SpeedFusion Whitepaper

How to Get the Most Out of SpeedFusion Technologies

Table of Contents

01

Chapter 1. What is SpeedFusion?

SpeedFusion is Peplink's patented system architecture that delivers unbreakable internet connections, enhanced speeds, and seamless, jitter-free video conferencing.

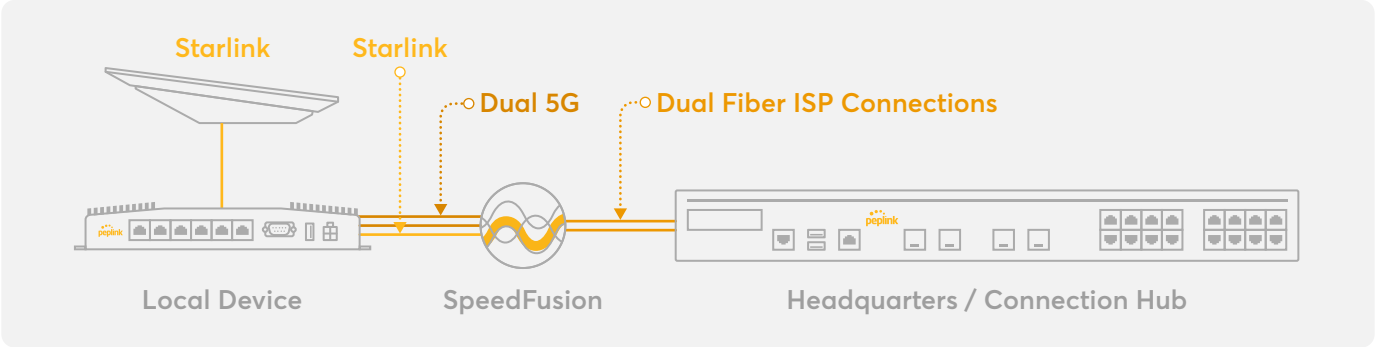
SpeedFusion achieves this by combining an end-to-end architecture with core technology integrated into every Peplink product. It includes software that allows users to configure, define, and prioritize the quality of service for their applications, along with cloud services that are available on-demand or through subscription.

For the purposes of this white paper, we will be discussing the overall architecture and capabilities of the SpeedFusion system.

Chapter 2. Technology Overview

Multiple connections are always better than one. If you need reliable connectivity SpeedFusion can help to achieve it. One of SpeedFusion's most powerful features is that it can use multiple WAN connections to create a single logical VPN tunnel between two endpoints. This allows SpeedFusion bonding to address common problems such as:

- Unreliable networks
- Highly variable and heavily congested wireless networks
- Real-time services that can be very sensitive to packet loss



SpeedFusion protocol uses all available WAN links working together as a single virtual network connection. If a WAN connection fails, SpeedFusion can detect this failure and seamlessly redirect traffic, at a packet level, across other available WAN connections. This WAN connection failure detection at the packet level allows SpeedFusion to provide highly reliable and resilient site-to-site connectivity.

Creating a SpeedFusion connection requires two endpoints. Peplink supports multiple bonding options which are suitable for home users or large organizations. The second endpoint can be:

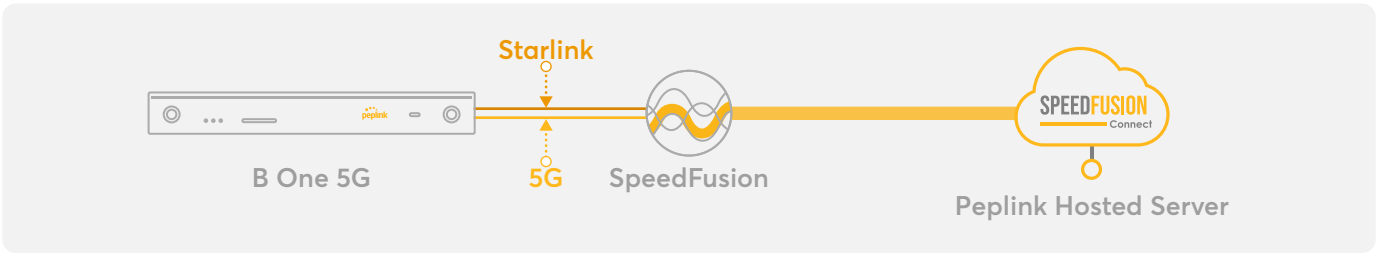
1. Peplink Hosted Server: Available through a Peplink SpeedFusion Connect subscription.
2. SpeedFusion-Compatible Peplink Device: Any Peplink device that supports SpeedFusion.
3. User-Hosted Cloud Instance: A cloud-native virtual instance (FusionHub or FusionHub Solo) managed by users on their own cloud infrastructure.

Peplink SpeedFusion supports the following topology options:

1. Point-to-Point

Three ways to achieve this:

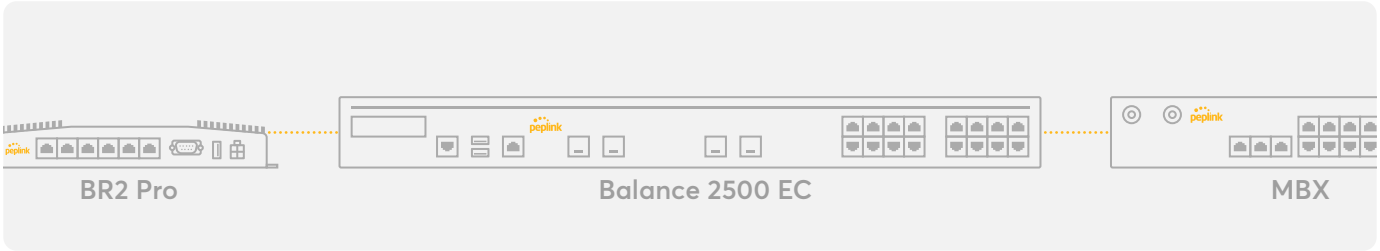
- Available through a Peplink SpeedFusion Connect subscription.
- Private server bonding with FusionHub Solo or FusionHub.
- Connect two devices that support SpeedFusion (a public IP address/hostname is needed).



2. Star Topology

Two ways to achieve this:

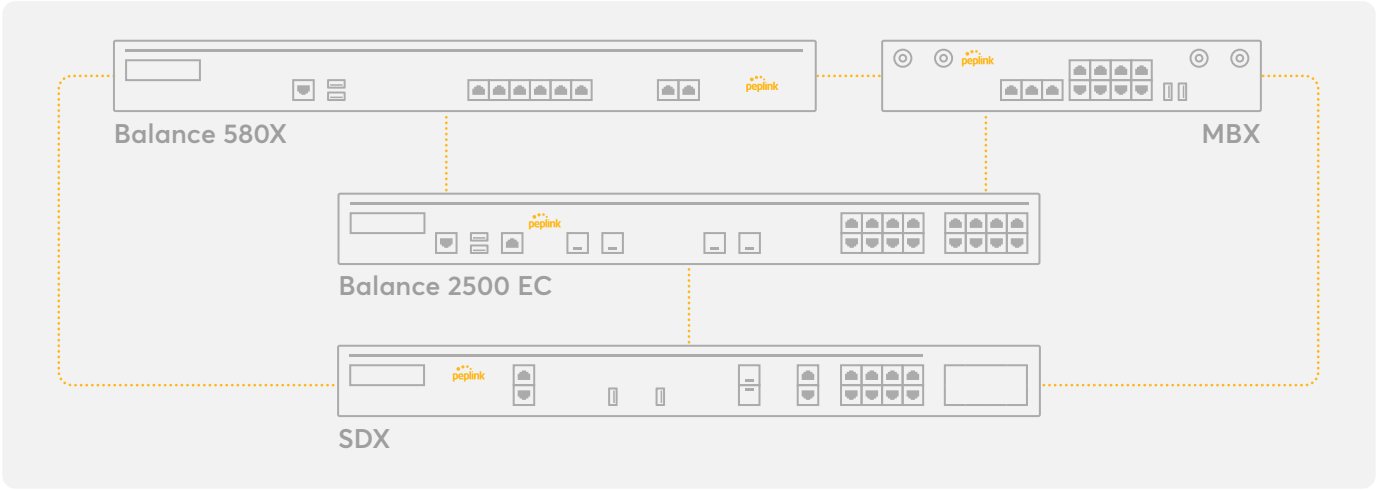
- Private server bonding with FusionHub.
- Connect two devices that support SpeedFusion (a public IP address/hostname is needed).



3. Mesh Topology

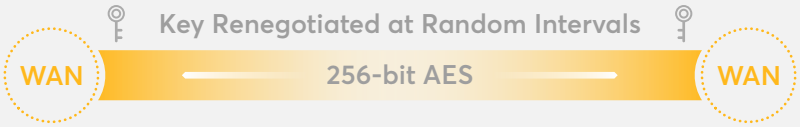
Two ways to achieve this:

- Private server bonding with FusionHub.
- Connect two devices that support SpeedFusion (a public IP address/hostname is needed).

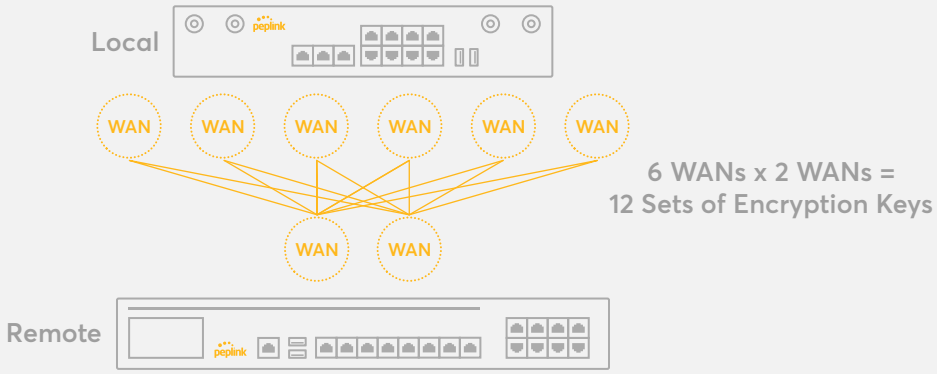


Chapter 3. SpeedFusion Security

In a SpeedFusion connection, a secure WAN-to-WAN link is established using the Diffie-Hellman key exchange protocol and public key cryptography. Perfect forward secrecy is ensured by deriving encryption keys from the exchanged master keys, which are renegotiated at random intervals. The data passing through the WAN-to-WAN links is protected by AES encryption, offering 256-bit security.



SpeedFusion utilizes all available WAN connections by creating WAN-to-WAN links. The number of tunnels depends on the WAN connections at both the local and remote sites. For example, if the local site has 4x 5G and 2x Starlink connections (6 WANs total) and the remote site has 2x Fiber connections, SpeedFusion will create 12 tunnels (6x2).



A single TCP/IP session is distributed across multiple links, with data broken down into packets. Since each WAN-to-WAN connection is encrypted separately, potential hackers would need to obtain the key for every connection before having any chance of accessing the data. This makes SpeedFusion technology virtually impervious to man-in-the-middle attacks.

Chapter 4. SpeedFusion Technologies

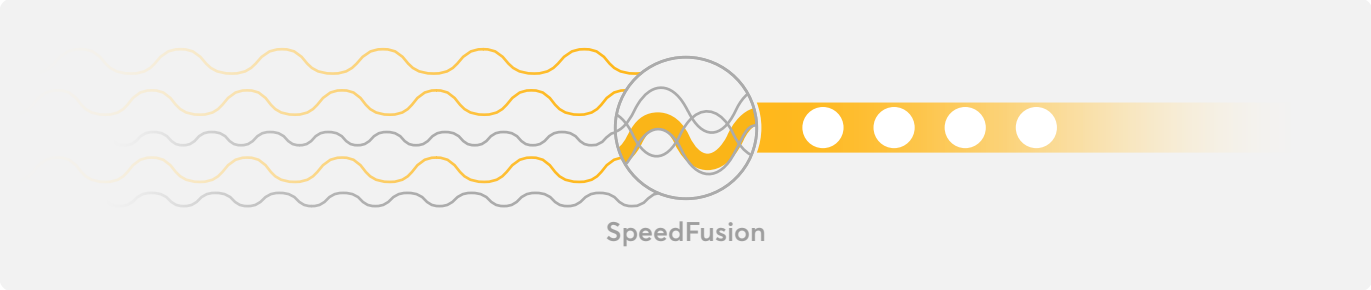
From consumer applications to enterprise branch offices, to remote industrial sites, to vehicular and maritime deployments, SpeedFusion technology requirements may vary. SpeedFusion protocol is versatile and can be optimized for the highest throughput or reliability.

Most common configuration scenarios are:

1. Bandwidth Bonding: Combine the speed and bandwidth of multiple WAN connections for enhanced performance.
2. WAN Smoothing/ Hot Failover: Combine multiple WAN connections to get the best possible latency and jitter-free data streams.

4.1 Bandwidth Bonding

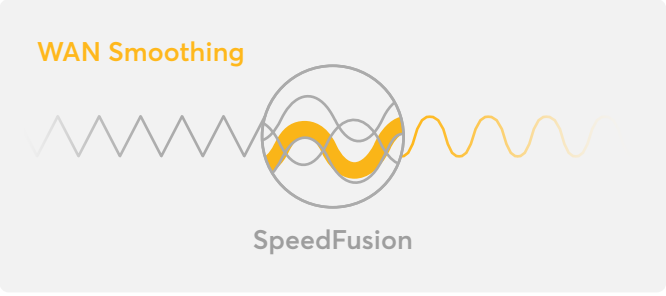
One of SpeedFusion's most powerful features is that it can use multiple WAN links to improve the maximum performance. Other solutions on the market typically offer load balancing while SpeedFusion uses a different approach by combining all WAN connections to create a single aggregated performance virtual WAN (a single logical VPN) connection.

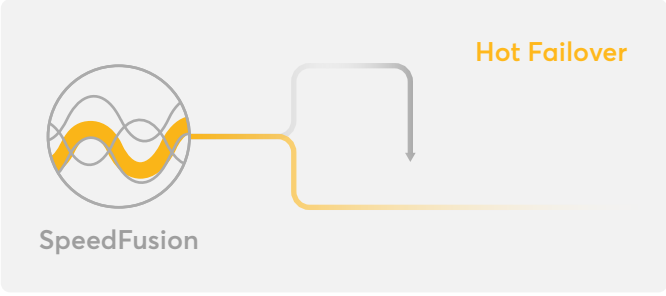


4.2 WAN Smoothing / Hot Failover

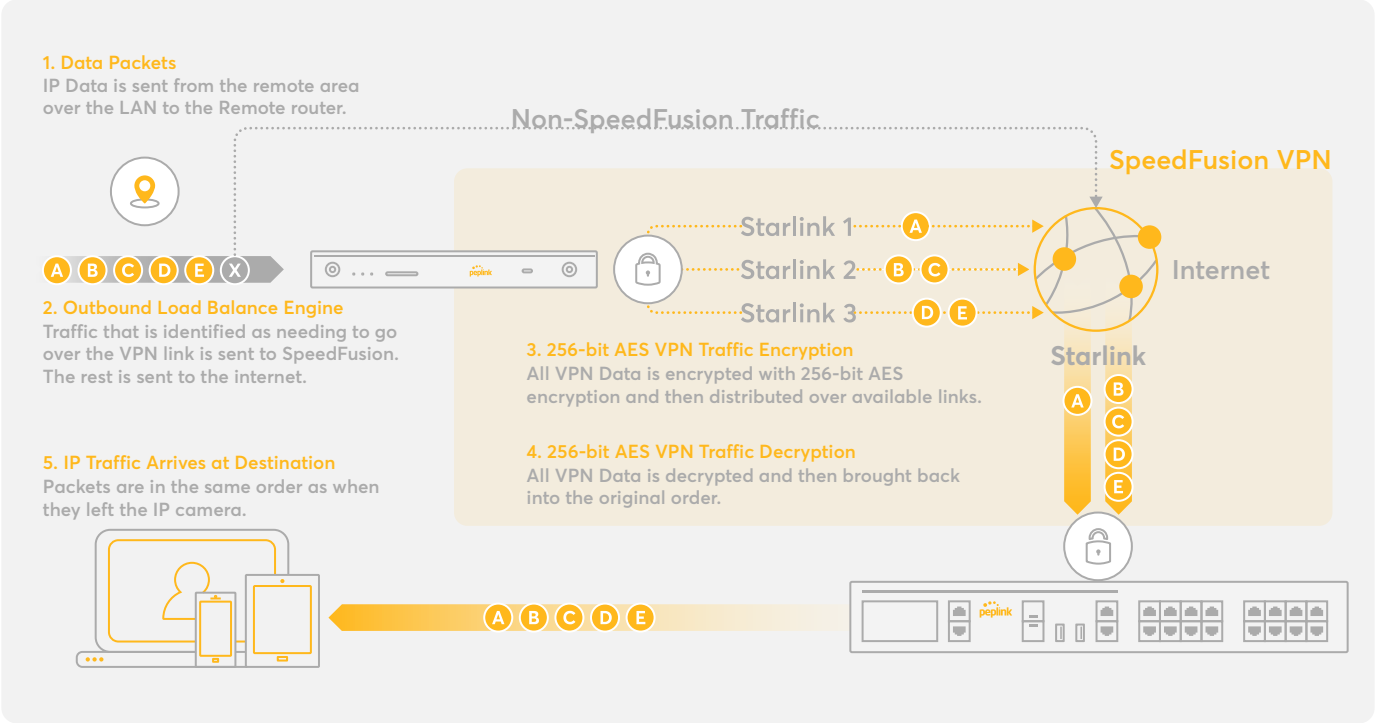
For streaming technologies such as VoIP or videoconferencing, the connection quality strongly influences the user experience. High latency breaks the flow of conversations, while packet loss leads to jitter and missed sentences. WAN Smoothing reduces the impact of packet loss and improves latency in exchange for extra bandwidth consumption via packet duplication.

Below is an example of how SpeedFusion can be used to transmit IP video from a remote location to a central data center over multiple WAN Links.





Below is an example of how SpeedFusion can be used to transmit IP video from a remote location to a central data center over multiple WAN Links.



The amount of WAN Smoothing can be configured based on your deployment's bandwidth sensitivity:

- Normal - The total bandwidth consumption will be at most 2x of the original data traffic.
- Medium - The total bandwidth consumption will be at most 3x of the original data traffic.
- High - The total bandwidth consumption will be at most 4x of the original data traffic.
- Maximum - The total bandwidth consumption depends on the no. of connected active WAN-to-WAN connection.

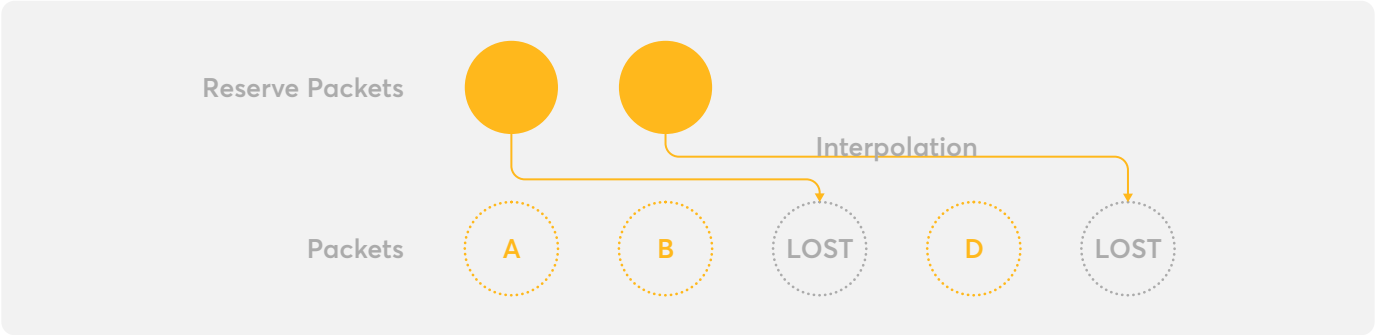
WAN Smoothing can also operate with a single WAN-to-WAN link, doubling the original data traffic consumption.

Additionally, two SpeedFusion peers can maintain session persistence while switching between multiple WAN connections—whether fixed lines, cellular, or a combination of both. This allows authenticated sessions and videoconferences to continue uninterrupted as connections change.

4.3 Adaptive Forward Error Correction (FEC)

This technology has been developed to provide packet loss protection while minimizing bandwidth consumption. While WAN Smoothing duplicates actual packets, Forward Error Correction (FEC) sends additional reserve packets which could be used to mitigate the effects of packet loss via interpolation.

For each SpeedFusion profile, you can either statically assign a Forward Error Correction (FEC) level—Low (13.3% bandwidth overhead) or High (26.7% bandwidth overhead)—or enable the new Adaptive FEC feature. Adaptive FEC dynamically adjusts bandwidth usage, ranging from 6.7% to 20%, to improve redundancy and efficiently recover lost packets.



4.4 Using SpeedFusion Behind a Firewall

By default, SpeedFusion uses TCP port 32015 and UDP port 4500 for establishing VPN connections and transmitting data. If IPsec or L2TP/IPsec services have been enabled, UDP port 32015 is available as well. However, you can change the Data Port assignment in your SpeedFusion profile to another value.

4.5 Using WAN Smoothing and FEC together

WAN Smoothing and Forward Error Correction (FEC) can be used together to enhance network performance. WAN Smoothing mitigates high latency events on a single WAN by duplicating traffic on another WAN link with more stable latency at that moment. A side effect is the same traffic is sent over different WAN connections and consumes data on all connections. FEC addresses packet loss by infilling lost packets on the fly from simultaneously transmitted parity data.

If you have a powerful enough router and sufficient bandwidth, it is recommended to use both WAN Smoothing and FEC simultaneously for optimal performance. The only exception is when streaming video on the move. In such cases, WAN Smoothing tends to work more reliably on its own, as WAN links can be highly variable and volatile from a latency perspective during video transmission.

4.6 Traffic Overflow - Prioritize Affordable Links

SpeedFusion can be configured to prioritize the order of WAN connections, which is particularly useful in deployments where some links are more affordable than others. You can set the cost-effective WANs as the first priority, while more expensive connections will only be used when the primary links reach their configured capacity.

Chapter 5. VPN Overhead Comparison

SpeedFusion traffic requires all transmitted data to be encapsulated in a special UDP stream. This stream contains additional packet headers with all the information needed to reconstruct the original data stream in the correct order at the remote location.

SpeedFusion adds an additional 80 bytes of data to each packet sent over a SpeedFusion connection, no matter what size the original data packet is. This compares well to the 58 bytes of overhead required by IPsec, especially considering that SpeedFusion is not only providing advanced routing and load balancing but 256 bit AES encryption within the tunnel too.

5.1 SpeedFusion and The Internet Mix (IMIX)

Internet Mix (IMIX) is a measurement of typical internet traffic passing through network equipment, such as routers, switches, or firewalls. When measuring equipment performance using an IMIX of packets, performance is assumed to resemble what can be seen in real life.

The IMIX traffic profile is used in the industry to simulate real-world traffic patterns and packet distributions. IMIX profiles are based on statistical sampling done on internet routers. More information about IMIX can be found here:

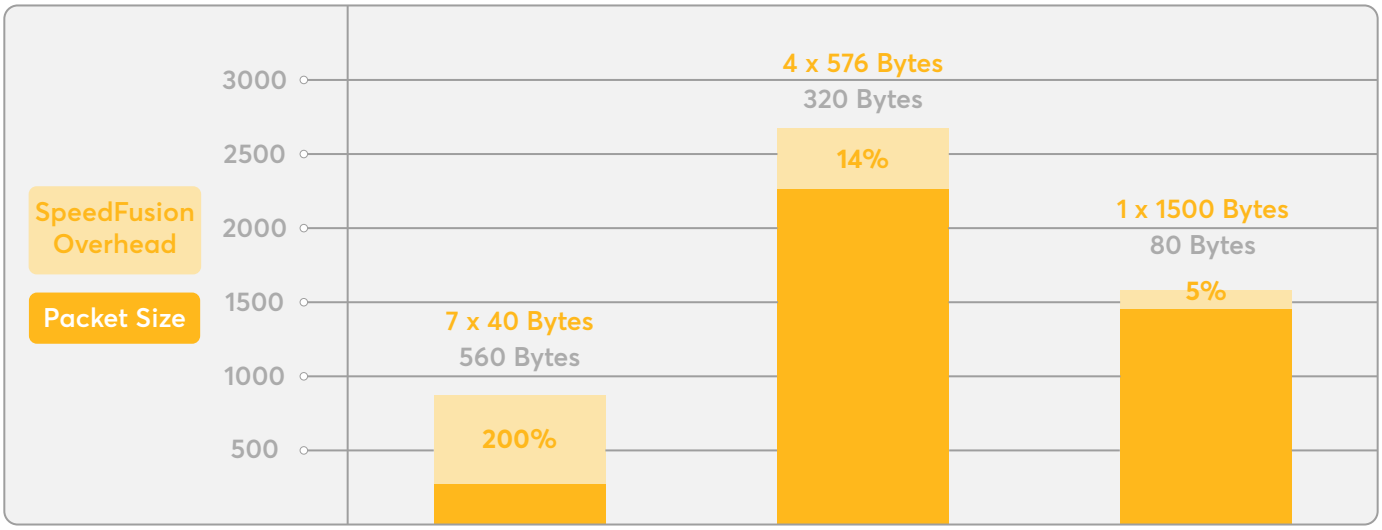
Packet Size	Packet #	Distribution in Packets	Bytes	Distribution
40	7	58%	280	7%
576	4	33%	2304	56%
1200	1	8%	1500	37%

A packet-level breakdown of the IMIX Standard. Total of 4084 Bytes (source: https://en.wikipedia.org/wiki/Internet_Mix).

5.2 The IMIX Standard and SpeedFusion Overhead

As the chart on the left shows, when a SpeedFusion VPN tunnel is used to transmit IMIX data (4084 bytes), an additional 960 bytes of SpeedFusion overhead is required.

The SpeedFusion overhead is 19% of the total transmitted data (IMIX + overhead). Since it uses a fixed number of bytes per packet transmitted (an additional 80 bytes), SpeedFusion Bandwidth Bonding is much more efficient when transmitting larger packet sizes. At packet sizes of 1500 bytes, SpeedFusion adds just 5% bandwidth overhead, but at packet sizes of 40 bytes, SpeedFusion overhead rises to 200%.



Effect of SpeedFusion Bandwidth Bonding overhead against the different packet sizes and quantities of the IMIX standard

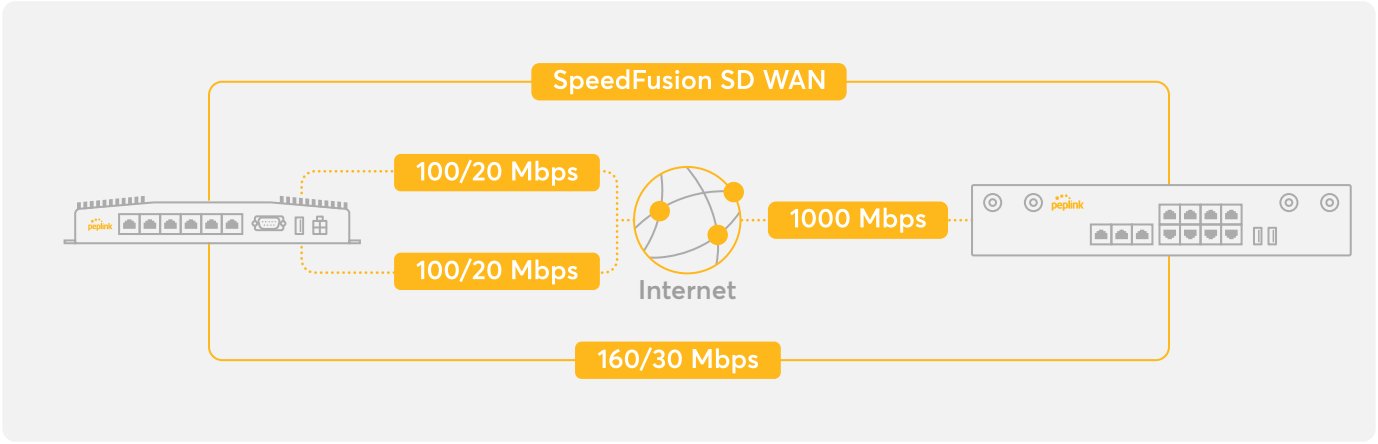
SpeedFusion (80 Bytes)
7x80 + 4x80 + 1 x80
(7x80 + 4x80 + 1 x80) = 19%
+
(280 + 2304 + 1500)

IPsec (58 bytes)
7x58 + 4x58 + 1 x58
(7x58 + 4x58 + 1 x58) = **14.6%**
+
(280 + 2304 + 1500)

In Summary...
For only 4.4% of additional overhead compared to IPsec, SpeedFusion includes bandwidth aggregation & WAN resilience.

Chapter 6. Calculating Bonding Overhead

SpeedFusion is often deployed to new customers who have just a single internet connection but want more bandwidth. So, they buy an additional connection with the intention of using SpeedFusion to bond the two links. Without understanding SpeedFusion bandwidth overhead, a user might be confused by bandwidth availability using SpeedFusion bonding across links. Consider the following configuration:



Accounting for SpeedFusion bandwidth overhead and assuming the traffic passing across the links is similar to the IMIX standard mentioned previously, we can calculate available real-world bandwidth at the remote site:

Download: 100Mbps + 100Mbps = 200Mbps x (1 - 19%) = 162Mbps
Upload : 20Mbps + 20Mbps = 40Mbps x (1 - 19%) = 32.4Mbps

It is important to explain SpeedFusion bandwidth overhead to your end users so that they understand why they will not get full 200Mbps/40Mbps bandwidth when using VPN bonding. Remember, **conventional VPN technology such as IPsec has an overhead of 14.6%. For only 4% of additional overhead, SpeedFusion provides bandwidth aggregation & WAN resilience.**

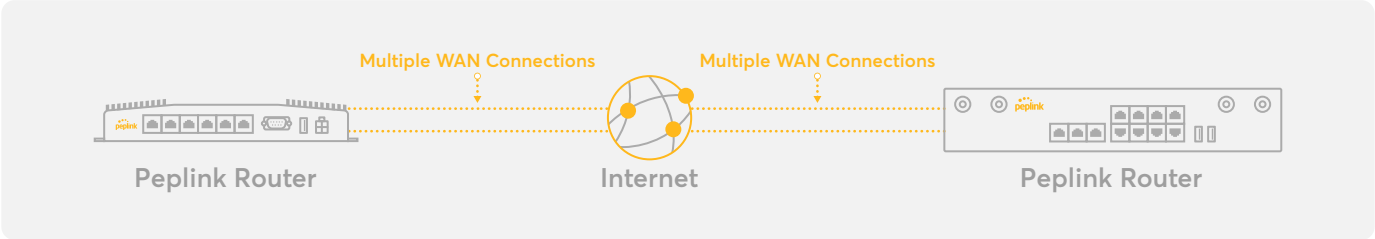
6.1 Bonding different WAN links

For the highest performance, we recommend the use of WAN links with similar bandwidth profiles (within 50% of each other) and similar latency characteristics (we recommend within 150ms of each other). Use WAN links from different ISPs to allow for the best possible SpeedFusion throughput. Using at least two different ISPs offers the benefit of provider diversity, which means less chance of a technical (or even accounting/billing) error causing a network outage. Provider diversity also lessens the impact of bandwidth sharing, a common problem when using multiple circuits from a single provider.

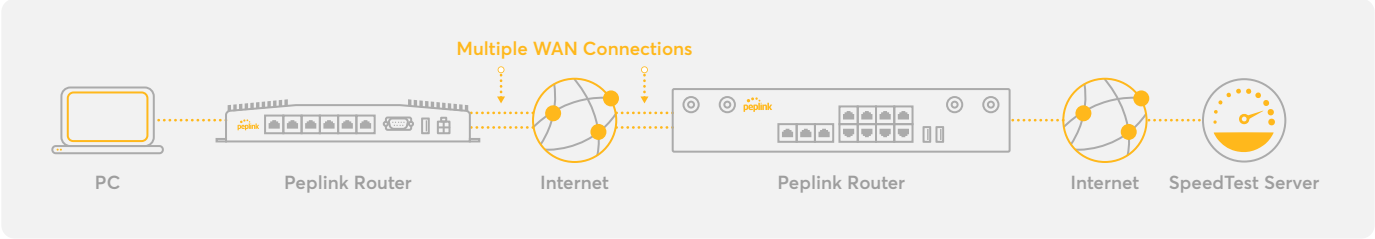
6.2 Measuring SpeedFusion tunnel capacity

When testing SpeedFusion, it is crucial to evaluate the throughput capacity between the SpeedFusion tunnel endpoints with and without SpeedFusion to make sure there is no bottleneck.

Below is a simplified topology:



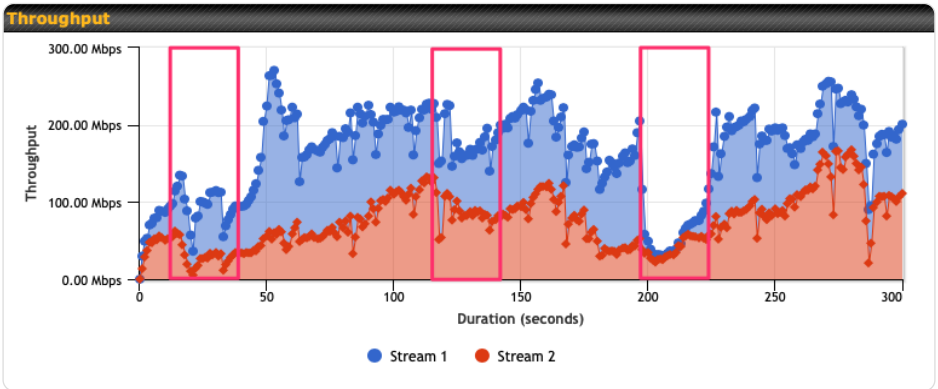
End users very often use SpeedTest to check the performance. The simplified setup is shown below:



We recommend using the WAN Analysis tool to measure performance between A and B points without SpeedFusion. Then use SpeedFusion VPN Test tool to test with SpeedFusion enabled.

By using both tools it allows to compare the performance with and without the SpeedFusion over the two endpoints. When using WAN analysis it is important to run throughput via all WAN simultaneously. Also another important factor is the TCP session number which is critical when testing over lossy wireless WANs such as 5G/LTE or Starlink/LEO..

Another important consideration is the duration of testing, which is especially critical for wireless networks like 5G/LTE and Starlink/LEO. In many cases, performance is only measured for 20-30 seconds, which can lead to inaccurate results due to significant fluctuations in performance, as shown in the image below.

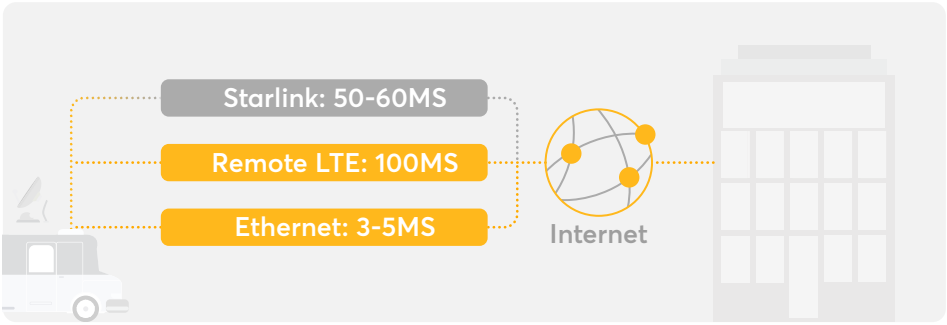


The example below shows a testing duration of 300 seconds or 5 minutes. If testing is conducted for only 20 seconds, it's difficult to account for the significant bandwidth variations that can occur on LEO, cellular, and other wireless networks.

From our observations, testing for 5-10 minutes provides a better average and more accurate performance results, allowing you to capture these fluctuations.

6.3 Addressing Latency and Packet Loss

SpeedFusion operates most optimally when the latency across connected WAN links is similar. If there are significant differences in latency between the WAN links, throughput can be negatively affected.



In the example above, all the connections are highly susceptible to packet loss. Because the latency across the SpeedFusion link is equalized to the highest latency link (50-60ms), if a packet is lost on the Ethernet link, SpeedFusion will take longer to spot the packet loss (50-60ms+).

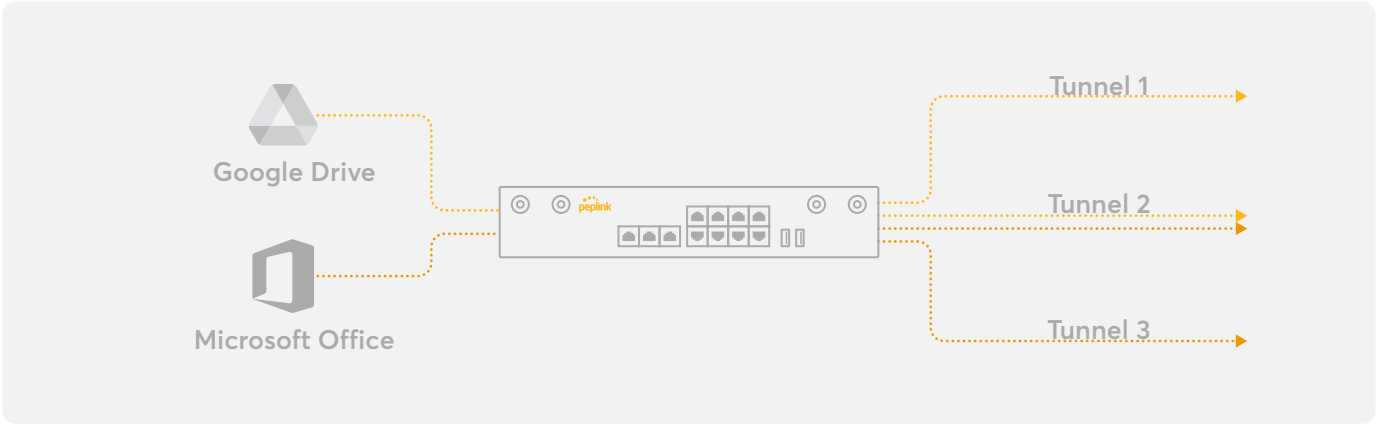
In cases of high variation in WAN link latency, the best approach (assuming there is enough bandwidth on low latency links) is to allocate lower latency links for SpeedFusion while setting higher latency links as failover connections.

Another approach is to use higher latency links for specific direct traffic types that are not as latency sensitive (like direct internet access) while reserving lower latency links for other important corporate traffic that needs to transit the VPN, such as live streaming, video conferencing, and online gaming.

Note: Using UDP traffic over SpeedFusion can provide higher throughput than TCP which has restrictive flow control and is very sensitive to packet loss.

Chapter 7. Outbound Policy - Traffic steering

Today's networks handle a diverse range of traffic, including video streams, VoIP calls, online gaming, cloud applications, WEB browsing and more. Depending on deployment needs, it might be helpful to prioritize specific types of traffic while throttling or blocking others to ensure optimal performance.. SpeedFusion gives you granular control of how different kinds of traffic travels within the network.

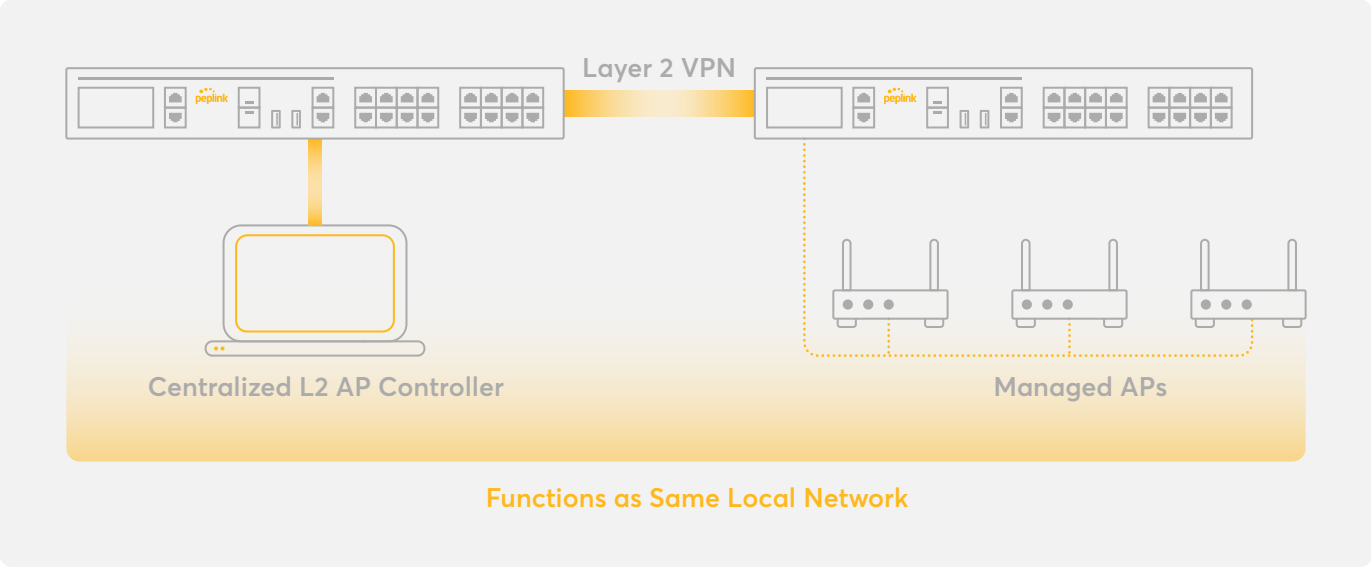


This is achieved through creating multiple SpeedFusion profiles within the same router. Each profile may use different WAN connections, have different bandwidth limits, as well as different configured WAN smoothing and FEC. Once these tunnels are defined, you can then use outbound policy to steer traffic based on protocol, ports, application, client type, etc.

Peplink has support for Deep Packet Inspection (DPI), which enables routers to scan passing packets and recognize traffic from various applications (such as Skype, Google Hangouts, Spotify, Dropbox, and BitTorrent). This information can then be used to define firewall rules, define outbound policy, and assign SpeedFusion Sub-Tunnels.

Chapter 8. Multi-Site Layer 2 VPN

By forming a SpeedFusion Layer 2 VPN, you can get devices from different sites to operate as if they were on the same local network. This technology is useful for remotely operating devices such as printers and access points, sharing network resources such as NAS drives, as well as facilitating server-to-server communications.



A Layer 2 VPN can be applied to the VLAN trunk or a specific VLAN. If it's' applied to the trunk, then all traffic from the router will go through the VPN and will be considered as LAN traffic by the remote site. Alternatively, if you define the Layer 2 VPN on a specific VLAN, then only traffic from that VLAN will be considered local traffic by the remote site.

Chapter 9. Cellular Application Considerations

Whatever WAN connections you are using, it is always a good idea to test each individually and repeatedly to discover its maximum throughput in both directions. Remember, bandwidth availability can vary throughout the day, especially if using cellular or fixed lines with variable contention.

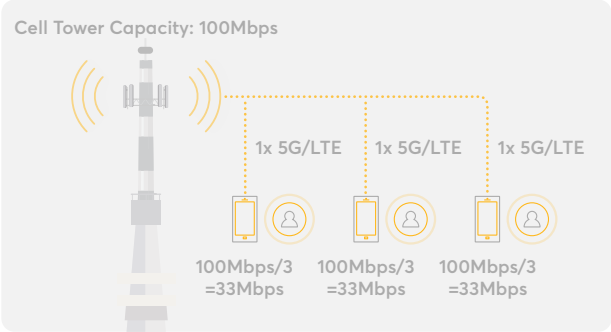
Customers are often surprised by how bandwidth availability can vary between links from the same ISP and how different actual bandwidth availability can be from that advertised by the ISP.

The amount of bandwidth available on a 5G/LTE connection is dependent on a number of factors:

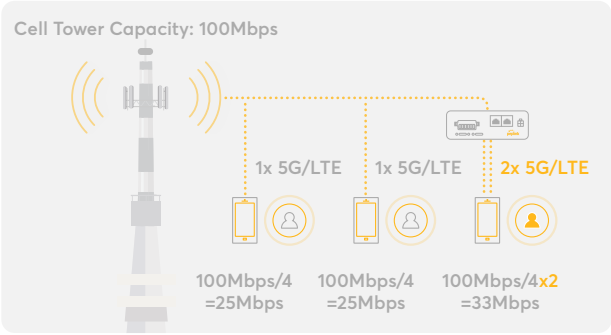
- Signal Strength - Determined by the distance to the nearest cellular tower and the subsequent signal quality received.
- Backhaul Bandwidth Availability – from the cellular tower to the ISP's core network.
- Congestion – Each cellular tower can serve a limited number of users. When too many users attempt to connect to a single cellular tower simultaneously, it affects performance.

9.1 Windowing / time-slicing

Windowing or time-slicing is a technique used by cellular carriers to manage and optimize the allocation of network resources. Each base station has a limited backhaul bandwidth to the ISP's backbone which under high traffic load can get saturated. This is important as the maximum router performance can be limited even if the cellular connectivity signal levels are good.



In the example picture below, three users with three 5G/LTE-enabled devices connect to a cell tower. Each gets 33Mbps, which is a third of the available bandwidth at the tower.



In the next example, the third user is using a Peplink router and has installed an additional 5G/LTE data SIM. There are now three devices with four 5G/LTE connections connected to the cell tower. The first two users get one quarter of the available bandwidth (25Mbps), but the third user gets two quarters (or half) of the available bandwidth using his two 5G/LTE connections.

Note: we frequently see users who have 100% signal strength yet get only a small percentage of the bandwidth. or additional Starlink/LEO connection.

Chapter 10. Starlink/LEO Applications Considerations

Starlink and other Low Earth Orbit (LEO) satellite connectivity solutions share some similarities with 5G/LTE cellular connectivity, despite operating in different environments - space versus terrestrial. The wireless links in Starlink/LEO are predominantly vertical, communicating between ground user terminals and orbiting satellites, while 5G/LTE links are primarily horizontal, connecting mobile devices to nearby cell towers. Both technologies are influenced by common performance factors such as signal strength and base station congestion. Each technology has its strengths and ideal use cases, with Starlink/LEO excelling in global coverage and remote access, and 5G/LTE providing low latency and high mobility in well-developed urban areas.

The amount of bandwidth available on Starlink/LEO connection is dependent on a number of factors:

- Signal Strength - signal strength is affected by satellite positioning, weather conditions, and obstructions, with proper terminal alignment being crucial.
- Congestion - it occurs when many Starlink/LEO users share the same satellite for connection. Each satellite has a limited backbone bandwidth.

Before you do any tests with SpeedFusion, we suggest measuring the throughput capacity. It is important to run a throughput test via Starlink/LEO connections simultaneously.

Very often we notice users are testing Starlinks one by one and then try to sum their performance, which is not the suggested approach.

Another important aspect of Starlink is that its satellites are continuously moving, requiring the user terminal to reconnect to a new satellite periodically (e.g., every 15-20 seconds). This can result in fluctuations in performance metrics such as throughput, latency, and packet loss, as traffic may be routed differently with each new satellite connection. Users with a single Starlink/LEO terminal dish often experience significant variations in upload bandwidth and latency over time, presenting a challenge. Adding a second dish and attempting to bond them introduces additional complexities, as the links may have different performance characteristics at any given moment. For example, one Starlink/LEO connection might have high latency while the other maintains lower latency, and packet loss can be highly variable. This can cause issues with SpeedFusion Dynamic Weighted Bonding, which may struggle to prioritize the lowest latency and packet loss paths effectively.

To mitigate these challenges, we recommend enabling Ignore **Packet Loss Event** in the SpeedFusion configuration (check chapter 12) and using Forward Error Correction (FEC) when bonding Starlink connections. This approach reduces jitter and packet loss, resulting in a more stable link and significantly improved user experience.

Chapter 11. Fixed Line Application Considerations

Most internet connections are provided as a contended service. This means that although your provider has advised you will get up to 1000Mbps broadband connectivity, depending on how oversubscribed your service is (literally how many people in your area are connected to the ISP's service), the bandwidth that's actually available at any given moment could be considerably less.

The amount of bandwidth available on a contented service can vary considerably over the period of a day (and even minute to minute), with higher speeds possible during working hours compared to the evenings when your neighbours are home and using the same internet service heavily.

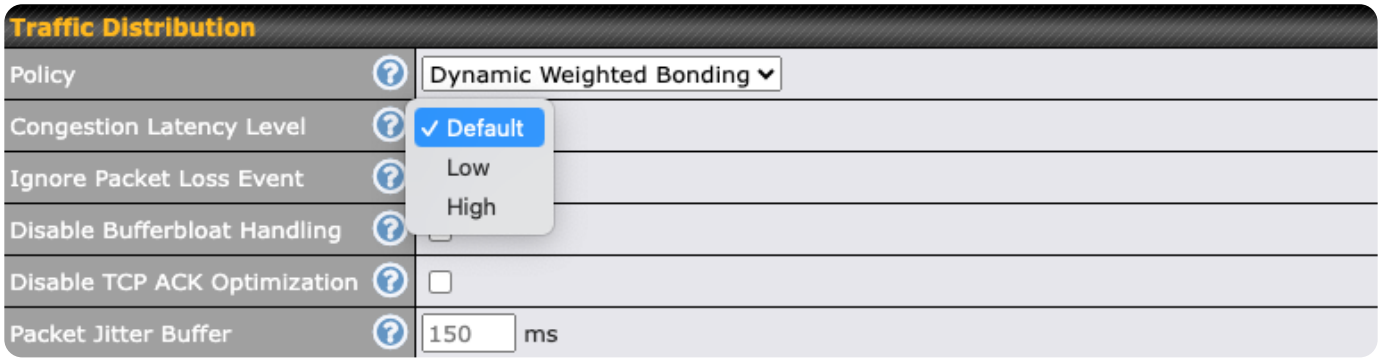
Adding an additional fixed line service from the same ISP can give you a 'bigger share' of the bandwidth that's available. 1 :1 contended services are available from ISPs to counter this issue but are naturally more expensive than 20:1 or 50:1 services.

Chapter 12. Fine tuning Dynamic Weighted Bonding algorithm

The biggest challenge with bonding with wireless connections is a constantly changing Wireless WAN connectivity performance. This may include changing throughput, jitter, packet loss variation, etc. To address all these challenges starting from firmware 8.2.0 Peplink introduced a new bonding algorithm - Dynamic Weighted Bonding (DWB). This new algorithm is adopted to dynamically changing WAN connection quality and helps to improve bonding performance.

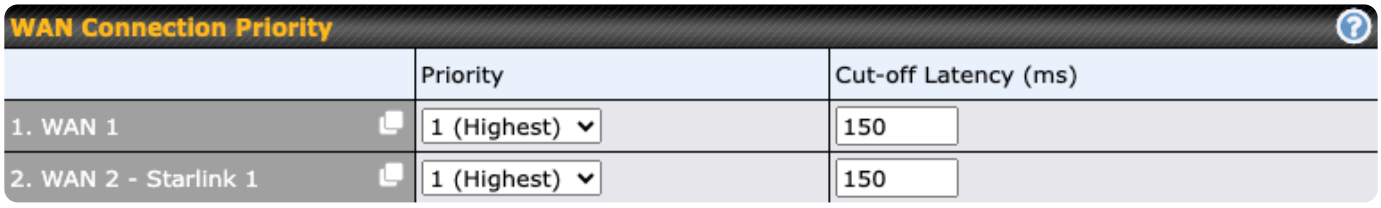
In most cases the DWB algorithm will adapt very well for changing WAN performance, but for certain cases we added additional options to fine tune the performance.

1. Congestion Latency Level



DWB treats a WAN as congested based on the latency change (compared to the idle latency). For example, when a WAN's latency is 30ms, DWB by default will treat it as congested when latency is higher than 60ms. However, this behavior may vary for different ISPs. Setting this option to Low will reduce the latency threshold (i.e., 45ms), and setting it to High will tolerate higher latency (i.e., 75ms).

If the above dynamic latency threshold doesn't fit, you can also define a hard threshold for each WAN in the Cut-off Latency field. In the above screenshot, WAN 1 and WAN 2 will be treated as congested when latency is higher than 150ms.



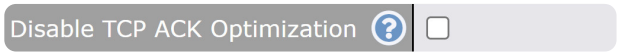
2. Ignore Packet Loss Event

If your WAN has a known packet loss problem and you don't want DWB to treat it as congestion, you can enable this option. Any packet loss on that WAN will be treated as normal and DWB won't reduce traffic sent to it.

3. Disable Bufferbloat Handling

DWB can monitor latency change and avoid bufferbloat issues by dropping packets before the latency gets worse. However for some scenarios you may prefer higher latency instead of packet loss, for example if you are doing video streaming, then enabling this option will disable the behavior.

4. Disable TCP ACK Optimization



DWB can duplicate TCP ACK packets, which are small in size so they won't consume a lot of bandwidth. However, duplicating them will guarantee a more reliable packet retransmission in a lossy tunnel. If you want to reduce this extra bandwidth consumption, you can enable this option then TCP ACK packets will not be duplicated.

5. Packet Jitter Buffer

This is a dynamic buffer that will try to reduce the packet jitter, default is 150ms which means a packet may be delayed for maximum 150ms (when required). Setting this value to 0 will disable the buffer. In the field trial we found that this buffer can help to improve bonding performance a lot when there are a lot of Wireless WANs combined (like more than 4 WANs), the default 150ms should be an optimized value when all WANs have similar latency, but if the latency difference is large (e.g. WAN1 is 30ms and WAN2 is 250ms), increasing this buffer may help to improve performance.

Note: modifying the value of "Packet Jitter Buffer" will disable "Receive Buffer", these two features cannot be enabled at the same time.