traffic etc. it's the Real Server's own IP address (the RIP). The service/process (e.g. IIS, httpd) must respond to both addresses.

- Each Real Server must be configured so that it does not respond to ARP requests for the VIP address – only the load balancer should do this.

Configuring the Real Servers in this way is referred to as 'Solving the ARP problem'. The steps required depend on the particular OS being used.

For detailed steps on solving the ARP problem for Linux, Windows and various other operating systems, please refer to the Administration Manual and search for “DR Mode Considerations”.
14. Document Revision History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Change</th>
<th>Reason for Change</th>
<th>Changed By</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0.0</td>
<td>14 January 2020</td>
<td>Initial document creation</td>
<td></td>
<td>IBG</td>
</tr>
</tbody>
</table>
About Loadbalancer.org

Loadbalancer.org's mission is to ensure that its clients' businesses are never interrupted. The load balancer experts ask the right questions to get to the heart of what matters, bringing a depth of understanding to each deployment. Experience enables Loadbalancer.org engineers to design less complex, unbreakable solutions - and to provide exceptional personalized support.