LOADBALANCER

DEPLOYMENT GUIDE

Load Balancing Remote Desktop Services

Version 2.3.0

Table of Contents

1. About this Guide

This guide details the steps required to configure a load balanced Microsoft Remote Desktop Services (RDS) environment utilizing Loadbalancer.org appliances. It covers the configuration of the load balancers and also any Microsoft Remote Desktop Services configuration changes that are required to enable load balancing. The guide focuses on Windows 2012 and later, although reference is made to 2008 R2 where appropriate.

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

2. Loadbalancer.org Appliances Supported

All our products can be used used with Remote Desktop Services. For full specifications of available models please refer to https://www.loadbalancer.org/products/enterprise.

Some features may not be available or fully supported in all cloud platforms due to platform specific limitations. For more details, please refer to the "Main Differences to our Standard (Non-Cloud) Product" section in the appropriate cloud platform Quick Start Guide or check with Loadbalancer.org support.

3. Software Versions Supported

3.1. Loadbalancer.org Appliance

- V8.9.1 and later

> **Note** The screenshots used throughout this document aim to track the latest Loadbalancer.org software version. If you're using an older version, or the very latest, the screenshots presented here may not match your WebUI exactly.

3.2. Microsoft Windows

- Windows 2008 R2 and later

4. Remote Desktop Services (RDS)

4.1. Introduction

Remote Desktop Services can be used to provide:

- Access to full remote desktops- this can be either session-based or VM-based and can be provided locally from PC's, laptops & thin clients or from virtually anywhere using mobile devices
- Access to applications - RemoteApp can be used to provide users with access to applications running on RD Session Host servers. These applications look and feel just like locally installed programs
- Secure remote access - Remote Desktop Gateway (RD Gateway) can be used to provide secure remote access to desktops and applications without the need for a VPN

4.2. Role Services

The following role services can be deployed as part of the RDS role.

Role Service	Purpose
RD Virtualization Host	This role service integrates with the Hyper-V role in Windows Server 2012 R2 to provide VMs that can be used as virtual desktops. The RD Virtualization Host role service also monitors and reports on established client sessions to the RD Connection Broker role service. This role service is responsible for managing the VMs that function as pooled and personal virtual desktops. If VMs are in a saved state, the RD Virtualization Host role service starts the VMs to prepare them for a user connection. For pooled virtual desktops, the RD Virtualization Host role service reverts the VMs to their initial state when users sign out. RD Virtualization Host role service is required in a VM-based deployment of RDS.
RD Session Host	This role service configures a server to provide session-based desktops and applications. Users can connect to an RD Session Host server and then run applications and use the network resources that the RD Session Host offers. RD Session Host is a required role service in a session-based desktop deployment of RDS.
RD Connection Broker	This role service manages connections to RemoteApp programs and virtual desktops, and it directs client connection requests to an appropriate endpoint. The RD Connection Broker role service also provides session re-connection and session load balancing. For example, when a user disconnects from a session and later establishes a connection, the RD Connection Broker role service ensures that the user reconnects to his or her existing session. RD Connection Broker is mandatory in all RDS deployments.
RD Web Access	This role service provides a web-based interface to RemoteApp programs, session-based virtual desktops, or VM-based virtual desktops. A webpage provides each user with a customized view of all RDS resources that have been published to that user. This role service supports organizing resources in folders, which enables administrators to group remote applications in a logical manner. It also publishes available RDS resources in an RDWeb feed, which can integrate with the Start screen on client devices. RD Web Access is a mandatory role service for each RDS deployment.

Role Service	Purpose
RD Licensing	This role service manages RDS client access licenses (RDS CALs) that are required for each device or user to connect to an RD Session Host server. You use RD Licensing to install, issue, and track RDS CAL availability on an RD Licensing server. You are not required to install this role service during an initial RDS deployment, but an RDS deployment without proper licensing ceases to function after 120 days.
RD Gateway	This role service allows authorized remote users to connect securely to RemoteApp programs and virtual desktops from outside the organization over the Internet. An RD Gateway server acts as a proxy for external users to connect to internal RDS resources. To increase compatibility with firewalls in public locations such as hotels, RDP traffic is encapsulated in Hypertext Transfer Protocol Secure (HTTPS) packets. Access is controlled by configuring Remote Desktop connection authorization policies (RD CAPs) and Remote Desktop resource authorization policies (RD RAPs). An RD CAP specifies who is authorized to make a connection, and an RD RAP specifies to which resources authorized users may connect. RD Gateway is an optional role service.

For much more information about RDS please refer to this URL.

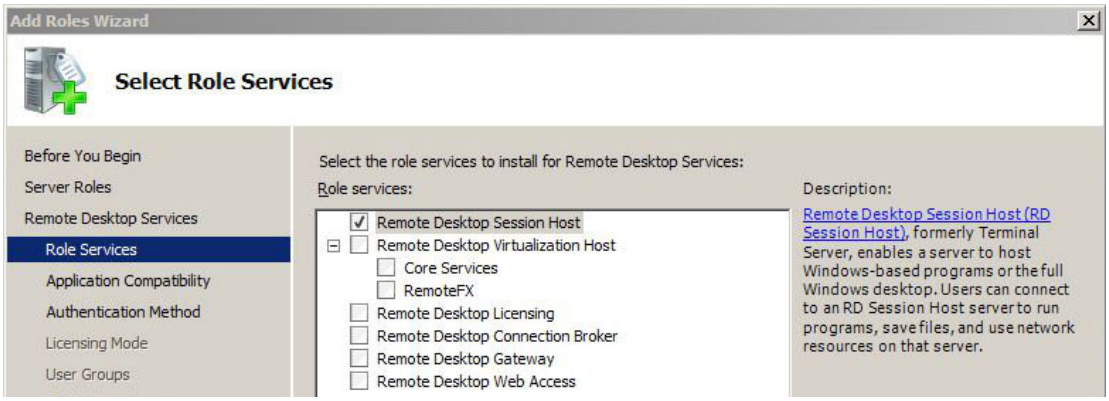
> **Note** It is possible to deploy just RD Session Host Servers & a Loadbalancer.org appliance without the complete RDS infrastructure. If you only require the ability to provide multiple full desktops then this approach may be appropriate. For more information, please refer to Scenario 4 - Load Balancing Stand alone Session Hosts.

4.2.1. Role Service – Server Location / Collocation

Depending on the number of users and the server specifications, role services can be collocated, although Microsoft recommends that whenever possible the Session Host and Connection Broker role services should be kept on dedicated servers. Typically, RD Gateway and RD Web Access are candidates for collocation.

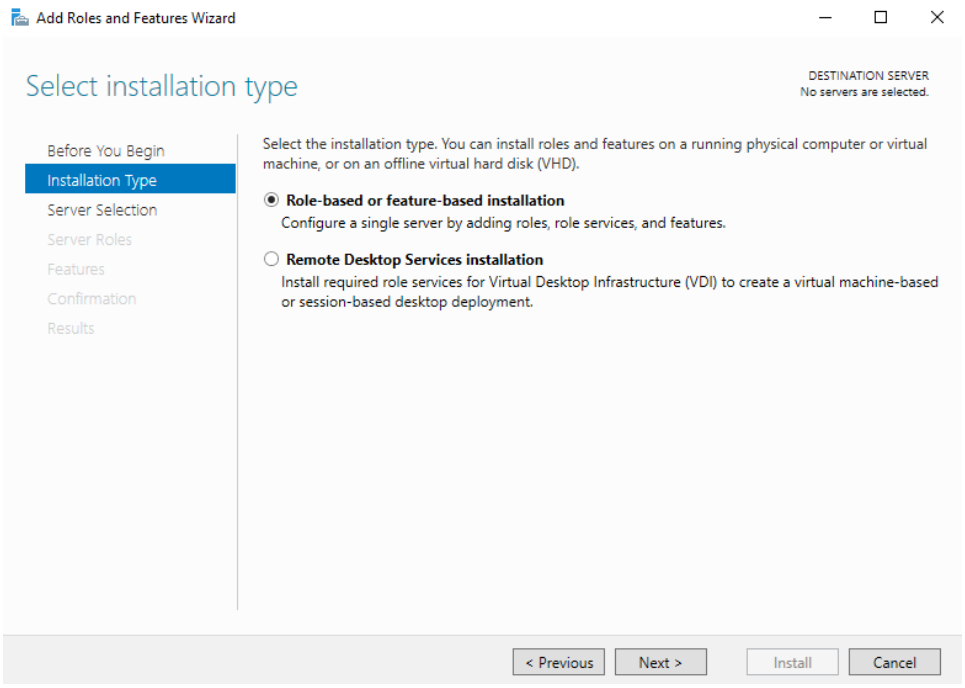
4.3. RDS Installation – Windows 2008 R2

Installation of RDS under Windows 2008 R2 uses the traditional role/service concept. The RDS infrastructure must be built by manually installing the required services on the various servers to build the desired infrastructure. The screenshot below shows the initial service selection screen for installing RDS under Windows 2008 R2.



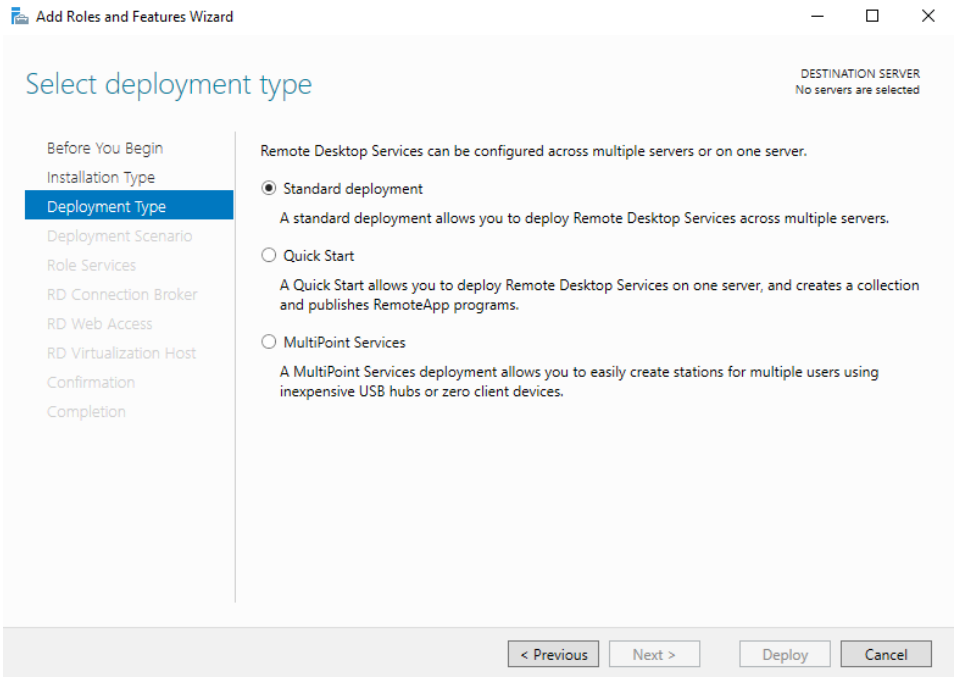
4.4. RDS Installation – Windows 2012 & 2016

Windows 2012 & 2016 provides two installation types as shown in the screenshot below:



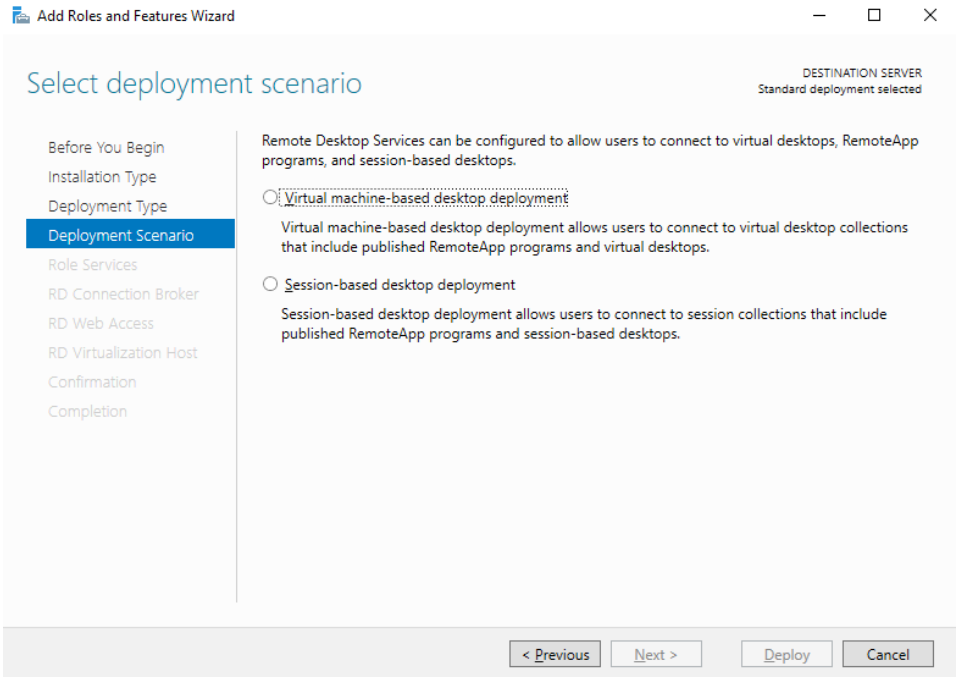
- **Role-based or feature-based** - Roles and services are installed on individual servers using standard role installation methods as per Windows 2008 R2
- **Remote Desktop Services Installation** - Centrally based RDS specific installation which enables all role services to be installed on multiple servers from a single management interface

When the option **Remote Desktop Services Installation** is selected, there are 3 deployment types to choose from as shown in the screenshot below:



- **Standard deployment** - Enables RDS to be deployed across multiple servers
- **Quick Start** - All services are deployed to a single server
- **Multipoint Services** (2016 & later only) - Designed for classroom-type deployments where more desktop control & monitoring functionality is required

When the option **Standard Deployment** is selected, there are 2 deployment scenario's to choose from as shown in the screenshot below:



- **Virtual machine-based desktop deployment** - Provides users with access to a full Windows client operating system that runs on a VM, for example, Windows 7 or Windows 10

- **Session-based desktop deployment** - A session based virtual desktop deployment the same as the traditional "Terminal Server" concept where multiple client sessions run on the same server

4.4.1. Choosing Between VM-Based & Session-Based Desktop Deployments

RDS has 2 deployment scenario's as mentioned above. You must decide which RDS deployment type is best for your environment based on various requirements. Consider whether the applications run correctly on windows Server and whether it works properly in a multi-user environment. Also, consider that a VM-based virtual desktop deployment typically requires a more powerful server infrastructure and more disk storage than a session-based virtual desktop deployment for the same number of users. Generally, Microsoft recommend session-based virtual desktops if possible. Session-based virtual desktops support a larger number of users than VM-based virtual desktops on the same hardware.

4.4.2. The Standard Deployment – Recommended by Microsoft

This kind of deployment is created using the **Remote Desktop Services Installation** option, selecting **Standard Deployment** and then selecting **Session-based Desktop Deployment**.

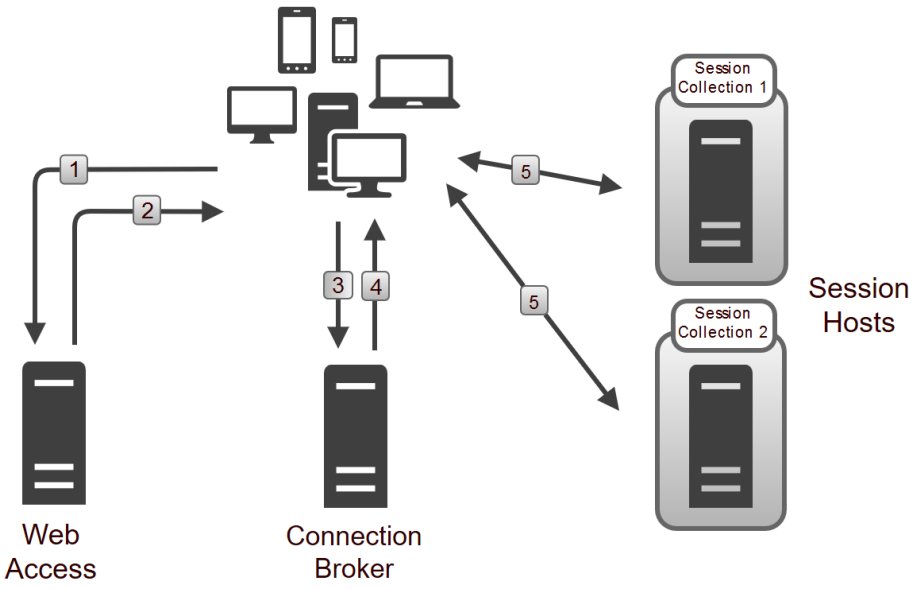
Using the Standard Deployment is considered best practice by Microsoft. When selected, it will start a deployment wizard that enables the following role services to be installed from a single management interface:

- 1 x RD Web Access
- 1 x RD Connection Broker
- 1 or more RD Session Hosts

RD Gateways, RD Licensing servers, additional Connection Brokers, additional Web Access servers and more Session Hosts can be added after initial deployment. As mentioned earlier, role services can be collocated, although Microsoft recommends that Session Hosts run as dedicated servers.

The Standard Deployment – How it Works

The diagram below shows the various role services, and how users interact with them when accessing the deployment:



The following process is used when clients connect to a session collection by using RD Web Access:

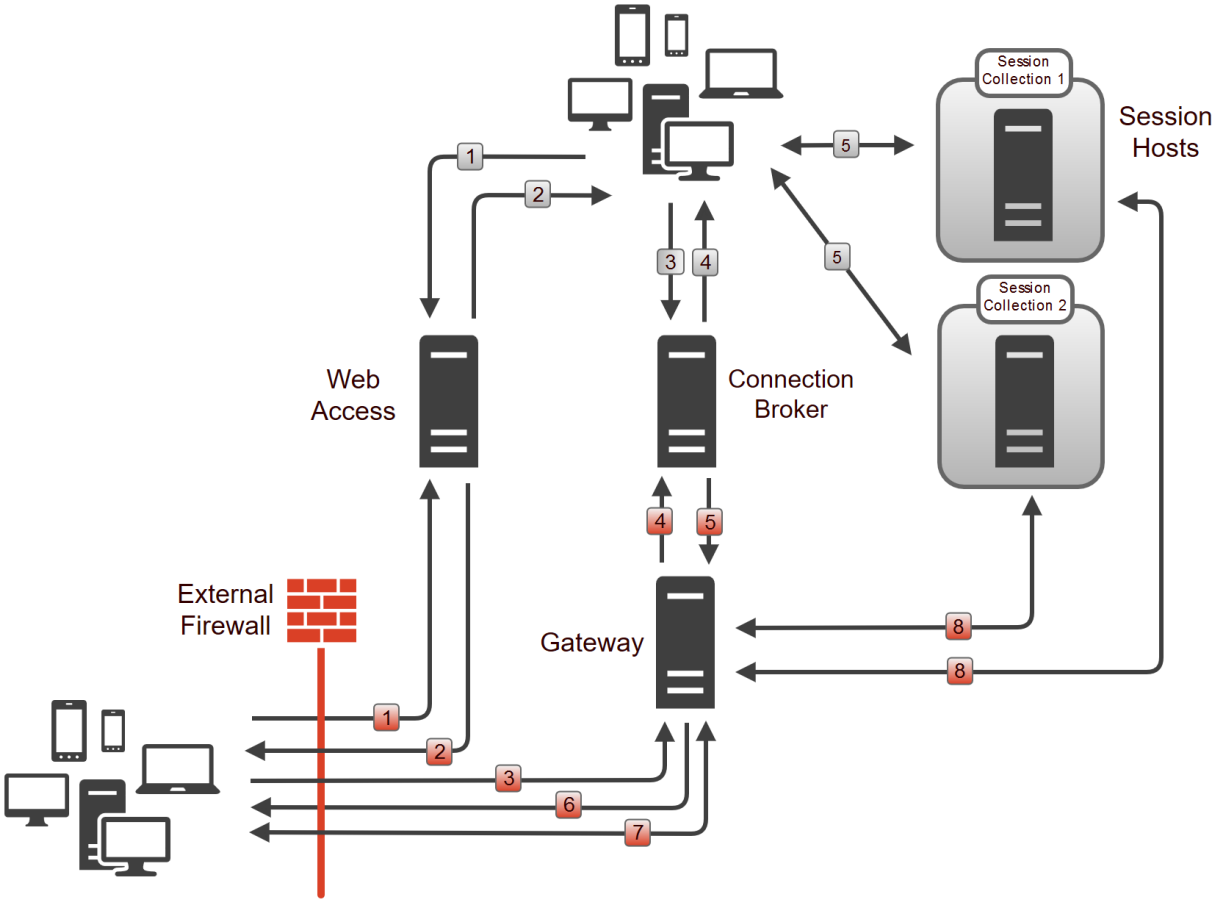
1. Users connect to the RD Web Access portal and identify the RDS resource to which they want to connect.
2. Users click the link on the RD Web Access portal for the RDS resource they want to access. This downloads the .RDP file, which contains information about the resource to which the user wants to connect.
3. RDC is launched, and it uses the information in the .RDP file to initiate a connection with the RD Connection Broker role service. After users authenticate to the RD Connection Broker role service, the RDC passes the request about the RDS resource to which the user wants to connect.

> **Note** Since Windows 2012, the RD Connection Broker is the ***default Initial connection point***.

4. The RD Connection Broker role service examines the request to find an available RD Session Host server in the desired collection and sends the connection information back to the RDC client. If the request matches a session that already is established for the associated user, RD Connection Broker redirects the client to the server in the collection where the session was established. If the user doesn't have an existing session in the collection, the client redirects to the server that is most appropriate for the user connection, based on the built-in RD Connection Broker load balancing algorithm.
5. The RDC client establishes a session with the RD Session Host server that RD Connection Broker provided.

Adding RD Gateway - Proving Secure Access from the Internet

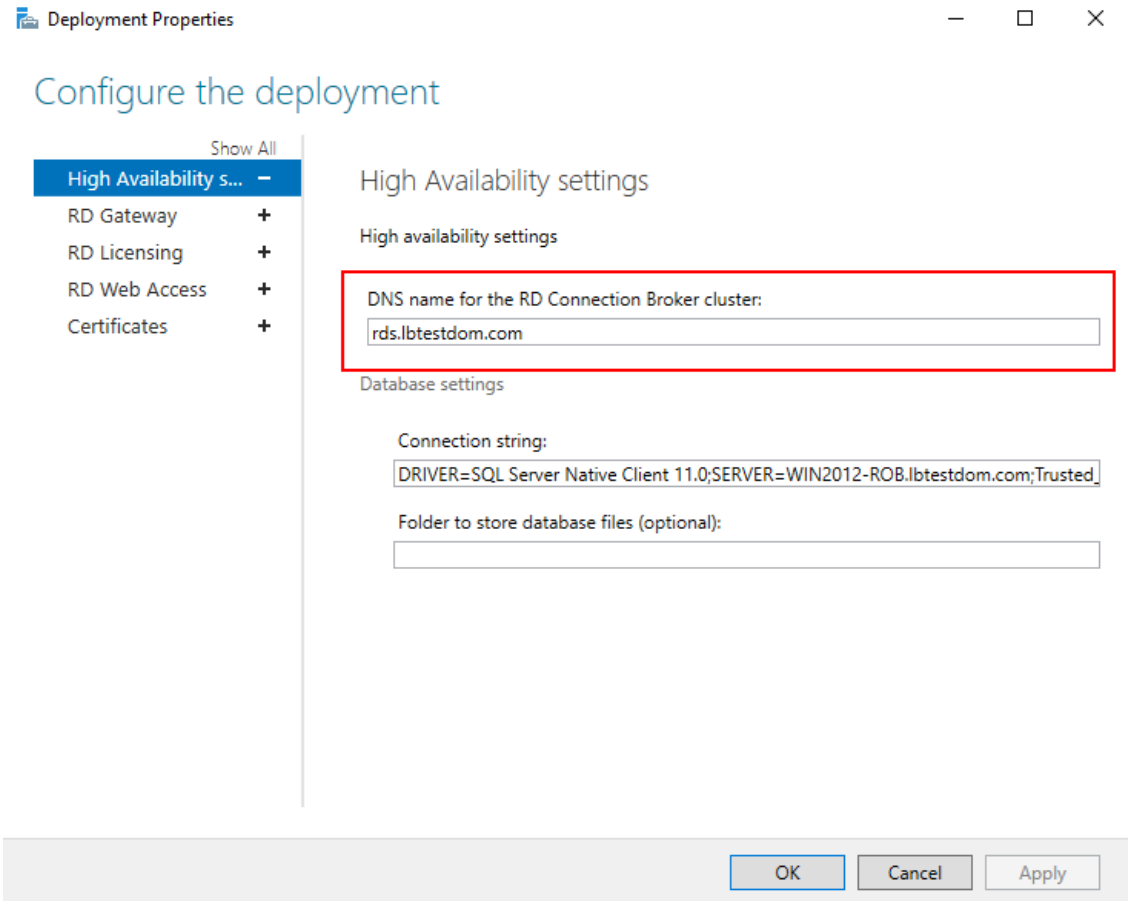
RD Gateway is used to provide secure access to the RDS deployment from the Internet.



The additional red numbers show the process when external Internet based users connect to the deployment. In this case, the RD Gateway acts as a proxy when accessing the Connection Brokers and the Session Hosts.

4.5. RDS Configuration – Deployment Properties

4.5.1. High Availability Settings



The FQDN specified in ***DNS name for the RD Connection Broker cluster*** is set during initial deployment and is the FQDN that clients use to connect to the deployment. This FQDN is written to the .RDP files created by Web Access. Once configured, it's not possible to change this via the Windows UI, Powershell must be used instead. For Windows 2016 this is documented here.

> **Note** When the Loadbalancer.org appliance is deployed, DNS must configured so that this FQDN points at the Virtual Service (VIP) on the load balancer as explained for the various scenarios in the following sections: Testing and Verification (scenario 2) Testing and Verification (scenario 3) and Testing and Verification (scenario 5).

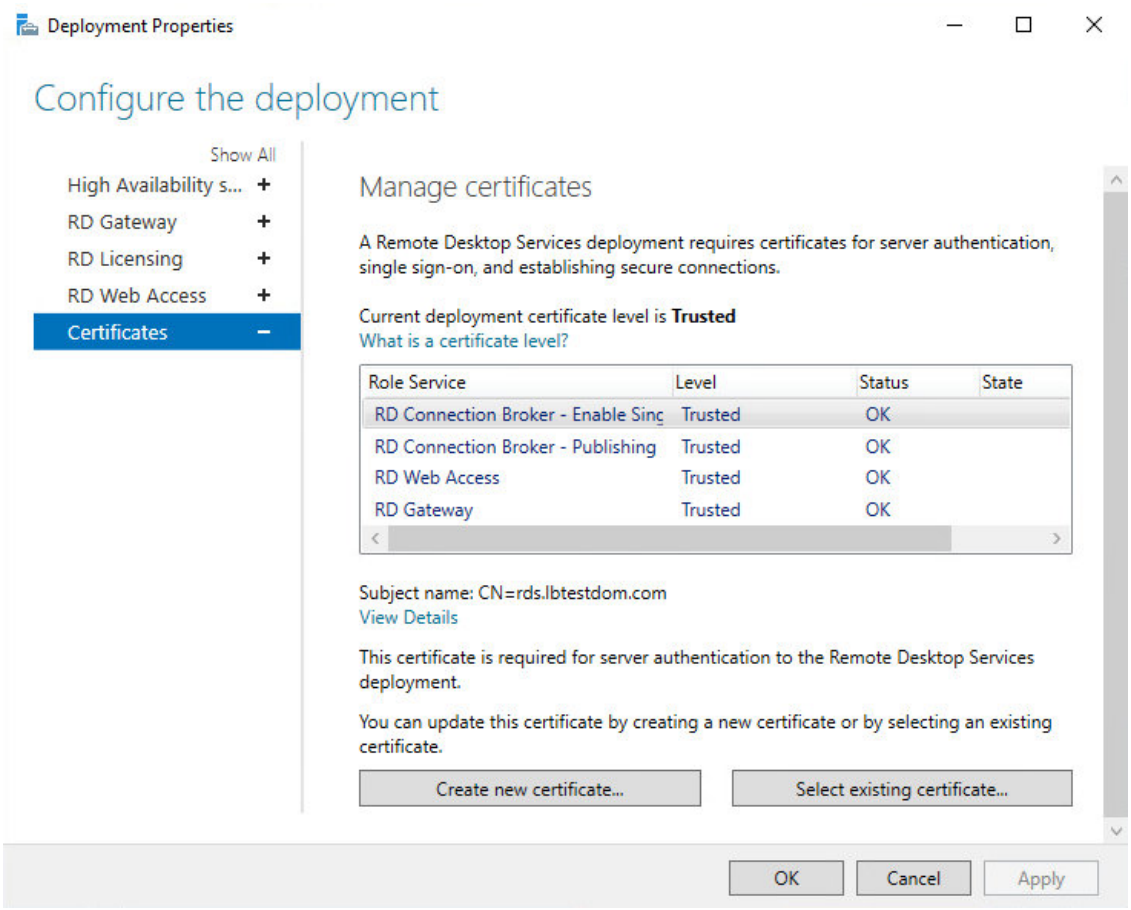
4.5.2. Certificates

From Windows 2012, RDS certificates are managed from the Certificates tab of Deployment Properties as shown below. Detailed information about RDS certificate requirements is available here.

Certificate used for this Guide

In the test environment used for this guide, a single certificate signed by an internal CA was used. The certificate was configured as follows:

Name = rds.lbtestdom.com
SAN1 = rdgateway.lbtestdom.com
SAN2 = *.lbtestdom.com (this SAN covers all individual Session Hosts in the deployment)



> **Note** If you're sending the Initial connection to load balanced Session Hosts rather than the default which is to send the Initial connections to load balanced Connection Brokers, you may receive certificate warnings due to the self-signed RD certificate on each Session Hosts. Please refer to RDS Certificate Handling for more information on how to deal with this.

5. Load Balancing RDS – Concepts

For HA, there should be at least 2 of each role service. These role services can then be load balanced.

> **Note** Whenever possible, it's highly recommended that you have a working RDS environment first before implementing the load balancer.

5.1. What About the built-in Load Balancing mechanism?

The built-in RDS load balancing mechanism is used to distribute client sessions to the Session Hosts when the Initial connection is handled by the Connection Brokers (the default Microsoft recommended method). It is not used by any other role service.

5.2. Which Role Services Should I Load Balance?

All role services that are deployed should be load balanced to provide HA:

- Connection Brokers
- Session Hosts

> **Note**
>
> It is possible to load balance Session Hosts using an external load balancer rather than the built-in RDS load balancing mechanism. In this case, the Initial connection is handled by the load balanced Session Hosts rather than the load balanced Connection Brokers.
>
> All load balancing scenarios are explained in Remote Desktop Services – Load Balancing Scenarios.

> **Note**
>
> It's not possible to load balance Connection Brokers and Session Hosts with an external load balancer at same time. If Connection Brokers are load balanced, clients are send directly or via the load balancer to a specific Session Host – specifying a load balanced FQDN for the Session Hosts is not possible. Likewise, if Session Hosts are load balanced, the Session Hosts refer directly to the Connection Brokers and specifying a load balanced FQDN for the Connection Brokers is not possible.

- Web Access Servers
- Gateways

5.3. Load Balanced Ports & Services

The following table shows the RDS ports and services that are load balanced:

Protocol	Port	Purpose / Role Service
TCP/HTTPS	443	HTTPS (RD Gateway, RD Web Access)
TCP/UDP/RDP	3389	RDP (UDP transport was added in RDP v8.0)
UDP	3391	RDP (RD Gateway)

5.4. Persistence (Server Affinity) Requirements & Options

Persistence means consistently sending a particular client to the same back-end server during a particular session. This must be enabled for some role services. The following table summarizes the requirements:

Service	LB.org Appliance Persistence Required?	Comments	LB.org Appliance Persistence Method(s)
Virtualization Hosts	N/A	Virtualization Hosts are not load balanced using the LB.org appliance. Connection Broker & the built-in load balancing mechanism is used to re-establish client / desktop sessions.	N/A
Session Hosts	Yes	When the Initial connection is handled by load balanced Connection Brokers (the default), Session Host load balancing is handled by the built-in load balancing mechanism. When the Initial connection is handled by load balanced Session Hosts, and you're happy for redirected sessions to go direct to the Session Hosts (the default). When the Initial connection is handled by load balanced Session Hosts, and you want to ensure all sessions (both new and redirected) pass via the load balancer. For a minimal deployment without Connection Broker with just the load balancer & 2 or more Session Hosts.	N/A **Source IP** - this is required to ensure that both TCP & UDP traffic for the RDP session is handled by the same Session Host for new connections. **MS Session Broker** - for this to work, all Session Hosts must be configured in ***Routing Token Redirection Mode***. In this mode, UDP transport for RDS is not supported because a Layer 7 VIP is required, which does not support UDP. Source IP or RDP Client Cookie
Connection Brokers	No	Persistence is not required since the load balancer only handles the Initial connection and not the active RDP session.	N/A

Service	LB.org Appliance Persistence Required?	Comments	LB.org Appliance Persistence Method(s)
Gateways	Yes	TCP connections for a session must go to the same Gateway and UDP connections the session must go to the same Gateway, but TCP and UDP can be handled by different Gateways.	Source IP
Web Access Servers	Yes	Uses IIS with authentication which is to a specific server.	Source IP or HTTP Cookie

5.4.1. MS Session Broker Persistence

This mode can be used only when the Initial connection is handled by load balanced Session Hosts. In this mode, the load balancer interacts with Connection Broker by enabling ***Routing Token Redirection Mode*** on the Session Hosts. This mode allows the reconnection of disconnected sessions by utilizing a ***Routing Token*** to enable the load balancer to re–connect the client to the correct Session Host. ***Routing Token redirection Mode*** works as follows:

1. The client connects to the VIP on the load balancer and is load balanced to one of the Session Hosts.
2. The Session Host authenticates the user and checks with one of the Connection Brokers if the user has an existing, disconnected session.
3. If there is an existing session, the IP address for the Session Host where the session is running is encoded in a ***Routing Token*** and returned to the client via the load balancer.
4. The client then reconnects to the load balancer presenting this ***Routing Token***, the load balancer then connects the client to the Session Host specified in the ***Routing Token***.
5. If there was no existing session, a new session is started on the Session Host where the user was originally load balanced.

For more information about redirection modes, please refer to this URL. For more information about Routing Tokens, please refer to page 25-27 of this Microsoft document.

> **Note**
>
> If this persistence method is used, all connections will pass via the load balancer, including those that have been redirected.
>
> Since this persistence method requires a layer 7 VIP, UDP is not supported, which means that Session Host connectivity for internal ***and*** external clients (via the Gateway) will only utilize TCP.

5.4.2. Source IP Persistence

This method is appropriate when each clients actual source IP addresses can be seen by the load balancer. This

will typically be the case within a LAN but in some situations – e.g. a remote office connecting via some kind of NAT device, all clients would appear to come from the same address and therefore load may not be evenly distributed between the RDS servers.

5.4.3. RDP Client Cookie Persistence

This method can be used with a simple deployment which does not have Connection Broker, just Session Hosts and the load balancer appliance. It utilizes the cookie sent from the client in the Connection Request PDU. This cookie is created when the username is entered at the first client login prompt (mstsc.exe). If the username is not entered here, the cookie is not created.

The cookie only supports up to 9 characters, so this method may have limited use, especially in cases where users login using the domain\username format. In this case, if the domain name was 9 characters in length, the RDP cookie would be the same for all users, resulting in all sessions being sent to the same Session Host. If users login using the UPN format (User Principle Name), i.e. **username@domain**, it's more likely to be unique.

> **Note** When RDP cookie persistence is selected, the load balancer will attempt to use RDP cookie persistence, but if a cookie is not found, source IP persistence will be used instead as a fallback.

5.5. Load Balancer Deployment Mode

The load balancer can be deployed in 4 fundamental ways: ***Layer 4 DR mode***, ***Layer 4 NAT mode***, ***Layer 4 SNAT mode*** and ***Layer 7 SNAT mode***. These modes are explained in detail in Load Balancer Deployment Modes.

5.5.1. Web Access Servers

Layer 7 SNAT mode is recommended for Web Access servers. If Web Access servers are collocated with Gateways, Layer 7 SNAT mode must also be used for the Gateway's TCP part. This is because it's not possible to configure a layer 7 SNAT mode VIP and a layer 4 SNAT mode VIP with ***the same real servers listening on the same port***.

5.5.2. Connection Brokers

Layer 7 SNAT mode is recommended for Connection Brokers although any other mode can be used if preferred.

5.5.3. Gateways

Layer 4 SNAT mode is recommended for the UDP part of RD Gateway and Layer 7 SNAT mode is recommended for TCP. Layer 4 SNAT mode can also be used for the TCP part, but when RD Gateway and Web Access are collocated, you'd also need to use layer 4 SNAT for Web Access for the reason mentioned in the Web Access servers section above.

5.5.4. Session Hosts

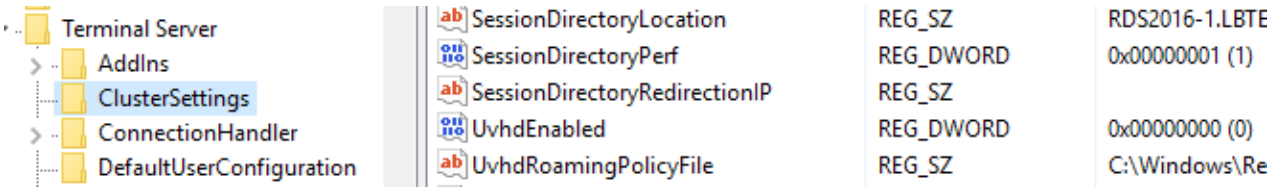
If you're load balancing Session Hosts and you require all sessions (both new and redirected) to pass via the load balancer, you must use Layer 7 SNAT mode with MS Session Broker persistence and you must enable ***Routing Token Redirection Mode*** on each Session Host. The downside here is that RDP over UDP will not work for internal clients and the external clients who pass via the Gateway. This is because layer 7 SNAT mode does not support UDP.

If you require UDP support for internal and external clients, one of the layer 4 methods must be used. Layer 4 SNAT mode is recommended since no real server changes are required.

If Layer 4 methods are used, it will not be possible to use ***Routing Token Redirection Mode***. The default method (***IP Address Redirection Mode***) must be used.

If you use NAT mode, the default gateway of the Session Hosts must be the load balancer.

If you use DR mode you'll need to solve the 'ARP problem' as explained in DR Mode Considerations. You'll also need to configure the following registry entry on each Session Host to ensure that the main interface IP address and not the loopback adapter address is passed back to the client for re-connection:




```
HKLM\SYSTEM\CurrentControlSet\Control\Terminal Server\ClusterSettings
```

SessionDirectoryRedirectionIP - set to the IP address to send to the client, this should be the main interface IP address of the Session Host

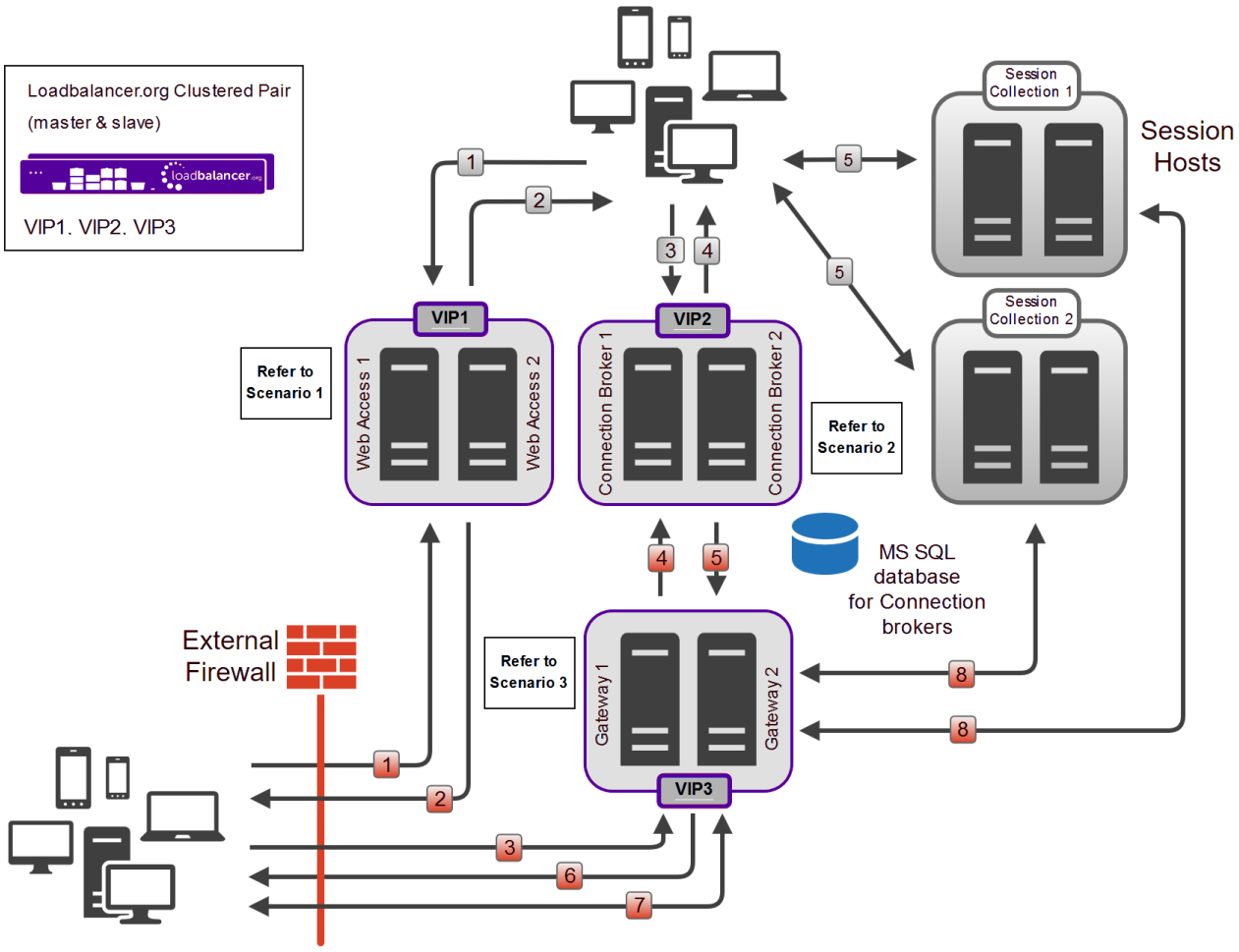
5.6. Deploying the Load Balancer – VIP Location

The following VIPs are normally configured on the load balancer when load balancing Remote Desktop Services:

- **VIP1** – the connection point for the load balanced Web Access Servers
- **VIP2** – the connection point for the load balanced Connection Brokers, DNS must be configured so that the FQDN specified in ***DNS Name for the RD Connection Broker Cluster*** (**Deployment Properties > High Availability**) resolves to this VIP
- **VIP3** - the connection point for the load balanced RD Gateway Servers

When using the default Microsoft-recommended deployment, Session Hosts are load balanced by the built in mechanism as described in The Standard Deployment – Recommended by Microsoft, so there is no VIP for the Session Hosts.

The following diagram illustrates where the load balancer is deployed when the Standard Microsoft deployment is used:



Notes

- The Initial connection is from RDP client to **Connection Broker** as recommended by Microsoft
- The Loadbalancer.org server feedback agent ***cannot be used*** in this case because the Session Hosts are load balanced by the built-in load balancing mechanism and not by the Loadbalancer.org appliance.

  If you want to use the Loadbalancer.org feedback agent, you'll need to send the Initial connection to the load balanced Session Hosts rather than the load balanced Connection Brokers as described in Scenario 5 - Load Balancing Session Hosts when Deployed with Connection Broker.

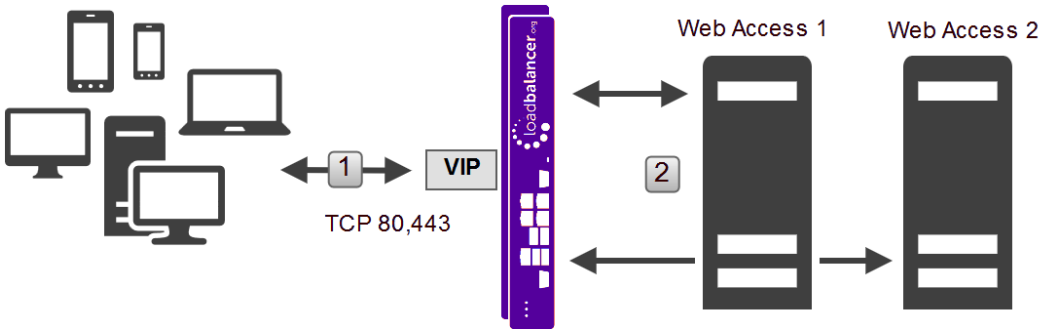
- A **Session Collection** is simply a way to group Session Hosts for load balancing, RemoteApp publishing, and common settings purposes. For example, if you set the Idle session limit to 3 hours in the properties of the collection, then all Session Hosts that are part of the collection will have a 3 hour idle timeout.

> **Note** If you want to use the appliance to load balance Session Hosts please refer to Scenario 5 - Load Balancing Session Hosts when Deployed with Connection Broker.

6. Remote Desktop Services – Load Balancing Scenarios

6.1. Scenario 1 - Load Balancing Web Access Servers

Scenario 1 is part of the Standard Deployment as illustrated in the Standard Deployment Diagram.



6.1.1. Client Connection Process

1. Client initiates session request to the VIP on the load balancer.
2. The load balancer forwards the request to one of the load balanced Web Access servers.
3. The client continues the session to the selected Web Access server via the load balancer (assuming a layer 7 SNAT configuration as used in this guide).

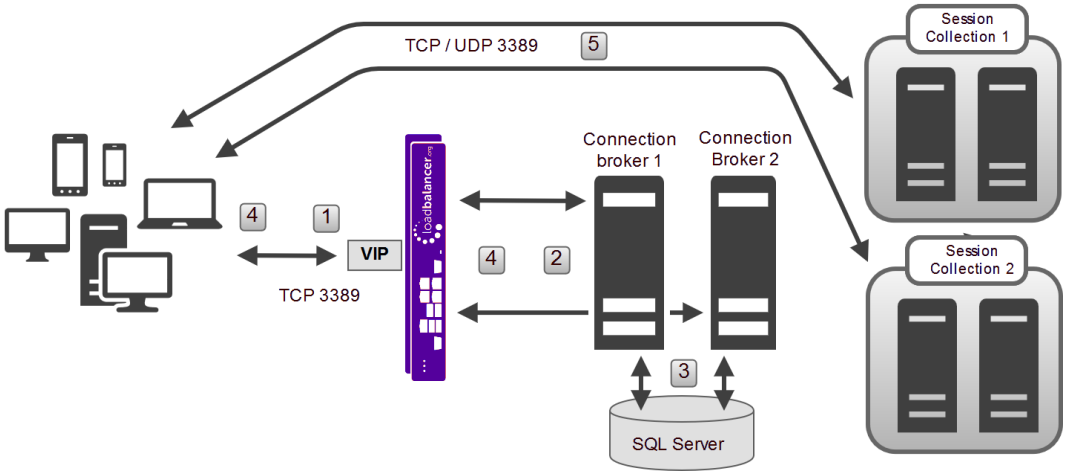
6.1.2. Scenario Notes

- Web Access servers use IIS so it's effectively the same as load balancing Microsoft Web Servers.
- Session persistence from client to Web Access server is based on client source IP address.
- The Web Access servers have a built in HTTP -→ HTTPS redirect, so the VIP also listens on port 80 to enable this to function correctly.
- Layer 7 SNAT mode is recommended and is used for the example in this guide. It's also possible to use Layer 4 DR, NAT or SNAT modes depending on your infrastructure and requirements (see Layer 4 DR Mode , Layer 4 NAT Mode and Layer 4 SNAT Mode for descriptions of these modes).
- Clients connect using a Web Browser.

> **Note** See Load Balancing Web Access Servers (Scenario 1) for load-balancer configuration steps and RDS configuration notes related to this scenario.

6.2. Scenario 2a - Load Balancing Connection Brokers with Session Hosts

Scenario 2 is part of the Standard Deployment as illustrated in the Standard Deployment Diagram.



6.2.1. Client Connection Process

1. Client initiates session request to the VIP on the load balancer.
2. The load balancer forwards the request to one of the load balanced Connection Brokers.
3. The Connection Broker checks the SQL database to determine if the user has an existing session, if yes the IP address for that server is selected, if no then the RDS built in load balancing mechanism selects a host/IP address where to start a new session.
4. The Connection Broker returns this IP address back to the client via the load balancer (assuming a Layer 7 configuration as used in this guide).
5. The client connects ***directly*** to the Session Host specified.

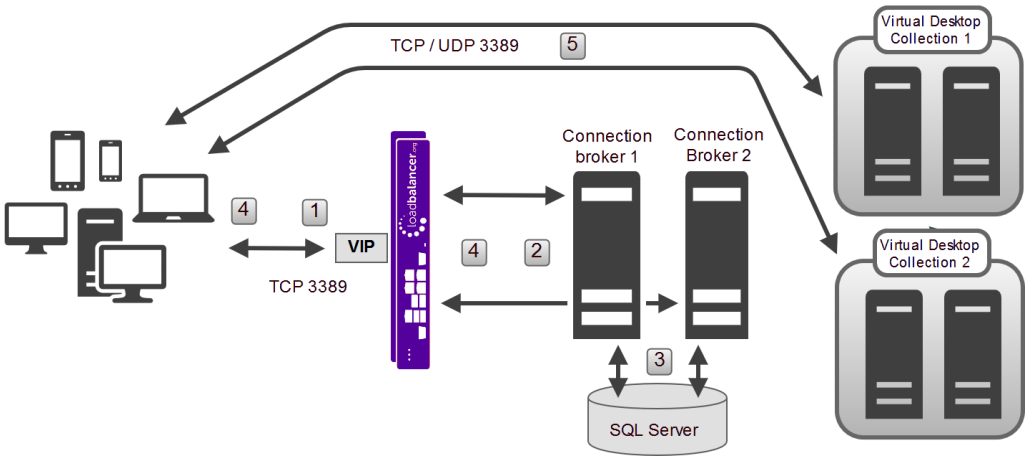
6.2.2. Scenario Notes

- In this scenario the Initial connection is to the Connection Brokers (via the load balancer).
- Session persistence from client to Connection Broker is not required because it handles the initial request and not active sessions.
- Layer 7 SNAT mode is recommended and is used for the example in this guide. It's also possible to use Layer 4 DR, NAT or SNAT modes depending on your infrastructure and requirements (see Layer 4 DR Mode , Layer 4 NAT Mode and Layer 4 SNAT Mode for descriptions of these modes).
- DNS must be configured so that the FQDN specified in ***DNS Name for the RD Connection Broker Cluster*** (**Deployment Properties > High Availability**) resolves to the Connection Broker VIP.
- Clients connect using RemoteAPP via RD Web Access or modified .RDP files and not just by specifying the DNS name or IP address of the Connection Brokers in mstsc.exe as explained here.

> **Note** See Load Balancing Connection Brokers (Scenarios 2a & 2b) for load balancer configuration steps and RDS configuration notes related to this scenario.

6.3. Scenario 2b - Load Balancing Connection Brokers with Virtualization Hosts

Scenario 2 is part of the Standard Deployment as illustrated in the Standard Deployment Diagram.



6.3.1. Client Connection Process

1. Client initiates session request to the VIP on the load balancer.
2. The load balancer forwards the request to one of the load balanced Connection Brokers.
3. The Connection Broker checks the SQL database to determine if the user has an existing session, if yes the IP address for that server is selected, if no then the RDS built in load balancing mechanism selects a host/IP address where to start a new session.
4. The Connection Broker returns this IP address back to the client via the load balancer (assuming a Layer 7 configuration as used in this guide).
5. The client connects ***directly*** to the virtualization host specified.

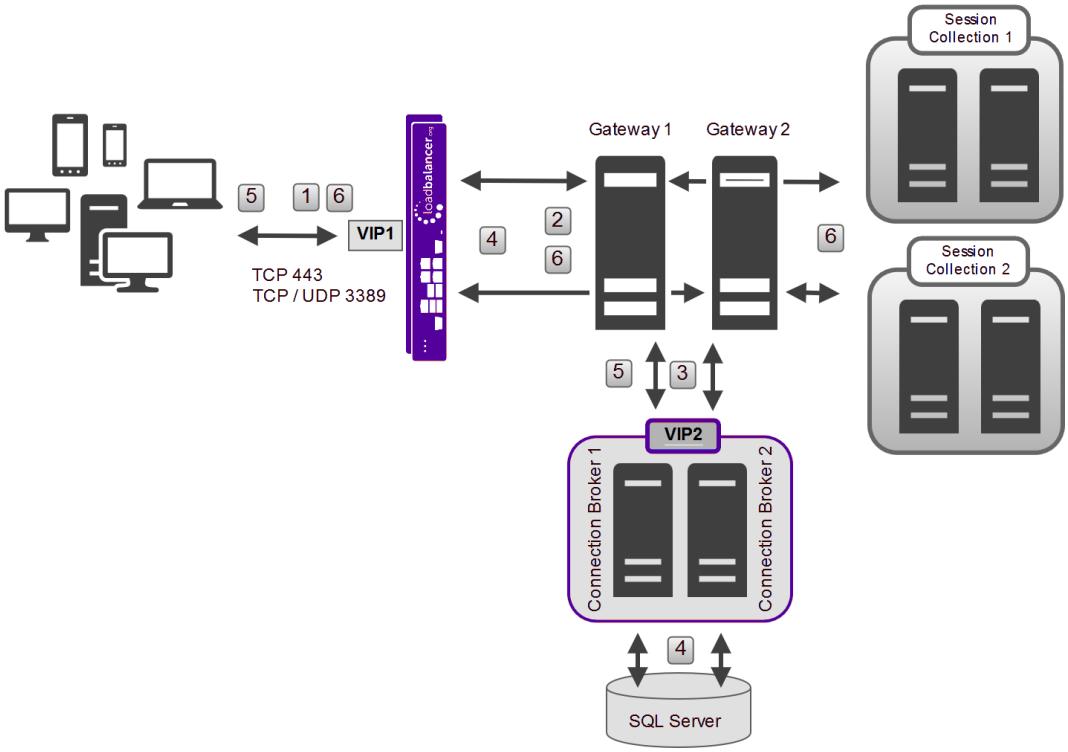
6.3.2. Scenario Notes

- In this scenario the Initial connection is to the Connection Brokers (via the load balancer).
- Session persistence from client to Connection Broker is not required because it handles the initial request and not active sessions.
- Layer 7 SNAT mode is recommended and is used for the example in this guide. It's also possible to use Layer 4 DR, NAT or SNAT modes depending on your infrastructure and requirements (see Layer 4 DR Mode , Layer 4 NAT Mode and Layer 4 SNAT Mode for descriptions of these modes).
- DNS must be configured so that the FQDN specified in ***DNS Name for the RD Connection Broker Cluster*** (**Deployment Properties > High Availability**) resolves to the Connection Broker VIP.
- Clients connect using RemoteAPP via RD Web Access or modified .RDP files and not just by specifying the DNS name or IP address of the Connection Brokers in mstsc.exe as explained here.

> **Note** See Load Balancing Connection Brokers (Scenarios 2a & 2b) for load-balancer configuration steps and RDS configuration notes related to this scenario.

6.4. Scenario 3 - Load Balancing Gateways

Scenario 3 is part of the Standard Deployment as illustrated in the Standard Deployment Diagram.



6.4.1. Client Connection Process

1. Client initiates session request to the VIP on the load balancer.
2. The load balancer forwards the request to one of the load balanced Gateways.
3. The selected Gateway proxies the request to the FQDN specified in ***DNS name for the RD Connection Broker*** in the deployment properties (this is normally the Connection Broker VIP as shown above, but if you're sending the Initial connection to the load balanced Session Hosts then this would be the load balanced Session Host VIP. If the load balanced Session Host VIP only supports TCP, then client connections via the Gateway will also support only UDP).
4. The Connection Broker checks the SQL database to determine if the user has an existing session. If yes, the IP address for that server is selected. If no, then the RDS built in load balancing mechanism selects a host/IP address on which it will start a new session.
5. The Connection Broker returns this IP address back to the client via the Gateway and load balancer.
6. The client connects via the Load Balancer & Gateway to the Session Host specified.

6.4.2. Scenario Notes

- Session persistence from client to Gateway is based on client source IP address.
- For each client session there are 1 or 2 HTTPS channels, and if they can be established, 1 or 2 additional UDP channels. The actual number of channels depends on the RDP client version being used and whether it's a 2012 or 2016 Gateway. The HTTPS channels need to be handled by the same Gateway as do the UDP channels, but HTTPS and UDP can be handled by different Gateways. For more information please refer to this link.

> **Note** If the Gateways are set up to proxy the RDP connections to load balanced Session Hosts (rather than the default, in which they are set up to proxy RDP connections to load balanced Connection

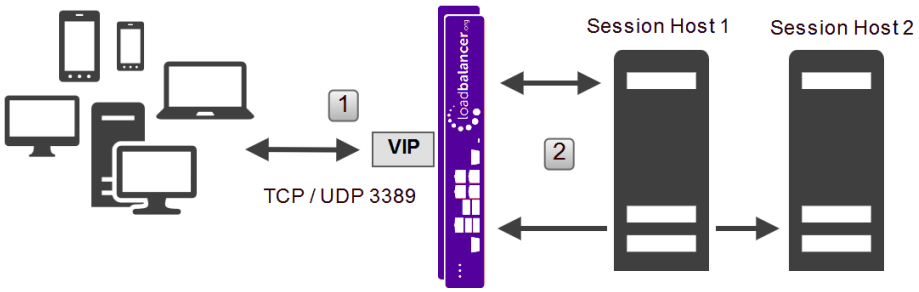
Servers), you must use a single VIP for the load balanced Gateways. This is required to ensure that both TCP and UDP are handled by the same RD Gateway. Then, when the VIP for the Session Hosts handles the connections, the source IP address is the same for both TCP & UDP and therefore both are forwarded to the same Session Host. If different VIPs were used for TCP and UDP, then it's possible that the UDP and TCP for the same session would be forwarded to different Session Hosts which would not work. The VIP configuration in this case is covered in Using a Single Layer 4 SNAT Mode VIP for Both TCP & UDP.

- Layer 7 SNAT mode is recommended for the TCP part and layer 4 SNAT mode is recommended for the UDP part , and are used for the example in this guide. It's also possible to use Layer 4 DR mode or layer 4 NAT mode depending on your infrastructure and requirements (see Layer 4 DR Mode and Layer 4 NAT Mode for descriptions of these modes).
- Clients connect using RemoteAPP via RD Web Access, modified .RDP files or via mstsc.exe.

Note See Load Balancing Gateways (Scenario 3) for load balancer configuration steps and RDS configuration notes related to this scenario.

6.5. Scenario 4 - Load Balancing Stand alone Session Hosts

Scenario 4 is ***NOT*** part of the Standard Deployment illustrated in the Standard Deployment Diagram. It offers a simple alternative to a full RDS deployment utilizing just Session Hosts and the load balancer.



6.5.1. Client Connection Process

1. Client initiates session request to the VIP on the load balancer.
2. If the client has connected previously, and the persistence (stick) table entry has not timed out, the load balancer forwards the request to the same Session Host that was used for the previous session. If the client has not connected previously or the stick-table entry has expired, the request is load balanced to one of the Session Hosts according to the load balancing algorithm selected.
3. The client continues the session to the selected Session Host via the load balancer (assuming a Layer 7 configuration as used in this guide).

6.5.2. Scenario Notes

- Appropriate for simple deployments that only require multiple full desktop sessions.
- In this scenario Connection Broker is not used.
- RemoteApp programs and Web Access are not available or supported.

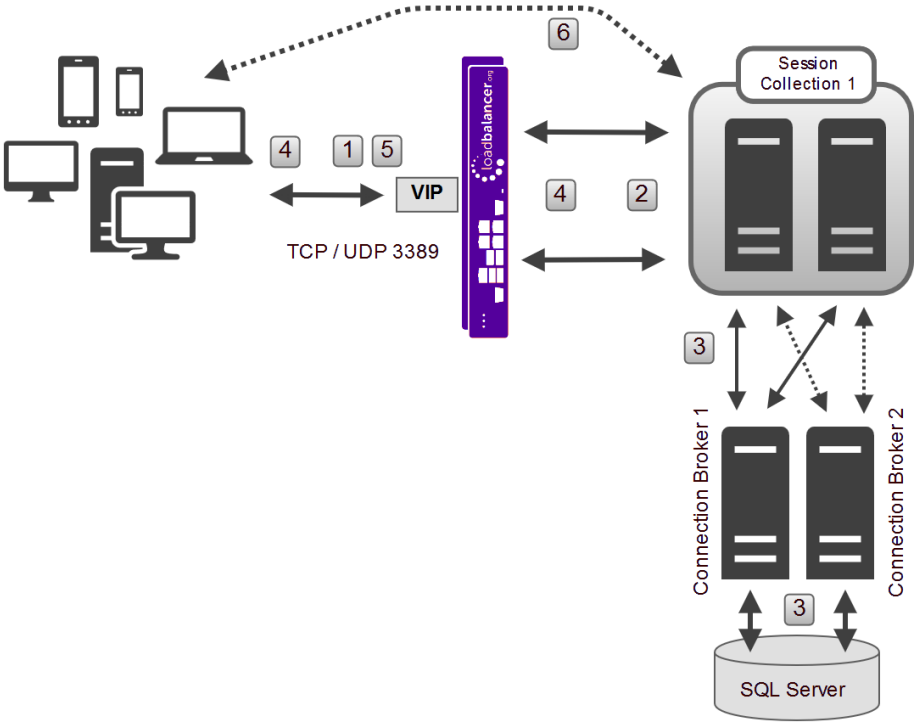
- In this scenario, session persistence can be based on client source IP address or the RDP cookie (mstshash – see RDP Client Cookie Persistence for more details) sent from the client in the Connection Request PDU.
- For Windows 2012 / 2016 It will not be possible to use Server Manager and/or most of the RDS Powershell commands to manage RDS. You will need to use group policy settings, WMI & registry edits.
- Layer 7 SNAT mode is recommended and is used for the example in this guide. It's also possible to use Layer 4 DR, NAT or SNAT mode depending on your infrastructure and requirements (see Layer 4 DR Mode , Layer 4 NAT Mode and Layer 4 SNAT Mode for descriptions of these modes).
- Clients connect using the Microsoft RDP client (mstsc.exe) or equivalent.

> **Note** For more details on using Session Host without Connection Broker, please refer to this URL.

> **Note** See Load Balancing Standalone Session Hosts (Scenario 4) for load balancer configuration steps and RDS configuration notes related to this scenario.

6.6. Scenario 5 - Load Balancing Session Hosts when Deployed with Connection Broker

Scenario 5 is ***NOT*** part of the Standard Deployment illustrated in the Standard Deployment Diagram. Here, the Session Hosts are load balanced by the load balancer appliance rather than the built-in mechanism of RDS.



6.6.1. Client Connection Process

For IP Address Redirection Mode (the default)

1. Client initiates session request to the VIP on the load balancer.

2. The load balancer forwards the request to one of the load balanced Session Hosts.
3. The Session Host checks with one of the active/active Connection Brokers to determine if there is an existing session.
4. If there is an existing session, the IP address for the Session Host where the session is running is passed to the client in the encrypted load balance packet.
5. -
6. The client then reconnects ***directly*** to the Session Host specified.

   If there was no existing session, a new session is started on the Session Host where the user was originally load balanced

For Routing Token Redirection Mode (configured via Group Policy)

1. Client initiates session request to the VIP on the load balancer.
2. The load balancer forwards the request to one of the load balanced Session Hosts.
3. The Session Host checks with one of the active/active Connection Broker to determine if there is an existing session.
4. If there is an existing session, the IP address for the Session Host where the session is running is encoded in a ***Routing Token*** and returned to the client via the load balancer.
5. The client then reconnects to the load balancer presenting this ***Routing Token***, the load balancer then connects the client to the Session Host specified in the ***Routing Token***.

   If there was no existing session, a new session is started on the Session Host where the user was originally load balanced.

> **Note** For detailed information about Routing Tokens and their format please refer to this Microsoft document.

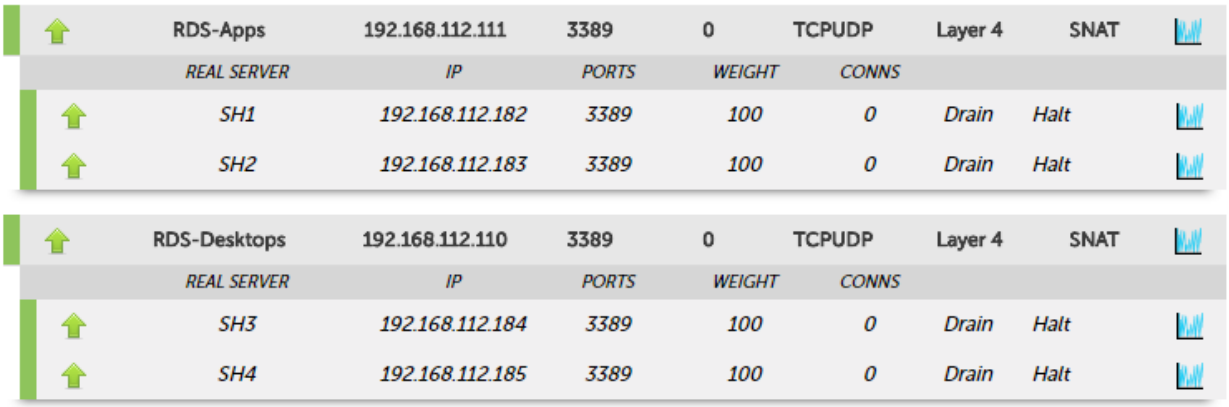
6.6.2. Scenario Notes

- In this scenario the Initial connection is handled by the load balanced Session Hosts, this is not the default Microsoft method. For Windows 2012 and later, the default is to send the Initial connection to the load balanced Connection Brokers as per Scenario 2a - Load Balancing Connection Brokers with Session Hosts and Scenario 2b - Load Balancing Connection Brokers with Virtualization Hosts.
- The built in load balancing mechanism must be ***disabled*** for all Session Hosts so that only the LB.org appliance is responsible for load balancing connections to the Session Hosts. This is achieved through Group Policy as described in Group Policy.
- In this scenario the Loadbalancer.org feedback agent can be used to modify the load balancing algorithm in real time based on Session Host RAM & CPU utilization.
- DNS must be configured so that the FQDN specified in ***DNS Name for the RD Connection Broker Cluster*** (**Deployment Properties > High Availability**) resolves to the Session Host VIP.
- If you have only one Connection Broker (not recommended), Web Access will not work since the FQDN

written to the .RDP files will be the FQDN of the Connection Broker server. In this case, you'll need to download and manually modify the .RDP files so that this FQDN can be replaced with the FQDN of your Session Host VIP.

- The load balancer is not aware of RDS Session Collections, so if the deployment consists of more than one Collection, multiple VIPs are needed segregating the Session Hosts according to Session Collection membership. Also, Web Access no longer works correctly and .RDP files need to be manually modified/created to ensure clients are sent to the correct VIP.

  For Example, if you have 2 Session Collections each with 2 Session Hosts, you would need to create 2 VIPs as shown below:



RDS-Apps	192.168.112.111	3389	0	TCPUDP	Layer 4	SNAT	
REAL SERVER	IP	PORTS	WEIGHT	CONNS			
SH1	192.168.112.182	3389	100	0	Drain	Halt	
SH2	192.168.112.183	3389	100	0	Drain	Halt	
RDS-Desktops	192.168.112.110	3389	0	TCPUDP	Layer 4	SNAT	
REAL SERVER	IP	PORTS	WEIGHT	CONNS			
SH3	192.168.112.184	3389	100	0	Drain	Halt	
SH4	192.168.112.185	3389	100	0	Drain	Halt	

  2 DNS records would be required that point to these VIPs , e.g. :

```
rds-apps.lbtestom.com ---> 192.168.112.111
```

```
rds-desktops.lbtestdom.com ---> 192.168.112.110
```

  And .RDP files would be need to configured for each VIP by modifying the original .RDP files generated by Web Access and replacing all occurrences of **rds.lbtestdom.com** with one of the above FQDN's, then distributing these to your clients.

- Using RPC, each Session Host ensures that it has an active connection with one of the Connection Brokers in the deployment. If for any reason that connection is lost, e.g. due to a failure of the first Connection Broker, a new connection is automatically established with one of remaining Connection Brokers.

> **Note**
>
> The settings **Select Active Connection Broker** in Windows 2012 and the equivalent setting **Select RD Management Server** in Windows 2016 have no effect on which Connection Broker is used by each Session Host, so there is no need to configure this setting in relation to load balancing. These settings are used to configure which Connection Broker is able to accept configuration changes made in either the Server Manager Console or via Powershell.

- The default redirection method is ***IP Address Redirection Mode***. The alternative method - ***Routing Token Redirection Mode*** can be selected by configuring Group Policy as described in Group Policy.

- When using ***Routing Token Redirection Mode***, Layer 7 SNAT mode configured with MS Session Broker persistence ***must*** be used to enable the ***Routing Tokens*** to be read. In this case, RDP over UDP will not work for both internal and external clients because the layer 7 Session Host VIP does not support UDP. If you require UDP transport for RDP, layer 4 SNAT mode is recommended.
- Clients connect using RemoteAPP via RD Web Access, modified .RDP files or via ***mstsc.exe***.

> **Note** See Load Balancing Session Hosts Deployed with Connection Broker (Scenario 5) for load balancer configuration steps and RDS configuration notes.

7. Loadbalancer.org Appliance – the Basics

7.1. Virtual Appliance

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, XEN and Nutanix AHV and has been optimized for each Hypervisor. By default, the VA is allocated 2 vCPUs, 4GB of RAM and has a 20GB 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 Virtual Appliance Installation and the ReadMe.txt text file included in the VA download for additional information on deploying the VA using the various Hypervisors.

> **Note** The VA has 4 network adapters. For VMware only the first adapter (**eth0**) is connected by default. For HyperV, KVM, XEN and Nutanix AHV all adapters are disconnected by default. Use the network configuration screen within the Hypervisor to connect the required adapters.

7.2. Initial Network Configuration

After boot up, follow the instructions on the appliance console to configure the management IP address, subnet mask, default gateway, DNS servers and other network and administrative settings.

> **Important** Be sure to set a secure password for the load balancer, when prompted during the setup routine.

7.3. Accessing the Appliance WebUI

The WebUI is accessed using a web browser. By default, users are authenticated using Apache authentication. Users can also be authenticated against LDAP, LDAPS, Active Directory or Radius - for more information, please refer to External Authentication.

> **Note** There are certain differences when accessing the WebUI for the cloud appliances. For details, please refer to the relevant Quick Start / Configuration Guide.

1. Using a browser, navigate to the following URL:

   https://<IP-address-configured-during-the-network-setup-wizard>:9443/lbadmin/

   > **Note** You'll receive a warning about the WebUI's SSL certificate. This is due to the default self signed certificate that is used. If preferred, you can upload your own certificate - for more information, please refer to Appliance Security Features.

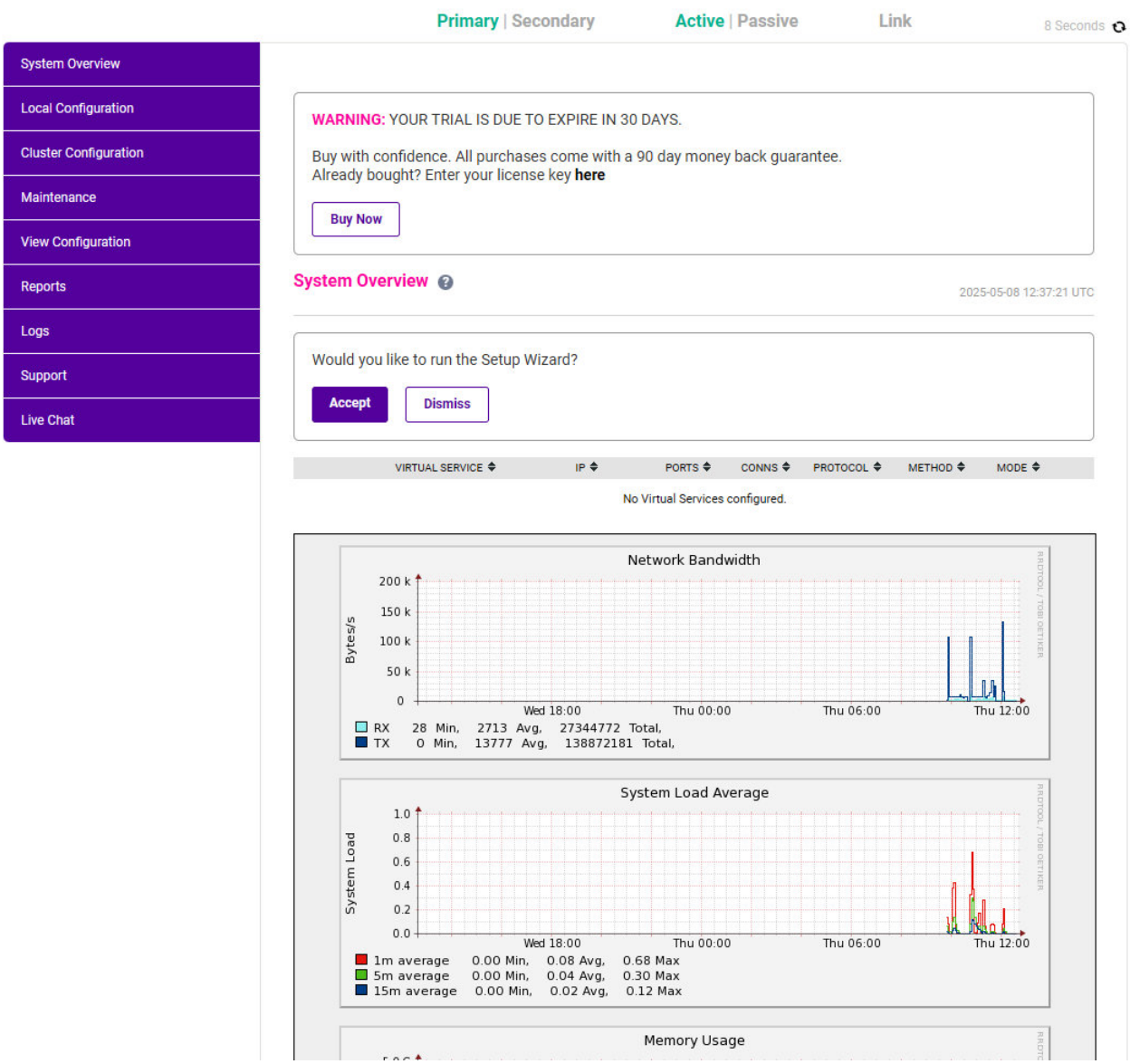
   > **Note** If you need to change the port, IP address or protocol that the WebUI listens on, please refer to Service Socket Addresses.

2. Log in to the WebUI using the following credentials:

   Username: loadbalancer
   Password: <configured-during-network-setup-wizard>

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

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



3. You'll be asked if you want to run the Setup Wizard. Click **Dismiss** if you're following a guide or want to configure the appliance manually. Click **Accept** to start the Setup Wizard.

> **Note** The Setup Wizard can only be used to configure Layer 7 services.

7.3.1. Main Menu Options

System Overview - Displays a graphical summary of all VIPs, RIPs and key appliance statistics
Local Configuration - Configure local host settings such as IP address, DNS, system time etc.
Cluster Configuration - Configure load balanced services such as VIPs & RIPs
Maintenance - Perform maintenance tasks such as service restarts and creating backups
View Configuration - Display the saved appliance configuration settings
Reports - View various appliance reports & graphs
Logs - View various appliance logs
Support - Create a support download, contact the support team & access useful links

Live Chat - Start a live chat session with one of our Support Engineers

7.4. Appliance Software Update

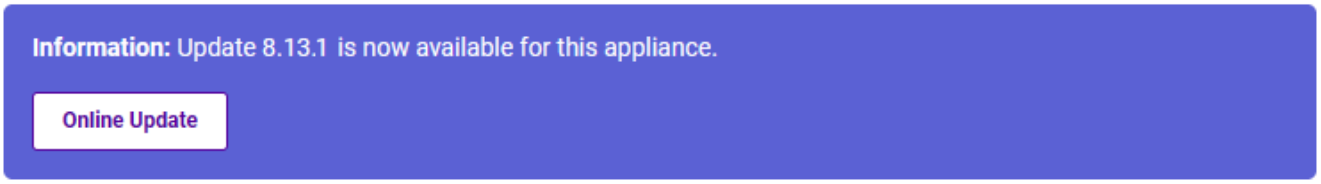
We recommend that the appliance is kept up to date to ensure that you benefit from the latest bug fixes, security updates and feature improvements. Both online and offline update are supported.

> **Note** For full details, please refer to Appliance Software Update in the Administration Manual.

> **Note** Services may need to be restarted/reloaded after the update process completes or in some cases a full appliance restart may be required. We therefore recommend performing the update during a maintenance window.

7.4.1. Online Update

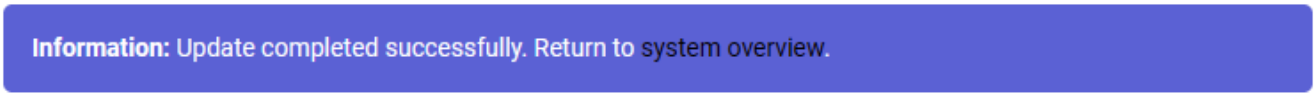
The appliance periodically contacts the Loadbalancer.org update server (**update.loadbalancer.org**) and checks for updates. This is the default behavior and can be disabled if preferred. If an update is found, a notification similar to the example below will be displayed at the top of the WebUI:



Click **Online Update**. A summary of all new features, improvements, bug fixes and security updates included in the update will be displayed. Click **Update** at the bottom of the page to start the update process.

> **Important** Do not navigate away whilst the update is ongoing, this may cause the update to fail.

The update can take several minutes depending on download speed and upgrade version. Once complete, the following message will be displayed:



If services need to be reloaded/restarted or the appliance needs a full restart, you'll be prompted accordingly.

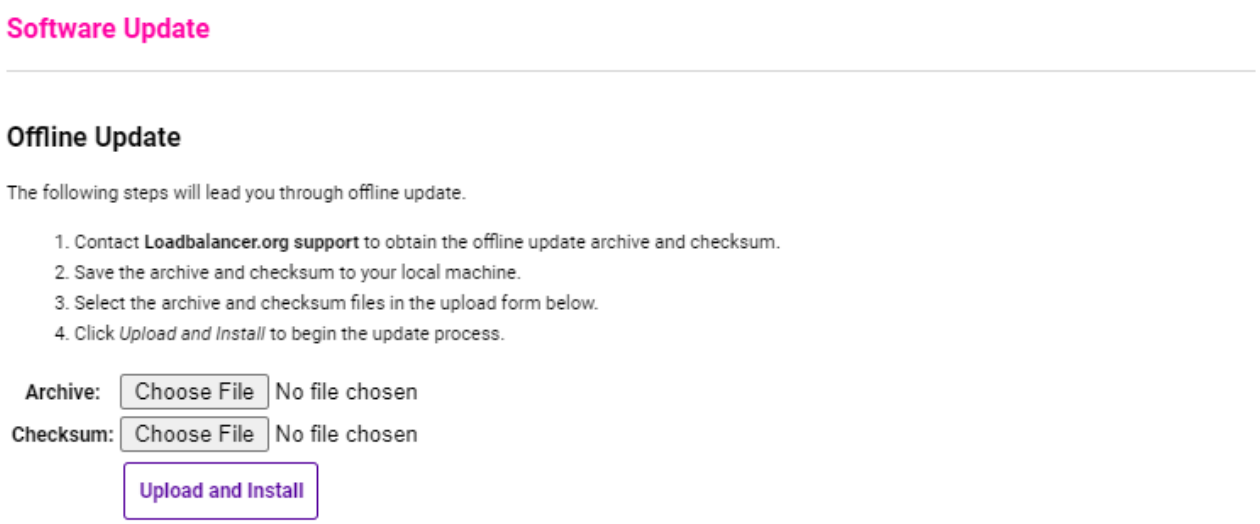
7.4.2. Offline Update

If the appliance does not have access to the Internet, offline update can be used.

To check for the latest version, please refer to our product roadmap page available here. To obtain the latest offline update files contact support@loadbalancer.org.

To perform an offline update:

1. Using the WebUI, navigate to: *Maintenance > Software Update*.
2. Select **Offline Update**.
3. The following screen will be displayed:



4. Select the *Archive* and *Checksum* files.
5. Click **Upload and Install**.
6. If services need to be reloaded/restarted or the appliance needs a full restart, you'll be prompted accordingly.

7.5. Ports Used by the Appliance

By default, the appliance uses the following TCP & UDP ports:

Protocol	Port	Purpose
TCP	22 *	SSH
TCP & UDP	53 *	DNS / GSLB
TCP & UDP	123	NTP
TCP & UDP	161 *	SNMP
UDP	6694	Heartbeat between Primary & Secondary appliances in HA mode
TCP	7778	HAProxy persistence table replication
TCP	9000 *	Gateway service (Centralized/Portal Management)
TCP	9080 *	WebUI - HTTP (disabled by default)
TCP	9081 *	Nginx fallback page
TCP	9443 *	WebUI - HTTPS
TCP	25565 *	Shuttle service (Centralized/Portal Management)

> **Note** The ports used for SSH, GSLB, SNMP, the WebUI, the fallback page, the gateway service and the shuttle service can be changed if required. For more information, please refer to Service Socket

Addresses.

7.6. 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 unit is covered in Configuring HA - Adding a Secondary Appliance.

8. Load Balancing Web Access Servers (Scenario 1)

Scenario 1 is part of the Standard Deployment as illustrated in the Standard Deployment Diagram. Please also refer to Scenario 1 - Load Balancing Web Access Servers for detailed notes on how the load balancer interacts with RDS in this scenario.

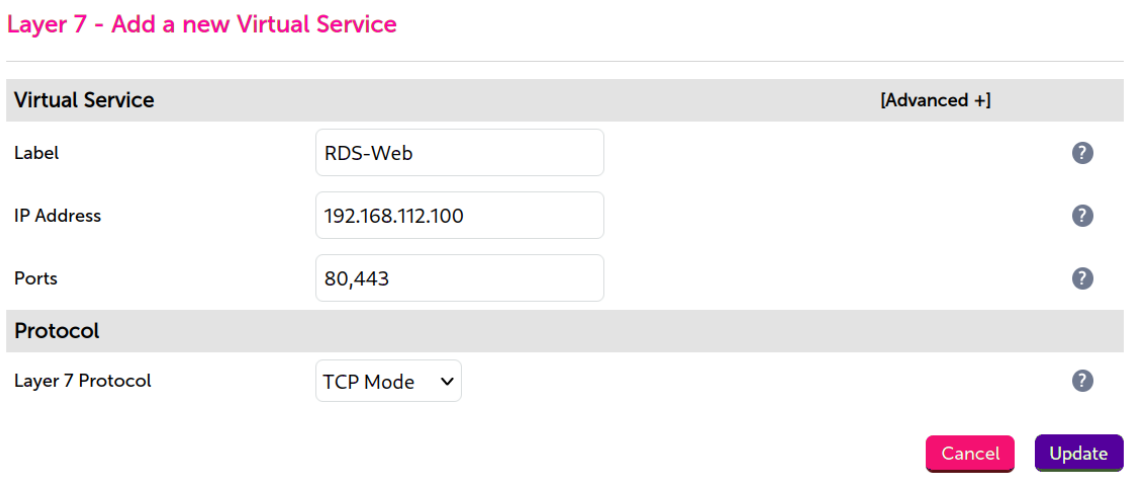
8.1. RDS Installation & Configuration

- Use the *Remote Desktop Services* installation type to perform a Standard deployment with 1 Connection Broker, 1 Web Access Server and the required number of Session Hosts / Virtualization Hosts.
- Add one or more Web Access Servers to the deployment.
- Configure RDS Certificates as mentioned in RDS Certificates.

8.2. Appliance Configuration

8.2.1. 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:

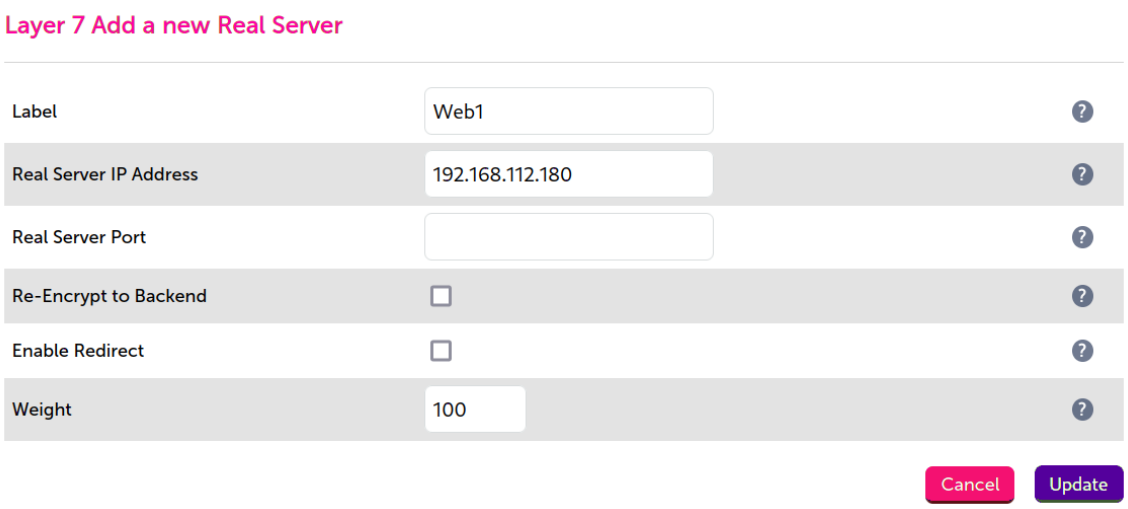


3. Enter an appropriate name (Label) for the Virtual Service, e.g. **RDS-Web**.
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.112.100**.
5. Set the *Virtual Service Ports* field to **80,443**.

6. Set the *Protocol* to **TCP Mode**.
7. Click **Update**.
8. Now click **Modify** next to the newly created Virtual Service.
9. Ensure that *Persistence Mode* is set to **Source IP**.
10. Set the *Persistence Timeout* to **2h** (i.e. 2 hours).
11. In the *Health Checks* section click **Advanced** to expand the section.
12. Configure the *Health Check* settings to look for an **HTTP 200 OK** response:
    - Set *Health Checks* to **Negotiate HTTPS (GET)**
    - Set *Check Port* to **443**
    - Leave all other fields **blank**
13. In the *Other* section click **Advanced** to expand the section.
14. Enable (check) the *Timeout* checkbox and set both *Client Timeout & Real Server Timeout* to **2h** (i.e. 2 hours).
15. Click **Update**.

8.2.2. 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 Web Access, e.g. **Web1**.
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.112.180**.
5. Leave the *Real Server Port* field **blank**.
6. Click **Update**.
7. Now repeat for your remaining Web Access server(s).

8.3. Testing & Verification

> **Note** For additional guidance on diagnosing and resolving any issues you may have, please also refer to Diagnostics & Troubleshooting.

Configure DNS so that the FQDN to be used for Web Access resolves to the VIP address. The load balanced Web Access servers should now be accessible via the load balancer.

Connect to the Web Access URL using your browser, e.g. :

```
https://rds.lbtestdom.com/RDweb
```

9. Load Balancing Connection Brokers (Scenarios 2a & 2b)

Scenario 2 is part of the Standard Deployment as illustrated in the Standard Deployment Diagram. Please also refer to Scenario 2a - Load Balancing Connection Brokers with Session Hosts and Scenario 2b - Load Balancing Connection Brokers with Virtualization Hosts for detailed notes on how the load balancer interacts with RDS in these scenarios.

9.1. RDS Installation & Configuration

- Use the ***Remote Desktop Services*** installation type to perform a Standard deployment with one Connection Broker, one Web Access Server and the required number of Session Hosts / Virtualization Hosts.
- Configure Connection Broker HA mode:
  - The SQL configuration for the LAB used for this guide is shown below:

    ```
    DRIVER=SQL Server Native Client 11.0; SERVER=WIN2012-TEST.lbtestdom.com;
    Trusted_Connection=Yes; APP=Remote Desktop Services Connection Broker; DATABASE=RDCB01
    ```

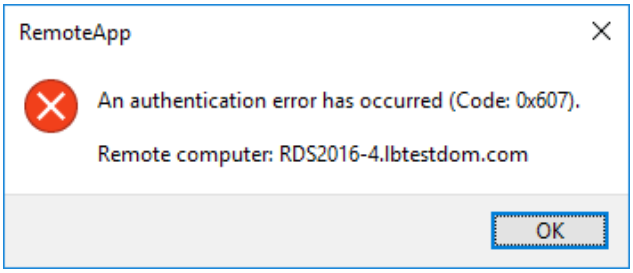
  - The native client can be downloaded here and must be installed on each Connection Broker.
- Add one or more Connection Brokers to the deployment.
- Configure RDS Certificates as mentioned in RDS Certificates.
- Session Host health checking is periodically performed by the Connection Brokers. The health check interval and other related settings can be changed using the following registry path on each Connection Broker server:

  ```
  HKLM/SYSTEM/CurrentControlSet/Services/Tssdis /Parameters
  ```

Name	Type	Data
(Default)	REG_SZ	(value not set)
DBConnString	REG_SZ	DRIVER=SQL Server
NumberFailedPingsBeforePurge	REG_DWORD	0x00000002 (2)
PingMode	REG_DWORD	0x00000000 (0)
RecoverWhenStart	REG_DWORD	0x00000001 (1)
TimeBetweenPings	REG_DWORD	0x0000001e (30)
TimeServerSilentBeforePing	REG_DWORD	0x0000003c (60)
TraceOutputMode	REG_DWORD	0x00000000 (0)
WorkingDirectory	REG_SZ	

The time-related settings in (brackets) are in seconds

- The following error can mean that a custom RDS certificate has been installed on the Session Hosts: **An authentication error has occurred (Code: 0x607)**.



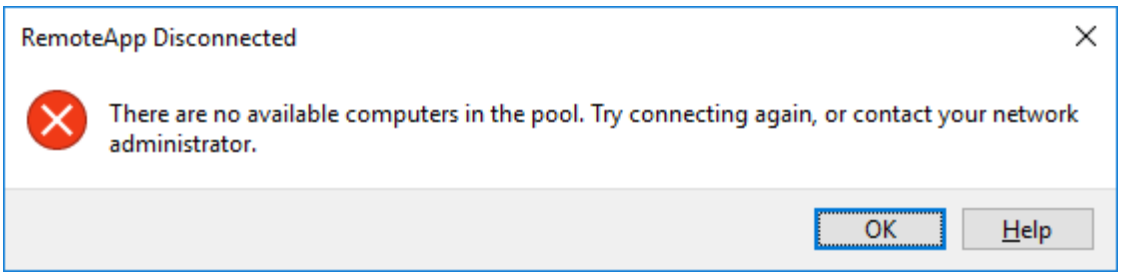
When the Initial connection is handled by the Connection Broker (the default for Windows 2012 & later), the client will authenticate the Connection Broker using a certificate (and/or Kerberos), and then the Broker will authenticate the target Session Host on behalf on the client.

Make sure that the default self-signed RDS certificate is being used on each Session Host.

> **Note** To revert to the default self-signed RDS certificate, please refer to Reverting to Default Self-Signed RDS Certificate.

- If you receive the following error:

There are no available computers in the pool. Try connecting again, or contact your network administrator.



- After adding the additional Connection Brokers, if you see multiple Event 1016's as shown below:

```
RD Connection Broker service denied the remote procedure call (RPC) from an unauthorized
computer 192.168.112.184
```

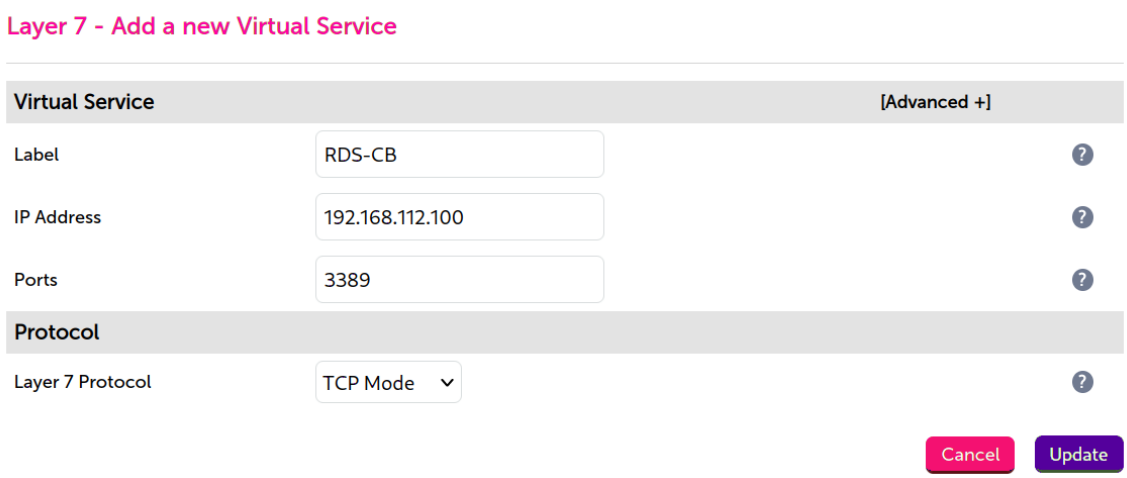
  Make sure that the **RDS Endpoint Servers** group on each Connection Broker server includes all Connection Broker servers in the deployment.

- Ensure there is a valid DNS entry for the HA Connection Broker defined in the deployment settings. e.g. configure a DNS entry for **rds.lbtestdom.com** pointing to the VIP address.

9.2. Appliance Configuration

9.2.1. Setting up the Virtual Service (VIP)

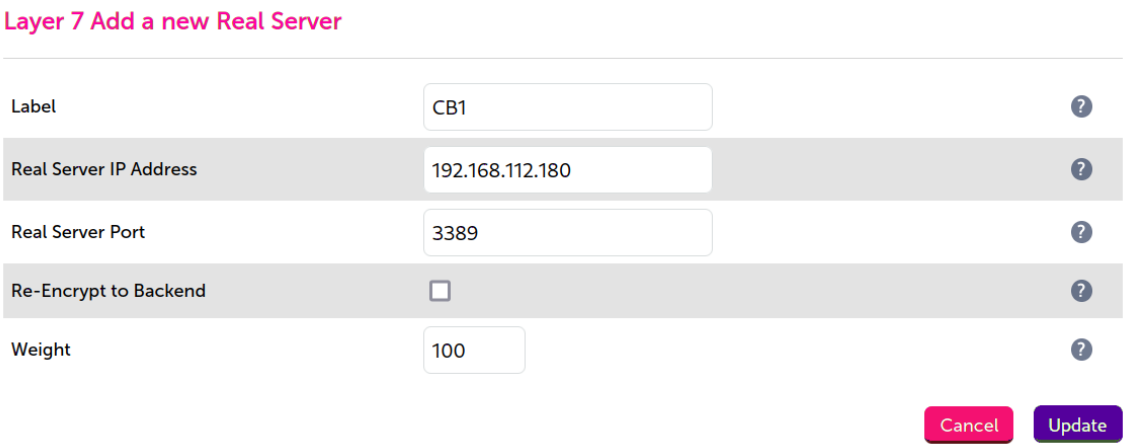
1. Using the WebUI, navigate to: *Cluster Configuration > Layer 7 – Virtual Services* and click **Add a New Virtual Service**.
2. Enter the following details:



3. Enter an appropriate name (Label) for the Virtual Service, e.g. **RDS-CB**.
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.112.100**.
5. Set the *Virtual Service Ports* field to **3389**.
6. Set the *Layer 7 Protocol* to **TCP Mode**.
7. Click **Update**.
8. Now click **Modify** next to the newly created Virtual Service.
9. Set *Persistence Mode* to **None**.
10. In the *Other* section click **Advanced** to expand the section.
11. Enable (check) the *Timeout* checkbox and set both *Client Timeout & Real Server Timeout* to **5m**.
12. Click **Update**.

9.2.2. 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 RDS server, e.g. **CB1**.
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.112.180**.
5. Set the *Real Server Port* field to **3389**.
6. Click **Update**.
7. Now repeat for your remaining Connection Broker server(s).

9.2.3. Applying the new Layer 7 Settings

1. Once the configuration is complete, use the **Reload HAProxy** button at the top of the screen to commit the changes.

9.3. Testing & Verification

> **Note** For additional guidance on diagnosing and resolving any issues you may have, please also refer to Diagnostics & Troubleshooting.

DNS must be configured so that the FQDN (e.g. **rds.lbtestdom.com**) specified in *DNS Name for the RD Connection Broker Cluster* (**Deployment Properties > High Availability**) resolves to the Connection Broker VIP. The load-balanced Connection Brokers should now be accessible via the load balancer.

Use Web Access / RemoteAPP to verify that published applications are available.

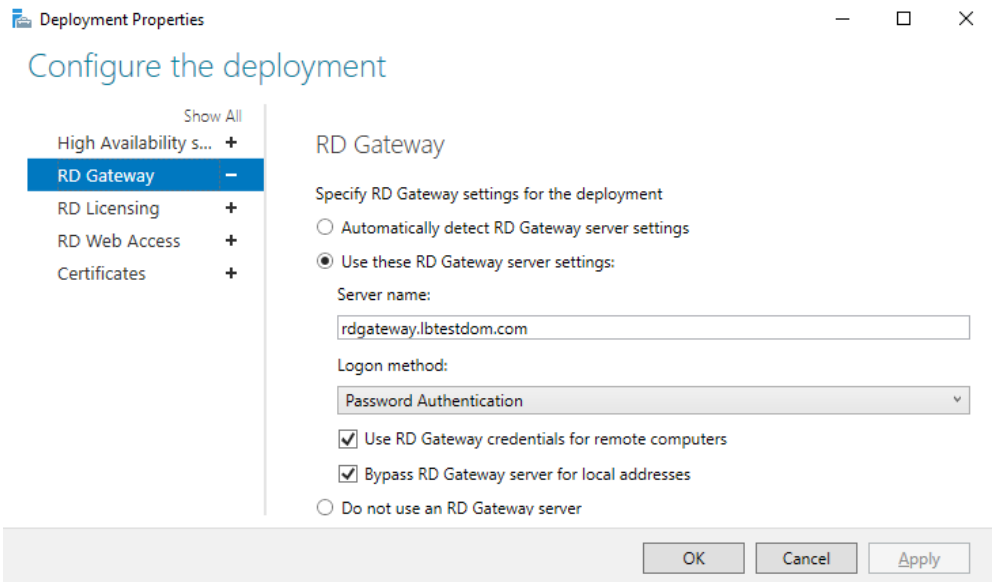
10. Load Balancing Gateways (Scenario 3)

Scenario 3 is part of the Standard Deployment as illustrated in the Standard Deployment Diagram. Please also refer to Scenario 3 - Load Balancing Gateways for detailed notes on how the load balancer interacts with RDS in this scenario.

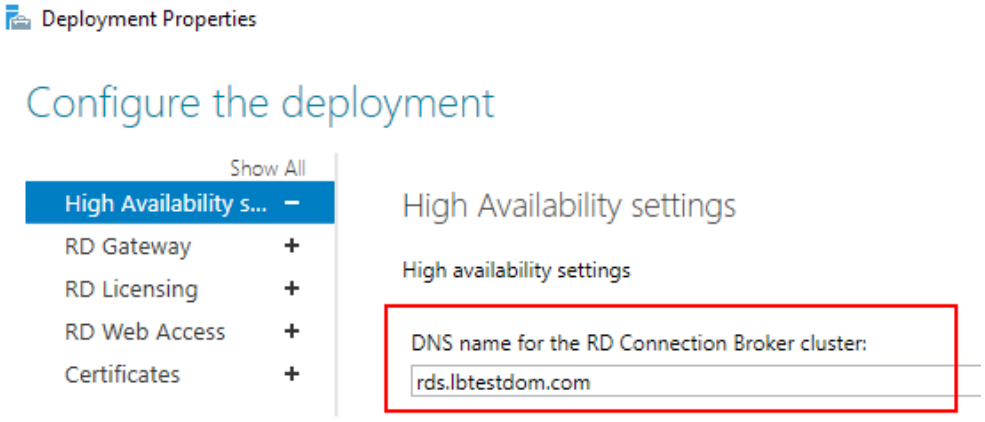
> **Note** It is possible to install and use RD Gateway without a complete RDS infrastructure that includes Connection Broker. A useful resource to help set this up is available here.

10.1. RDS Installation & Configuration

- Use the ***Remote Desktop Services*** installation type to perform a Standard deployment with 1 Connection Broker, 1 Web Access Server and the required number of Session Hosts / Virtualization Hosts.
- Add 2 or more Gateways to the deployment.
- Configure RDS Certificates as mentioned in RDS Certificates.
- Ensure that the RD Gateway settings are configured according to your requirements:

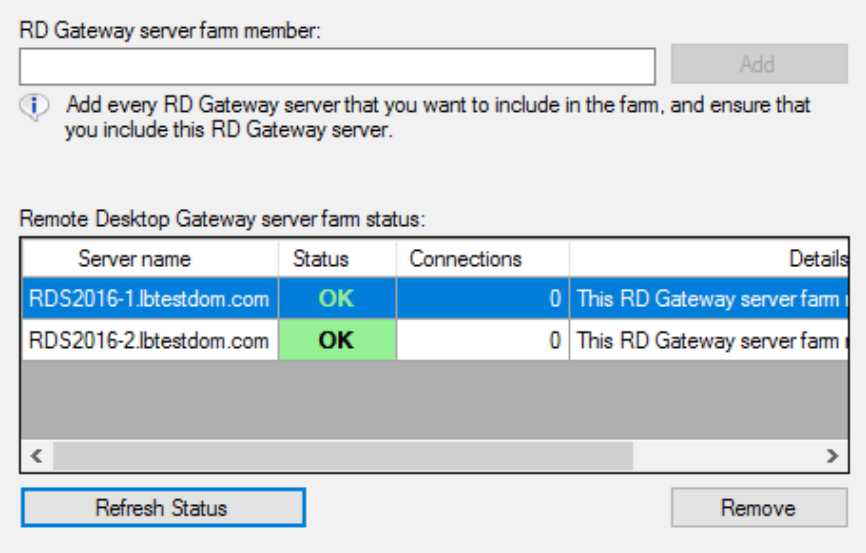


- Ensure that clients can resolve the FQDN for the load balanced Gateways (**rdgateway.robstest.com** in the above example). This should point to the load balanced Gateway VIP (see Using 2 VIPs – One for TCP & One for UDP for details on configuring this VIP).
- Ensure that the *DNS name for the RD Connection Broker cluster* (**Deployment Properties > High Availability**) is configured correctly to suit your environment, e.g. :



- Ensure that all load balanced RD Gateways are members of the same RD Gateway server farm as shown in

the example below:



- Ensure that the CAP & RAP policies are configured correctly to specify which users can connect to the RDS deployment and which resources they can access. By default all users in the domain are granted access to all computers in the domain.

  Also make sure that the FQDN used to access your deployment is included. In Windows 2016 the FQDN specified in *DNS name for the RD Connection Broker cluster* (**Deployment Properties > High Availability**) is automatically added to the default RAP **RDG_HighAvailabilityBroker_DNS_RR**.

10.2. Appliance Configuration

- If the Gateways proxy their connections to load balanced Connection Brokers (the default) then two VIPs are used – one for TCP/HTTPS on port 443, the second is for UDP on port 3391. This enables different Gateways to be used for the TCP & UDP parts of the Session. For configuration steps, please refer to the section: Using 2 VIPs – One for TCP & One for UDP below.
- If the Gateways proxy their connections to load balanced Session Hosts, a single VIP ***must*** be used to ensure that both TCP and UDP are handled by the same RD Gateway. Then, when the VIP for the Session Hosts handles the connections, the source IP address is the same for both TCP & UDP and therefore both are forwarded to the same Session Host. For configuration steps, please refer to the section: Using a Single Layer 4 SNAT Mode VIP for Both TCP & UDP.

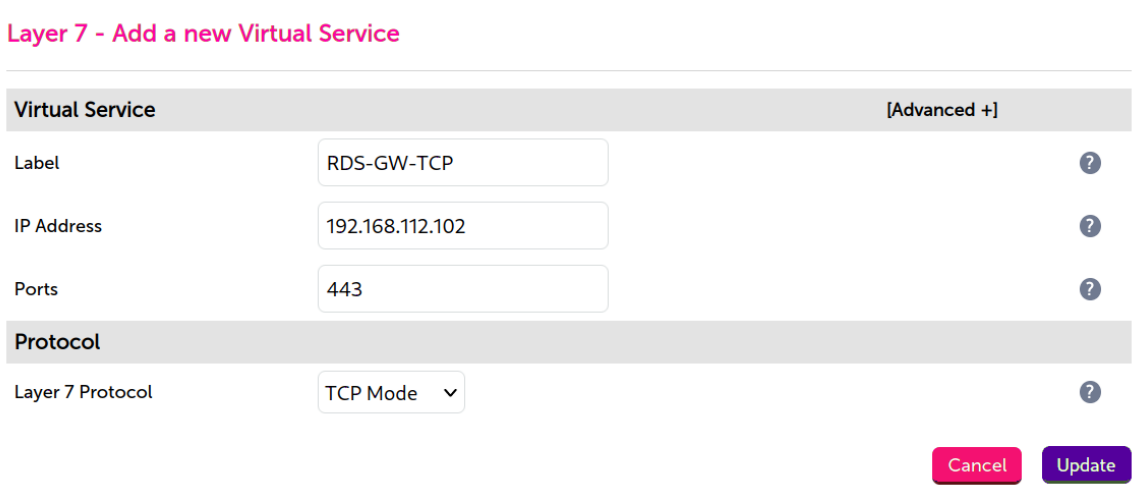
> **Note** If a single layer 4 SNAT mode VIP is used and your deployment has a single Session Collection and RD Gateway is collocated with Web Access, then the Web Access VIP described in Load Balancing Web Access Servers (Scenario 1) must be configured using layer 4 SNAT mode rather than layer 7 SNAT mode.

10.2.1. Using 2 VIPs – One for TCP & One for UDP

Setting up the Virtual Service (VIP) for TCP / HTTPS

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

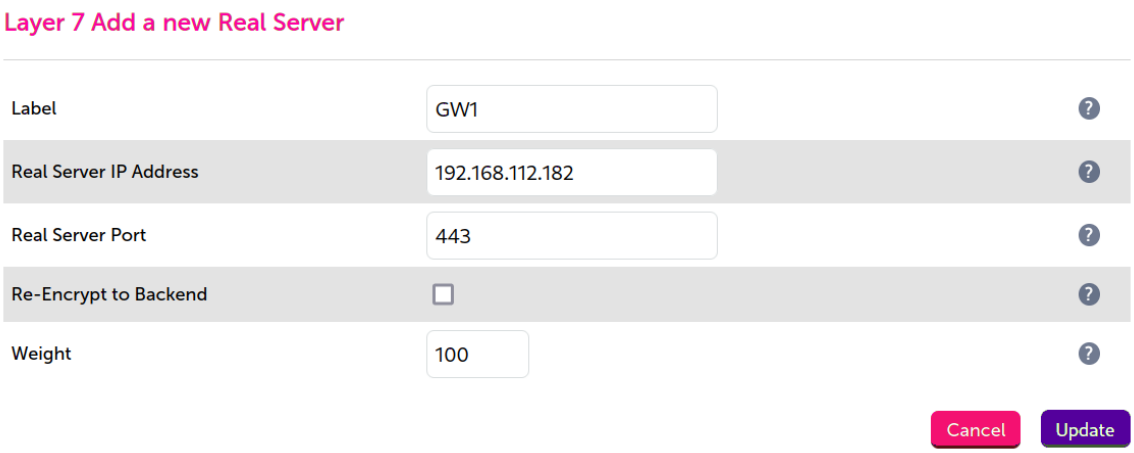
2. Enter the following details:



3. Enter an appropriate name (Label) for the Virtual Service, e.g. **RDS-GW-TCP**.
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.112.102**.
5. Set the *Virtual Service Ports* field to **443**.
6. Set the *Layer 7 Protocol* to **TCP Mode**.
7. Click **Update**.
8. Now click **Modify** next to the newly created Virtual Service.
9. Enable (check) *TCP keep-alive*.
10. Ensure that *Persistence Mode* is set to **Source IP**.
11. Leave the *Persistence Timeout* set to to **30** (i.e. 30 minutes).
12. In the *Other* section click **Advanced** to expand the section.
13. Enable (check) the Timeout checkbox and set both *Client Timeout* and *Real Server Timeout* to **30m** (i.e. 30 minutes).
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:



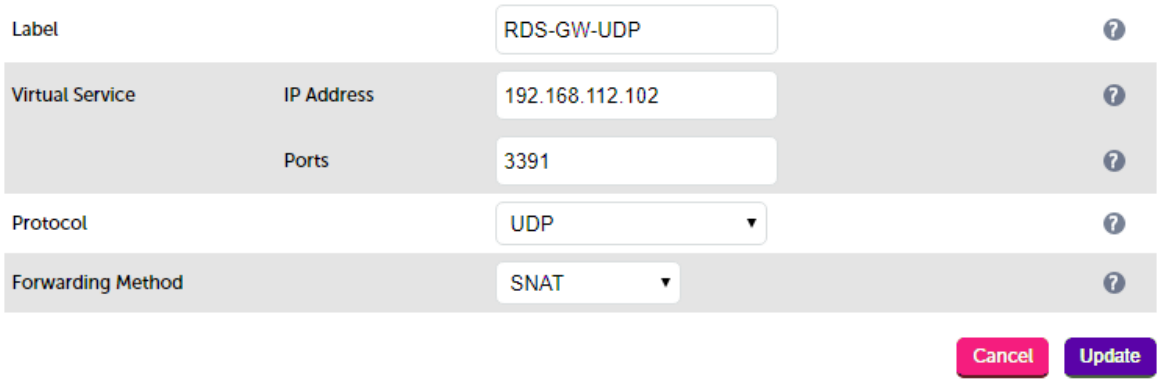
3. Enter an appropriate name (Label) for the first RD Gateway, e.g. **GW1**.
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.112.182**.
5. Set the *Real Server Port* field to **443**.
6. Click **Update**.
7. Now repeat for your remaining RD Gateway(s).

Applying the new Layer 7 Settings

1. Once the configuration is complete, use the **Reload HAProxy** button at the top of the screen to commit the changes.

Setting up the Virtual Service (VIP) for UDP

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

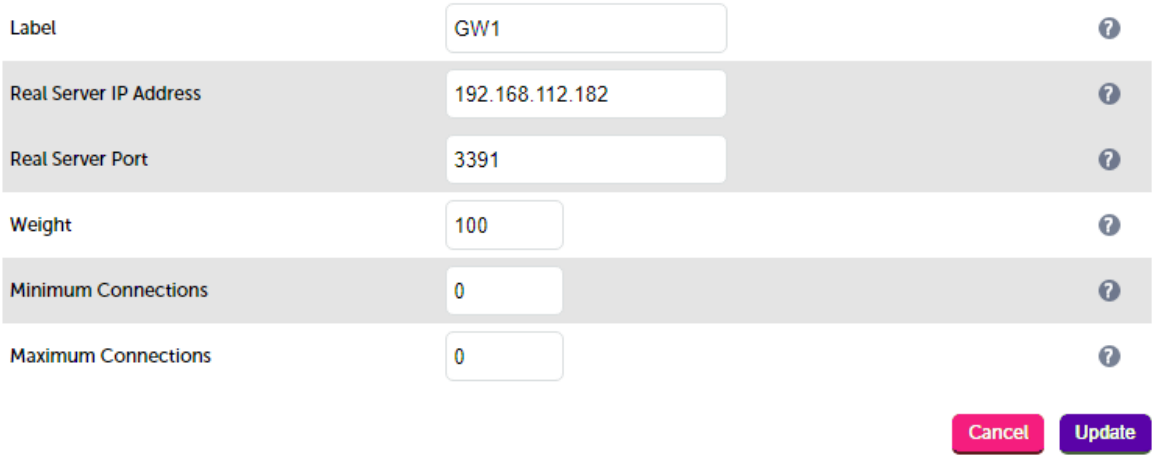


3. Enter an appropriate name (Label) for the Virtual Service, e.g. **RDS-GW-UDP**.
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.112.102**.
5. Set the *Virtual Service Ports* field to **3391**.
6. Set the *Protocol* to **UDP**.

7. Set the *forwarding Method* to **SNAT**.
8. Click **Update**.
9. Now click **Modify** next to the newly created Virtual Service.
10. Ensure that the *Persistent* check-box is enabled (checked).
11. Leave the *Persistent Timeout* is set to **300** (i.e. 5 minutes).
12. Ensure the *Health Checks Check Type* is set to **Ping Server**.
13. 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:

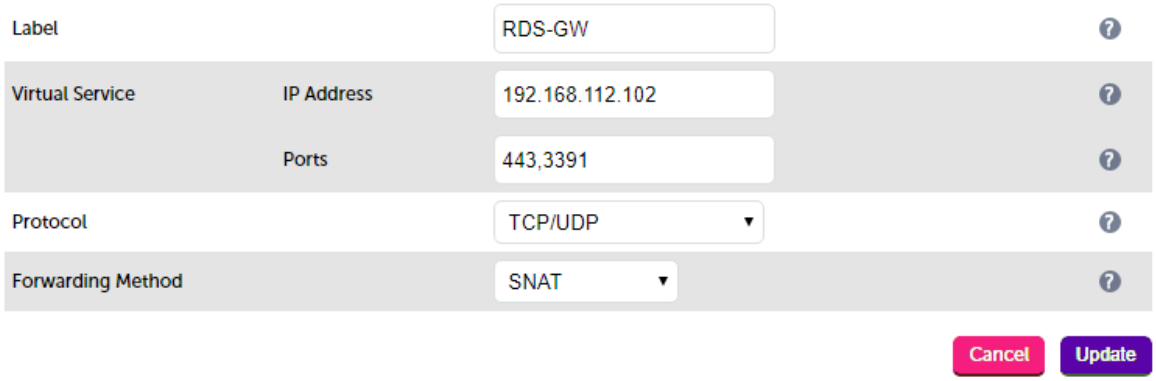


3. Enter an appropriate name (Label) for the first RD Gateway, e.g. **GW1**.
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.112.182**.
5. Leave other values at the default values.
6. Click **Update**.
7. Now repeat for your remaining RD Gateway(s).

10.2.2. Using a Single Layer 4 SNAT Mode VIP for Both TCP & UDP

Setting up the Virtual Service (VIP)

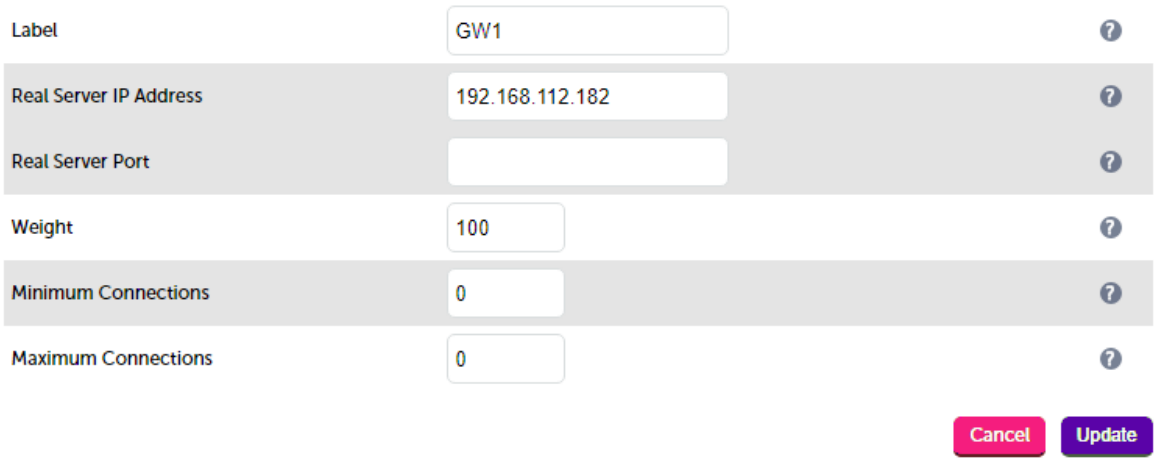
1. Using the WebUI, navigate to: *Cluster Configuration > Layer 4 – Virtual Services* and click **Add a New Virtual Service**.
2. Enter the following details:



3. Enter an appropriate name (Label) for the Virtual Service, e.g. **RDS-GW**.
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.112.102**.
5. Set the *Virtual Service Ports* field to **443,3391**.
6. Set the *Protocol* to **TCP/UDP**.
7. Set the *forwarding Method* to **SNAT**.
8. Click **Update**.
9. Now click **Modify** next to the newly created Virtual Service.
10. Ensure that the *Persistent* check-box is enabled (checked).
11. Leave the *Persistent Timeout* set to **300** (i.e. 5 minutes).
12. Leave the *Health Checks Check Type* is set to **Connect to port**.
13. Set the *Check Port* to **443**.
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:



3. Enter an appropriate name (Label) for the first RD Gateway, e.g. **GW1**.
4. Change the ***Real Server IP Address*** field to the required IP address, e.g. **192.168.112.182**.
5. Leave the ***Real Server Port*** field blank.
6. Click **Update**.
7. Now repeat for your remaining RD Gateway(s).

10.3. Testing & Verification

> **Note** For additional guidance on diagnosing and resolving any issues you may have, please also refer to Diagnostics & Troubleshooting.

Configure DNS so the FQDN to be used for RD Gateway (e.g. **rdgateway.lbtestdom.com**) resolves to the VIP address. Also ensure that the Gateways can resolve the FQDN for the load balanced Connection Brokers (e.g. **rds.lbtestdom.com**).

Use Web Access / RemoteAPP to verify that published applications are available via the load balancer / Gateways.

11. Load Balancing Standalone Session Hosts (Scenario 4)

Scenario 4 is ***NOT*** part of the Standard Deployment illustrated in the Standard Deployment Diagram. It offers a simple alternative to a full RDS deployment utilizing just Session Hosts and the load balancer. Please refer to Scenario 4 - Load Balancing Stand alone Session Hosts for detailed notes on how the load balancer interacts with RDS in this scenario.

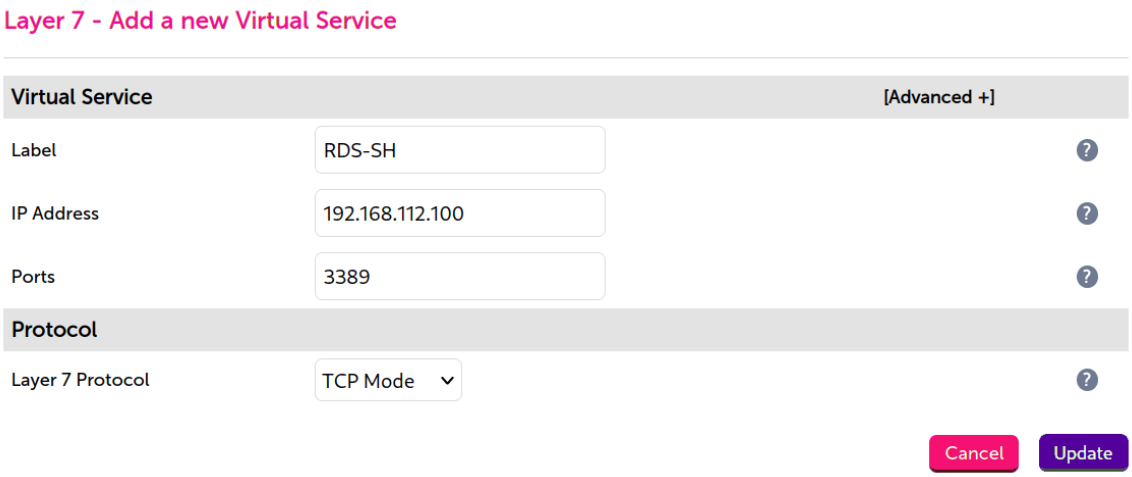
11.1. RDS Installation & Configuration

- Use the ***Role-based or feature-based*** installation type to install the Session Host role service on multiple servers.
- For Windows 2012 / 2016 It will not be possible to use Server Manager and/or most of the RDS Powershell commands to manage RDS. You will need to use group policy settings, WMI & registry edits.

11.2. Appliance Configuration

11.2.1. 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. **RDS-SH**.
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.112.100**.
5. Set the *Virtual Service Ports* field to **3389**.
6. Set the *Layer 7 Protocol* to **TCP Mode**.
7. Click **Update**.
8. Now click **Modify** next to the newly created Virtual Service.
9. Set *Persistence Mode* to either **Source IP** or **RDP Client Cookie** depending on your requirements.

> **Note** Please refer to Source IP Persistence for more details of these persistence methods.

10. Set *Persistence Timeout* to an appropriate value, e.g. **120** (i.e. 2 hours).
11. In the *Other* section click **Advanced** to expand the section.
12. Enable (check) the *Timeout* checkbox and set both *Client Timeout & Real Server Timeout* to **2h**.

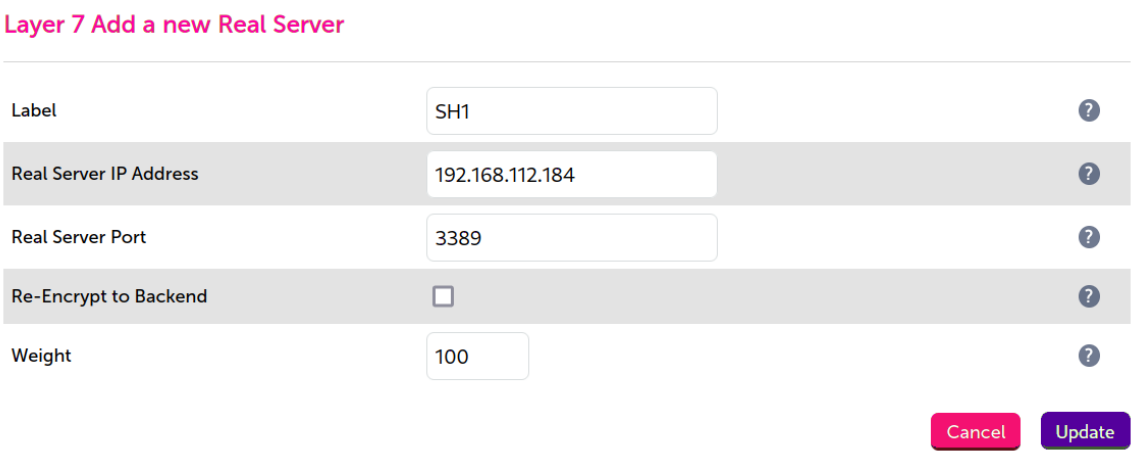
> **Note** If persistence is set to *RDP Client Cookie*, and the timeout values are left blank, they will be automatically set to 12h. Also, for this persistence mode, TCP Keep-alive is automatically enabled.

> **Note** The *Persistence Timeout*, *Client Timeout* and *Real Server Timeout* should be set to the same value as the idle session timeout on your Session Hosts.

13. Click **Update**.

11.2.2. 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 RDS server, e.g. **SH1**.
4. Change the ***Real Server IP Address*** field to the required IP address, e.g. **192.168.112.184**.
5. Set the ***Real Server Port*** field to **3389**.
6. Click **Update**.
7. Now repeat for your remaining Session Host server(s).

11.2.3. Applying the new Layer 7 Settings

1. Once the configuration is complete, use the **Reload HAProxy** button at the top of the screen to commit the changes.

11.3. Testing & Verification

> **Note** For additional guidance on diagnosing and resolving any issues you may have, please also refer to Diagnostics & Troubleshooting.

Configure DNS so that the FQDN to be used for your Session Hosts resolves to the VIP address. The load-balanced Session Hosts should now be accessible via the load balancer.

Connect to this address from the Microsoft RDP client (mstsc.exe) or equivalent.

12. Load Balancing Session Hosts Deployed with Connection Broker (Scenario 5)

Scenario 5 is ***NOT*** part of the Standard Deployment illustrated in the Standard Deployment Diagram. Here, the Session Hosts are load balanced by the load balancer appliance rather than the built-in mechanism of RDS. Please refer to Scenario 5 - Load Balancing Session Hosts when Deployed with Connection Broker for detailed notes on how the load balancer interacts with RDS in this scenario.

12.1. RDS Installation & Configuration

- Use the ***Remote Desktop Services*** installation type to perform a Standard deployment with 1 Connection

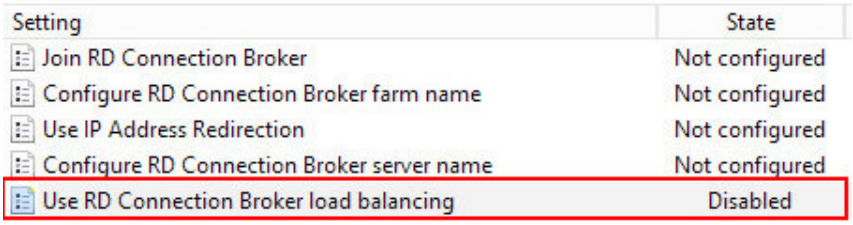
Broker, 1 Web Access Server and the required number of Session Hosts / Virtualization Hosts.

> **Note**
>
> If your deployment has multiple Session Collections, Web Access will not work correctly as mentioned in the notes for Scenario 5 in Scenario 5 - Load Balancing Session Hosts when Deployed with Connection Broker. However, in these cases it's still useful to install Web Access because it's a useful source for properly-configured .RDP files based on the current configuration of the deployment. This comes in handy when want to manually distribute .RDP files to clients.

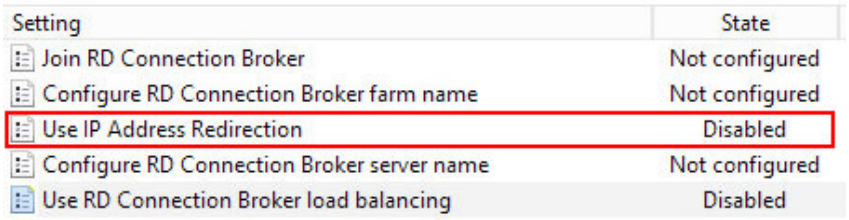
- Configure RDS Certificates as mentioned in RDS Certificates.
- On all Session Hosts, disable the built in load balancing mechanism:

  Using either a Group Policy Object that applies to all Session Hosts or by configuring each server individually using local group policy, disable ***Use RD Connection Broker load balancing***. This settings can be accessed here:

  Computer Configuration | Administrative Templates | Windows Components | Remote Desktop Services | Remote Desktop Session Host | RD Connection Broker

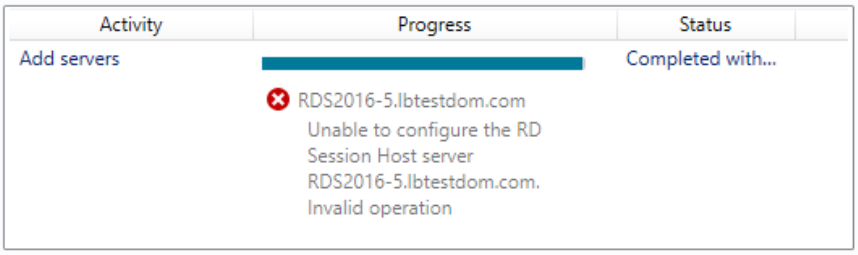


- If you want to use ***Routing Token Redirection Mode***, you'll also need to disable IP Address Redirection from the same Group Policy section as shown below:



> **Note**
>
> Make sure that the Session Hosts are already added to the relevant Session Collection before configuring these settings. If Session Hosts are added to collections afterwards, you may receive the following error:

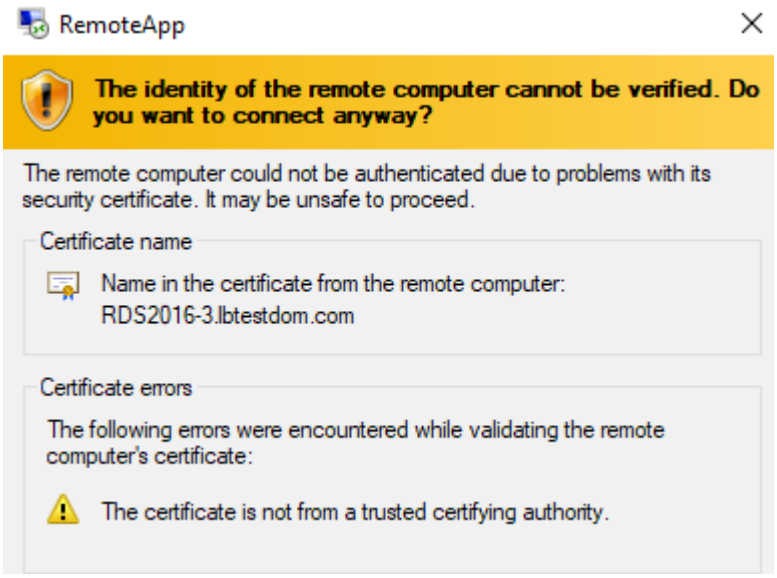
Unable to configure the RD Session Host sever. Invalid operation.



> **Note** Please refer to Configuring Win 2008 R2 for Routing Token Redirection Mode for configuring Windows 2008 R2 for ***Routing Token Redirection Mode***.

- Each Session Host has a self-signed RDS certificate. As mentioned here, when the Initial connection is handled by the Connection Broker (the default for Windows 2012 & later), the client will authenticate the Connection Broker using a certificate (and/or Kerberos), and then the Broker will authenticate the target Session Host on behalf on the client.

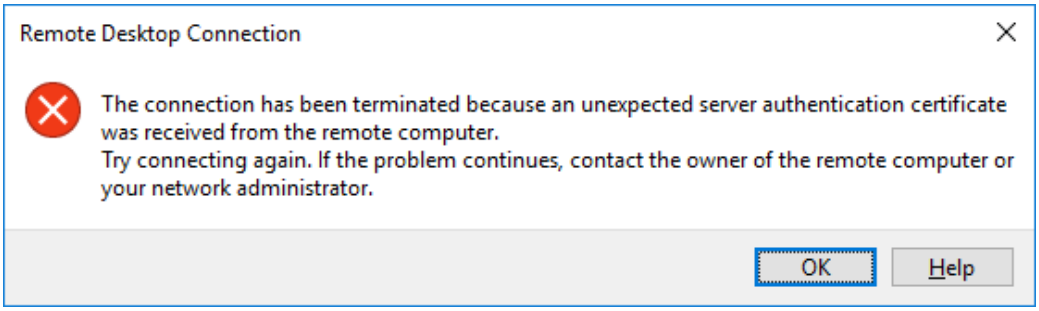
  When the Initial connection is handled by the Session Hosts the client may receive the following certificate warning:



  Under certain circumstances clients may also receive this error:

  The connection has been terminated because an unexpected server authentication certificate was received from the remote computer.

  Try connecting again. If the problem continues, contact the owner of the remote computer or your network administrator.



To prevent these warnings and errors, the self-signed certificate on the Session Hosts must be replaced with a trusted certificate signed by your CA.

To do this, perform the following steps:

1. Ensure that your RDS deployment certificate includes a SAN for the Session Hosts. The easiest way to achieve this is to add a wild card SAN such as ***.lbtestdom.com** (see RDS Certificates for more information on certificate requirements).
2. Import the certificate into the Local Machine Personal Certificate Store on each Session Host.
3. Run the following command from a PowerShell prompt on each Session Host:

```
wmic /namespace:\\root\cimv2\TerminalServices PATH Win32_TSGeneralSetting Set
SSLCertificateSHA1Hash="THUMBPRINT"
```

   - Enter this as a single command.
   - Replace THUMBPRINT with the thumbprint from your certificate, make sure you remove the spaces from the thumbprint and leave the double quotes in the command.
4. Restart the Remote Desktop Services service.

12.1.1. To remove this certificate and revert to the default self-signed RDS certificate

- Run Regedit
- Navigate to: HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Terminal Server\WinStations\RDP-Tcp
- Delete the **SSLCertificateSHA1Hash** registry value.
- Restart the Remote Desktop Services service.

> **Note**
>
> The settings **Select Active Connection Broker** in Windows 2012 and the equivalent setting **Select RD Management Server** in Windows 2016 have no effect on which Connection Broker is used by each Session Host, so there is no need to configure this setting in relation to load balancing. These settings are used to set which Connection Broker is able to accept configuration changes made in either the Server Manager Console or via Powershell.

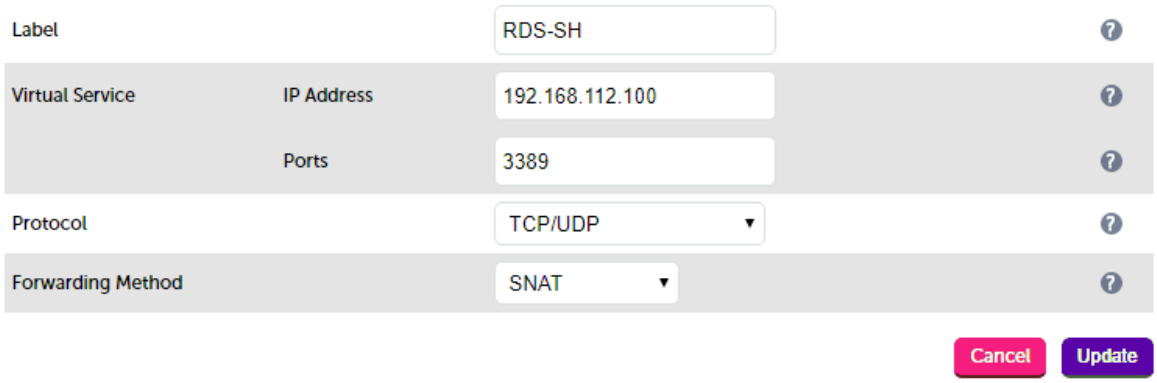
12.2. Appliance Configuration

- If you require UDP transport for RDP you'll have to use a layer 4 VIP that supports both TCP and UDP. In this case it will not be possible to use ***Routing Token Redirection Mode*** where all connections (new and redirected) pass via the load balancer. For configuration steps, please refer to the section: ***Using Layer 4 SNAT Mode (Required for UDP Transport)*** below.
- If you require all connections (new and redirected) to pass via the load balancer, you must use ***Routing Token Redirection Mode*** which requires a layer 7 VIP. In this case it will not be possible to use UDP transport for RDP. For configuration steps, please refer to the section: Using Layer 7 SNAT Mode (Required for Token Redirection Mode).
- If you have multiple Session Collections you'll need to configure multiple VIPs as explained in the "scenario notes" of Scenario 5 - Load Balancing Session Hosts when Deployed with Connection Broker.

12.2.1. Using Layer 4 SNAT Mode (Required for UDP Transport)

Setting up the Virtual Service (VIP)

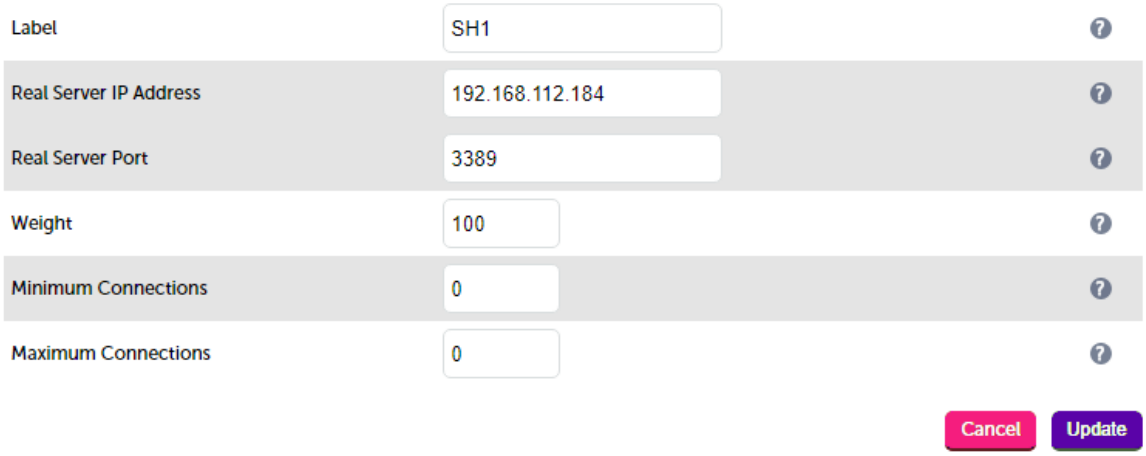
1. Using the WebUI, navigate to: ***Cluster Configuration > Layer 4 – Virtual Services*** and click *Add a New Virtual Service *.
2. Enter the following details:



3. Enter an appropriate name (Label) for the Virtual Service, e.g. **RDS-SH**.
4. Set the ***Virtual Service IP address*** field to the required IP address, e.g. **192.168.112.100**.
5. Set the ***Virtual Service Ports*** field to **3389**.
6. Set the ***Protocol*** to **TCP/UDP**.
7. 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:

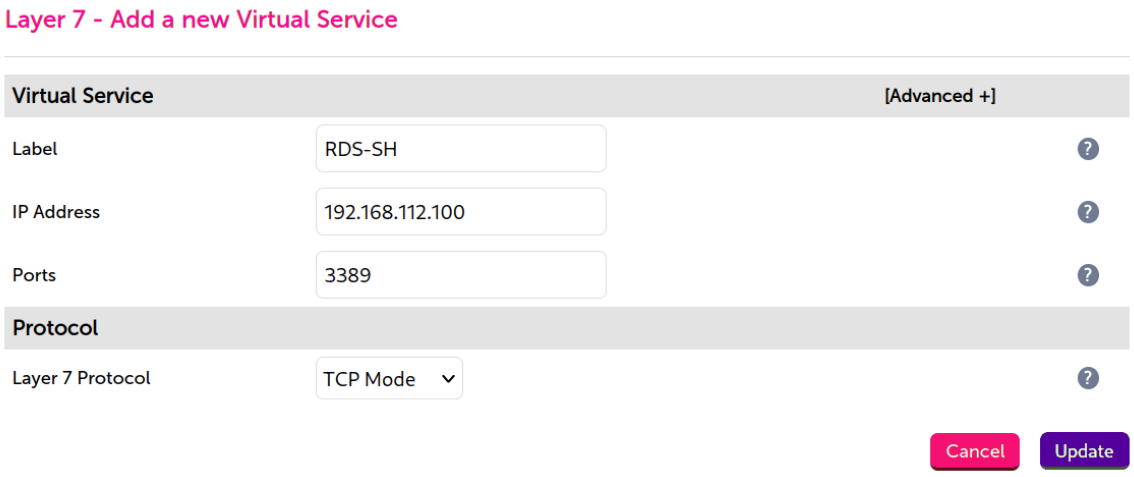


3. Enter an appropriate name (Label) for the first RDS server, e.g. **SH1**.
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.112.184**.
5. Set the *Real Server Port* field to **3389**.
6. Click **Update**.
7. Now repeat for your remaining Session Host server(s).

12.2.2. Using Layer 7 SNAT Mode (Required for Token Redirection 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. **RDS-SH**.
4. Set the *Virtual Service IP address* field to the required IP address, e.g. **192.168.112.100**.
5. Set the *Virtual Service Ports* field to **3389**.
6. Set the *Layer 7 Protocol* to **TCP Mode**.

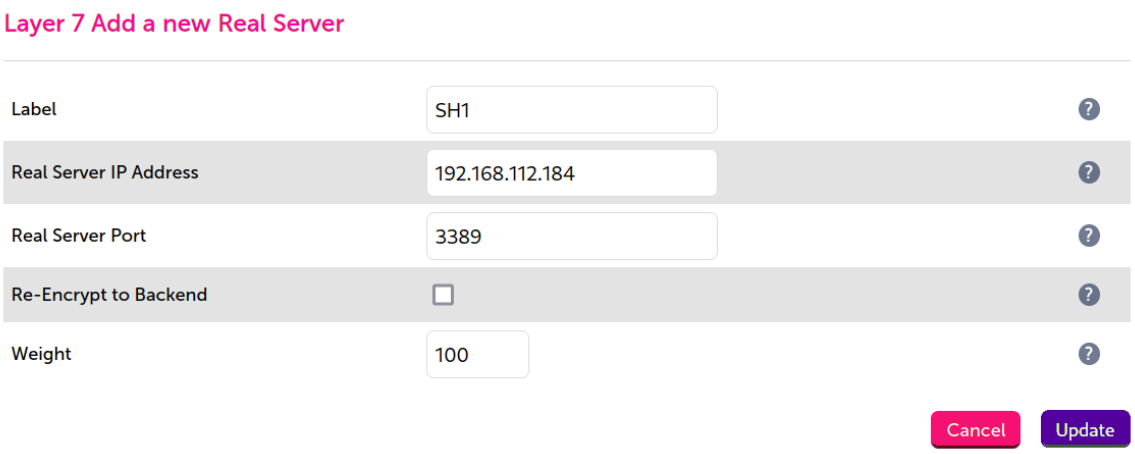
7. Click **Update**.
8. Now click **Modify** next to the newly created Virtual Service.
9. Set the *Persistence Mode* to **MS Session Broker**.

> **Note** When the Persistence Mode is set to **MS Session Broker**, TCP Keep-alive is automatically enabled.

10. In the *Other* section click **Advanced** to expand the section.
11. Enable (check) the *Timeout* checkbox and set both *Client* and *Real Server Timeout* to **30m** (i.e. 30 minutes).
12. 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 RDS server, e.g. **SH1**.
4. Change the *Real Server IP Address* field to the required IP address, e.g. **192.168.112.184**.
5. Set the *Real Server Port* field to **3389**.
6. Click **Update**.
7. Now repeat for your remaining Session Host server(s).

Applying the New Layer 7 Settings

1. Once the configuration is complete, use the **Reload HAProxy** button at the top of the screen to commit the changes.

12.3. Testing & Verification

> **Note** For additional guidance on diagnosing and resolving any issues you may have, please also refer to Diagnostics & Troubleshooting.

DNS must be configured so that the FQDN (e.g. **rds.lbtestdom.com**) specified in ***DNS Name for the RD Connection Broker Cluster*** (**Deployment Properties > High Availability**) resolves to the Session Host VIP. The load balanced Session Hosts should now be accessible via the load balancer.

Connect to this address from Web Access / RemoteAPP if your RDS deployment has a single Session Collection, or via modified .RDP files if you only have a single Connection Broker or there are multiple Session Collections.

13. 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.

14. Further Documentation

For additional information, please refer to the Administration Manual.

15. Appendix

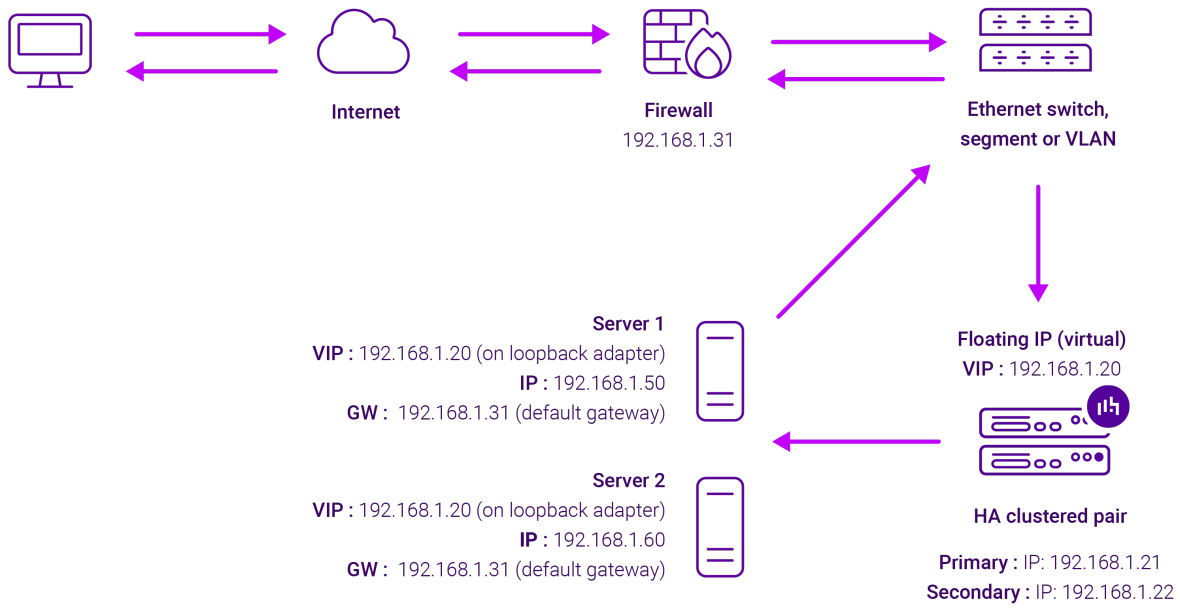
15.1. Load Balancer Deployment Modes

The load balancer can be deployed in one of 4 fundamental ways; *Layer 4 DR mode*, *Layer 4 NAT mode*, *Layer 4 SNAT mode* or *Layer 7 SNAT mode*. These are described below.

15.1.1. Layer 4 DR Mode

Layer 4 DR (Direct Routing) mode is a very high performance solution that requires little change to your existing infrastructure. The image below shows an example network diagram for this mode.

> **Note** Kemp, Brocade, Barracuda & A10 Networks call this *Direct Server Return* and F5 call it *nPath*.

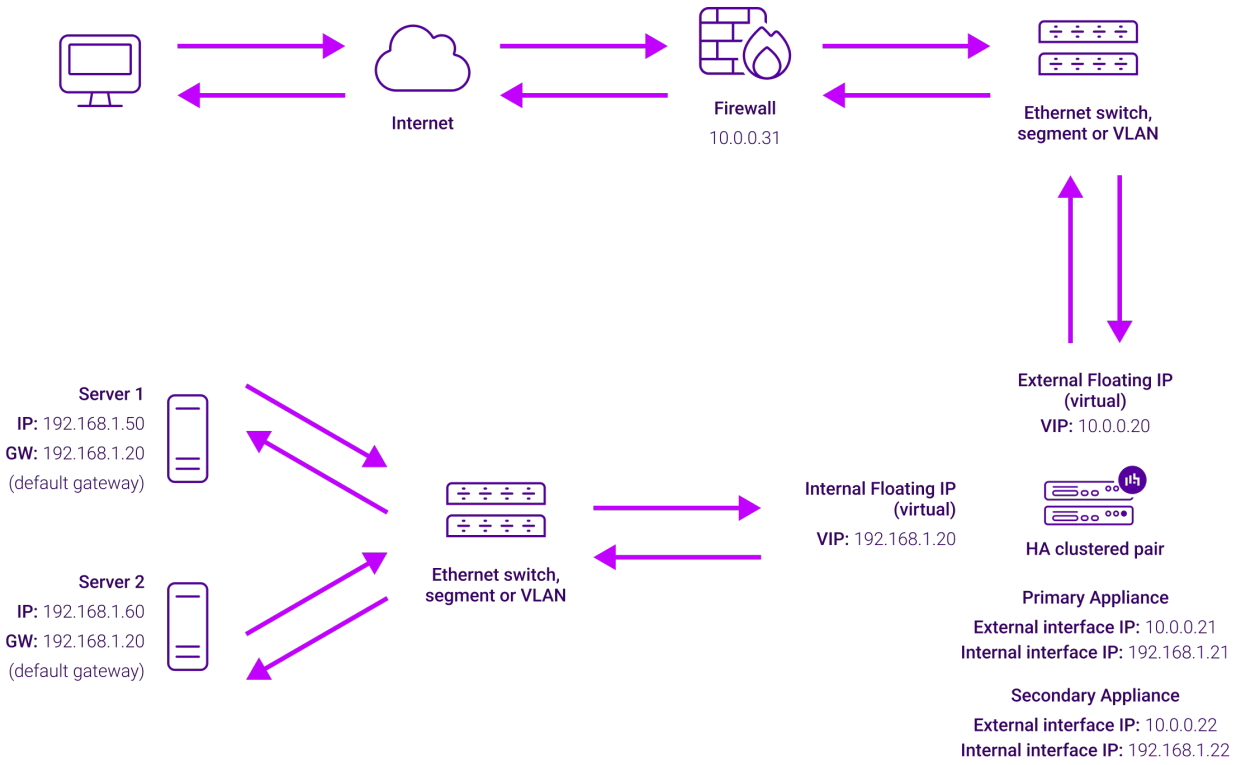


- DR mode works by changing the destination MAC address of the incoming packet to match the selected Real Server on the fly which is very fast.
- When the packet reaches the Real Server it expects the Real Server to own the Virtual Services IP address (VIP). This means that each Real Server (and the load balanced application) must respond to both the Real Server's own IP address and the VIP.
- The Real Server should not respond to ARP requests for the VIP. Only the load balancer should do this. Configuring the Real Server in this way is referred to as "Solving the ARP Problem". For more information please refer to DR Mode Considerations.
- On average, DR mode is 8 times quicker than NAT mode for HTTP and much faster for other applications such as Remote Desktop Services, streaming media and FTP.
- The load balancer must have an interface in the same subnet as the Real Servers to ensure layer 2 connectivity which is required for DR mode to operate.
- The VIP can be brought up on the same subnet as the Real Servers or on a different subnet provided that the load balancer has an interface in that subnet.

- Port translation is not possible with DR mode, e.g. VIP:80 → RIP:8080 is not supported.
- DR mode is transparent, i.e. the Real Server will see the source IP address of the client.

15.1.2. Layer 4 NAT Mode

Layer 4 NAT mode is a high performance solution, although not as fast as layer 4 DR mode. This is because real server responses must flow back to the client via the load balancer rather than directly as with DR mode. The image below shows an example network diagram for this mode.



- The load balancer translates all requests from the Virtual Service to the Real Servers.
- NAT mode can be deployed in the following ways:
  - **Two-arm (using 2 Interfaces)** (as shown above) - Here, 2 subnets are used. The VIP is located in one subnet and the load balanced Real Servers are located in the other. The load balancer requires 2 interfaces, one in each subnet.

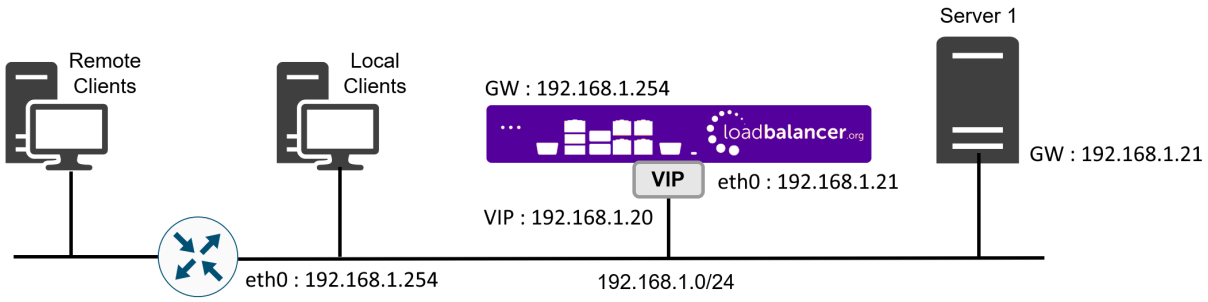
> **Note** This can be achieved by using two network adapters, or by creating VLANs on a single adapter.

    - Normally **eth0** is used for the internal network and **eth1** is used for the external network, although this is not mandatory since any interface can be used for any purpose.
    - If the Real Servers require Internet access, *Auto-NAT* should be enabled using the WebUI menu option: *Cluster Configuration > Layer 4 - Advanced Configuration*, the external interface should be selected.
    - The default gateway on the Real Servers must be set to be an IP address on the load balancer.

> **Note** For an HA clustered pair, a floating IP should be added to the load balancer and used as the Real Server's default gateway. This ensures that the IP address can "float" (move) between Primary and Secondary appliances.

- Clients can be located in the same subnet as the VIP or any remote subnet provided they can route to the VIP.

- **One-arm (using 1 Interface)** - Here, the VIP is brought up in the same subnet as the Real Servers.



- To support remote clients, the default gateway on the Real Servers must be an IP address on the load balancer and routing on the load balancer must be configured so that return traffic is routed back via the router.

> **Note** For an HA clustered pair, a floating IP should be added to the load balancer and used as the Real Server's default gateway. This ensures that the IP address can "float" (move) between Primary and Secondary appliances.

- To support local clients, return traffic would normally be sent directly to the client bypassing the load balancer which would break NAT mode. To address this, the routing table on the Real Servers must be modified to force return traffic to go via the load balancer. For more information please refer to One-Arm (Single Subnet) NAT Mode.

- If you want Real Servers to be accessible on their own IP address for non-load balanced services, e.g. RDP, you will need to setup individual SNAT and DNAT firewall script rules for each Real Server or add additional VIPs for this.
- Port translation is possible with Layer 4 NAT mode, e.g. VIP:80 → RIP:8080 is supported.
- NAT mode is transparent, i.e. the Real Server will see the source IP address of the client.

NAT Mode Packet re-Writing

In NAT mode, the inbound destination IP address is changed by the load balancer from the Virtual Service IP address (VIP) to the Real Server. For outbound replies the load balancer changes the source IP address of the Real Server to be the Virtual Services IP address.

The following table shows an example NAT mode setup:

Protocol	VIP	Port	RIP	Port
TCP	10.0.0.20	80	192.168.1.50	80

In this simple example all traffic destined for IP address 10.0.0.20 on port 80 is load-balanced to the real IP address 192.168.1.50 on port 80.

Packet rewriting works as follows:

1) The incoming packet for the web server has source and destination addresses as:

Source	x.x.x.x:34567	**Destination**	10.0.0.20:80

2) The packet is rewritten and forwarded to the backend server as:

Source	x.x.x.x:34567	**Destination**	192.168.1.50:80

3) Replies return to the load balancer as:

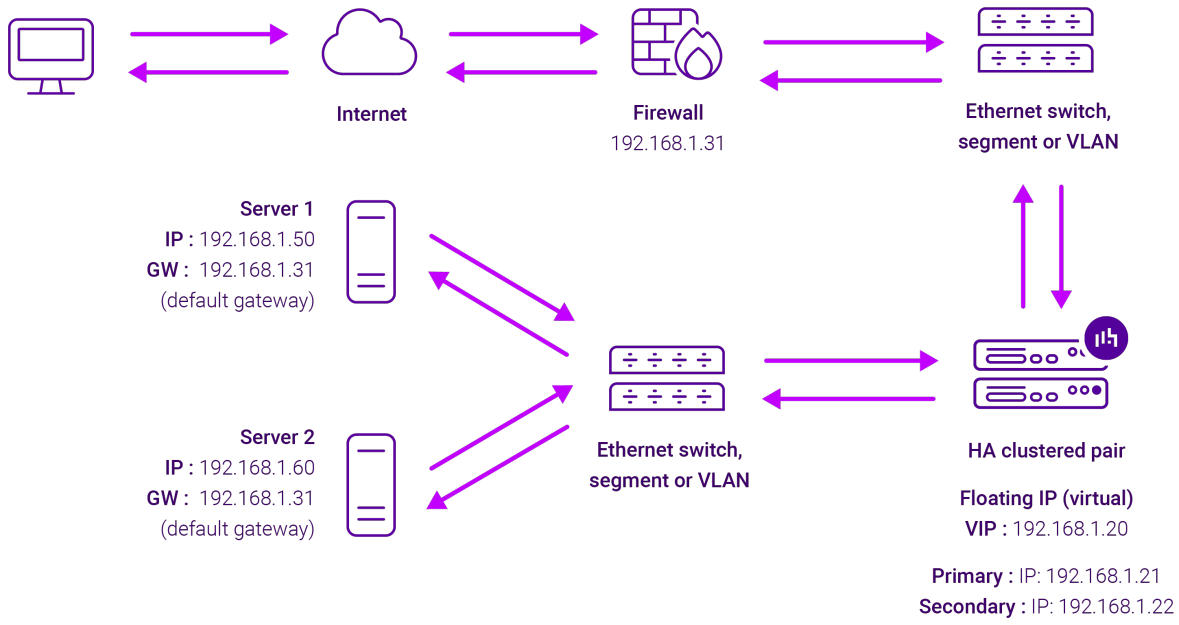
Source	192.168.1.50:80	**Destination**	x.x.x.x:34567

4) The packet is written back to the VIP address and returned to the client as:

Source	10.0.0.20:80	**Destination**	x.x.x.x:34567

15.1.3. Layer 4 SNAT Mode

Layer 4 SNAT mode is a high performance solution, although not as fast as Layer 4 NAT mode or Layer 4 DR mode. The image below shows an example network diagram for this mode.

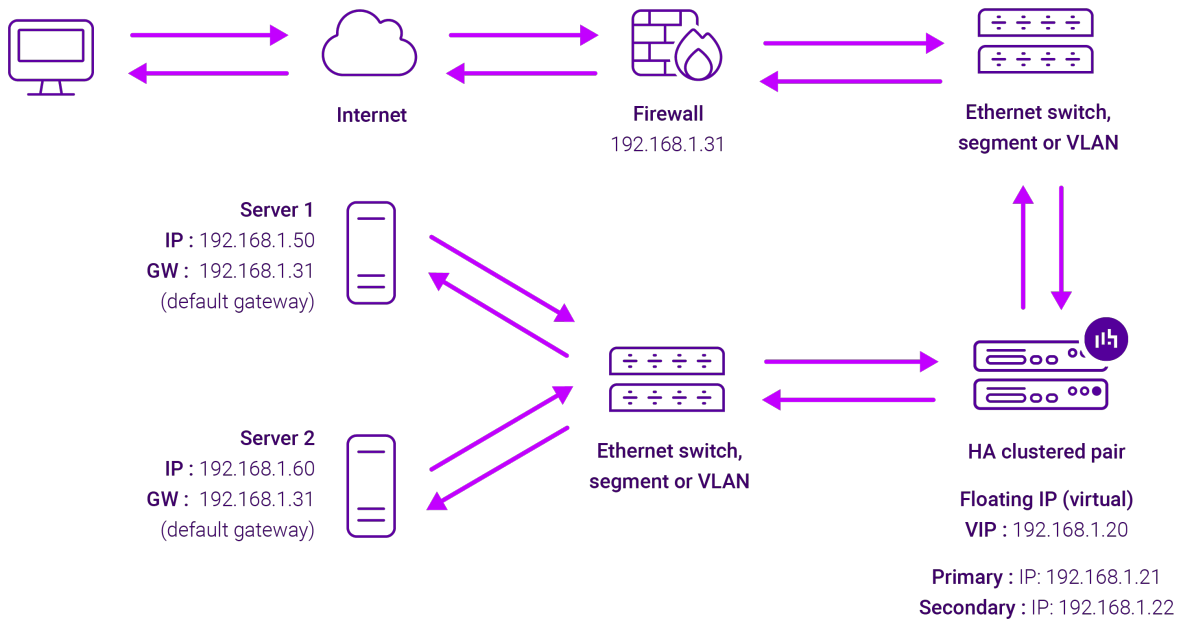


- Real Servers in the cluster can be on any accessible network including across the Internet or WAN.
- Layer 4 SNAT mode is not transparent, an iptables SNAT rule translates the source IP address to be the load balancer rather than the original client IP address.

- Layer 4 SNAT mode can be deployed using either a one-arm or two-arm configuration. For two-arm deployments, **eth1** is typically used for client side connections and **eth0** is used for Real Server connections, although this is not mandatory since any interface can be used for any purpose.
- Requires no mode-specific configuration changes to the load balanced Real Servers.
- Port translation is possible with Layer 4 SNAT mode, e.g. VIP:80 → RIP:8080 is supported.
- You should not use the same RIP:PORT combination for layer 4 SNAT mode VIPs and layer 7 SNAT mode VIPs because the required firewall rules conflict.

15.1.4. Layer 7 SNAT Mode

Layer 7 SNAT mode uses a proxy (HAProxy) at the application layer. Inbound requests are terminated on the load balancer and HAProxy generates a new corresponding request to the chosen Real Server. As a result, Layer 7 is typically not as fast as the Layer 4 methods. Layer 7 is typically chosen when either enhanced options such as SSL termination, cookie based persistence, URL rewriting, header insertion/deletion etc. are required, or when the network topology prohibits the use of the layer 4 methods. The image below shows an example network diagram for this mode.



- Because layer 7 SNAT mode is a full proxy, Real Servers in the cluster can be on any accessible network including across the Internet or WAN.
- Layer 7 SNAT mode is not transparent by default, i.e. the Real Servers will not see the source IP address of the client, they will see the load balancer's own IP address by default, or any other local appliance IP address if preferred (e.g. the VIP address). This can be configured per layer 7 VIP. If required, the load balancer can be configured to provide the actual client IP address to the Real Servers in 2 ways. Either by inserting a header that contains the client's source IP address, or by modifying the Source Address field of the IP packets and replacing the IP address of the load balancer with the IP address of the client. For more information on these methods please refer to Transparency at Layer 7.
- Layer 7 SNAT mode can be deployed using either a one-arm or two-arm configuration. For two-arm deployments, **eth1** is typically used for client side connections and **eth0** is used for Real Server connections,

although this is not mandatory since any interface can be used for any purpose.

- Requires no mode-specific configuration changes to the load balanced Real Servers.
- Port translation is possible with Layer 7 SNAT mode, e.g. VIP:80 → RIP:8080 is supported.
- You should not use the same RIP:PORT combination for layer 7 SNAT mode VIPs and layer 4 SNAT mode VIPs because the required firewall rules conflict.

15.2. Server Feedback Agent

The load balancer can dynamically modify the weight of each Real Server by gathering data from either a custom feedback agent or a HTTP server. Reducing the weight of a server compared to others in the pool will reduce the amount of traffic it receives.

For layer 4 VIPs, both the agent and HTTP Server methods can be used, for Layer 7 VIPs, only the agent method is supported.

By default, the agent listens on TCP port 3333, although this can be changed if required.

A telnet to port 3333 on a Real Server with the agent installed returns the current idle value as an integer value between 0 and 100. By default, the idle value is based on current CPU utilization. This can also be based on RAM utilization and the number of current connections or a combination of all three.

This can be configured by modifying the XML configuration file located in the agent's installation folder - by default C:\ProgramData\LoadBalancer.org\LoadBalancer. The file can be edited directly or by clicking the **Configuration** button in the agent monitor program - see "Controlling the Agent" below.

The load balancer uses the formula **(idle value/100) * initial weight** to find the new dynamic weight.

> **Note**
>
> The "initial weight" is the weight set in the WebUI for each Real Server.
>
> For more information about the feedback agent, please refer to this blog.

15.2.1. Windows Agent

The latest Windows feedback agent (v4.6.0) can be downloaded here.

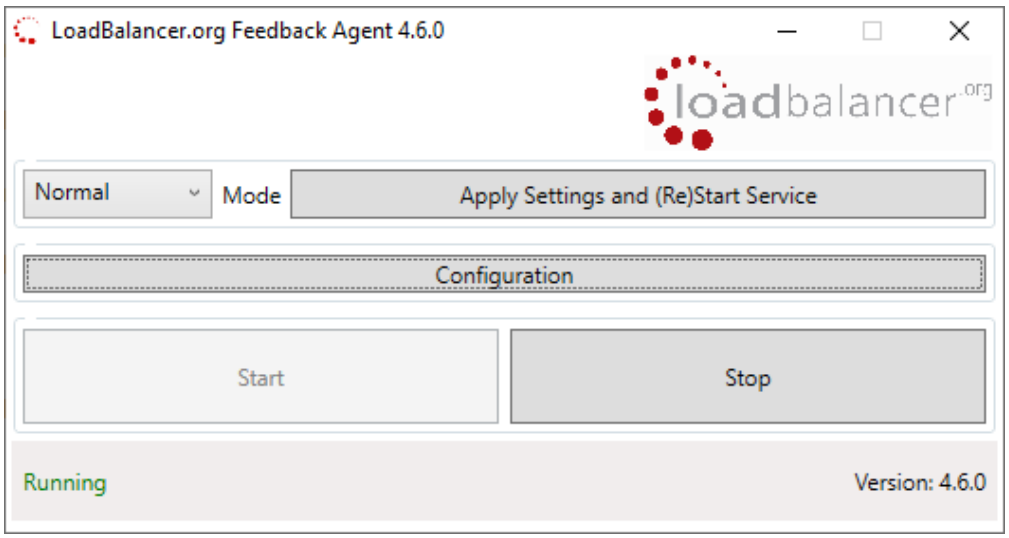
Installing the Agent

To install the agent, run **loadbalanceragent.msi**. Once the installation is complete, the Feedback Agent service is started automatically.

> **Note** The agent must be installed on each Real Server.

Controlling the Agent

The Feedback Agent service can be controlled and configured using the *Loadbalancer.org Feedback Agent* monitor program. By default this can be accessed from: *Windows Start Menu > Loadbalancer.org*.



15.2.2. Linux/Unix Agent

The Linux feedback agent files can be downloaded using the following links:

readme file: https://downloads.loadbalancer.org/agent/linux/v4.1/readme.txt
xinetd file: https://downloads.loadbalancer.org/agent/linux/v4.1/lb-feedback
feedback script: https://downloads.loadbalancer.org/agent/linux/v4.1/lb-feedback.sh

Installation & Testing

Install xinetd - if not already installed:

```
apt-get install xinetd
```

Insert the following line into /etc/services:

```
lb-feedback 3333/tcp # Loadbalancer.org feedback daemon
```

Then run the following commands:

```
cp lb-feedback.sh /usr/bin/lb-feedback.sh
chmod +x /usr/bin/lb-feedback.sh
cp lb-feedback /etc/xinetd.d/lb-feedback
chmod 644 /etc/xinetd.d/lb-feedback
/etc/init.d/xinetd restart
```

To test the agent:

```
telnet 127.0.0.1 3333
Trying 127.0.0.1...
Connected to 127.0.0.1.
Escape character is '^]'.
95%
```

```
Connection closed by foreign host.
```

> **Note** The agent files must be installed on each Real Server.

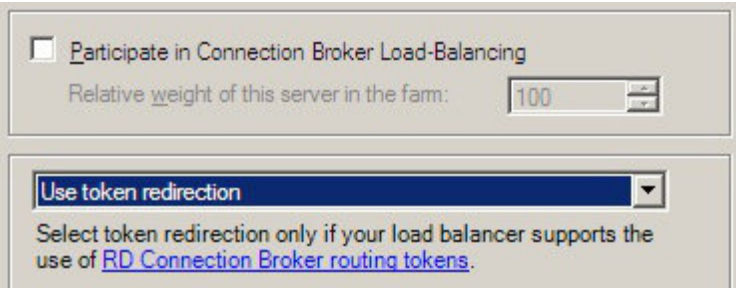
15.2.3. HTTP Server

You can use any HTTP server responding on port 3333 to give feedback information to the load balancer. The format of this information must be an integer number of 0-100 without any header information. Using this method, you can generate a custom response based on your application's requirements.

15.3. Configuring Win 2008 R2 for Routing Token Redirection Mode

Install Connection Broker on the server designated to hold the Connection Broker role. Then on each RDS to be included in the cluster/Farm:

1. Open Remote Desktop Host Session Configuration.
2. Right–click 'Member of farm in RD Connection Broker' and select Properties.
3. Click Change Settings.
4. Select Farm Member, enter the DNS name of the server running the Connection Broker role service and the name of the farm (all servers within the same farm require the same name to be specified) and click OK.
5. Leave **Participate in Connection Broker Load-Balancing** un-checked and select **Use token redirection** from the drop down as shown below:



15.4. Configuring HA - Adding a Secondary Appliance

Our recommended configuration is to use a clustered HA pair of load balancers to provide a highly available and resilient load balancing solution. We recommend that the Primary appliance is fully configured first, then the Secondary appliance can be added to create an HA pair. Once the HA pair is configured, load balanced services must be configured and modified on the Primary appliance. The Secondary appliance will be automatically kept in sync.

> **Note** For Enterprise Azure, the HA pair should be configured first. For more information, please refer to the Azure Quick Start/Configuration Guide available in the documentation library

The clustered HA pair uses Heartbeat to determine the state of the other appliance. Should the active device (normally the Primary) suffer a failure, the passive device (normally the Secondary) will take over.

15.4.1. Non-Replicated Settings

A number of settings are not replicated as part of the Primary/Secondary pairing process and therefore must be manually configured on the Secondary appliance. These are listed by WebUI menu option in the table below:

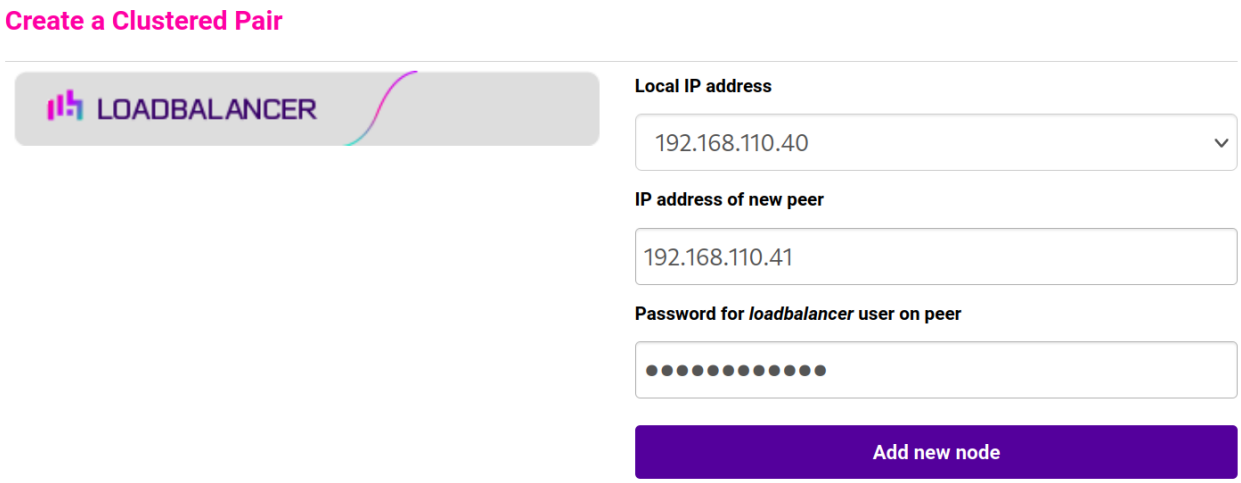
WebUI Main Menu Option	Sub Menu Option	Description
Local Configuration	Hostname & DNS	Hostname and DNS settings
Local Configuration	Network Interface Configuration	Interface IP addresses, bonding configuration and VLANs
Local Configuration	Routing	Default gateways and static routes
Local Configuration	System Date & time	Time and date related settings
Local Configuration	Physical – Advanced Configuration	Various appliance settings
Local Configuration	Portal Management	Portal management settings
Local Configuration	Security	Security settings
Local Configuration	SNMP Configuration	SNMP settings
Local Configuration	Graphing	Graphing settings
Local Configuration	License Key	Appliance licensing
Maintenance	Backup & Restore	Local XML backups
Maintenance	Software Updates	Appliance software updates
Maintenance	Fallback Page	Fallback page configuration
Maintenance	Firewall Script	Firewall (iptables) configuration
Maintenance	Firewall Lockdown Wizard	Appliance management lockdown settings

> **Important** Make sure that where any of the above have been configured on the Primary appliance, they're also configured on the Secondary.

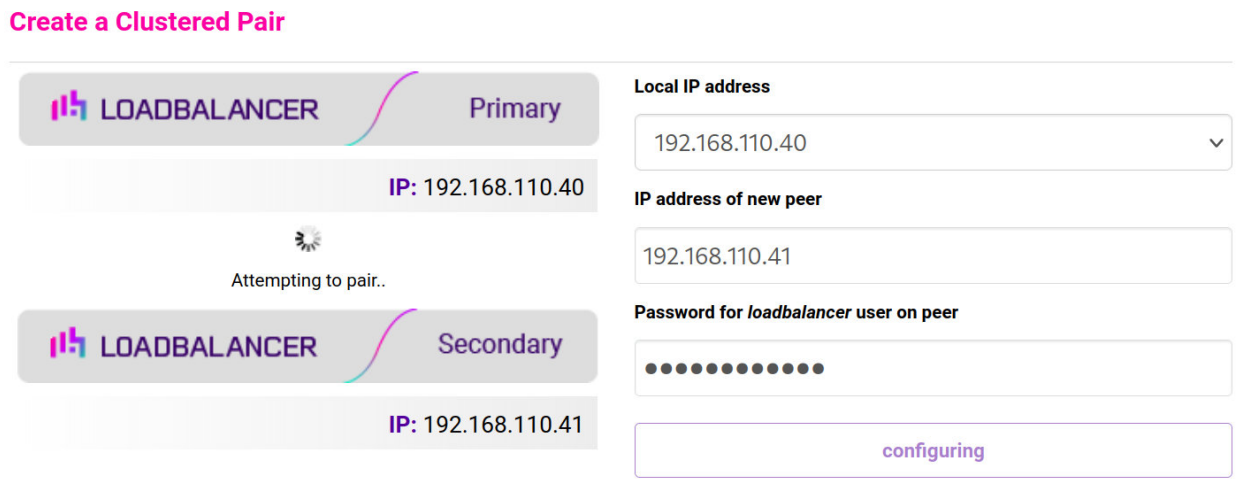
15.4.2. Configuring the HA Clustered Pair

> **Note** If you have already run the firewall lockdown wizard on either appliance, you'll need to ensure that it is temporarily disabled on both appliances whilst performing the pairing process.

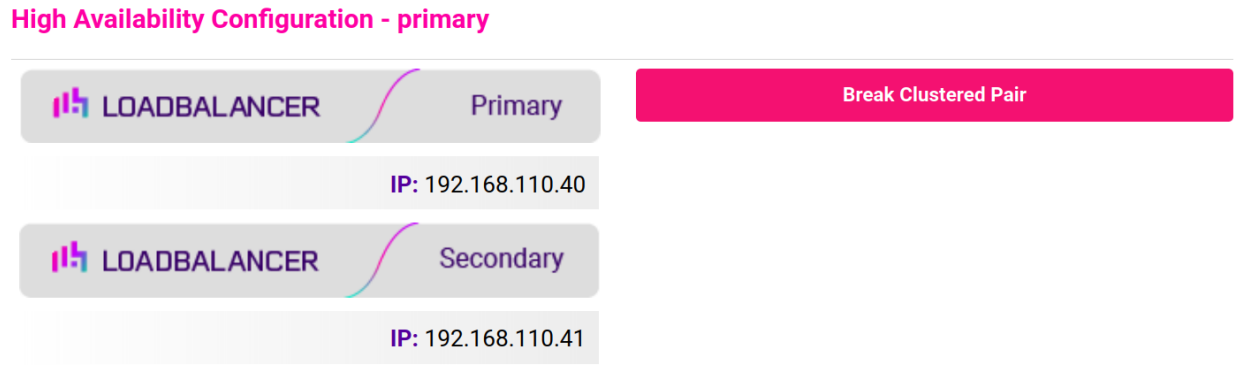
1. Deploy a second appliance that will be the Secondary and configure initial network settings.
2. Using the WebUI on the Primary appliance, navigate to: ***Cluster Configuration > High-Availability Configuration***.



3. Specify the IP address and the *loadbalancer* user's password for the Secondary (peer) appliance as shown in the example above.
4. Click **Add new node**.
5. The pairing process now commences as shown below:



6. Once complete, the following will be displayed on the Primary appliance:



7. To finalize the configuration, restart heartbeat and any other services as prompted in the "Commit changes" message box at the top of the screen.

> **Note** Clicking the **Restart Heartbeat** button on the Primary appliance will also automatically restart heartbeat on the Secondary appliance.

> **Note** For more details on configuring HA with 2 appliances, please refer to Appliance Clustering for HA.

> **Note** For details on testing and verifying HA, please refer to Clustered Pair Diagnostics.

16. Document Revision History

Version	Date	Change	Reason for Change	Changed By
2.1.0	9 August 2019	Styling and layout	General styling updates	RJC
2.1.1	10 June 2020	New title page Updated Canadian contact details Additional instructions for setting client and server timeout settings	Branding update Change to Canadian contact details Changes to the appliance WebUI	AH
2.2.0	1 November 2021	Converted the document to AsciiDoc	Move to new documentation system	AH, RJC, ZAC
2.2.1	20 April 2022	Removed dead link	Resource retired and removed by Microsoft	AH
2.2.2	28 September 2022	Updated layer 7 VIP and RIP creation screenshots	Reflect changes in the web user interface	AH
2.2.3	5 January 2023	Combined software version information into one section Added one level of section numbering Added software update instructions Added table of ports used by the appliance Reworded 'Further Documentation' section	Housekeeping across all documentation	AH
2.2.4	2 February 2023	Updated screenshots	Branding update	AH
2.2.5	7 March 2023	Removed conclusion section	Updates across all documentation	AH
2.3.0	24 March 2023	New document theme Modified diagram colours	Branding update	AH

LOADBALANCER

Visit us: www.loadbalancer.org

Phone us: +44 (0)330 380 1064

Phone us: +1 833 274 2566

Email us: info@loadbalancer.org

Follow us: @loadbalancer.org

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.