appia the cloud at work®

Ensuring the Quality of Real-Time Traffic

Contents

Summary

The main reason for real-time traffic issues is *oversubscription*; that is, the bandwidth available is not sufficient for the demand. Oversubscription leads to poor user experience: audio and video calls sound and look choppy, and even disconnect.

The best way to ensure the quality of real-time traffic is to use separate connections for real-time traffic and all other traffic, and to choose a provider with direct connections to diverse internet service providers.

Definitions

Packets

Packets are pieces of information that are sent across your internet connection. For example, when you visit a website, the information on that site is sent to you in packets, which your computer then reassembles into a display of the web page.

Packet Loss

This is a packet that does not reach its destination. Packet loss causes voice and video calls to sound "choppy."

Latency

Latency refers to the time it takes for packets to get from one point to another over the internet. Ideally, latency is zero. For real-time traffic, it needs round-trip latency needs to be no more than 30 milliseconds.

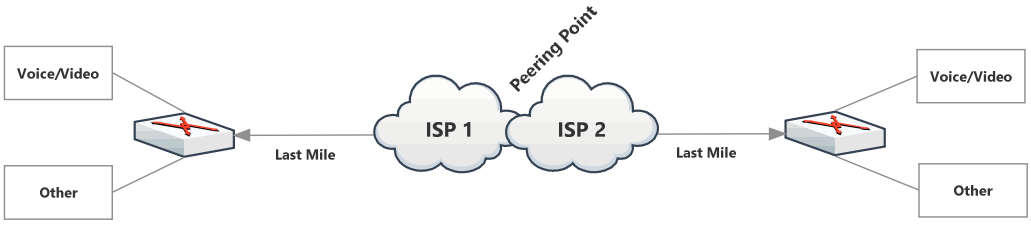
Jitter

Jitter is the difference between the minimum and maximum latency occurring in a series of packet transmissions. Jitter causes "clicks" or buzzing sounds during a phone call.

Oversubscription

This is the main cause of packet loss, delay, and jitter. Oversubscription can occur regularly when there is not enough bandwidth for regular use. It can also occur intermittently; for example, when someone is uploading or downloading a large file or streaming content.

Let's begin by taking a look at how the internet works. The drawing below shows an example of the path packets take from senders to receivers over an internet connection.



As the drawing indicates, the sender's traffic travels:

- Over the sender's data or voice/video network
- Up the "last mile" -- the connection between the sender and the Internet Service Provider (ISP)
- From one ISP to another via peering points
- Down the receiver's last mile
- Over the receiver's data or voice/video network.

Note that in the example, there is one connection to the ISP for both voice/video and other traffic. This is typical in most organizations today. This approach is problematic for the reasons discussed below.

Causes of Oversubscription

Unless there are widespread network outages or severe malfunctions, oversubscription mainly occurs at:

- The ISP's shared connection
- The peering points
- The last mile.

The ISP's Shared Connection

Cable and DSL subscribers share connections to an ISP. The subscriber's last mile might not be oversubscribed, but the connection all of the users share can be. This is becoming more common as the internet is being used for video.

As cable and DSL providers upgrade their networks to fiber, this issue will be less common, provided that the capacity they install keeps pace with internet use.

Peering Points

Most often, senders and receivers are on two networks. In the case of a conference call or video conference, they may be on three or more networks. So the ISPs involved must pass traffic to one another. This is called "peering."

Peering involves an enormous amount of internet traffic, which requires very powerful routers and continuous upgrading of router capacity as internet traffic grows. If the owner of one or more of the networks decides not to invest in such upgrades, the peering point can be so oversubscribed that internet traffic is be delayed as the traffic waits for outdated routers to process it.

The Last Mile

How does oversubscription of the last mile happen? In addition to the traffic one sender is generating, others in the organization are also sending and receiving traffic, the total of which exceeds the connection's capacity. As noted, last mile oversubscription can be continuous or intermittent, as when users upload or download large files or stream video.

Solutions

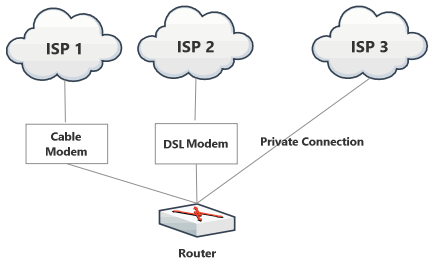
There is little that can be done if the ISP's shared connection is oversubscribed, short of waiting for the ISP to upgrade its infrastructure.

Some service providers, such as Appia, have invested in direct connections to various ISPs, including cable and DSL providers, to address this issue. Our experience is that doing goes a long way to minimizing peering point issues.

The remaining solutions concern the last mile.

Multiple Internet Connections

One solution is to have three (or more) internet connections.

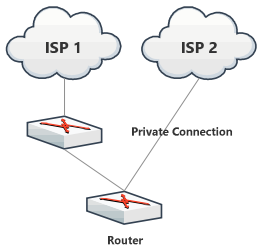


- One connection is for real-time traffic
- The second is for all other traffic
- The third is for fail-over in the event one of the other two connections fails.

Each of these connections should be to a different ISP, and ideally the real-time connection would be a private connection to your cloud service provider.

This option may seem costly. However, with the internet connection options available today, it can be cost-effective, especially in light of its benefits for user experience.

There is also an option to have two internet connections – one for real-time traffic and the other for all other traffic, with failover in the event one of the connections goes down.



However, this solution runs the risk of poor user experience if both real-time and other traffic are forced to share the same connection.

Single Internet Connection with Quality of Service

The final option is to have one internet connection with quality of service (QoS) enabled. This option has the risks associated with a connection outage. However, with QoS, real-time traffic has priority over other traffic, so even if the connection is degraded, the impact on real-time traffic will be minimized.

Single Internet Connection with Overprovisioned Bandwidth

The last option is to have a single connection that has more bandwidth available than both real-time and other traffic require.

This option has the risk of a connection outage, of course. It also requires monitoring to ensure that there is a sufficient cushion, especially at peak times. Even then, large bursts of traffic can still create temporary oversubscription.

It also requires development and enforcement of internet usage policies; e.g., forbidding content streaming, etc. When we have reported how bandwidth is being used, customers are often surprised at how much bandwidth is consumed by Netflix and other high-bandwidth applications.

Frequently Asked Questions

We have run speed tests, which indicate that we have enough bandwidth.

Speed tests are helpful, but they are valid only for a moment in time. We have a simple tool you can use to measure bandwidth over a more sustained period so that you can see the effects of bursts in traffic.

We called our provider and they said that our bandwidth is adequate.

Providers tend to focus on total bandwidth, not on how the bandwidth is being used or variations in bandwidth usage during the day. The bandwidth you have may indeed be adequate for non-real-time traffic; as noted, you will rarely notice the effects of oversubscription (packet loss, delay and jitter) when sending or receiving email, but you will notice it in voice and video.

Also, providers are looking at the service they provision for you over the last mile. They are generally not forthcoming in regard to oversubscription or peering points. It is also important to remember that cable

and DSL providers offer services on a “best efforts” basis. This is why they almost always say that the upload and download speeds are “up to” whatever values they advertise.