

Gateway Portal Load Balancing

Pre-requisites

We advise you to start by reading our [Load Balancing overview](#).

Generated Clients and Web Access

There are two ways to connect to a Load Balanced cluster:

- Using a **Generated Client** (executable program, created by the [Connection Client Generator](#))
- Using a **Gateway Portal Web Access** (web page in the browser, created by the [Webmaster Toolkit](#))

Please note that the Gateway Portal is only enabled in Terminal Service Plus Corporate edition.

This page describes how to use the Load Balancing Manager for use with a Gateway Portal.

[For help on the Generated Client Load Balancing specific settings, please click here](#)

Gateway and Application Server Roles

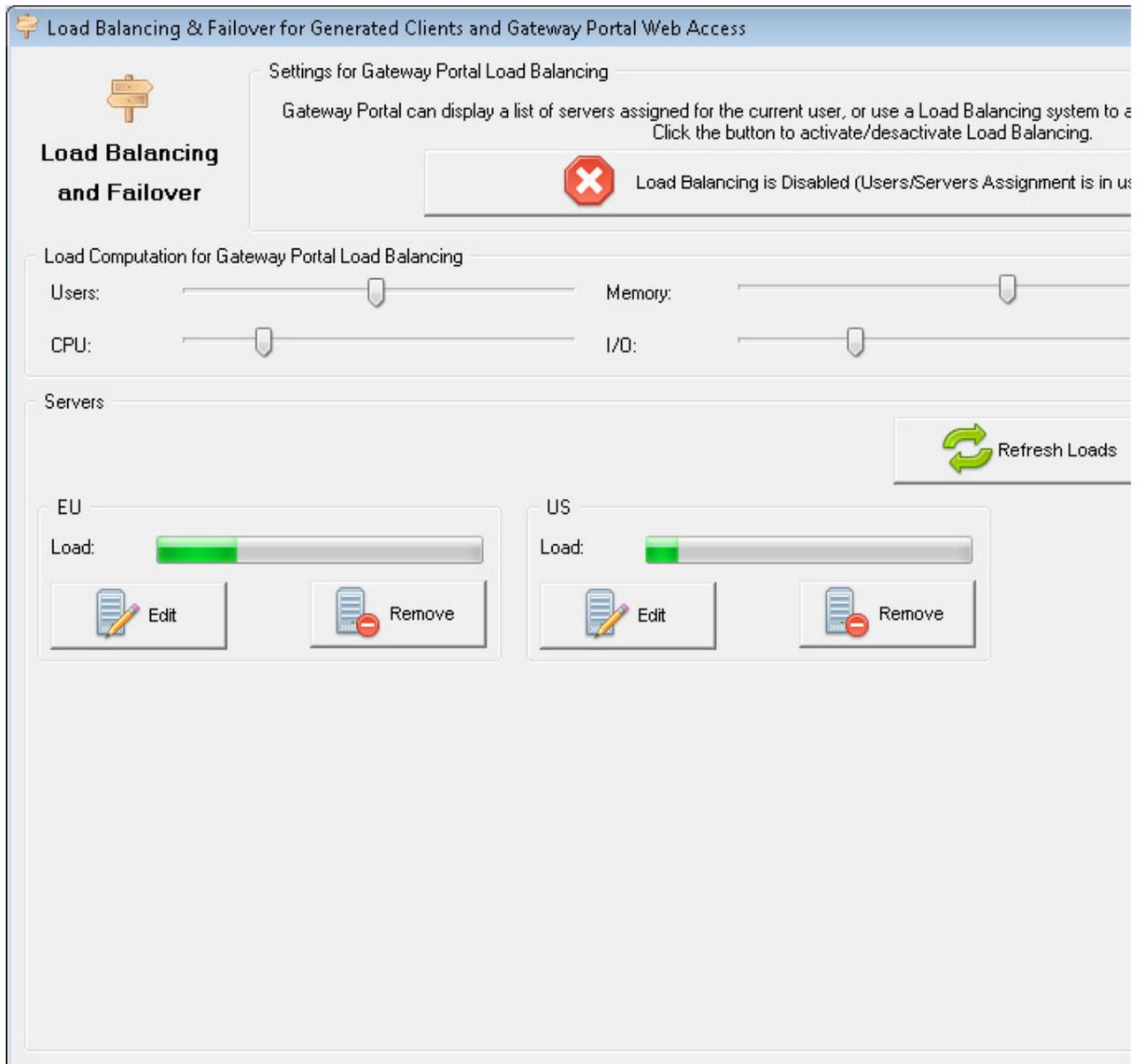
The Load Balancing Manager tool **must be used on the server having the Gateway role**. Using it on an other server of the load balanced cluster will have no effect on the load balancing.

Load Balancing Main Window

The main window display a lot of data and allows you to configure your Gateway Portal Load Balancing.

It lists all the servers in your Load Balanced cluster, and allows you to add, edit or remove them. The progress bars show the current load of each server. It also confirms you that they are well configured and can be contacted by the Gateway.

It also allows you to enable, disable and configure the Load Balancing (more on this below).



Enabling/Disabling Load Balancing

On top of the Gateway Portal Load Balancing window, you will see a big button:

- This button display the current state of the Load Balancing.
- If you click on it, it will enable or disable the Load Balancing depending on its current state.

Here is the button when the Load Balancing is **Disabled**. A click on it will activate it, and disable the [Users/Servers Assignment](#).



Here is the button when the Load Balancing is **Enabled**. A click on it will deactivate it, and enable the [Users/Servers Assignment](#).



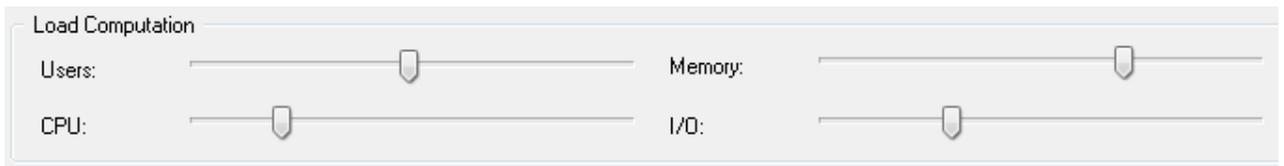
Configuring the Computation of Servers' Loads

When Load Balancing is enabled, each user will be sent to the less loaded server at the time of his connection.

To determine which is the less loaded server, the load of every server is computed using a weighted average between several performance indicators:

- number of connected users
- processor usage
- memory usage
- disk usage

You can modify the weight (importance) of these indicators by using the following sliders:



Using these sliders, you will be able to fine tune the Load Balancing, and to optimize it to your own needs. For instance if your users launch a business application with big memory requirements, it might be a good idea to increase the impact of the RAM in the load computation using the "Memory" slider.

You can also reset these values by clicking on the "Back to Default Settings" button.

How is Server's Load computed?

The load of each server is computed when needed in order to decide on which server the user must be sent.

For this computation, we use a **weighted average** between 4 hardware metrics.

The 4 sliders allow you to give more (or less) weight on each of these metrics, which are:

- Users : number of connected users
- CPU : percentage of non-idle processor time
- Memory : percentage of used memory
- I/O : percentage of non-idle disk time

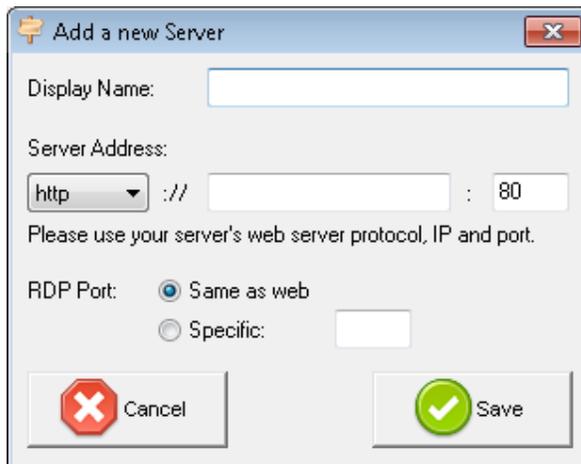
For example, if you put the "Memory" slider on its right side and all the other sliders on their left side, the load of each servers will be computed using almost only the percentage of used memory - and your users will be sent to the server having the most available memory.

- When a slider is completely on the left, the weight used in load computation will be 1.
- When a slider is completely on the right, the weight used in load computation will be 100.

Moreover, the load of a given server depends on the ratio [used ressources / total ressources] ; so if a server A is twice as powerful as a server B, 2 times more users should be sent to server A than to server B (all other things being equals).

Adding a new Server

To add a new server, simply click on the "Add a new Server" button. The following window will be displayed:



The screenshot shows a dialog box titled "Add a new Server". It contains the following fields and options:

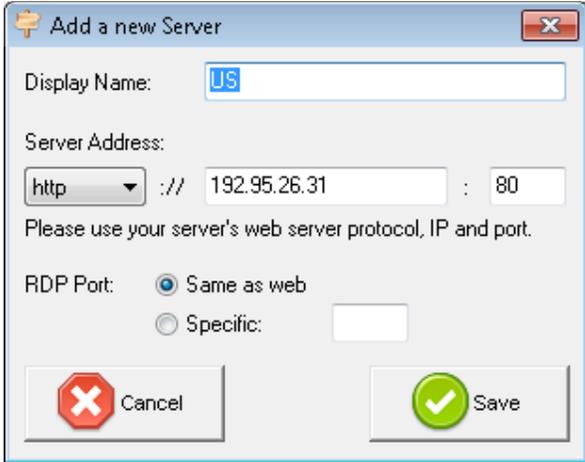
- Display Name:** A text input field.
- Server Address:** A dropdown menu showing "http", followed by "://", a text input field for the IP address, and a port field showing "80".
- Instructions:** "Please use your server's web server protocol, IP and port."
- RDP Port:** Two radio buttons: "Same as web" (which is selected) and "Specific:" followed by a text input field.
- Buttons:** "Cancel" (with a red X icon) and "Save" (with a green checkmark icon).

The "Display Name" is the title that will be shown to your users on the Web Access HTML page. It is supposed to be more user-friendly than a technical value (such as an IP), for instance "US Server" or "Blue Zone".

When using the Gateway Portal Load Balancing, the "RDP Port" will not be used. It is only used when connecting using a Generated Client. We advise you to keep the default setting ("Same as web").

Editing an existing Server

To modify an existing server, simply click on the "Edit" button of the server you want to update. The following window will be displayed:



The screenshot shows a dialog box titled "Add a new Server" with a close button in the top right corner. The dialog contains the following fields and options:

- Display Name:** A text input field containing "US".
- Server Address:** A dropdown menu set to "http", followed by a text input field containing "192.95.26.31" and a port input field containing "80".
- Instructions:** The text "Please use your server's web server protocol, IP and port." is displayed below the address fields.
- RDP Port:** Two radio button options: "Same as web" (which is selected) and "Specific:" followed by an empty text input field.
- Buttons:** A "Cancel" button with a red 'X' icon and a "Save" button with a green checkmark icon.