**TCP window scale option**

From Wikipedia, the free encyclopedia

[Jump to navigation](https://en.wikipedia.org/wiki/TCP_window_scale_option#mw-head) [Jump to search](https://en.wikipedia.org/wiki/TCP_window_scale_option#searchInput)

The **TCP window scale option** is an option to increase the [receive window](https://en.wikipedia.org/wiki/Receive_window) size allowed in [Transmission Control Protocol](https://en.wikipedia.org/wiki/Transmission_Control_Protocol) above its former maximum value of 65,535 bytes. This TCP option, along with several others, is defined in IETF [RFC 1323](https://tools.ietf.org/html/rfc1323) which deals with [long fat networks](https://en.wikipedia.org/wiki/Long_fat_network) (LFNs).

**TCP windows**

The throughput of a communication is limited by two windows: the [congestion window](https://en.wikipedia.org/wiki/Congestion_window) and the receive window. The congestion window tries not to exceed the capacity of the network ([congestion control](https://en.wikipedia.org/wiki/Congestion_control)); the receive window tries not to exceed the capacity of the receiver to process data ([flow control](https://en.wikipedia.org/wiki/Flow_control_(data))). The receiver may be overwhelmed by data if for example it is very busy (such as a Web server). Each TCP segment contains the current value of the receive window. If, for example, a sender receives an ack which acknowledges byte 4000 and specifies a receive window of 10000 (bytes), the sender will not send packets after byte 14000, even if the congestion window allows it.

**Theory**

TCP window scale option is needed for efficient transfer of data when the [bandwidth-delay product](https://en.wikipedia.org/wiki/Bandwidth-delay_product) (BDP) is greater than 64K. For instance, if a [T1](https://en.wikipedia.org/wiki/Digital_Signal_1) transmission line of 1.5 Mbit/second was used over a satellite link with a 513 millisecond [round trip time](https://en.wikipedia.org/wiki/Round_trip_time) (RTT), the bandwidth-delay product is 1 , 500 , 000 × 0.513 = 769 , 500 {\displaystyle \scriptstyle 1,500,000\times 0.513=769,500}  bits or about 96,187 bytes. Using a maximum buffer size of 64 KiB only allows the buffer to be filled to (65,535 / 96,187) = 68% of the theoretical maximum speed of 1.5 Mbits/second, or 1.02 Mbit/s.

By using the window scale option, the receive window size may be increased up to a maximum value of 1 , 073 , 725 , 440   ( = ( 2 16 − 1 ) × ( 2 14 ) = 65 , 535 × 16 , 384 ) ) {\displaystyle 1,073,725,440\ \scriptstyle \left(=(2^{16}-1)\times (2^{14})=65,535\times 16,384)\right)}  bytes. This is done by specifying a two byte shift count in the header options field. The true receive window size is left shifted by the value in shift count. A maximum value of 14 may be used for the shift count value. This would allow a single TCP connection to transfer data over the example satellite link at 1.5 Mbit/second utilizing all of the available bandwidth.

Essentially, not more than one full transmission window can be transferred within one round-trip time period. The window scale option enables a single TCP connection to fully utilize an LFN with a BDP of up to 1 GB, e.g. a 10 Gbit/s link with round-trip time of 800 ms.

**Possible side effects**

Because some firewalls do not properly implement TCP Window Scaling, it can cause a user's Internet connection to malfunction intermittently for a few minutes, then appear to start working again for no reason. There is also an issue if a firewall doesn't support the TCP extensions.[[1]](https://en.wikipedia.org/wiki/TCP_window_scale_option#cite_note-1)

**Configuration of operating systems**

**Windows**

TCP Window Scaling is implemented in [Windows](https://en.wikipedia.org/wiki/Microsoft_Windows) since [Windows 2000](https://en.wikipedia.org/wiki/Windows_2000).[[2]](https://en.wikipedia.org/wiki/TCP_window_scale_option#cite_note-2)[[3]](https://en.wikipedia.org/wiki/TCP_window_scale_option#cite_note-3) It is enabled by default in Windows Vista / Server 2008 and newer, but can be turned off manually if required.[[4]](https://en.wikipedia.org/wiki/TCP_window_scale_option#cite_note-4) [Windows Vista](https://en.wikipedia.org/wiki/Windows_Vista) and [Windows 7](https://en.wikipedia.org/wiki/Windows_7) have a fixed default TCP receive buffer of 64 kB, scaling up to 16 MB through "autotuning", limiting manual TCP tuning over [long fat networks](https://en.wikipedia.org/wiki/Long_fat_network).[[5]](https://en.wikipedia.org/wiki/TCP_window_scale_option#cite_note-5)

**Linux**

|  |  |
| --- | --- |
|  | This section **contains** [**instructions, advice, or how-to content**](https://en.wikipedia.org/wiki/Wikipedia:What_Wikipedia_is_not#NOTHOWTO). The purpose of Wikipedia is to present facts, not to train. Please help [improve this article](https://en.wikipedia.org/w/index.php?title=TCP_window_scale_option&action=edit) either by rewriting the how-to content or by [moving](https://meta.wikimedia.org/wiki/Help:Transwiki) it to [Wikiversity](https://en.wikiversity.org/wiki/), [Wikibooks](https://en.wikibooks.org/wiki/) or [Wikivoyage](https://en.wikivoyage.org/wiki/). *(February 2016)* |

[Linux kernels](https://en.wikipedia.org/wiki/Linux_kernel) (from 2.6.8, August 2004) have enabled TCP Window Scaling by default. The configuration parameters are found in the [/proc filesystem](https://en.wikipedia.org/wiki/Procfs), see pseudo-file /proc/sys/net/ipv4/tcp\_window\_scaling and its companions /proc/sys/net/ipv4/tcp\_rmem and /proc/sys/net/ipv4/tcp\_wmem (more information: man tcp, section sysctl).[[6]](https://en.wikipedia.org/wiki/TCP_window_scale_option#cite_note-linuxsysctl-6)

Scaling can be turned off by issuing the following command.

$ sudo sysctl -w "net.ipv4.tcp\_window\_scaling=0"

To maintain the changes after a restart, include the line "net.ipv4.tcp\_window\_scaling=0" in /etc/sysctl.conf (or /etc/sysctl.d/99-sysctl.conf as of systemd 207).

**FreeBSD, OpenBSD, NetBSD and Mac OS X**

Default setting for [FreeBSD](https://en.wikipedia.org/wiki/FreeBSD), [OpenBSD](https://en.wikipedia.org/wiki/OpenBSD), [NetBSD](https://en.wikipedia.org/wiki/NetBSD) and [Mac OS X](https://en.wikipedia.org/wiki/Mac_OS_X) is to have window scaling (and other features related to [RFC 1323](https://tools.ietf.org/html/rfc1323)) enabled.  
To verify their status, a user can check the value of the "net.inet.tcp.rfc1323" variable via the [sysctl](https://en.wikipedia.org/wiki/Sysctl) command:

$ sysctl net.inet.tcp.rfc1323

A value of 1 (output "net.inet.tcp.rfc1323=1") means scaling is enabled, 0 means "disabled". If enabled it can be turned off by issuing the command:

$ sudo sysctl -w net.inet.tcp.rfc1323=0

This setting is lost across a system restart. To ensure that it is set at boot time, add the following line to */etc/sysctl.conf*: net.inet.tcp.rfc1323=0

However, on macOS 10.14 this command provides an error

sysctl: unknown oid 'net.inet.tcp.rfc1323'

**Sources**

 [*"Network connectivity may fail when you try to use Windows Vista behind a firewall device"*](http://support.microsoft.com/kb/934430)*. Support.microsoft.com. Retrieved July 11, 2019.*

  [*"Description of Windows 2000 and Windows Server 2003 TCP Features"*](http://support.microsoft.com/kb/224829)*. Support.microsoft.com. Retrieved July 11, 2019.*

  [*"TCP Receive Window Size and Window Scaling"*](https://web.archive.org/web/20080101144824/http:/msdn2.microsoft.com/en-us/library/ms819736.aspx)*. Archived from* [*the original*](http://msdn2.microsoft.com/en-us/library/ms819736.aspx) *on January 1, 2008.*

  [*"Network connectivity fails when you try to use Windows Vista behind a firewall device"*](http://support.microsoft.com/kb/934430)*.* [*Microsoft*](https://en.wikipedia.org/wiki/Microsoft)*. July 8, 2009.*

  [*"MS Windows"*](https://fasterdata.es.net/host-tuning/ms-windows/)*. Fasterdata.es.net. Retrieved July 11, 2019.*

 [*"/proc/sys/net/ipv4/\* Variables"*](https://www.kernel.org/doc/Documentation/networking/ip-sysctl.txt)*.*