loadbalancer.org

**Load Balancing
Microsoft Exchange
2010**

**Deployment Guide
v1.9.2**

Table of Contents

1. About this Guide

This guide details the steps required to configure a load balanced Microsoft Exchange 2010 environment utilizing Loadbalancer.org appliances. It covers the configuration of the load balancers and also any Microsoft Exchange 2010 configuration changes that are required to enable load balancing.

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 Exchange 2010. The complete list of models is shown below:

Discontinued Models	Current Models *
Enterprise R16	Enterprise R20
Enterprise VA R16	Enterprise MAX
Enterprise VA	Enterprise 10G
Enterprise R320	Enterprise R330
	Enterprise VA R20
	Enterprise VA MAX
	Enterprise AWS **
	Enterprise AZURE **

* For full specifications of these models please refer to: http://www.loadbalancer.org/products/hardware
** Some features may not be supported, please check with Loadbalancer.org support

3. Loadbalancer.org Software Versions Supported

* v7.6.4 and later

4. Microsoft Exchange Software Versions Supported

* Microsoft Exchange 2010 – all versions

5. Exchange Server 2010

Exchange 2010 is Microsoft's enterprise level messaging and collaboration server.

6. Exchange 2010 Server Roles

System functionality is split into five role as shown in the following table. Mandatory roles are Mailbox, Client Access and Hub Transport. The Edge Transport and Unified Messaging roles are optional and depend on the infrastructure and operational requirements.

Role	Purpose
Mailbox Server	This server hosts mailboxes and public folders.
Client Access Server	This is the server that hosts the client protocols, such as Post Office Protocol 3 (POP3), Internet Message Access Protocol 4 (IMAP4), Secure Hypertext Transfer Protocol (HTTPS), Outlook Anywhere, Availability service, and Autodiscover service. The Client Access Server also hosts Web services. *Note: A number of issues have been seen with IOS-7 on the iPhone when used with ActiveSync. Upgrading to IOS-8 resolved these issues.*
Unified Messaging Server	This is the server that connects a Private Branch exchange (PBX) system to Exchange 2010.
Hub Transport Server	This is the mail routing server that routes mail within the Exchange organization.
Edge Transport Server	This is the mail routing server that typically sits at the perimeter of the topology and routes mail in to and out of the Exchange organization.

CLIENT ACCESS SERVER

The Client Access Server Role also known as CAS, provides Exchange connectivity for all clients regardless of client type or protocol including Outlook Web App (aka OWA), ActiveSync, POP3, IMAP4, RPC Client Access (MAPI) and Outlook Anywhere (previously known as RPC over HTTP). Exchange now has a single common path through which all data access occurs.

Therefore, due to the critical nature of this role, it's common practice to implement load balancing and redundancy technologies to ensure availability.

HUB TRANSPORT SERVER

For internal server to server mail traffic, HT servers are automatically load balanced by Exchange 2010 and there is no need to configure any type of load balancing mechanism to load balance the mail submission traffic among Exchange servers.

However, some sites may decide not to deploy an ET server. In this scenario, inbound SMTP mail is typically forwarded from a third party smart host directly to the HT server. Also, internal applications and systems often need to send email via Exchange and typically are only able to do so using an SMTP connection. To provide redundancy in these cases, additional load balancing & HA techniques are required to ensure availability of the HT role.

MAILBOX SERVER/DATABASE AVAILABILITY GROUP'S (DAG)

Exchange 2010 brings the ability to combine both CAS and HT roles on a mailbox server that is also

configured as a DAG member. This permits a highly available solution using just two Exchange servers and one or two (configured as a clustered pair for added redundancy) Loadbalancer.org appliances. Another server is needed to act as the witness server, but this doesn't need to be an Exchange server. It could be any Windows 2003/2008 file server within the environment.

> Note:
>
> DAG's utilize Microsoft Clustering Services which cannot be enabled on the same server as Microsoft Network Load Balancing (NLB). Therefore, using Microsoft NLB is not an option in this case. Using a Loadbalancer.org appliance provides an ideal solution.

7. Load Balancing Exchange 2010

> Note:
>
> It's highly recommended that you have a working Exchange 2010 environment first before implementing the load balancer.

WHICH ROLES?

The CAS role does not have any built-in load balancing functionality. The HT role does provide load balancing functionality for server to server mail traffic, but not external SMTP traffic that arrives from other applications or from outside the organization directly to the HT server. Therefore, it is a common requirement to load balance both the CAS and HT roles. In some cases only the CAS role is load balanced. The exact load balancing requirements depend on the number of servers in use and how/where the roles are deployed.

PERSISTENCE (AKA SERVER AFFINITY)

Some Exchange 2010 protocols require affinity and others do not. For more details please refer to the following Microsoft Technet article: http://technet.microsoft.com/en-us/library/ff625248.aspx

For additional information on the various affinity options, please refer to the following Microsoft Technet article: http://technet.microsoft.com/en-us/library/ff625247.aspx#affinity

Summary of Persistence Requirements:

Persistence – Required	Persistence – Recommended	Persistence – Not Required
Outlook Web App	Outlook Anywhere	Offline Address Book
Exchange Control Panel	ActiveSync	AutoDiscover
Exchange Web Service	Address Book Service	POP3
RPC Client Access Service	Remote PowerShell	IMAP4

For simplicity and consistency we recommend that source IP persistence is used for all protocols that require persistence between client and back-end server.

> Note:
>
> If your Outlook Web App users pass through a NAT device to reach the load balancer then IP based persistence may not be appropriate since the source IP address may be the same for all users. This would cause all OWA sessions to be directed to the same backend CAS. In this

> situation, HTTP cookie persistence can be used. This requires the HTTPS traffic to be terminated on the load balancer to allow the cookie to be read. For more details on this, please refer to the Appendix.

VIRTUAL SERVICE (VIP) REQUIREMENTS

There are a number of options when deciding on the number of VIPs required for the CAS and HT roles. This deployment guide presents two options as shown below:

Option 1 – Four VIPs (Used for the example configuration in this guide)

This method uses three VIPs for the CAS role, and one VIP for the HT role as follows:

- CAS role – HTTPS & HTTP services
- CAS role – RPC services
- CAS role – IMAP4 or POP3 services (*if used/required*)
- HT role – SMTP services

This method allows the settings for each VIP to be customized (e.g. persistence/affinity options) to suit the service being load balanced and also ensures more granular health-checks.

> Note:
>
> IMAP4 and POP3 are not typically used. Therefore the IMAP4 and POP3 VIPs are not generally required.

Option 2 – Two VIPs (not recommended for production deployments)

This method uses two VIPs - one VIP for all CAS services, and one VIP for the HT role. This is useful for rapid deployments and is only recommended for evaluation & testing purposes. For details of this, please refer to section 1 in the Appendix on page 26.

PORT REQUIREMENTS

The following table shows the port list that must be load balanced for the CAS and HT roles. Note that some services such as IMAP4 or POP3 may not be used in your environment.

TCP Port	Role(s)	Uses
25	HT	SMTP
80	CAS	HTTP – various
110	CAS	POP3 clients
135	CAS	RPC end point mapper
143	CAS	IMAP4 clients
443	CAS	HTTPS – various
993	CAS	Secure IMAP4 clients
995	CAS	Secure POP3 clients
60200*	CAS	Static port for RPC client access service

60201*	CAS	Static port for Exchange address book service

- **HT** = Hub Transport Server, **CAS** = Client Access Server

- * These ports have been chosen as the static RPC ports. Microsoft recommends that any port within the range 59531 to 60554 should be used, and that the same ports should be used on all Client Access Servers within the same AD site.

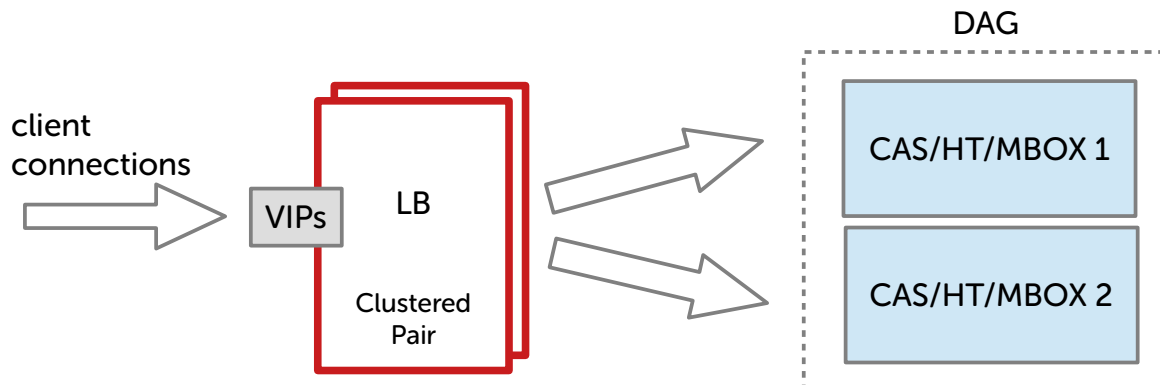- For a full Exchange Server 2010 port list, please refer to the following Microsoft Technet article: http://technet.microsoft.com/en-us/library/bb331973.aspx

> **Note:**
>
> If you use Microsoft ISA Server or TMG, you may need to disable the RPC Filter to allow Outlook client RPC related communication to work correctly.

8. Deployment Concept

There are multiple ways to deploy Exchange, but in this example two servers are used. Each server hosts the CAS & HT roles, as well as the Mailbox role in a DAG configuration. This provides high availability for these three key Exchange roles and uses a minimum number of Exchange servers.

Clients then connect to the Virtual Services (VIPs) on the load balancer rather than connecting directly to one of the Exchange servers. These connections are then load balanced across the Exchange servers to distribute the load according to the load balancing algorithm selected.

VIPs = **V**irtual **IP** Addresses

> **Note:**
>
> The load balancer can be deployed as a single unit, although Loadbalancer.org recommend a clustered pair for resilience & high availability. Please refer to section 6 in the appendix on page 29 for more details on configuring a clustered pair.

9. 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: **http://192.168.2.21:9080**
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*

Method 3 - Using Linux commands

At the console, set the initial IP address using the following command:
```
ip addr add <IP address>/<mask> dev eth0
```

At the console, set the initial default gateway using the following command:
```
route add default gw <IP address> <interface>
```

At the console, set the DNS server using the following command:
```
echo nameserver <IP address> >> /etc/resolv.conf
```

> **Note:**
>
> If method 3 is used, you must also configure these settings using the WebUI, otherwise the settings will be lost after a reboot.

ACCESSING THE WEB USER INTERFACE (WEBUI)

The WebUI can be accessed via HTTP at the following URL: **http://192.168.2.21:9080/lbadmin**
* Note the port number → **9080**

The WebUI can be accessed via HTTPS at the following URL: **https://192.168.2.21:9443/lbadmin**
* Note the port number → **9443**

(replace 192.168.2.21 with the IP address of your load balancer if it's been changed from the default)

Login using the following credentials:

> **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 on the following page:

(shows v8.2.x)

HA 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 6 of the Appendix on page 29.

LOAD BALANCER DEPLOYMENT METHOD

As with several other Microsoft applications, the load balancer for Exchange 2010 is deployed in one-arm SNAT mode (*Source Network Address Translation*) at layer 7 using HAProxy. This mode is recommended by Microsoft and also has the advantage that it requires no changes to the Exchange 2010 servers.

NOTE: Source IP Address Transparency

It's important to remember that when using HAProxy, the source IP address of packets reaching the Exchange servers will be the IP address of the load balancer and not the source IP address of the client.

If this is an issue, please refer to section 4 in the Appendix on page 27 for details on using TProxy. TProxy enables the original source IP address to be maintained, but requires that separate subnets are used, and also requires that the load balancer becomes the default gateway for the Exchange Servers. Enabling TProxy is a global setting and therefore effects all Virtual Services configured on the load balancer which may not always be desirable.

Transparency is normally only an issue for SMTP traffic at the receive connector. System Administrators typically want to lock down receive connectors to accept SMTP connections only from a controlled set of devices such as external smart mail hosts, printers, networked photocopiers etc.

If transparency for SMTP is the only issue, there are a couple of options available to address this:

Option 1 – Use a Layer 7 VIP for SMTP as detailed on page 23 of this guide and also enable the load balancers on-board firewall to lock down inbound SMTP connections rather than the receive connector. This is covered in section 2 of the Appendix on page 26.

Option 2 – Configure a layer 4 Virtual Service for SMTP rather than a layer 7 (HAProxy) based Virtual Service. Layer 4 is transparent by default so the source IP address is maintained. This is covered in section 5 of the Appendix on page 28.

10. Exchange 2010 Configuration for Load Balancing

The Exchange 2010 need to be configured in preparation for being load balanced. This is detailed in the steps below.

STEP 1 – CONFIGURE THE CAS ARRAY & INTERNAL/EXTERNAL URL'S

THE CAS ARRAY

To enable multiple CAS servers to work with the load balancer, a CAS array must be configured in your Exchange environment using the 'New-ClientAccessArray' command as detailed below. Exact configuration details obviously depend on the specific environment.

- Install the CAS 2010 servers

- Create a DNS record for the CAS Array, this should be the same as the Virtual Service's IP address, e.g. cas.domain.com (also refer to the Load Balancer configuration section starting on page 12)

- Create a new CAS array object using the New-ClientAccessArray command in the Exchange 2010 management shell:

    New-ClientAccessArray –Name "CAS-array" –FQDN "cas.domain.com" -Site "YourSiteName"

> **Note:**
>
> Change "**YourSiteName**" to the AD site appropriate for your Client Access Servers
> Change "**cas.domain.com**" to the FQDN of the CAS array
> To determine your site name:
>
> - °   On an Exchange server use: **get-adsite**
> - °   On a Domain Controller use: **nltest /server:<servername> /dsgetsite**

- If the mail database already existed before creating the array, you'll also need to run the following command to relate the new CAS array to the database:

  Set-MailboxDatabase "NameofDatabase" -RpcClientAccessServer "cas.domain.com"

  > **Note:**
  >
  > Change "**cas.domain.com**" to the FQDN of the CAS array

- To verify the configuration of the CAS array, use the following commands from the Exchange Shell:

  - °   to list the available Client Access Servers: **get-ClientAccessServer**
  - °   to list the Client Access Array and its members: **get-ClientAccessArray**

THE INTERNAL & EXTERNAL URL'S

Once the CAS Array has been created it's important to remember that clients should then connect using the CAS Array based address rather than individual CAS servers. This applies to both internal and external URL's. The following list provides a number of examples that illustrate how the various URL's can be checked/configured:

```
Get-OABvirtualdirectory | fl identity,internalurl,externalurl
Set-OABvirtualDirectory -Identity "CAS01\OAB (Default Web Site)" -ExternalUrl
"https://cas.domain.com/OAB"
```

```
Get-ActiveSyncVirtualDirectory | fl identity,internalurl,externalurl
Get-AutodiscoverVirtualDirectory -server CAS01 | Set-AutodiscoverVirtualDirectory -ExternalUrl
https://cas.Domain.local/Autodiscover/Autodiscover.xml
```

```
Get-ClientAccessServer | fl Identity,AutoDiscoverServiceInternalUri
Set-ClientAccessServer -Identity CAS01 -AutoDiscoverServiceInternalUri
https://cas.domain.com/Autodiscover/Autodiscover.xml
```

```
Get-Webservicesvirtualdirectory | fl identity,internalurl,externalurl
Set-WebServicesVirtualDirectory -Identity "CAS01\EWS (Default Web Site)" -ExternalUrl
https://cas.domain.com/ews/exchange.asmx
```

Get-OWAvirtualdirectory | fl identity,internalurl,externalurl

Set-OWAvirtualDirectory -Identity "CAS01\owa (Default Web Site)" -ExternalUrl
https://cas.domain.com/owa

Get-ECPvirtualdirectory | fl identity,internalurl,externalurl

Set-ECPvirtualDirectory -Identity "CAS01\ECP (default web site)" -Externalurl https://cas.domain.com/ecp

Get-ActiveSyncVirtualDirectory | fl identity,internalurl,externalurl

Set-ActiveSyncVirtualDirectory -Identity "CAS01\Microsoft-Server-ActiveSync (Default Web Site)"
-ExternalUrl https://cas.domain.com/Microsoft-Server-ActiveSync

Please refer to the following Microsoft article for further details:

http://social.technet.microsoft.com/wiki/contents/articles/5163.managing-exchange-2010-externalinternal-url-s-via-powershell.aspx

STEP 2 – CONFIGURE STATIC RPC PORTS

By default the RPC Client Access service and the Address Book Service on an Exchange 2010 Client Access Server uses the TCP End Point Mapper port (TCP/135) and the dynamic RPC port range (6005-59530) for outgoing connections when an Outlook clients establishes a connection to Exchange.

Since this would add complexity to the load balancers configuration, and would also uses substantially more on-board memory, it's recommended to configure static ports as described below.

> Note:
>
> If you later apply a service pack to your Exchange servers, re-check that the settings described in this section are still valid and if required re-configure them.
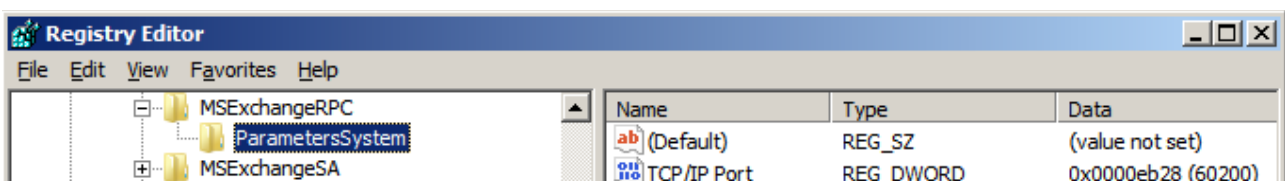
CONFIGURE THE RPC CLIENT ACCESS SERVICE

To set a static port for the RPC Client Access Service, open the registry on each CAS and navigate to:

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\services\MSExchangeRPC

Here, you need to create a new key named **ParametersSystem**, and under this key create a new **DWORD (32-bit) Value** named **TCP/IP Port** as shown below. The Value for the **DWORD** should be the port number you want to use. Microsoft recommends you set this to a unique value between 59531 and 60554 and use the same value on all CAS. In this deployment guide, the port used is 60200.

IMPORTANT!! Make sure you use a **DWORD Value** for this key

> Note:
>
> Once this registry change has been made, restart the RPC Client Access Service to apply the

loadbalancer.org

> new setting. This process must be completed on all CAS.
>
> If there is a possibility that these ports are already in use, for example if the server was serving clients prior to implementing these changes, then a reboot is recommended rather than a service restart.

> Note:
>
> If you have separate Mailbox Servers, the above steps must also followed on these servers. This ensures that static ports are also used for the RPC Client Access service that runs on the Mailbox Servers.

A list of services running on the various roles can be found here:

http://technet.microsoft.com/en-us/library/ee423542(v=exchg.141).aspx
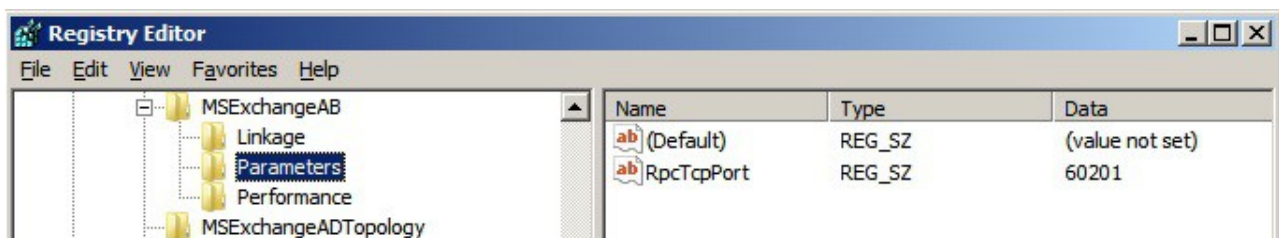
CONFIGURE THE EXCHANGE ADDRESS BOOK SERVICE (SP1 & LATER)

To set a static port for the Address Book Service, open the registry on each CAS and navigate to:

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\services\MSExchangeAB

Here, you need to create a new key named **Parameters**, and under this key create a new **String Value** named **RpcTcpPort** as shown below. Microsoft recommends you set this to a unique value between 59531 and 60554 and use the same value on all CAS. In this deployment guide, the port used is 60201.

IMPORTANT!! Make sure you use a **STRING Value** for this key

> Note:
>
> Once this registry change has been made, restart the Address Book Service to apply the new setting. This process will need to be completed on all CAS.
>
> If there is a possibility that these ports are already in use, for example if the server was serving clients prior to implementing these changes, then a reboot is recommended rather than a service restart.

CONFIGURE THE EXCHANGE ADDRESS BOOK SERVICE (PRE SP1)

For Exchange 2010 without SP1, the static port for the Exchange Address Book service is configured in a different way. First, navigate to the following folder:

C:\Program Files\Microsoft\Exchange Server\V14\Bin

Using Notepad, open the file **microsoft.exchange.addressbook.service.exe.config**

Now change the value for the key **RpcTcpPort** to the port you want to use. The ports specified must be different than the port used for the RPC Client Access Service. In this deployment guide the port used is 60201.

> **Important:**
>
> Once the settings listed in this section have have been configured and the services have been restarted, verify that all servers are listening on these newly configured ports by using the following command in a command window on each Exchange server: **netstat -an -p tcp**

For more information on configuring & verifying static ports, please refer to the following URL:

http://social.technet.microsoft.com/wiki/contents/articles/configure-static-rpc-ports-on-an-exchange-2010-client-access-server.aspx

STEP 3 – CONFIGURE SEND & RECEIVE CONNECTORS

In cases where there is no Edge Transport server, the Hub Transport server must be configured to accept and send mail. It is possible to send and receive directly to/from the Internet, although a more secure and typical configuration would be to use a 3rd party external smart host. To establish mail flow to and from the Internet through a Hub Transport server the basis steps required are:

1. *Create a Send connector on the Hub Transport server to send e-mail to the Internet*
2. *Modify the default Receive connector to allow anonymous connections*

CONFIGURE THE SEND CONNECTOR (USING THE NEW-SENDCONNECTOR CMDLET)

- To configure a new send connector, open the Exchange Management Shell and run the following command on each server:

  New-SendConnector -Name "<Name for this send connector>" -Usage Internet -AddressSpaces "*" -SourceTransportServers "<Hub Transport Server Name>" -DNSRoutingEnabled:$true -UseExternalDNSServersEnabled:$true

CONFIGURE THE RECEIVE CONNECTOR (USING THE SET-RECEIVECONNECTOR CMDLET)

- To change the permissions of the default receive connector, open the Exchange Management Shell and run the following command on each server:

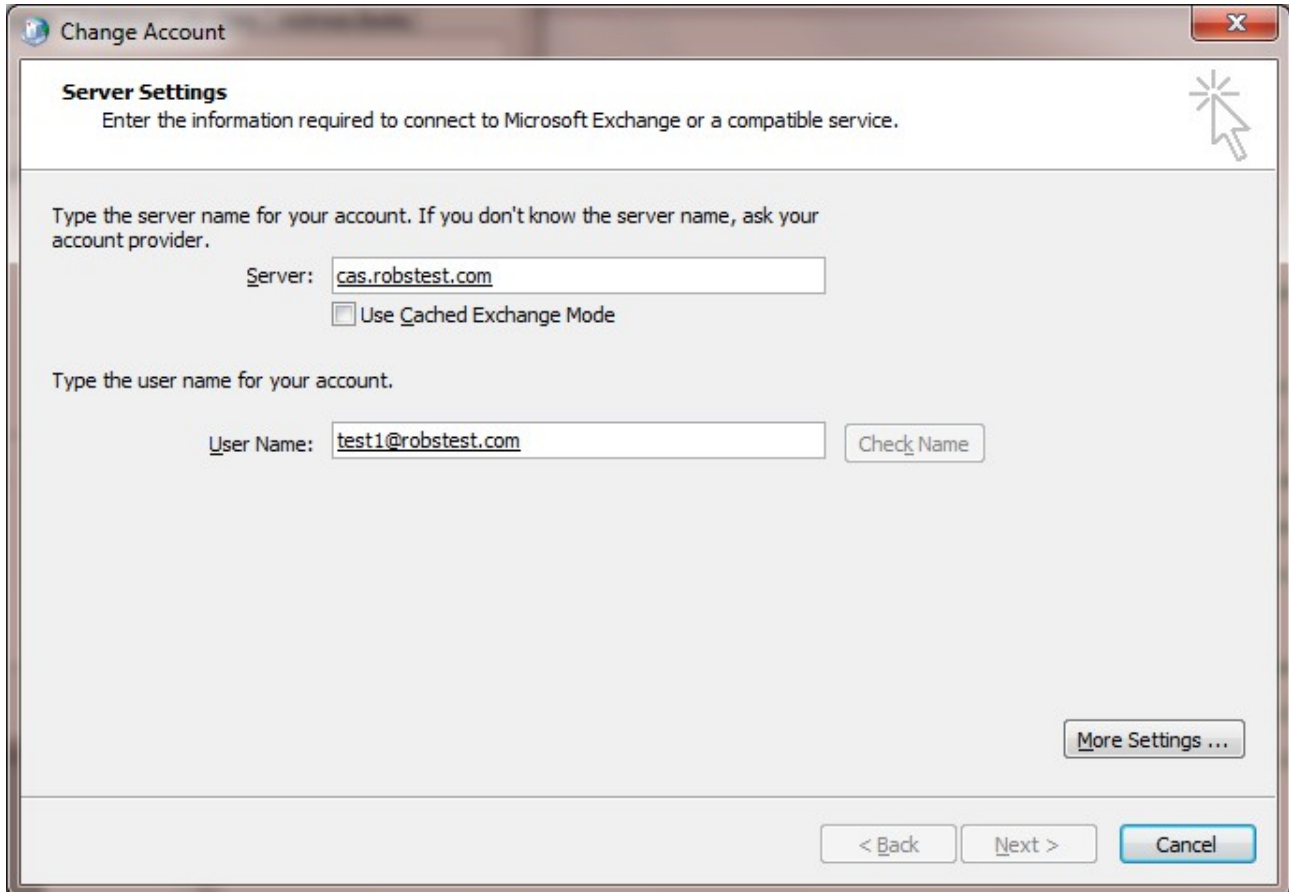 **Set-ReceiveConnector -Identity "Default <ServerName>" -PermissionGroups AnonymousUsers**

  > **Important:**
  >
  > The exact configuration steps required depend on your environment. The commands listed above are provided as examples only.

STEP 4 – MICROSOFT OUTLOOK CLIENT CONFIGURATION

All Outlook clients must be configured to connect to the CAS array rather than an individual Client Access Server. To do this, the Exchange Server Connection settings must be modified. If Autodiscover is enabled this configuration should occur automatically, if Autodiscover is not enabled specify the FQDN of the CAS array configured and enter a valid email account in the User Name field.

For example:

11. Appliance Configuration for Exchange 2010

STEP 1 – CONFIGURE LAYER 7 GLOBAL SETTINGS

To ensure that client connections remain open during periods of inactivity, the Client and Real Server Timeout values must be changed from their default values of 43 seconds and 45 seconds respectively to 1 hour. To do this follow the steps below:

1. Using the WebUI, navigate to: *Cluster Configuration > Layer 7 – Advanced Configuration*

2. Change *Client Timeout* to **3600000** as shown above (i.e. 1 hour)
   *Note: You can also enter **1h** rather than 3600000*
3. Change *Real Server Timeout* to **3600000** as shown above (i.e. 1 hour)
   *Note: You can also enter **1h** rather than 3600000*
4. Click the **Update** button to save the settings

> **Note:**
>
> This step must be completed to avoid Outlook client timeout & reconnection issues that would otherwise occur within around 45 seconds with unchanged default values.

STEP 2 – CONFIGURING THE VIRTUAL SERVICES & REAL SERVERS

VIP1 – CAS ROLE HTTP & HTTPS SERVICES

a) Setting up the Virtual Service

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

3. Enter an appropriate label for the VIP, e.g. **CAS-WEB**
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.30.10**
5. Set the *Virtual Service Ports* field to **80,443**
6. Set *Layer 7 Protocol* to **TCP Mode**
7. Click **Update**
8. Now click **Modify** next to the newly created VIP
9. Set *Balance mode* to **Weighted Round Robin**

> **Note:**
>
> Microsoft recommends that 'Round Robin' rather than 'Least Connection' should be used to help prevent over loading servers when they are brought online. This could occur if Least Connection was selected, since the load balancer would try to balance the number of connections across all real severs and therefore send all new requests to the new server. The trade off here is that using Round Robin will mean that server load may remain unbalanced for some time after bringing a new server into the active pool.

10. Set *Persistence mode* to **Source IP**
11. Change *Persistence Timeout* to 60 (i.e. 1 hour)
12. Click **Update**

b) Setting up the Real Servers

1. Using the WebUI, navigate to: *Cluster Configuration > Layer 7 – Real Servers* and click **Add a new Real Server** next to the newly created VIP
2. Enter the following details:

3. Enter an appropriate label for the RIP, e.g. **CAS1**
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.30.20**
5. Leave the *Real Server Port* field <u>blank</u>
6. Click **Update**
7. Repeat the above steps to add your other CAS Server(s)

> **Note:**
>
> Because SNAT is a full proxy, any server in the cluster can be on any accessible subnet including across the Internet or WAN.

VIP2 – CAS ROLE RPC SERVICES

a) Setting up the Virtual Service

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

Label		CAS-RPC	❓
Virtual Service	IP Address	192.168.30.10	❓
	Ports	135,60200,60201	❓
Layer 7 Protocol		TCP Mode ▾	❓
Manual Configuration		☐	❓

<div align="right">Cancel   Update</div>

3. Enter an appropriate label for the VIP, e.g. **CAS-RPC**
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.30.10**
5. Set the *Virtual Service Ports* field to **135,60200,60201**
6. Set *Layer 7 Protocol* to **TCP Mode**
7. Click **Update**
8. Now click **Modify** next to the newly created VIP
9. Set *Balance mode* to **Weighted Round Robin**

> **Note:**
>
> Microsoft recommends that 'Round Robin' rather than 'Least Connection' should be used to help prevent over loading servers when they are brought online. This could occur if Least Connection was selected, since the load balancer would try to balance the number of connections across all real severs and therefore send all new requests to the new server. The trade off here is that using Round Robin will mean that server load may remain unbalanced for some time after bringing a new server into the active pool.

10. Set *Persistence mode* to **Source IP**

11. Change *Persistence Timeout* to 60 (i.e. 1 hour)
12. Set *Check Port* to **60200** (i.e. the static port used for the RPC client access service)
13. Click **Update**

b) Setting up the Real Servers

1. Using the WebUI, navigate to: *Cluster Configuration > Layer 7 – Real Servers* and click **Add a new Real Server** next to the newly created VIP
2. Enter the following details:

3. Enter an appropriate label for the RIP, e.g. **CAS1**
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.30.20**
5. Leave the *Real Server Port* field blank
6. Click **Update**
7. Repeat the above steps to add your other CAS Server(s)

> Note:
>
> Because SNAT is a full proxy, any server in the cluster can be on any accessible subnet including across the Internet or WAN.

VIP3 – CAS ROLE IMAP4 OR POP3 SERVICES

a) Setting up the Virtual Service

> Note:
>
> These steps show IMAP4 settings, for POP3 change the port numbers from 143 & 993 to 110 & 995.

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

3. Enter an appropriate label for the VIP, e.g. **CAS-IMAP4**
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.30.10**
5. Set the *Virtual Service Ports* field to **143,993**
6. Set *Layer 7 Protocol* to **TCP Mode**
7. Click **Update**
8. Now click **Modify** next to the newly created VIP
9. Set *Balance mode* to **Weighted Round Robin**

> **Note:**
>
> Microsoft recommends that 'Round Robin' rather than 'Least Connection' should be used to help prevent over loading servers when they are brought online. This could occur if Least Connection was selected, since the load balancer would try to balance the number of connections across all real severs and therefore send all new requests to the new server. The trade off here is that using Round Robin will mean that server load may remain unbalanced for some time after bringing a new server into the active pool.

10. Click **Update**

> **Note:**
>
> Persistence is not required for IMAP or POP3.

b) Setting up the Real Servers

1. Using the WebUI, navigate to: *Cluster Configuration > Layer 7 – Real Servers* and click **Add a new Real Server** next to the newly created VIP
2. Enter the following details:

3. Enter an appropriate label for the RIP, e.g. **CAS1**
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.30.20**
5. Leave the *Real Server Port* field blank
6. Click **Update**
7. Repeat the above steps to add your other CAS Server(s)

> **Note:**
>
> Because SNAT is a full proxy, any server in the cluster can be on any accessible subnet including across the Internet or WAN.

VIP4 – HT ROLE SMTP SERVICES

NOTE: Source IP Address Transparency

It's important to remember that when using HAProxy, the source IP address of packets reaching the Exchange servers will be the IP address of the load balancer and not the source IP address of the client.

If this is an issue, please refer to section 4 in the Appendix on page 27 for details on using TProxy. TProxy enables the original source IP address to be maintained, but requires that separate subnets are used, and also requires that the load balancer becomes the default gateway for the Exchange Servers. Enabling TProxy is a global setting and therefore effects all Virtual Services configured on the load balancer which may not always be desirable.

Transparency is normally only an issue for SMTP traffic at the receive connector. System Administrators typically want to lock down receive connectors to accept SMTP connections only from a controlled set of devices such as external smart mail hosts, printers, networked photocopiers etc. If transparency for SMTP is the only issue, there are a couple of options available to address this:

Option 1 – Use a Layer 7 VIP for SMTP as detailed below and also enable the load balancers on-board firewall to lock down inbound SMTP connections rather than the receive connector. This is covered in section 2 of the Appendix on page 26.

Option 2 – Configure a layer 4 Virtual Service for SMTP rather than a layer 7 (HAProxy) based Virtual Service. Layer 4 is transparent by default so the source IP address is maintained. This is covered in section 5 of the Appendix on page 28.

a) Setting up the Virtual Service

1. Using the WebUI, navigate to: *Cluster Configuration > Layer 7 – Virtual Service* and click **Add a New Virtual Service**

2. Enter the following details:

3. Enter an appropriate label for the VIP, e.g. **HT-SMTP**
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.30.10**
5. Set the *Virtual Service Ports* field to **25**
6. Set *Layer 7 Protocol* to **TCP Mode**
7. Click **Update**

> Note:
>
> Persistence is not required for SMTP.

b) Setting up the Real Servers

1. Using the WebUI, navigate to: *Cluster Configuration > Layer 7 – Real Servers* and click **Add a new Real Server** next to the newly created VIP
2. Enter the following details:

3. Enter an appropriate label for the RIP, e.g. **HT1**
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.30.20**
5. Change the *Real Server Port* field to **25**
6. Click **Update**
7. Repeat the above steps to add your other HT Server(s)

> **Note:**
>
> Because SNAT is a full proxy, any server in the cluster can be on any accessible subnet including across the Internet or WAN.

STEP 3 – FINALIZING THE CONFIGURATION

To apply the new settings, HAProxy must be restarted as follows:

- Go to *Maintenance > Restart Services* and click **Restart HAProxy**

12. Microsoft Exchange Testing Tool

The Exchange Remote Connectivity Analyzer tool, available at https://www.testexchangeconnectivity.com/ is a useful Web-based Microsoft tool designed to help IT Administrators troubleshoot connectivity issues with their Exchange Server deployments. The tool simulates several client logon and mail flow scenarios. When a test fails, many of the errors have troubleshooting tips to assist the IT Administrator in correcting the problem.

13. Technical Support

For more details about configuring the appliance and assistance with designing your deployment please don't hesitate to contact the support team using the following email address: support@loadbalancer.org

14. Further Documentation

The Administration Manual contains much more information about configuring and deploying the appliance. It's available here: http://pdfs.loadbalancer.org/loadbalanceradministrationv8.pdf

15. Conclusion

Loadbalancer.org appliances provide a very cost effective solution for highly available load balanced Exchange 2010 environments.

16. Appendix

1 – CONFIGURING THE LOAD BALANCER USING A SINGLE VIP FOR ALL CAS SERVICES

For testing purposes it's also possible to configure the load balancer with a single VIP for all CAS services. The basic steps to create the VIP & associated RIPs are the same as described previously, the only difference is that all ports would be specified in one VIP as shown below:

> **Note:**
>
> This configuration is not recommended for production deployments.

Label		CAS-AllServices	?
Virtual Service	IP Address	192.168.30.20	?
	Ports	80, 110, 135, 143, 443, 993,	?
Layer 7 Protocol		TCP Mode ▼	?
Manual Configuration		☐	?
		Cancel Update	

Define all CAS ports in a single VIP

> **Note:**
>
> The full list of ports in the 'Virtual Service Ports' field is: 80, 110, 135, 143, 443, 993, 995, 60200, 60201.

2 – LIMITING INBOUND SMTP CONNECTIONS USING FIREWALL RULES

Since layer 7 is not transparent by default, it's not possible to filter inbound SMTP connections by IP address at the receive connector on the Hub Transport Server. One way to address this is to add firewall rules to the load balancer to limit which hosts can connect inbound on port 25.

Rules can be added using the WebUI option: *Maintenance > Firewall Script*. Simply copy/paste/edit the examples below into the firewall script then click **Update**.

EXAMPLES:

1) to limit inbound SMTP connections to a specific smart host:

VIP1="192.168.30.10"
SRC1="192.168.30.50"
iptables -A INPUT -p tcp --src $SRC1 --dst $VIP1 --destination-port 25 -j ACCEPT
iptables -A INPUT -p tcp --dport 25 -j DROP

These rules will only allow SMTP traffic from the host 192.168.30.50 to reach the 192.168.30.10 VIP.

2) to limit inbound SMTP connections to a range of smart hosts:

VIP1="192.168.30.10"

SRC1="192.168.30.50-60"

iptables -A INPUT -p tcp -m iprange --src-range $SRC1 --destination $VIP1 --destination-port 25 -j ACCEPT

iptables -A INPUT -p tcp --dport 25 -j DROP

These rules will only allow SMTP traffic from hosts in the range 192.168.30.50 through 192.168.30.60 to reach the 192.168.30.10 VIP.

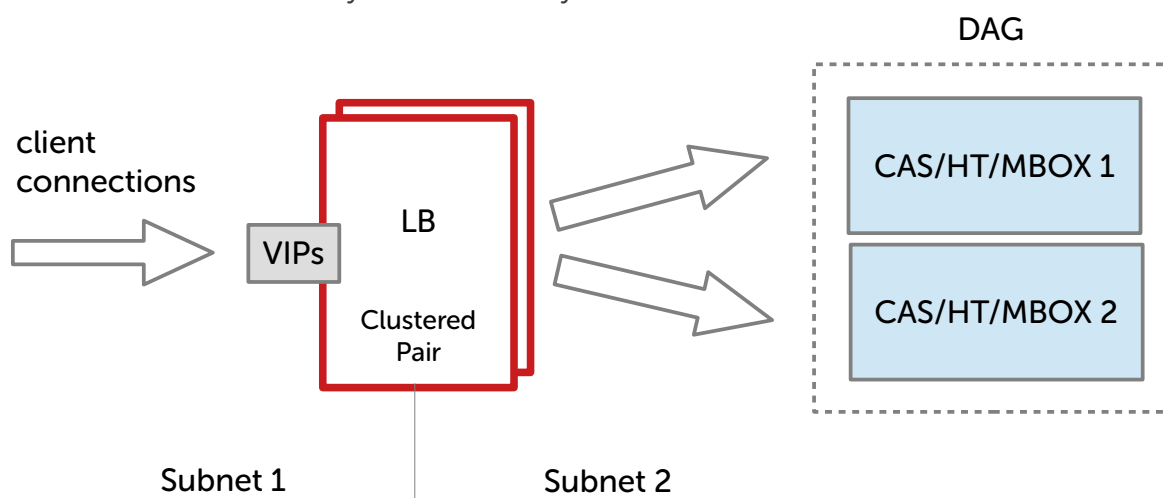## 3 – USING HTTP COOKIE PERSISTENCE FOR OWA USERS

If IP persistence cannot be used for OWA (e.g. if clients pass through a NAT device), HTTP cookie persistence can be used as an alternative. To use cookie persistence, the SSL stream must be decrypted on the load balancer to enable the cookies to be inserted/read. For more details on setting this up, refer to configuration/deployment example 3 in the Administration Manual.

> Note:
>
> SSL termination on the load balancer can be very CPU intensive. Therefore this is only advised when IP persistence is not possible. In most cases, for a scalable solution terminating SSL on the real servers is the best option.

4 – ENABLING FULL TRANSPARENCY USING TPROXY

If a fully transparent configuration is required, TProxy can be used. The main point to note is that two subnets must be used for TProxy to work correctly.

Key points to consider:

- The Exchange servers must be on a different subnet to the VIP – this can achieved by using 2 IP addresses assigned to a single interface, or two separate interfaces (eth0 & eth1)

- The default gateway on the Exchange servers **must** be configured to be an IP address on the load balancer. For a clustered pair of load balancers, It's best to add an additional floating IP for this to allow failover to the slave

- TProxy must be enabled. Using the WebUI go to *Cluster Configuration > Layer 7 – Advanced Configuration* , check **Transparent Proxy** and click **Update**

5 – USING A LAYER 4 VIRTUAL SERVICE FOR THE HT ROLE

This guide uses Layer 7 HAProxy based Virtual Services for all load balanced services. Layer 7 Virtual Services are not transparent by default which can be an issue for the HT role. One option In these cases is to use a Layer 4 VIP - there are two possibilities: **DR** (Direct Return) mode and **NAT** (Network Address Translation) mode.

Layer 4 – DR Mode

DR mode requires that the '*ARP Problem*' is solved on each Real Server.

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) 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 the various versions of Windows, please refer to the Administration Manual and search for "*DR Mode Considerations*".

Layer 4 – NAT Mode

NAT mode works by changing the destination address of inbound packets from the VIP address to be one of the load balanced Real Server IP addresses, and by changing Real Server response packets from the real address back to the VIP address. NAT mode requires the VIP and associated RIPs to be in different subnets and also the default gateway on each Real Server must be set to be the load balancer.

For examples of setting up layer 4 Virtual Services, please refer to the configuration/deployment examples in the Administration Manual.

6 – 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
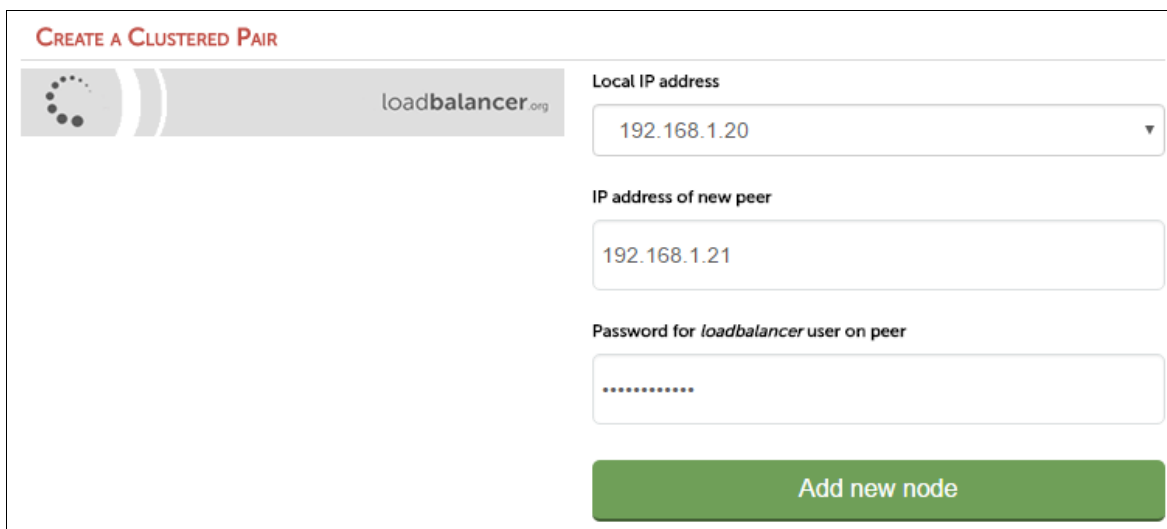> - Software updates

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:

- Deploy a second appliance that will be the slave and configure initial network settings
- 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

- Click **Add new node**

- The pairing process now commences as shown below:

- Once complete, the following will be displayed:

- 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.

7 - COMPANY CONTACT INFORMATION

Website	URL: www.loadbalancer.org
North America (US)	Loadbalancer.org, Inc. 4250 Lancaster Pike, Suite 120 Wilmington DE 19805 USA Tel: +1 888.867.9504 Fax: +1 302.213.0122 Email (sales): sales@loadbalancer.org Email (support): support@loadbalancer.org
North America (Canada)	Loadbalancer.org Ltd 300-422 Richards Street Vancouver, BC V6B 2Z4 Canada Tel: +1 866.998.0508 Fax: +1 302.213.0122 Email (sales): sales@loadbalancer.org Email (support): support@loadbalancer.org
Europe (UK)	Loadbalancer.org Ltd. Compass House North Harbour Business Park Portsmouth, PO6 4PS UK Tel: +44 (0)330 3801064 Fax: +44 (0)870 4327672 Email (sales): sales@loadbalancer.org Email (support): support@loadbalancer.org
Europe (Germany)	Loadbalancer.org GmbH Tengstraße 27 D-80798 München Germany Tel: +49 (0)89 2000 2179 Fax: +49 (0)30 920 383 6495 Email (sales): vertrieb@loadbalancer.org Email (support): support@loadbalancer.org