

Proto Balance – Enterprise Load Balancer

<http://www.protonet.co.za/>

User Manual

Proto Balance, Proto Balance Advanced - User Manual

March 28, 2010

Contents

1. Installation	3
2. QuickStart Proto Balance	3
2.1. All Unix systems, 3	◇
2.2. Microsoft Windows, 3	◇
2.3. Apple Mac OS X, 3	◇
2.4. Ensuring that Proto Balance restarts after your machine has rebooted, 4	◇
2.5. Proto Balance command-line options and flags, 4.	
3. Configuration	5
3.1. Logging in and setting your password, 5	◇
3.2. Creating a forwarder to divert TCP connections to another host, 5	◇
3.3. Explanation of traffic lights, 6	◇
3.4. Creating a cluster, 7	◇
3.5. Removing a box from a cluster, 8.	
4. Connection Options	8
5. Advanced Options	9
5.1. Box Retry, 9	◇
5.2. Load Balancing Algorithms, 9	◇
5.3. Connection rate limits, denial of service limits, 10	◇
5.4. Idle Timeout, 10	◇
5.5. X-Forwarded-For HTTP header, 10	◇
5.6. Box affinity, 11	◇
5.7. Client connection dictionary, 11	◇
5.8. Box fallover, 12	◇
5.9. Speed factor, 12	◇
5.10. Close connection if... , 13.	
6. Multiple Clusters	13
7. Protocols that Don't Load Balance	14
8. Traffic Graphs	15
9. Monthly client logs	15
10. Security	16
11. OS-specific Event Based I/O	16
12. Operating System Considerations and Tuning	17
12.1. Summary of limits, 17	◇
12.2. Maximum number of socket handles / file descriptors, 17	◇
12.3. Maximum number of threads (-pthread option), 18	◇
12.4. Shared memory and Semaphore limits (-shared shmat option), 18	◇
12.5. TCP port limits, 18	◇
12.6. Limits of Physical Memory, 19	◇
12.7. Other limits, 19.	
13. System Settings for High Performance with Your Operating System	19
13.1. Windows, 20	◇
13.2. Linux, 22	◇
13.3. MacOS, 23	◇
13.4. HP-UX, 23	◇
13.5. Solaris, 25	◇
13.6. FreeBSD, 25.	
14. Frequently Asked Questions	26

1 Installation

Proto Balance can be installed by running the installation executable for your operating system downloaded from <http://www.protonet.co.za/>

This program must be run as superuser on Unix systems, and with admin privileges on Windows. On Windows you may get an alert "Windows firewall has blocked some features of this program". In this case, click on "Unblock". On Unix you must first log in as the "root" user.

Proto Balance comes as an installation binary "install.bin" or "install.exe". Simply run the binary and follow the instructions. See below for Mac specific instructions.

2 QuickStart Proto Balance

2.1 All Unix systems

If it does not already exist, create the directory where Proto Balance can store data and configuration files:

```
mkdir /var/balance/
```

Note that if you use Proto Balance with the mmap option (see below) you will need 500MB of free disk space under this directory.

The simplest way to invoke Proto Balance is with a shell command as follows:

```
protobalance -pthread
```

Use the TERM signal (i.e. 15) to shutdown Proto Balance.

Please go to the section "System Settings for High Performance" if you intend running Proto Balance with a large number of concurrent connections.

2.2 Microsoft Windows

On Windows it is best to start Proto Balance from your Start menu.

If you are an advanced system administrator you may want to use Proto Balance from the command-line. In this case read on:

A directory would have already been created for Proto Balance to store data and configuration files. On Windows XP this directory is as follows. (Check by right clicking on "Proto Balance" in the Start menu and selecting "Properties" – see "Start in:"). This is usually:

```
C:\Documents and Settings\All Users\Application Data\Proto Balance
```

Invoke Proto Balance at the command prompt as follows:

```
protobalance -directory "C:\Documents and Settings\All Users\Application Data\Proto Balance"
```

Please go to the section "System Settings for High Performance" if you intend running Proto Balance with a large number of concurrent connections.

2.3 Apple Mac OS X

Proto Balance does not install itself like a Mac desktop application so it will not appear in any of your desktop menus. To install Proto Balance goto "Utilities" - "Terminal". Inside the Terminal window type,

```
su - root
```

then enter the system administrator password for your machine. You can now run the install executable depending on where you downloaded the file "install.bin". For example, if I am logged in as user "sally", I can enter:

```
chmod a+x /Users/sally/Desktop/install.bin
```

Followed by:

```
/Users/sally/Desktop/install.bin
```

Continue with the instructions for Unix above.

2.4 Ensuring that Proto Balance restarts after your machine has rebooted

On Unix, your operating system documentation should explain how to add the "protobalance" command to your start-up scripts. You will need to have system administrator privileges to make this change however. Use the "-daemon" and "-logfile" options to make Proto Balance run in the background and write to a log file.

On Windows go to "Scheduled Tasks" in your "Control Panel" and click on the "Add Scheduled Task" wizard. Select "Proto Balance" from the list of Applications and then select the "When my computer starts" radio button as follows:



The remaining wizard options can be left as default.

On Mac OS X, edit the file /etc/rc.local or create it if it does not exist. Give it execute permissions with the command "chmod a+x /etc/rc.local". This file should contain something like:

```
#!/bin/sh
/bin/protobalance -shared shmat \
    -directory /var/balance \
    -logfile /var/log/protobalance.log \
    -daemon
```

2.5 Proto Balance command-line options and flags

Proto Balance displays a comprehensive explanation of command-line options when invoked with the -help option. Please run

```
protobalance -help
```

to view options available on your operating system.

3 Configuration

Proto Balance's usefulness comes from its ease of configuration. This section will explain how to set up a simple cluster.

To start, point your web browser to the URL where Proto Balance is running. If Proto Balance is running on the local machine, this URL will be `http://127.0.0.1:8080/`. If you have started Proto Balance with the `-configport` option, a port of other than 8080 must be specified in the URL. If you are configuring Proto Balance from a remote machine, its IP address must replace "127.0.0.1" in the URL.

(Any web browser will work. The Proto Balance configuration web page does not use JavaScript.)

3.1 Logging in and setting your password

When you access the configuration page for the first time, the login password will be empty - hence just click on "Login" without entering a password.

For security you should set your password after logging in for the first time - click on "Change Password" under the "Login" tab.

Proto Balance will remember the IP address of the machine your web browser is running on. If you access the configuration web page from a different machine, you will be forced to log in again.

3.2 Creating a forwarder to divert TCP connections to another host

The simplest configuration is where Proto Balance listens on a port and forwards all data to a port on a different host. For instance, you can run Proto Balance on a firewall and forward all incoming mail to a mail server behind the firewall as follows:

1. Create a new cluster by clicking on "Add Cluster".
2. Enter the settings as follows. Replace "196.43.34.190" with the address of your outside (i.e. public) network interface. Note that "Unique Identifier" is a free text string that should briefly explain your cluster's purpose.

[Login](#) | [Info](#) | [Clusters](#) | [Save Config](#) | [Add Cluster](#) | [Contact](#) |

(Hint: add multiple clusters to listen on multiple ports)

Unique Identifier (alpha-numeric, no white-space)	<input type="text" value="MAIL SERVER"/>
Listen interface address	<input type="text" value="196.43.34.190"/>
TCP listen port 1-65535	<input type="text" value="25"/>
Comment	<input type="text" value="Incoming mail"/>
Box connect timeout in milliseconds or 0 for none	<input type="text" value="10000"/>
Maximum number of connections	<input type="text" value="10000"/>

[Show advanced options](#)

Clicking on the "Clusters" page will show your current settings:

[Login](#) | [Info](#) | [Clusters](#) | [Save Config](#) | [Add Cluster](#) | [Contact](#) |

1: **MAIL SERVER** [Add box](#) [Modify](#) [Delete](#)

TCP listen port	25
Listen interface address	196.43.34.190
Number of OS threads or processes	1
Total connections	0
Max allowed total connections	10000
Default connection timeout	10000 milliseconds
Box affinity	(disabled)
Box load balancer algorithm	box with least connections
No boxes	

3. Now add a physical mail server to which incoming connections will be forwarded. Click on "Add box":

[Login](#) | [Info](#) | [Clusters](#) | [Save Config](#) | [Add Cluster](#) | [Contact](#) |

Unique Identifier (alpha-numeric, no white-space)	<input type="text" value="MAIL MACHINE A"/>
TCP connect address	<input type="text" value="192.168.1.9"/>
TCP connect port 1-65535	<input type="text" value="25"/>
Comment	<input type="text"/>
Last fallover	
Box speed factor	<input type="text" value="1"/>
Box connect timeout in milliseconds or 0 for none	<input type="text" value="10000"/>
Fallover Box	<input type="text"/>
If box falls over, retry in (seconds) or 0 for never	<input type="text" value="0"/>

Clicking on "Clusters" will now give:

[Login](#) | [Info](#) | [Clusters](#) | [Save Config](#) | [Add Cluster](#) | [Contact](#) |

1: **MAIL SERVER** [Add box](#) [Modify](#) [Delete](#)

TCP listen port	25
Listen interface address	196.43.34.190
Number of OS threads or processes	1
Total connections	0
Max allowed total connections	10000
Default connection timeout	10000 milliseconds
Box affinity	(disabled)
Box load balancer algorithm	box with least connections

[Collapse](#)

 MAIL MACHINE A : 192.168.1.9:25,
connections:0, conn-timeout:10000, factor:1 [delete](#)

Finally, click on the green traffic light to enable traffic to the physical machine (box). Your mail server will now receive mail.

3.3 Explanation of traffic lights

The traffic lights are a unique feature of Proto Balance. At any time, it is easy to see the state of all physical machines (boxes) in the cluster. A square border around a light indicates that the light is active.

The traffic lights have the following meaning:

1. Red: The box is not receiving connections. This would be because either the box has failed, or because the box is newly created, or because the box has been explicitly disabled. You can explicitly disable a box by clicking on the red traffic light – which immediately kills all current connections to the box.
2. Yellow: The box will not receive any new connections but existing connections will not be terminated. Any existing connections will be allowed to complete and terminate naturally. Clicking on the yellow traffic light disables the box but does not kill any existing connections.
3. Green: The box is up and ready to receive new connections.
4. A broken connection icon: Connection to the box has failed as shown here:



Note that in this example it is useful to click on "MAIL MACHINE A" to see the "Last fallover" reason.

3.4 Creating a cluster

Begin by creating the port forwarder example from the section "Creating a forwarder..." above. Now add a second box by clicking on "Add box" again:

Login Info Clusters Save Config Add Cluster Contact	
Unique Identifier (alpha-numeric, no white-space)	<input type="text" value="MAIL MACHINE B"/>
TCP connect address	<input type="text" value="192.168.1.10"/>
TCP connect port 1-65535	<input type="text" value="25"/>
Comment	<input type="text"/>
Last fallover	
Box speed factor	<input type="text" value="1"/>
Box connect timeout in milliseconds or 0 for none	<input type="text" value="10000"/>
Fallover Box	<input type="text"/>
If box falls over, retry in (seconds) or 0 for never	<input type="text" value="0"/>
<input type="button" value="Add Box"/>	

Your clusters page will now look like:

[Login](#) | [Info](#) | [Clusters](#) | [Save Config](#) | [Add Cluster](#) | [Contact](#) |

1: MAIL SERVER [Add box](#) [Modify](#) [Delete](#)

TCP listen port	25
Listen interface address	196.43.34.190
Number of OS threads or processes	1
Total connections	0
Max allowed total connections	10000
Default connection timeout	10000 milliseconds
Box affinity	(disabled)
Box load balancer algorithm	box with least connections

[Collapse](#)

MAIL MACHINE A : 192.168.1.9:25, cor

MAIL MACHINE B : 192.168.1.10:25, cc

Don't forget to enable your new box by clicking on the green traffic light. Also remember to click on "Save Config" so that your configuration persists across reboots.

The above cluster will load balance between the mail machines 192.168.1.9 and 192.168.1.10. I.e. each new connection will be forwarded to the machine with the current lowest number of total connections so that the two machines evenly share the incoming mail load.

It is easy to add additional machines to the cluster - even while the cluster is running.

3.5 Removing a box from a cluster

You can remove a box from the cluster even while Proto Balance is in full operation.

To remove a box, carry out the following steps:

1. Click on the yellow traffic light for the box. This will prevent new connections from going to the box.
2. Periodically click on the "Clusters" tab to refresh it. Watch the field "connections:" in the box information text. If the number of connections is not zero it means this box is still serving a client. You must then make a decision about the impact of severing the connection, versus how long you are prepared to wait for the connection to terminate of its own accord.
3. Click on the red traffic light for the box. Any remaining connections will now be forcefully terminated.
4. Click on the "delete" link to the right of the box information text.

4 Connection Options

The options "Box connect timeout in milliseconds" and "Maximum number of connections", under the "Add Cluster" page, will be discussed in this section.

TCP connections have a default "connect timeout" of 180 seconds. This means that if a remote client is trying to connect to a Proto Balance cluster through to a machine that has become disconnected from the physical network, Proto Balance will only give up trying the machine after 3 minutes. The client will wait unhappily for 3 full minutes before Proto Balance diverts it to an alternative machine in the cluster.

In practice, Proto Balance should sensibly assume – even after a few seconds – that the machine is down if no response comes. This is because cluster network connections are high-speed connections that ordinarily respond within a few milliseconds (under moderate loading conditions). Settings for "Box connect timeout

in milliseconds" might be in the range of 500 to 20000, with the high end of this range (20000) being the more conservative and safer setting.

The setting "Maximum number of connections" instructs Proto Balance to drop new client connections when the total number of connections exceeds this setting. This option is useful for preventing Proto Balance from consuming resources that other applications might need, as well as to prevent your cluster from being overwhelmed. Most operating systems become unstable under this kind of extremely high load, hence you should set this value to the maximum number of connections that you are sure your cluster can cope with.

5 Advanced Options

This section discusses the settings under "Show advanced options" of the "Add Cluster" page, as well as further settings of individual machines in the cluster.

5.1 Box Retry

Normally when a box is unavailable, Proto Balance shows a broken-connection icon (see "Explanation of traffic lights" above.) until such time as you might re-enable the box by clicking on the green traffic light. The following setting causes the box to revert to "green" state after a given number of seconds; in this example 30 seconds:

If box falls over,
retry in (seconds) or 0 for never

This means that, should the box fail, then new connections will be routed to another box for 30 seconds, and the failed box will show the broken-connection icon. After 30 seconds, the box will automatically revert to "green" state. Should the box still be unavailable, the first new connection that discovers this will transparently be routed to another box in a way that is invisible to the client, and the box will show the broken connection icon as usual.

If you use this option you should also set the option "Box connect timeout in milliseconds" to a low value because it may take this long to discover a failure - latency that may be unacceptable to your client application and may occur every 30 seconds.

Should there be no new client connections after 30 seconds, the box will stay in "green" state indefinitely. This can be confusing as it would look like the box is working perfectly when in fact it is unavailable. Therefore you should always check your box's traffic graphs to see whether it really is working.

5.2 Load Balancing Algorithms

There are several ways to distribute clients across a cluster to ensure that all boxes in a cluster are evenly loaded. Proto Balance supports all commonly used load balancing algorithms. The possible settings are:

1. Box with least connections:

A new connection goes to the box with the lowest connection count. Here, the connection count is calculated as the box's number of current connections plus the the number of new connections serviced by the box in the past second. Note that using this setting has an odd side-effect: if you get less than one connection per second, and your connections last less than a second, then all traffic will go to the first box on the cluster. Similarly if you never get more than two connections at the same time, only the first two boxes will ever get any load.

This setting is useful if you are trying to limit the number of connections to each box - for instance if you know your box's performance degrades linearly with the number of connections.

2. Box with least traffic:

This has similar behavior to the previous setting except that the loading is calculated as the cumulative traffic in and out of the box over the last second. The box with the least loading by this measure is the chosen box.

This setting is useful if you are trying to limit the total traffic to each box - for instance if you know your box's performance degrades linearly with the number of bytes served.

3. Box with least custom function:

This allows you to define a loading function of your choice, and chooses the box loaded least by this calculation.

4. Purely random box:

The box is chosen using a cryptographically strong random number generator seeded with the time in milliseconds and the process ID.

5. Next box by round robin:

The next box in sequence is chosen.

6. Box determined by hashing the client address:

This is explained in the section "Box affinity" below.

5.3 Connection rate limits, denial of service limits

(Proto Balance Advanced and Proto Balance SSL only)

The cluster setting "Maximum number of per-client connections per ten second period" works as follows:

For each new successive connection attempt from a client, the client's new statistic is checked against that in the current ten second period. If the client would exceed the allotted quota, the connection is dropped. A line is also logged of the form "max connections of NNN per ten seconds reached for client NN.NN.NN.NN - dropping". If the client has not exceeded the allotted quota, the stats are incremented by one. If the ten-second window has rolled over, it is reset back to one.

Similarly, the cluster setting "Maximum number of per-client concurrent connections" will set the quota on the total number of simultaneous connections a single client can have open. Connections that would exceed this limit are dropped, and a line logged of the form "max connections of NNN per ip reached for client NN.NN.NN.NN - dropping".

5.4 Idle Timeout

Proto Balance Advanced and Proto Balance SSL have a feature to forcefully close connections that have been idle for too long. These would be connections over which no traffic has been transferred in either direction for some time. Set the configuration option "Client idle timeout in milliseconds" to a value in seconds. A value of zero disables this feature.

5.5 X-Forwarded-For HTTP header

Proto Balance effectively hides the IP address of the connecting client. This is because, from your web server's point of view, the connection is coming from Proto Balance and not from a remote client. To log information about the connecting client, Proto Balance Advanced and Proto Balance SSL have the additional configuration setting, "Add X-Forwarded-For to HTTP header". Setting this to "Enabled" will harmlessly insert the IP address of the client into the request header sent to your web server.

You will need your web server to log this field. In Apache this is easy: edit the file `httpd.conf` and change your log format to include the field,

```
\ "%{X-Forwarded-For}i\ "
```

For example, the `www.protonet.co.za` web site has the common log format as:

```
LogFormat "\ "%{Referer}i\ " \ "%{X-Forwarded-For}i\ " %h %l %u %t \ "%r\ " %>s %b \ "%{User-Agent}i\ " common
```

Note that there is a slight performance penalty when you do this because Proto Balance Advanced/SSL has to analyze the client request content.

Protocols other than HTTP do not have the facility to add such headers.

5.6 Box affinity

There are some cluster environments where you would prefer a client to connect to the same physical box each time. This might be because each box caches local client data, whereas it is expensive to migrate that data to a sister box. For example, if a client happens to be directed to box 13 of 20 on first login, then the same client should also be directed to box 13 on all subsequent connection attempts.

With Proto Balance there are two ways of doing this. The first method requires the following setting:

New connections go to

This setting works as follows: Consider a client connecting from the IP address `41.241.74.39`. Proto Balance performs a cryptographic hash of the IP address modulus the number of boxes (See also "Box speed factor"). The result of this computation is that if there are, say, 20 boxes then `41.241.74.39` will map to the number 13 and the client will be directed to box 13. The same IP address will always map to the number 13 hence this client will continue to connect to box 13 as long as its IP address stays `41.241.74.39`.

Because the cryptographic hash is pseudo-random, load will be evenly spread over all boxes in the cluster.

The second method requires the following setting:

Try ensure clients reconnect
to the same box by remembering the

This setting works as follows: Consider a client connecting for the first time. The client has an IP address of `41.241.74.39`. A box, say box 13, is chosen for a client via one of the load-balancing algorithms discussed above. Proto Balance saves into a lookup table the client's IP address and chosen box. The next time the client connects, Proto Balance will find its IP address in the lookup table and then direct it to the same box as before.

5.7 Client connection dictionary

To view the current state of the lookup table discussed in "Box affinity" above, click on "Show client connection dictionary" in the "Info" tab.

This list shows five fields:

1. "IP": The IP address of the connecting client.
2. "Box": The name of the box that last serviced this client.

3. "Date": The GMT date and time of the last connection.
4. "Current": (Proto Balance Advanced and Proto Balance SSL) The total number of concurrent connections originating from this client.
5. "Accumulated": The total number of connection attempts over all time.
6. "Recent": The number of connection attempts in the last ten seconds.

With Proto Balance Advanced and Proto Balance SSL, the dictionary has an upper limit on the number of clients after which records of the oldest clients are deleted. This limit can be set in the configuration web page - see "Max dictionary size in megabytes" under the Info tab. Note that each client consumes approximately 160 bytes therefore the default setting of 40 (megabytes) is sufficient to hold 250000 IP addresses. To enlarge this limit, increase your shared memory settings.

With Proto Balance Advanced and Proto Balance SSL, the dictionary entries also have an expiry time. The setting "Delete an IP address from the dictionary after N seconds of inactivity" will set the upper age of a dictionary entry that has had no further connections from the same client IP address. The default setting of 604800 seconds equals seven days.

5.8 Box fallover






An option "Fallover Box" is useful in conjunction with box affinity explained above. This specifies the preferred box to route clients to in the event a box is unavailable. For instance if you have setup your cluster such that each box has a twin that contains replicas of its clients' data, then that twin would be the preferred box to send the client and twins would have each others' names in their "Fallover Box" setting.

Here is an example:

1: **MONSTER CLUSTER** [Add box](#) [Modify](#) [Delete](#)

TCP listen port	81
Listen interface address	41.241.74.39
Number of OS threads or processes	1
Total connections	0
Max allowed total connections	10000
Default connection timeout	5000 milliseconds
Box affinity	client IP4/IP6 address
Box load balancer algorithm	box with least connections

[Collapse](#)

-  PAIR 01 - A : 192.168.5.31:80, connections:0, fallover to PAIR 01 - B
-  PAIR 01 - B : 192.168.5.32:80, connections:0, fallover to PAIR 01 - A
-  PAIR 02 - A : 192.168.5.33:80, connections:0, fallover to PAIR 02 - B
-  PAIR 02 - B : 192.168.5.34:80, connections:0, fallover to PAIR 02 - A
-  PAIR 03 - A : 192.168.5.35:80, connections:0, fallover to PAIR 03 - B
-  PAIR 03 - B : 192.168.5.36:80, connections:0, fallover to PAIR 03 - A
-  PAIR 04 - A : 192.168.5.37:80, connections:0, fallover to PAIR 04 - B
-  PAIR 04 - B : 192.168.5.38:80, connections:0, fallover to PAIR 04 - A

5.9 Speed factor

A clusters may have machines of differing CPU speed or other performance factor. For example, you may have a cluster made up both of machines of 500MHz CPUs and as well as machines of 750MHz CPUs. If you would like to direct proportionately more connections to the faster machines. You can set the speed factor to "2" for the 500Mhz machines and to "3" for the 750Mhz machines.

You should use small ranges for this setting - i.e. do not use speed factor values of "100" or "200" when "1" or "2" is sufficient.

5.10 Close connection if...

Proto Balance supports an obscure feature useful for IIOP connections. IIOP is the TCP protocol used by CORBA. If you are unsure what these terms mean then feel free to skip this section.

There are many vendors that implement CORBA/IIOP: Borland(Visibroker/Janeva), BEA(WebLogic), Orbix, among others. There are some interoperability problems between certain versions of these implementations. Namely, some implementations don't respond to a CORBA "CloseConnection" request resulting in your application server being overwhelmed with connections. Proto Balance gives you the opportunity to force closing of a connection by setting this option as follows:

Close connection if passing data matches

Hex data

In this case Proto Balance will forcibly shutdown both the client and server side of the connection whenever it finds a packet matching the above binary sequence. Note that the above sequence is just an example - you should use a packet sniffer like <http://www.ethereal.com/> to grab an example packet from your server and verify the problem - the packet can differ slightly depending on the application server architecture. Also be aware that this kind of forced shutdown introduces a race condition, albeit one that is rarely noticeable.

6 Multiple Clusters













Proto Balance allows you to create an arbitrary number of clusters, with an arbitrary number of boxes per cluster. The question is often asked: "How do I listen on more than one port?" This can easily be done by creating multiple "clusters" each listening on a different port.

Note that all clusters share the same client connection dictionary (lookup table) discussed above. Hence if you are using the "remembering the client IP4/IP6 address" option in, say, two clusters, and it is foreseeable that the same client could connect to both of these clusters, then this is potentially a problem. You can avoid confusion by setting up both clusters with the same number of boxes having the same box names. As follows:

1: **ORDINARY WEB** [Add box](#) [Modify](#) [Delete](#)

TCP listen port	80
Listen interface address	41.241.74.39
Number of OS threads or processes	1
Total connections	0
Max allowed total connections	10000
Default connection timeout	5000 milliseconds
Box affinity	client IP4/IP6 address
Box load balancer algorithm	box with least traffic

[Collapse](#)

-     HTTP/HTTPS A : 192.168.5.30:80,
-     HTTP/HTTPS B : 192.168.5.31:80,
-     HTTP/HTTPS C : 192.168.5.32:80,

2: **SECURE WEB** [Add box](#) [Modify](#) [Delete](#)

TCP listen port	443
Listen interface address	41.241.74.39
Number of OS threads or processes	1
Total connections	0
Max allowed total connections	10000
Default connection timeout	5000 milliseconds
Box affinity	client IP4/IP6 address
Box load balancer algorithm	box with least traffic

[Collapse](#)

- HTTP/HTTPS A : 192.168.6.60:443
- HTTP/HTTPS B : 192.168.6.61:443
- HTTP/HTTPS C : 192.168.6.62:443

Alternatively, you can run multiple instances Proto Balance each on a different machine.

(Note that under no circumstances should you ever try to startup more than one Proto Balance program on the same machine using the "-shared" option.)

In general, you should not create more than 50 clusters, since this can degrade performance.

7 Protocols that Don't Load Balance

Most protocols can be proxied/redirected through Proto Balance and will work well. Examples of such protocols are those shown in the above examples, such as HTTP, SMTP, POP, etc. However there are protocols that encode TCP port and IP information into their data stream: To the server it appears as though the connection is originating from Proto Balance whereas the data itself says that the connection is originating from a remote client. Load balancing such a protocol will not work. Examples of non-working protocols are BEA WebLogic "t3://" connections (with HTTP tunnelling disabled) and FTP connections.

For the case of BEA WebLogic "t3://..." connections, change your client's connection setting to "http://..." instead of "t3://...", and check "Enable Tunnelling" on your WebLogic console as shown here:

The screenshot shows the BEA WebLogic console configuration page for Protocols. The 'Advanced Options' section is expanded, showing the following settings:

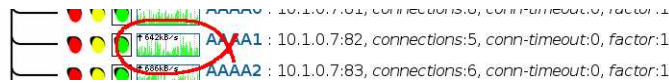
- Complete Message Timeout:** 60 seconds. Description: The maximum amount of time (between 0 and 480 seconds) spent waiting for a complete message to be received.
- Idle Connection Timeout:** 65 seconds. Description: The maximum amount of time (in seconds) that a connection is allowed to be idle before it is closed by this server.
- Enable Tunnelling:** . Description: Specifies whether tunneling for the T3, T3S, HTTP, HTTPS, IIOP, and IIOPS protocols should be enabled for this server.

BEA has more information about "HTTP tunneling" in their documentation.

Another way to load balance t3:// connections is to use the same listening port as the WebLogic server (albeit on a different interface or different machine). This works because the t3:// protocol only encodes the port number and not the client address information. This means that, if your WebLogic server listens on port 7000, then all other WebLogic servers in your cluster must also listen on port 7000, and Proto Balance too must listen on port 7000.

8 Traffic Graphs

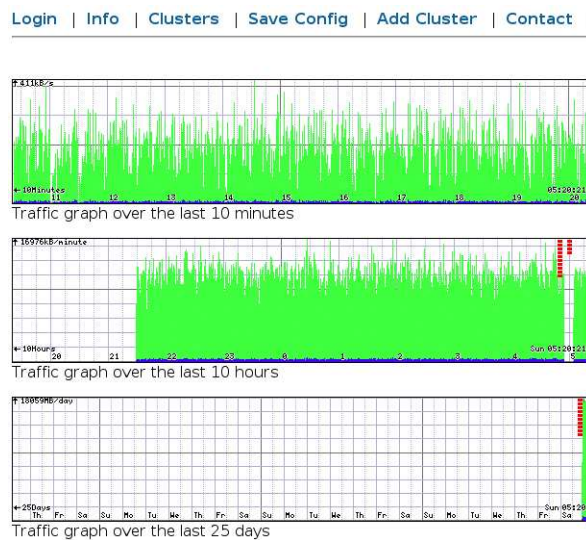
Proto Balance provides comprehensive traffic graphs which show failure histories as well as input and output traffic. Each box also has a thumbnail graph which shows traffic processed in the past minute:



If a box fails, then red tokens will appear on the graph. The following box had five pending connections at the time of failure; therefore the graph shows five tokens (even though box failed only once):



Clicking on the thumbnail will take you to a set of three graphs, each showing traffic over a different time scale:



Note that green represents bytes downloaded from your box to the client, whereas blue represents bytes uploaded from the client to your box. The above traffic is static HTTP content, hence there is little uploaded traffic compared to downloaded traffic. In this example the box was booted at 9:30pm. Maintenance was performed on the box at approximate 5am for about 10 minutes.

9 Monthly client logs

Proto Balance can create logs of all the clients that connect, to log files named according to the current year and month. The command-line option

`-monthly <folder-path>`

specifies a folder where these logs are to be written, which folder must previously have been created and have the necessary access writes to allow Proto Balance to create log files. Log file names have the format YYYY-MM for example,

`-monthly C:\MyMonthlyLogs\`

will create the files,

```
C:\MyMonthlyLogs\2008-09
C:\MyMonthlyLogs\2008-10
C:\MyMonthlyLogs\2008-11
```

for the months September through November 2008. Proto Balance does not delete log files, therefore you will have to create a means of ensuring disk space is not used up.

Example log lines are:

```
2008-11-30 Sun 18:30:01.154: connect from 111.25.26.27 to port 80

2008-11-30 Sun 18:30:04.850: disconnect of 111.25.26.27 to port 80

upload/download=18/279 bytes
```

The figures for upload/download are in bytes. Each and every connection attempt is logged. If the X-Forwarded-For option (see the section X-Forwarded-For HTTP header) is enabled, then HTTP access logs will be included, for example:

```
2008-11-30 Sun 18:35:21.204: 111.25.26.27: GET /index.html HTTP/1.1

Host: protonet.co.za Referer: google.com User-Agent: IE Accept-Language: en
```

10 Security

The Proto Balance configuration web page is not secure. You should use the command-line option `-listeninterface` to force Proto Balance to accept connections from only one interface. The interface you choose should be one that is not accessible from untrusted machines. For example, if you run Proto Balance on an Internet firewall, then `-listeninterface` should be the interface of your internal network.

Proto Balance can improve the security of exposed networks when used to restrict access to only those services that Proto Balance is responsible for. For example, you may configure your firewall to forward no incoming Internet traffic, and at the same time configure Proto Balance to forward specific services to particular internal machines. Thus Proto Balance behaves like a discrete set of firewalling rules.

11 OS-specific Event Based I/O

(This section is technical background.)

Proto Balance Mach is able to use a high performance "event-based" I/O model specific to your particular operating system. An industry standard for high performance event-based I/O does not exist, and therefore each OS vendor has developed their own. The list is as follows:

Windows: IO Completion Ports

Solaris/HP-UX: /dev/poll

FreeBSD/MacOS: kqueues

Linux: epoll

Proto Balance supports all of these interfaces. The advantage of these interfaces over traditional programming models, is that Proto Balance can scale to arbitrarily large numbers of concurrent connections without incurring performance degradation.

To disable "event-based" I/O, use the command-line option "-poll-method select" option.

Consult the operating system specific sections below for more about "-poll-method" under "System Settings for High Performance with Your Operating System".

12 Operating System Considerations and Tuning

Proto Balance uses a unique software architecture to be able to cope with so many concurrent connections while using as little operating system (OS) resources as possible. Little or no OS configuration is required. However, for enterprise-level environments where a high number of concurrent connections are expected, you should be aware of possible problems caused by limitations in your operating system.

Proto Balance creates as many OS processes or threads as there are the-square-root-of the number of connections. This means that the number of connections handled by one OS process or thread is the square root of the total number of connections. Hence for many system resources, Proto Balance raises the OS limit by a power of 2. This gives Proto Balance exceptional scalability on any platform.

For the discussions below, note that each client consumes two connections and hence two socket handles: one from the client, and one to the box.

12.1 Summary of limits

The following table shows considerations for each of the three methods of invocation:

Option	Limitation to consider					
	Number of socket handles per process	Shared memory	Semaphores	Unable to start more than one Proto Balance	Creates large temporary file	TCP timeout and port settings
-pthread	Yes	No	No	No	No	Yes
-shared shmat	No	Yes	Yes	Yes	No	Yes
-shared mmap	No	No	Yes	Yes	Yes	Yes

Note that for large numbers of concurrent connections, your operating system's settings will most certainly need customization. See the section "System Settings for High Performance..." below.

12.2 Maximum number of socket handles / file descriptors

Operating systems have a limit on the number of socket handles ("file descriptors" in Unix terminology) that a single process or thread can have open. Most OSs will have a means to increase these limits - consult your operating system (OS) documentation. Potential limits are:

1. The number of file descriptors per process.

This limit usually comes into effect when using the -pthread option, since this option tells Proto Balance to start only one process and to run multiple child threads. If your OS's limit is 1024 (a typical value) then Proto Balance will top out at 1024 concurrent connections and will then start dropping connections. This will happen even though each thread only has about 32 sockets open.

This limitation does not apply when using the -shared option.

2. The number of file descriptors per thread.

This limitation does not apply.

3. The total number of file descriptors for all processes.

This limitation will come into effect when you are using Proto Balance with thousands of concurrent connections. If you reach this limit then the load balancing box itself will probably freeze since no processes anywhere on that machine will be able to open a file descriptor at all. Therefore it is important to limit the "Maximum number of connections" option to well less than this limit.

If you get error messages like "accept on fd N failed: File table overflow", or "Too many open files", then you have exceeded this limit.

12.3 Maximum number of threads (-pthread option)

On some OSs, the Linux kernel in particular, there is a default limit on the maximum number of threads that a process can create. This means that when using the -pthread option, trying to create more than 200 (approx) clusters will result in a fatal error "pthread_create(): Cannot allocate memory".

This limitation will usually come into effect when there are a large number of clusters with a large number of connections each. Should you have few clusters (say 4) and should you expect less than, say, 500 concurrent clients, then this limit will never be reached. Further, the per-process file descriptor limit will usually be reached first.

12.4 Shared memory and Semaphore limits (-shared shmat option)

When Proto Balance is invoked with the option "-shared shmat" it starts multiple OS processes which communicate with each other using shared memory and semaphores (technically known as "System V IPC").

Some Unix operating systems have overly conservative shared memory limits - even modern ones. Because it can sometimes require a kernel rebuild to enlarge these limits, you may instead want to use the "-shared mmap" option - this will always work on all operating systems. Unfortunately, the option "-shared mmap" creates a cumbersome 500MB data swap file. Semaphore limits will come into affect with either -shared option, but are quite rare.

If there is a problem with your settings, then you may get the error "shmat(): Too many open files", "shmat(): Out of memory" or "shmget(nnnnnn): Invalid argument" and Proto Balance will abort. This can be catastrophic. Errors of the form "semop(): No space left on device" or other "sem.." errors similarly indicate a lack of semaphore resources.

Shared memory has maximum limits imposed by the operating system, and a process may not exceed these. These limits are in terms of the allowed (a) number and (b) size of memory blocks, both (c) system-wide and (d) per-process. The same applies to limits on the number of semaphores. Proto Balance creates a small number of successively larger shared memory blocks. You must tune your OS to allow at least: 32 blocks per process, 1024 blocks system-wide and with a maximum block size of 128 megabytes (134217728 bytes) or more. All semaphore limits similarly need increasing to 128 through 256 per process.

Each operating system has a different means of setting these limits. Consult the section "System Settings for High Performance..." below.

Note that FreeBSD has a system resource PMAP_SHPGPERPROC which may need to be increased (pertaining to certain "collecting PV entries" error messages). A web search on "PMAP_SHPGPERPROC" will point you to the correct help.

12.5 TCP port limits

This section is very important if you intend to handle high rates of new connections.

All operating systems allow a maximum of 65535 TCP ports per network interface (IP address). Each connection requires a unique port. Hence a physical machine can only handle 65535 connections times the number of network interfaces. (The number of network interfaces is usually 1 or 2.)

This limit is further reduced as will be explained.

TCP does not dispose of the port once the connection has died - it holds the port for a considerable period of time - this is called the TIME-WAIT state of TCP. The TIME-WAIT state is a precautionary measure: the remote machine may have lost track of the state of the connection because of dropped packets, and may demand a response even when the connection is supposed to have ended.

Officially, the TIME-WAIT state is supposed to last for four minutes.

Now consider that most operating systems do not use the full range of ports - artificially limiting port allocation to, say, 40000 through 65000. Hence, if an application sustains a high rate of new connections, then OS port allocation may top-out at, say, 100 new connections per second - an unacceptably low rate.

There are several TCP parameters that are always tunable on all operating systems that solve these problems:

1. The number of ports the OS is willing to serve is always adjustable by either high or low port limits or both.
2. The TIME-WAIT state can be lowered to 30 seconds or less for high speed networks.
3. Certain high-speed extensions to TCP (known as RFC-1323 - which includes the TCP "Timestamps" extension) are supported by all modern operating systems, but are usually disabled by default. It is best to enable them.

Then there are various timeouts, adjustable on some operating systems but not on others.

12.6 Limits of Physical Memory

Proto Balance consumes kernel and user resources of 20 megabytes of physical memory per 1000 concurrent connections. This figure varies widely for different operating systems. If you intend to handle 10'000 concurrent connections, then 512 to 1024 megabytes is recommended.

12.7 Other limits

Each OS has its own set of arbitrary limits when it comes to running high-performance software. The section "System Settings for High Performance..." contains more information.

13 System Settings for High Performance with Your Operating System

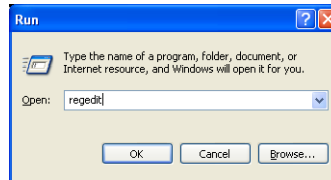
Proto Balance is designed to work efficiently with a large number of concurrent connections. With correct OS tuning Proto Balance can cope well with 10'000 concurrent connections. The operating system's effectiveness toward 10'000 concurrent connections depends on the hardware and the OS version. Proto Balance has by no means been tested to these extremes on every version of every OS, however where particular problems were encountered, these have been documented in the sections below.

If you are not in an environment of thousands of concurrent connections, with hundreds of new connections per second, we recommend you do not adjust any TCP settings. Other settings always apply however.

An attempt has been made to gather all the settings in one place for each supported OS. Please use these settings with care since the performance of other applications may be negatively impacted. Many of these settings are liable to render your system unstable, insecure, and/or unbootable if not done with the proper expertise. You are advised to follow your OS vendor's recommendations rather than any setting suggested in this document.

13.1 Windows

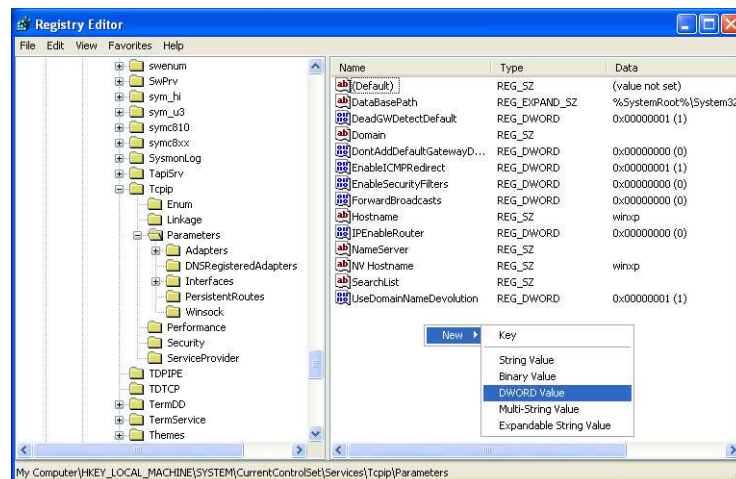
Goto your "Start" menu and select "Run...". Enter "regedit" as follows:



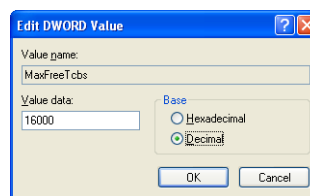
Navigate through the levels:

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\Tcpip\Parameters\

Right click on the white area to create a new "DWORD" registry entry "MaxFreeTcbs":



Double click on "MaxFreeTcbs" and select Decimal text, then enter "16000":



For Windows XP there are 6 new registry entries you must create in the same way:

MaxFreeTcbs: 16000 – See note below

MaxHashTableSize: 16384

```

MaxUserPort: 65534
Tcp1323Opts: 3
TcpNumConnections: 32000
TcpTimedWaitDelay: 30

```

If you are using Windows Server some of these entries do not apply. For Windows Server 2003 the default TcpNumConnections is 262144 and therefore need not be set. For Windows Server 2003 the values are:

```

MaxUserPort: 65534
Tcp1323Opts: 3
TcpTimedWaitDelay: 30

```

If you are using Windows Vista and Windows Server 2008 the values are:

```

MaxUserPort: 65534
Tcp1323Opts: 2
TcpTimedWaitDelay: 30
TcpFinWait2Delay: 60

```

When you are done creating the registry entries, restart Windows.

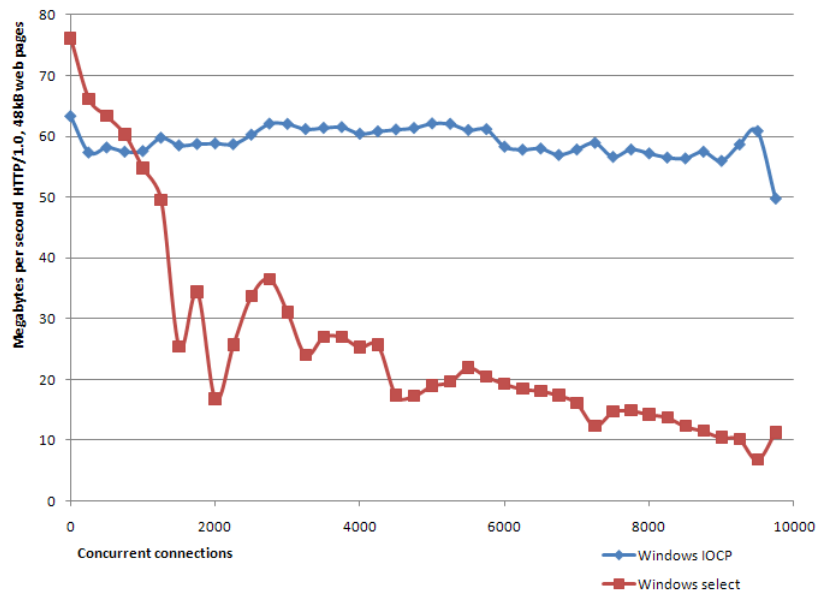
In addition, Windows Vista and Windows Server 2008 do not seem to be able to use the full range of ports without the "netsh" command executed under the Command Prompt. See <http://support.microsoft.com/kb/929851/> for more information if you are using Proto Balance in conditions of high connection rates.

Note that Windows XP prior to service pack 2 will not handle more than 3500 concurrent connections.

When using "-poll-method iocp", Windows will use an "IO Completion Port" architecture and will have near-perfectly scalability up to 10000 concurrent connections. When using "-poll-method select", Windows performance degrades substantially with large numbers of concurrent connections.

(If your version of Proto Balance does not support the option "-poll-method" it means "iocp" is unsupported in that version and "select" is the default.)

The following graph shows performance for both architectures (hardware is a commodity two core CPU at 1.2GHz):



13.2 Linux

Login as root. Add these commands early in your system startup scripts:

```
echo "1" > /proc/sys/net/ipv4/tcp_tw_recycle
echo "1024 65535" > /proc/sys/net/ipv4/ip_local_port_range
echo "1" > /proc/sys/net/ipv4/tcp_timestamps
echo "60" > /proc/sys/net/ipv4/tcp_fin_timeout
echo "3" > /proc/sys/net/ipv4/tcp_orphan_retries
echo "1" > /proc/sys/net/ipv4/tcp_rfc1337

echo "134217728" > /proc/sys/kernel/shmmax
echo "2097152" > /proc/sys/kernel/shmall
echo "4096" > /proc/sys/kernel/shmmni

echo 250 256000 100 1024 > /proc/sys/kernel/sem

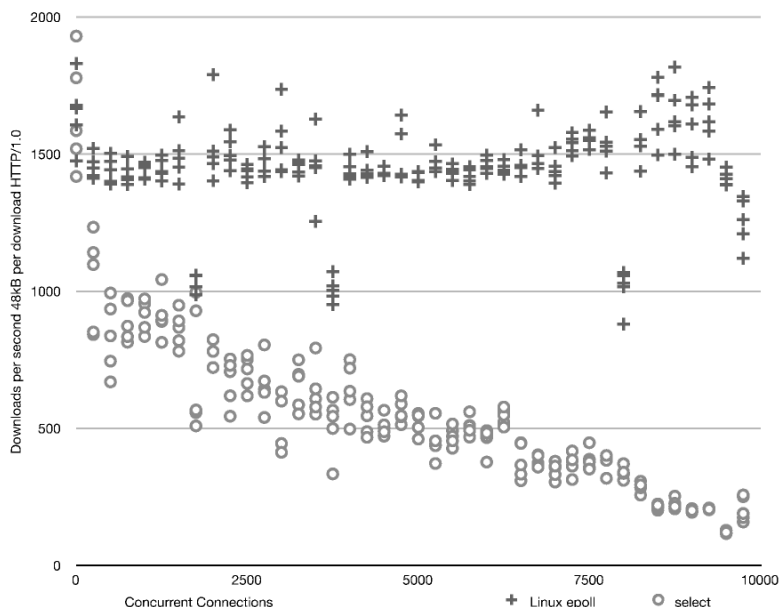
echo "32768" > /proc/sys/fs/file-max
```

(Note the setting for `tcp_fin_timeout` which has changed as of the present writing.)

When using `"-poll-method epoll"`, Linux will use an "epoll" event-based architecture and will have near-perfectly scalability up to 10000 concurrent connections. When using `"-poll-method select"`, Linux performance degrades substantially with large numbers of concurrent connections.

(If your version of Proto Balance does not support the option `"-poll-method"` it means "epoll" is unsupported in that version and "select" is the default.)

The following graph shows performance for both architectures (hardware is a commodity two core CPU at 1.2GHz):



13.3 MacOS

Login as root.

Place the following lines into the file `/etc/sysctl.conf` (create it if it does not exist).

```
kern.maxvnodes=16384
kern.sysv.shmmax=134217728
kern.sysv.shmmni=4096
kern.sysv.shmseg=32
kern.sysv.shmall=32767

kern.maxfiles=32768
kern.maxfilesperproc=32768
kern.maxprocperuid=256

kern.sysv.semume=256
kern.ipc.somaxconn=256

net.inet.tcp.rfc1323=1
net.inet.tcp.tcbhashsize=16384
```

Otherwise run the following commands on startup:

```
sysctl -w kern.maxvnodes=16384
sysctl -w kern.sysv.shmmax=134217728
sysctl -w kern.sysv.shmmni=4096
sysctl -w kern.sysv.shmseg=32
sysctl -w kern.sysv.shmall=32767

sysctl -w kern.maxfiles=32768
sysctl -w kern.maxfilesperproc=32768
sysctl -w kern.maxprocperuid=256

sysctl -w kern.sysv.semume=256
sysctl -w kern.ipc.somaxconn=256

sysctl -w net.inet.tcp.rfc1323=1
sysctl -w net.inet.tcp.tcbhashsize=16384
```

MacOS has maximum limits on many of these values that cannot be overridden. Therefore you may not be able to create a large number of boxes or cluster. Yet in spite of this problem, MacOS 10.3 (Darwin kernel version 7.0.0) successfully starts 10'000 concurrent connections.

When using `"-poll-method kqueue"` (the default), MacOS will use an "kqueue" event-based architecture and will have near-perfectly scalability up to 10000 concurrent connections. When using `"-poll-method select"`, MacOS performance degrades substantially with large numbers of concurrent connections.

(If your version of Proto Balance does not support the option `"-poll-method"` it means "kqueue" is unsupported in that version and "select" is the default.)

13.4 HP-UX

Login as root. Run the commands:

```
/bin/ndd -set /dev/tcp tcp_time_wait_interval 15000
/bin/ndd -set /dev/tcp tcp_tw_cleanup_interval 15000
```

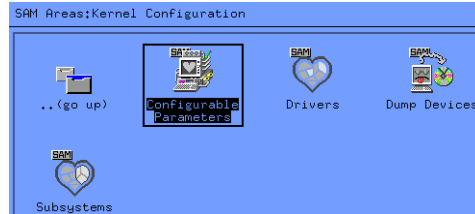
```

/bin/ndd -set /dev/tcp tcp_smallest_anon_port 1024
/bin/ndd -set /dev/tcp tcp_largest_anon_port 65535
/bin/ndd -set /dev/tcp tcp_ts_enable 2
/bin/ndd -set /dev/tcp tcp_fin_wait_2_timeout 60000
/bin/ndd -set /dev/tcp tcp_conn_request_max 16384

```

(Note the setting for tcp_fin_wait_2_timeout which has changed as of the present writing.)

Start sam and goto Kernel Configuration ⇒ Configurable Parameters:



Pending Kernel Based Upon: Current Kernel

Configurable Parameters 1 of 166 selected

Name	Current Value	Pending Value	Type	Associated Module	Description
semvmx	32767	32767	Static	N/A	Semaphore Maximum Value
sendfile_max	0	0	Static	N/A	Max No. of Buffer Cache Pages Usable by sendfile E
shmem	1	1	Static	N/A	Enable Sys V Shared Memory
shmmax	67108864	67108864	Dynamic	N/A	Max Shared Mem Segment (Bytes)
shmmni	200	200	Static	N/A	Number of Shared Memory Identifiers
shmseg	120	120	Dynamic	N/A	Shared Memory Segments Per Process
st_ats_enabled	0	0	Static	N/A	Boolean: IF Set, SCSI Tape Open Will Try To Reserv
st_fail_overruns	0	0	Static	N/A	Boolean: IF Set, SCSI Tape Read Resulting In Data
st_large_recs	0	0	Static	N/A	Boolean: IF Set, Enables Large Record Support For
st_san_safe	0	0	Static	N/A	Unknown

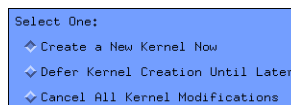
The parameters you will need to set are:

```

max_thread_proc: 512
max_files: 512
maxuprc: 150
nfile: 32768
shmmax: 134217728
shmmni: 1024
shmseg: 32
shmem: 1
semmap: 258
semnmi: 256
semnms: 256
semnmu: 256
semume: 256
tcphashsz: 65536

```

You will then need to create and boot a new kernel. These options appear when you Exit the Configuration Parameters window:



(Note that an old HP-UX model 9000/778/B180L won't perform well with over 1000 concurrent connections. With 100 concurrent connections this machine can consume 30% of the saturation of a 100 megabit full-duplex Ethernet interface, while Proto Balance consumes 95% CPU.)

When using "-poll-method dev_poll" (the default), HP-UX will use an "/dev/poll" event-based architecture and will have near-perfectly scalability up to 10000 concurrent connections. When using "-poll-method select", HP-UX performance degrades substantially with large numbers of concurrent connections.

(If your version of Proto Balance does not support the option "-poll-method" it means "kqueue" is unsupported in that version and "select" is the default.)

13.5 Solaris

Login as root. Run the commands:

```
/usr/sbin/ndd -set /dev/tcp tcp_time_wait_interval 15000
/usr/sbin/ndd -set /dev/tcp tcp_smallest_anon_port 1024
/usr/sbin/ndd -set /dev/tcp tcp_largest_anon_port 65535
/usr/sbin/ndd -set /dev/tcp tcp_tstamp_always 1
/usr/sbin/ndd -set /dev/tcp tcp_fin_wait_2_flush_interval 60000
```

(Note the setting for tcp_fin_wait_2_flush_interval which has changed as of the present writing.)

On Solaris 9 edit the file /etc/system, and append the lines:

```
set shmsys:shminfo_shmmax=134217728
set shmsys:shminfo_shmmni=16384
set shmsys:shminfo_shmseg=32
set semsys:seminfo_semmap=256
set semsys:seminfo_semmni=256
set semsys:seminfo_semmns=256
set semsys:seminfo_semmnu=256
set semsys:seminfo_semume=256
```

On Solaris 10 edit the file /etc/system, and append the lines:

```
forceload:sys/semsys
forceload:sys/shmsys
set shmsys:shminfo_shmmax=134217728
set shmsys:shminfo_shmmni=16384
set shmsys:shminfo_shmseg=32
set semsys:seminfo_semmap=256
set semsys:seminfo_semmni=256
set semsys:seminfo_semmns=256
set semsys:seminfo_semmnu=256
set semsys:seminfo_semume=256
```

Then reboot.

When using "-poll-method dev_poll" (the default), Solaris will use an "/dev/poll" event-based architecture and will have near-perfectly scalability up to 10000 concurrent connections. When using "-poll-method select", Solaris performance degrades substantially with large numbers of concurrent connections.

(If your version of Proto Balance does not support the option "-poll-method" it means "kqueue" is unsupported in that version and "select" is the default.)

13.6 FreeBSD

Login as root. Edit the file /boot/loader.conf and append the lines:

```
kern.ipc.shm_use_phys=1
kern.ipc.shmmax=134217728
kern.ipc.shmall=32768
kern.ipc.shmmni=4096
kern.ipc.shmseg=32

kern.maxfiles=32768
kern.ipc.somaxconn=256
kern.ipc.maxsockets=32768

net.inet.tcp.rfc1323=1
net.inet.tcp.tcbhashsize=16384

kern.ipc.semmap=1024
kern.ipc.semmni=1024
kern.ipc.semmns=1024
kern.ipc.semmnu=256
kern.ipc.semume=1024
```

For Proto Balance Mail, use:

```
kern.ipc.shmmax=34359738368
kern.ipc.shmall=8388608
```

Then reboot. It would seem that the above `kern.ipc.shmmax=134217728` setting sometimes does not take effect, resulting in an "shmget(nnnnnn): Invalid argument" error. Proto Balance then aborts. In this case you need to rebuild your kernel with a larger `SYSPSIZE` setting, and a larger `SHMALL` and `SHMSEG` setting. Consult FreeBSD resources for information on how to do this. Be sure to use the latest stable FreeBSD release on 64-bit architecture.

When using "-poll-method kqueue" (the default), FreeBSD will use an "kqueue" event-based architecture and will have near-perfectly scalability up to 10000 concurrent connections. When using "-poll-method select", FreeBSD performance degrades substantially with large numbers of concurrent connections.

(If your version of Proto Balance does not support the option "-poll-method" it means "kqueue" is unsupported in that version and "select" is the default.)

14 Frequently Asked Questions

- What is the performance of Proto Balance?

Proto Balance can perform between 1500 and 3000 new connections per second per CPU Gigahertz on all supported hardware platforms. Therefore to calculate the low bound, simply take the real CPU clock speed in GHz and multiple it by 1500.

- How many downloads/sec can I expect using HTTP 1.1?

Proto Balance can perform between 5000 and 8000 HTTP 1.1 requests per second per CPU Gigahertz on all supported hardware platforms. Therefore to calculate the low bound, simply take the real CPU clock speed in GHz and multiple it by 5000.

- How new SSL connections per second and Proto Balance SSL handle?

On x86 64-bit Unix systems (like AMD64 and Intel Core2 64-bit and later), Proto Balance SSL can perform about 400 new SSL negotiations per second per CPU Gigahertz.

On x86 64-bit Windows XP x64 Professional, Proto Balance SSL can perform about 400 new SSL negotiations per second per CPU Gigahertz.

On SPARC 64 platforms Proto Balance SSL can perform about 200 new SSL negotiations per second per CPU Gigahertz.

HP-UX PARISC 2.0 64-bit has performance in a similar range.

- What are the minimum hardware requirements for running Proto Balance?
Proto Balance can handle 10000 concurrent connections using under 512 megabytes of RAM. Proto Balance consumes only a few megabytes of disk space.
- Is there a licensing limit to the number of back-end servers over which I may distribute load?
There is none. You can have hundreds of back-end servers.
- Is there a licensing limit to the number of clients that may connect to Proto Balance? Are there traffic or connection-rate limits?
There is none. Thousands of concurrent clients may connect. There are no traffic or connection-rate limits.
- What license is Proto Balance distributed under?
The Proto Balance license is a once-off purchase that allows you to install Proto Balance on a single machine with an unlimited number of back-end servers and an unlimited number of clients. See the section "Licensing Terms".
- I want to deploy a great many Proto Balance instances throughout my organization. What should I do?
Please contact us. We can provide bulk or organization-wide licensing on request.
- Will Proto Balance allow me to do hot deployments? What is a hot deployment?
Hot deployments are software deployments that incur no downtime for the client users. With Proto Balance it is easy to switch from one back-end server to another. So you can re-deploy your server machines in sequence without ever having a moment of total unavailability. The redeployment is completely transparent to the client users.
- Will Proto Balance allow me to add and remove back-end machines on-the-fly?
Yes. With Proto Balance you can add and remove back-end server machines even during peek load.
- Will Proto Balance use multiple processors on a multi-processor machine?
Yes. Proto Balance starts many concurrent threads/processes, so these will be distributed to different processor cores.
- Is Proto Balance a hardware load balancer? What is a hardware load balancer?
A hardware load balancer is a dedicated network appliance that performs packet routing at the Ethernet or IP level. Proto Balance is more suitable than a hardware load balancer for most applications because it provides more comprehensive and controlled monitoring of the boxes in your cluster. Proto Balance also runs on all popular platforms, which are easy to upgrade, or replace in the event of failure.
- What operating systems does Proto Balance run on? Is my platform supported?
Proto Balance is available for the following systems:
 - For the PC (80386/80486/Pentium and later):
 - * Windows
 - * Linux
 - * Solaris
 - * FreeBSD
 - Other Unix systems:
 - * Solaris SPARC

* HP-UX PA RISC (HP 9000 and variants)

* MacOS X PowerPC

- o My OS is not supported. What can I do?
Contact us to request support for other platforms.
- o When does Proto Balance fail a box? What does a box have to do for Proto Balance to think it has fallen over?
Any service running on a box should receive or send at least one byte of data before closing the connection. If a connection error or close occurs before 1 byte of data has been transferred, then the box is deemed to have failed.
Note that HTTP protocols (included "Soap", "Web Services" and ordinary web) return "500 Internal Server Error" and "503 Service Unavailable" when the server is overloaded or undergoing maintenance. In this case the client is NOT redirected to another box. (For Proto Balance to redirect would require that it cache the entire HTTP request - which could be large - and resubmit it to another box.) The client is served the 500 or 503 error as any ordinary web page. The 500 or 503 error is however detected by Proto Balance and Proto Balance fails that box. The box traffic light goes to red status immediately. The "Last fallover" message in the box configuration will show this HTTP error message.
- o How can I install Proto Balance on 64-bit Intel/AMD Linux systems?
If you get the error "No such file or directory" when trying to run the installer, and you are using a 64-bit distribution of Linux (such as Ubuntu-x86_64) then you need to install the 32-bit compatibility libraries. The Ubuntu/Debian package is "ia32-libs". Simply type "apt-get install ia32-libs".
- o Does Proto Balance disable the Nagle Algorithm? What is the Nagle Algorithm?
The Nagle Algorithm queues short TCP packets and merges them before sending. This can increase communication latency and is inappropriate for a load balancer. Proto Balance disables the Nagle Algorithm.
- o How do I listen on multiple ports?
See the section "Multiple Clusters".
- o How many clusters can I create?
Proto Balance has no limit to the number of clusters, and should scale to thousands of clusters. However, testing has only been done with a small number of clusters hence please proceed with caution.
- o How many boxes can I create?
There is no limit to the number of boxes that you can create per cluster. However, testing has only been done with a moderate number of boxes. Proceed with caution over 200 boxes in total.
- o What is the MTTF (Mean Time To Failure) of Proto Balance?
Proto Balance has no known bugs, memory leaks, or other resource leaks. It should run indefinitely under all the platforms on which it is supported, provided that its operation does not exceed configurable or inherent OS limits. The MTTF is therefore that of the hardware and OS platform under which it is running.
- o How do I see the IP address of the connecting client?
See the section "X-Forwarded-For HTTP header" under "Advanced Options".
- o What is the maximum number of unique IP addresses I can have with server affinity mode?
This refers to the feature "remembering the client IP4/IP6 address" as used on Unix systems. With this feature enabled, Proto Balance has memory space for recording the access of 1.5 million

clients. Versions of Proto Balance before 200 do not expire these entries. If you expect more than say, 1 million, unique IP addresses visiting your site then do not enable this feature. Instead, see the section "Box affinity" in the user manual, and enable "Box determined by hashing the client address" - this will have a similar affect and ensure that clients do connect to the same physical machine each time.

Versions of Proto Balance after 200 (Proto Balance Advanced and Proto Balance SSL) expire the oldest clients from memory according to configurable parameters. See the Info tab in the configuration web page.

You may also increase shmmax over 128 megabytes, see "System Settings for High Performance" in the user manual. A value of 134217728 is recommended, but can be increased to 268435456 or 536870912 if your operating system allows it, and if your machine has sufficient physical memory. Remember that you need at least twice the physical memory as the value of shmmax.

Note that the option "-pthread" does not use shared memory, therefore has no such limitation. Also, the option "-shared mmap" is limited by other parameters.

- o What SysV shared memory key does Proto Balance use? Semaphore key?
The shared memory keys are from 27920 (0x6D10 in hex) to 27936 (0X6D20 in hex). The semaphore lock key is 49565 (0xC19D in hex) to 49571 (0XC1A3 in hex). These are not configurable.
- o How should I terminate Proto Balance? Can I use signal 9 (KILL) on Unix?
You should NEVER terminate Proto Balance with signal 9 (KILL or SIGKILL). Termination can be done with any proper terminating signal such as TERM, QUIT or INT (SIGTERM/SIGQUIT/SIGINT). Alternatively, go to the "Info" tab of the Proto Balance configuration web page and click on "Force balancer to shutdown".
- o What is the maximum throughput possible with Proto Balance?
Proto Balance makes efficient use of your OS, however operating systems vary in performance. Proto Balance should saturate your network interface if it is at all possible to do so in user-space for your hardware+OS combination.
At very large numbers of concurrent connections, Proto Balance total throughput will be degraded. This can be quite pronounced on certain operating systems.
- o On Windows I get the error "there is already another application listening on port 8080 on interface ALL..."? What do I do?
This means either that Proto Balance is already running, or that another application is listening on this port. If there is another application using port 8080, you will have to stop it first. Try un-installing any Web Service, Oracle, and application-server software.
Otherwise you can use the -configport option as explained in the Proto Balance top-level window.
- o On Windows I get the error "Only one usage of each socket address..."? What do I do?
You are running out of certain TCP resources. Follow the instructions under "System Settings for High Performance with Your Operating System".
- o Does Proto Balance use any custom kernel modules, kernel plug-ins, tap devices, low-level networking interception?
None. Proto Balance is an ordinary user-space executable.
- o What kind of testing has Proto Co Networking done on Proto Balance?
Besides the performance testing above, Proto Balance has been run for hundreds of hours continuously at network-saturation load. Proto Balance has been verified under memory profiling tools such as Valgrind and Purify. Proto Balance incurred no errors or warnings by either of these tools and had flat OS resource consumption.
The Proto Co Networking team uses an internally developed TCP-scenario automated test harness, and on-site hardware for all supported platforms.
Proto Balance is self-monitoring with over 750 different internal consistency tests.

Portions of this software:

Copyright (c) 1990, 1993, 1994 The Regents of the University of California. All rights reserved.

This code is derived from software contributed to Berkeley by Margo Seltzer.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. All advertising materials mentioning features or use of this software must display the following acknowledgement:
This product includes software developed by the University of California, Berkeley and its contributors.
4. Neither the name of the University nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE REGENTS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE REGENTS OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.