SIGNIANT

The Bandwidth Conundrum

Why a Bigger Pipe Alone Doesn't Lead to Faster File Transfers and How to Get the Most Out of Your Network

Fast file movement is vital for success in today's M&E industry. More files, bigger files, and "I-need-it-now" files push delivery boundaries every day. And with higher resolutions, new distribution platforms and content globalization, these challenges are growing exponentially.

For many, the first solution considered when file transfers become an issue is to buy additional bandwidth. After all, wouldn't a bigger pipe mean faster file movement?

As it turns out, there's much more to it.

For more than 15 years, M&E organizations have relied on Signiant's intelligent file transfer software to move petabytes of high-value content every day, from businesses of all sizes, between all storage types, and across a wide variety of networking environments. Through this experience, we've learned a lot about the variables that impact transfer performance and how to help companies move any type and size of file under any condition.

Of course, bandwidth is important, but before you invest in a bigger pipe, here are some important considerations to help you understand what impacts large file transfers, and some tips to help get the most out of your network.

❯ Bandwidth vs. Throughput

First, network bandwidth and network throughput are not the same. Available network bandwidth determines the potential maximum speed that data can move, whereas throughput is the actual speed at which the data moves.

When moving media files over long distances or congested networks without the right software, throughput can be dramatically lower than bandwidth because standard internet transfers use single-stream TCP (Transmission Control Protocol), which (as discussed below) has several significant limitations. Upgrading a network without acceleration software is like going from a country road to a major highway while driving a scooter; one could go super-fast if their transport mechanism allowed it, but as it stands they're just going the same speed in a wider lane.

SIGNIANT

As such, many organizations will spend money on additional bandwidth, then become frustrated when their transfer speeds only increase marginally.

❯ TCP vs. UDP

To understand why this is, let's get into the tech weeds a bit and look at the underlying protocols commonly used to move files over the internet.

TCP (Transmission Control Protocol) is what provides a reliable stream of data from one point to another during standard file transfers over the internet. For the large majority of internet traffic, TCP works fine. But for large files and data sets — especially when sent over long distance — TCP breaks down.

One fundamental problem with traditional TCP is that it uses a relatively unsophisticated sliding window mechanism, only sending a certain amount of data over the network before it expects an acknowledgement of receipt on the other end. As TCP receives acknowledgements, it advances its window and sends more data. If the data doesn't get through or an acknowledgement is lost, TCP will time out and retransmit from the last acknowledged point in the data stream.

There are a number of problems with this, such as retransmitting data that may have already been received, or long stalls in data sent while waiting on acknowledgements. Modern versions of TCP have addressed these challenges in different ways, including scalable windows.

Scalable window sizes allow the amount of data in flight to be greater than the original 64KB or 32KB supported by the protocol. This means that a system administrator can configure TCP to have a bigger window size and most systems today do so by default. Still, TCP controls a stream of data between two endpoints and will only send a limited amount before pausing the server to wait for acknowledgement that data was received on the other end. As such, this sliding window mechanism creates a lot of back and forth, with associated latency for every roundtrip.

SIGNIANT

UDP (User Datagram Protocol) was originally developed to send messages or datagrams over the internet on a best effort basis, making standard UDP a less reliable mechanism for transferring data. While it's ultimately faster than TCP, this is largely because it doesn't take the time to establish a connection, check for errors, and order packets as they are sent. Rather, UDP simply sends packets in random order and doesn't offer flow control, meaning that a packet will just be dropped if it's not received.

Modern software, as discussed below, is able to bring together the speed of UDP with the reliability of TCP along with several other benefits to performance and reliability.

Network Latency, Packet Loss & Congestion

Network latency refers to the time it takes for a packet of data to make the journey from one system to another. A number of things impact latency, but the largest factors are the actual distance the data is travelling and the amount of congestion on the network. Because standard TCP transfers small amounts of data before pausing and waiting to receive acknowledgment from the destination system, each roundtrip for a given transfer accumulates latency very quickly.

As a result, bandwidth or connection speed can be a misleading number. For example, a 1Gb/s pipe might seem like it would transfer one gigabit every second, but that's only if the network is completely clear — like a highway without traffic — and if the transfer software in use can take advantage of the available bandwidth, which is not the case with standard TCP. With standard TCP, adding some traffic (congestion) slows everything down. Add a lot of traffic and you get total gridlock.

SIGNIANT

Any activity on the network contributes to filling up bandwidth and slowing down data movement. With traditional protocols like TCP, data traveling short distances uses disproportionately higher amounts of bandwidth compared to data travelling long distances on the same network. Although bandwidth and latency are independent factors, when one combines high latency and high bandwidth a number of problems emerge that make it difficult for standard protocols to use all of the bandwidth. TCP, for example, only utilizes a fraction of the available bandwidth as to not overwhelm a network. You need the right highway and the right vehicle.

Handling Large Data Sets with Millions of Smaller Files

In media, dealing with massive files is the norm, but it's not uncommon that an organization works with massive datasets consisting of tens of thousands or even millions of smaller files. With frame-by-frame formats, such as those often used in VFX, moving folders with millions of files presents different challenges and traditional file transfer methods don't handle that well either, if at all.

The problem with application-level protocols such as FTP and HTTP is that each file typically requires at least one roundtrip. This necessitates a pause-and-send effect at the application layer on a per file basis similar to the TCP window size problem. Using multiple parallel streams is one technique that will reduce the impact of this pausing. For large data sets made up of many small files, settings that maximize the speed of switching from one file to the next can make a big difference in the overall time to job completion. Another approach is to pipeline sending files to eliminate the per file blocking roundtrip for acknowledgement. There are different challenges with sending a few large files when compared to sending a large number of smaller files, even if the total amount of data being moved is the same.

SIGNIANT

Getting the Most Out of Your Network

On the surface, it may seem that moving files from one location to another is relatively simple and that a faster pipe is all you need; but the challenges of the M&E industry are unique. Working with huge file sizes and data sets, across a global supply chain that involves people, systems and cloud services constantly interacting with each other creates complexity. Fortunately, there's modern software available to deal with exactly these challenges, and to help you get the most of your network.

Signiant's Patented Acceleration Technology

"With the other solutions, we couldn't utilize the speed we had or the speed a client had. We have a one gig pipe coming into our office and we move a lot of media, so this was a major concern."

~Michael Ball, Managing Editor, Accord Productions
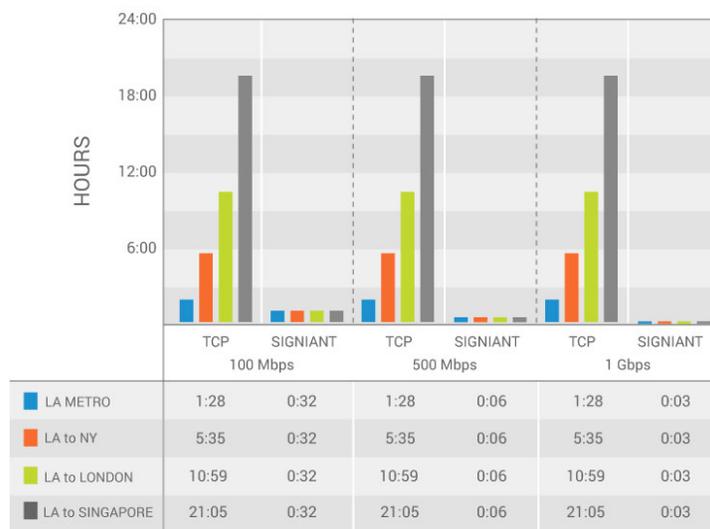
Signiant's Unique UDP Acceleration

Signiant's core acceleration technology harnesses the speed of UDP, but with added performance benefits and TCP-like reliability. To make UDP more reliable, Signiant adds functionality similar to TCP, but with a transfer control protocol implemented in a far more performant way, using:

→ Flow control, which makes sure data is transmitted at the optimal rate for the receiver,

→ Congestion control, which detects when the network is being overloaded and adapts accordingly, and

→ Reliability mechanisms, which ensure that data loss due to congestion or other network factors is compensated for, and that the order of the stream of data is maintained.

Signiant constantly measures effective throughput, network latency and loss, and builds a history that catalogues how all of these factors are changing over time. By analyzing the frequency of changes, we can locate network congestion. This affords us far more efficiency than congestion control algorithms that react to simple point-in-time packet loss, which is a problem even with modern TCP.

TIME TO TRANSFER 1 HR. OF HD CONTENT
(50 Mbps AVC)

	TCP 100 Mbps	SIGNIANT 100 Mbps	TCP 500 Mbps	SIGNIANT 500 Mbps	TCP 1 Gbps	SIGNIANT 1 Gbps
LA METRO	1:28	0:32	1:28	0:06	1:28	0:03
LA to NY	5:35	0:32	5:35	0:06	5:35	0:03
LA to LONDON	10:59	0:32	10:59	0:06	10:59	0:03
LA to SINGAPORE	21:05	0:32	21:05	0:06	21:05	0:03

Parallel Streams

Signiant's transport technology leverages parallel streams in certain cases to improve performance. This approach has two primary benefits:

→ Allows the load of a single transfer to be spread across more computing resources (e.g. across multiple CPU cores on a single machine or across multiple machines).

→ Parallel streams can, in aggregate, get past throughput limitations of any single stream.

Signiant's implementation of parallel streams works with both TCP or our proprietary UDP acceleration. Although the capacity of a single TCP stream and a single UDP stream are different (with UDP streams being much more performant in the presence of latency and/or packet loss), there are still benefits to using multiple UDP streams in some situations.

Parallel streams can be used with both large files and large sets of smaller files. With a large file, our technology will break it into parts and transfer the parts over separate streams. With a large number of smaller files, Signiant technology will use "pipelining" where multiple small files can be sent along a single stream, thereby avoiding the overhead of setting up a separate stream per file. Although the overhead of setting up a stream is extremely low, it typically involves at least one roundtrip on the network. When files are small, roundtrips that only require fractions of seconds can add up quickly, especially when moving files over long distances.

Intelligent Transport Using Machine Learning

In early 2019, Signiant introduced its patent-pending intelligent transport technology. The new architecture replaces standard TCP with our proprietary UDP-based acceleration protocol, and automatically deploys multiple parallel streams with either TCP or our UDP acceleration, adapting automatically to network conditions. In order to take full advantage of available bandwidth, Signiant's machine learning algorithm examines past history and current conditions, and optimally configures application and transport-level transfer parameters for both file-based and live media transfers. Machine learning is also used to determine the ideal number of streams and the optimal amount data to transfer on each stream based on a variety of inputs observed by the system. Not only does this ensure the best result without expensive and error-prone manual tuning, but results will continue to improve over time as the system learns.

Moving Data Between File Storage and Object Storage

There are two major types of storage that most organizations will interact with: file storage and object storage. Currently, most cloud storage available in the market uses an object storage architecture while most on-prem storage remains file based, although there is recent growth in the use of on-premises object storage.

Given that, a common challenge today is achieving fast, reliable transfers to and from public cloud storage. Even with the reliable, high-performance networks maintained by big cloud providers, such as AWS, Azure, and GCS, we've observed congestion that varies with time of day and geographic region.

SIGNIANT

As the industry moves to a hybrid cloud/multi-cloud world where file storage, public cloud storage, and on-premises object storage are all potentially in play, Signiant's commitment to storage independence is an important consideration as our software allows for data transfers between any type of storage whether on-prem or in the cloud, whether file or object based. Even if you're not using object storage today, Signiant's software provides a nice abstraction layer giving you the agility to introduce new storage types from any provider without disrupting operations.
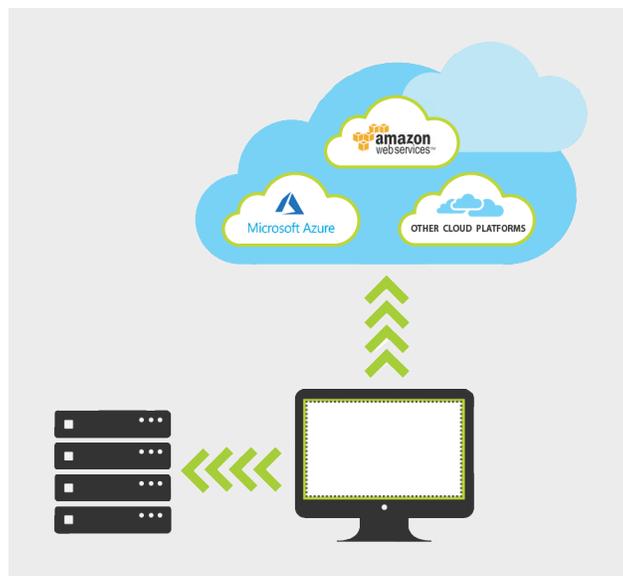
More Bandwidth + Signiant Software = Conundrum Solved

Spending money on a faster pipe that isn't used efficiently — whether because your transfer software holds it hostage with additional fees, or because your software simply isn't designed to handle it — is wasteful, and leaves you with the exact same problem, just short what you paid. Why shell out for something that you can't take advantage of when you can make the most of what you already have?

This is exactly what Signiant's accelerated file transfer solutions are designed to do: optimize transfers based on file and storage type, distance, and network conditions, to make sure you get the most out of all available bandwidth.

ABOUT SIGNIANT

Signiant's enterprise software provides the world's top content creators and distributors with fast, reliable, secure access to large media files, regardless of physical storage type or location. By enabling authorized people and processes to seamlessly exchange valuable content — within and between enterprises — Signiant connects the global media supply chain. **Find out more at www.signiant.com**.

Media Shuttle

THE EASIEST WAY TO SEND AND SHARE FILES ANYWHERE FAST

Media Shuttle is a hybrid SaaS solution that allows media professionals to quickly transfer any size file anywhere in the world.

Learn more:
www.signiant.com/mediashuttle

Jet

EASY, FAST SYSTEM-TO-SYSTEM FILE TRANSFERS

Jet is a powerfully simple SaaS solution to automate recurring file transfer jobs between global locations and partners.

Learn more:
www.signiant.com/jet

Manager+Agents

AUTOMATED SYSTEM-TO-SYSTEM FILE MOVEMENT

Manager+Agents is an enterprise software solution for automated delivery of large files across geographically dispersed locations.

Learn more:
www.signiant.com/manageragents

Flight

MOVE FILES INTO AND OUT OF CLOUD OBJECT STORAGE

Flight is an auto-scaling SaaS utility that accelerates the transfer of large data sets into and out of cloud object storage.

Learn more:
www.signiant.com/flight

SIGNIANT