FileCatalyst®
reinventing file transfer

Accelerating File Transfers

Increase File Transfer Speeds in Poorly-Performing Networks

www.filecatalyst.com

Contents

Introduction

According to a Hilty Moore & Associates study, companies have increased their reliance on digital file transfer via File Transfer Protocol (FTP) by 66% from 2006 to 2007.

__FTP__ resulted in incomplete delivery as much as __1 in 5 times__.

These companies were in not only high-tech, but also health, financial, manufacturing and service industries. Most respondents experienced delivery stoppages or incomplete transmissions with FTP as much as 1 out of 5 times. A more reliable alternative is needed.

This paper addresses technical issues faced by enterprises when transferring large files. It will also present a business case behind switching to a non-FTP digital delivery system, and guides the enterprise to find the optimal solution for its requirements.

Common File Delivery Methods

Companies with large data sets continue to use physical methods, recording to DVD, Flash, or HDD and using a delivery service. This works on an ad-hoc basis but can lead to high costs.

Standard digital alternatives include email or instant messaging services, but these have issues including file size limitations or reliance on the same TCP/IP backbone causing FTP to fail.

Alternatives should __overcome__ limitations of FTP, __maximize__ existing infrastructure, and __be simple.__

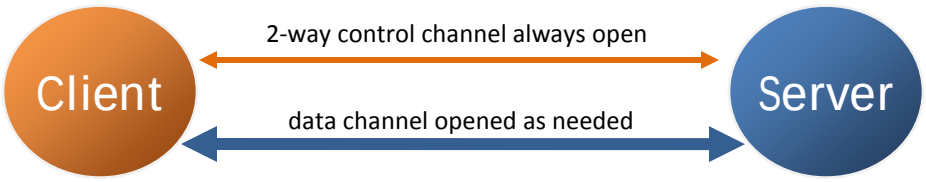
Digital delivery is still the ultimate solution for most enterprises. However, a solution should overcome the stoppages and slowdowns common to FTP. It should allow organizations to maximize their existing infrastructure, saving money that would otherwise have been wasted acquiring higher bandwidths. Finally, it should be easy to implement and—more importantly—easy to use.

Understanding FTP

*FTP has existed since the **1970s**, and is ubiquitous*

File Transfer Protocol (FTP) is one of the oldest protocols on the Internet, but is ubiquitous for file transfer. FTP resulted from an RFC first drafted in the 1970s, and has since gone through relatively few changes. FTP was designed to be completely agnostic—independent from operating systems, file storage systems and file types.

FTP runs on TCP/IP protocol, using two TCP connections. The first is a control channel used to exchange instructions, and is always open. The other is a data channel, used for file transfer and directory listings. The channel is opened or closed as needed.



TCP/IP is the most widely deployed protocol on the internet; consequently, FTP is the de facto standard for file transfer. However, it also inherits the drawbacks of TCP.

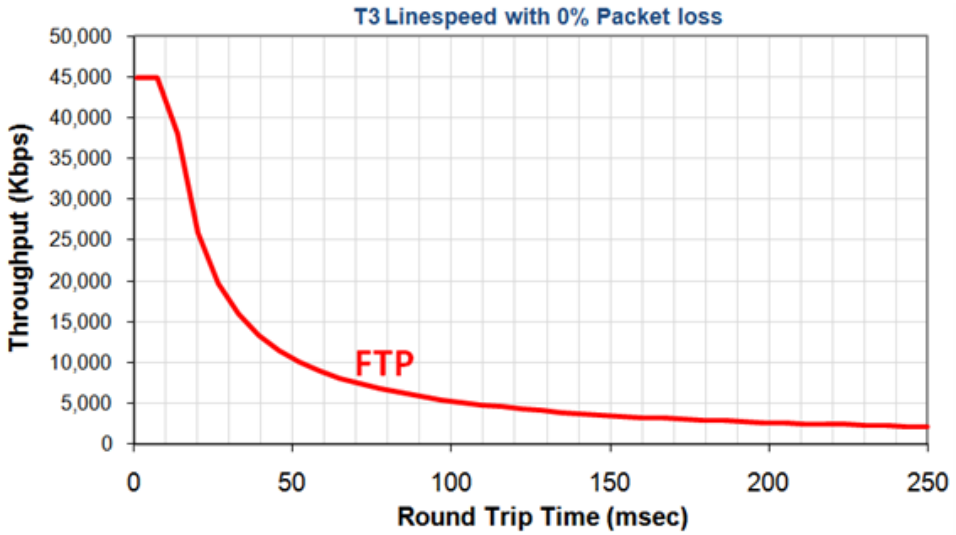
Latency and its effect on FTP

In order to provide reliable data transmission, TCP/IP requires the receiver to acknowledge each packet being sent, in sequential order. Each such communication is measured as round trip time (RTT), or the time it takes for a packet to be sent and acknowledged.

***Latency** (RTT) causes FTP to **slow down** considerably*

TCP responds to latency by adjusting the amount of unacknowledged data that can be on the link before waiting for a reply. The optimal amount of unacknowledged data en route should equal the end-to-end bandwidth multiplied by the RTT, also called the bandwidth-delay product. TCP continually estimates this value, setting a "TCP window" to control how much data should be sent. TCP has limits on the size of this value, so when the bandwidth-delay product exceeds a certain threshold, the result is a lot of waiting or "dead air". Satellite connections can be into the hundreds or even thousands of milliseconds of RTT.

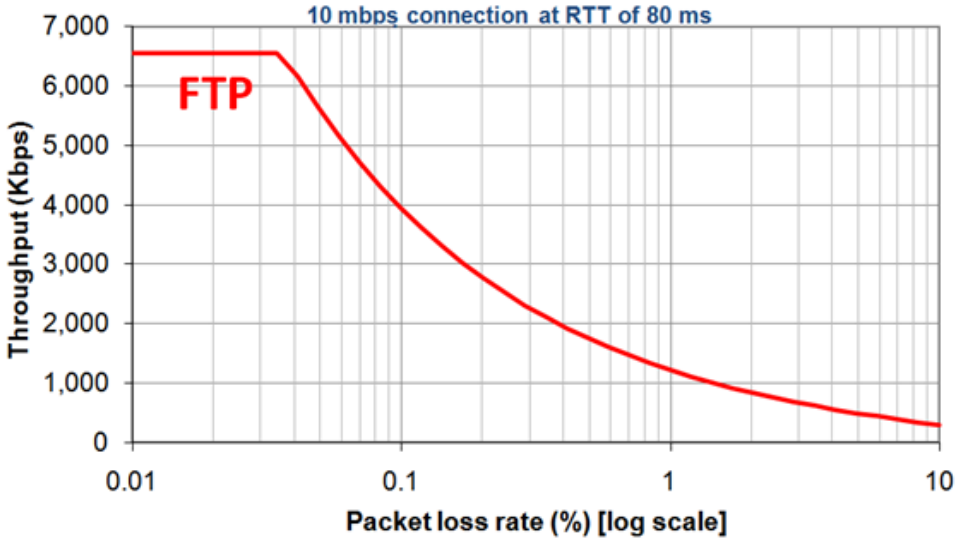
*Throughput is **reduced dramatically** at only 80ms RTT to just over 5,000kbps*



Packet loss and its effect on FTP

Network congestion typically causes buffer overflows of intermediate routers, causing packet loss. Since packets are sent sequentially, this can cause a hold-up in the cycle. Unacknowledged packets also cause the TCP window to shrink or even close completely for periods of time. Wireless and satellite transfers can have even higher packet loss due to sources of interference such as clouds or physical structures. Packet loss combined with high latency creates even worse performance for FTP transfers.

*After **0.03%** the throughput **drops exponentially.** Public internet averages **0.5%** packet loss.*



Existing File Transfer Processes

TCP is used as the backbone behind many existing corporate transfer processes. Other methods of sharing or distributing files include email attachments, shipping physical media, and instant messenger (IM) transfers. These are used independently or concurrently for getting data from source to sink. Each of these has potential limitations.

FTP Server Solution

FTP servers are often set up by media and design agencies to share large files. ***Maintenance*** *is difficult* ***and*** *FTP suffers from* ***TCP/IP drawbacks***

Including transfer speed issues, there are other potential problems with FTP:

- Learning Curve: The average end-user is not familiar with FTP clients.
- Firewall and Security: Not all IT networks allow for unrestricted FTP access; some users may not be able to connect to the FTP server.
- FTP Administration: IT administrators have to manage end-users on the FTP server and provide support and maintenance to the FTP server.
- TCP/IP Performance: Poorly-performing networks compromise transfer speed and encumber an organization's workflow.

Email File Attachments

Email is perhaps the most common method of sharing files others, both at work and at home. However, there are some considerable limitations:

Email servers simply ***cannot handle large file transfers*** *effectively or reliably. Archiving is* ***expensive****.*

- Reliability: Email servers are configured to handle attachments up to a certain size; usually 10 MB or less. Users cannot depend on their email arriving without being "bounced" for exceeding this limit.
- Archival: Emails are archived by the mail server. Attachments add to the storage requirements of the mail server, increasing IT cost.
- Poor Network Utilization: Email uses Simple Mail Transfer Protocol (SMTP) for transfers, which is not the most efficient protocol to transfer large files over a network.

Instant Messaging

According to PC World , Instant Messaging has become the most common form of electronic communications, surpassing email. There are some issues to be aware of:

__Interoperability,__ TCP/IP drawbacks and __security risks__ make IM less than ideal for enterprises.

- Interoperability: Although different IM clients are often able to share text messages with one another, file transfer is not usually possible between different clients.
- Security: Popular IM clients continue to be targets for trojans and exploits. Transmissions are not usually over secure channels.
- No tracking: Transferring files over IM makes it difficult or impossible to track files moving between different endpoints.
- Poor Network utilization: TCP/IP is typically used as the transport layer. The same network issues facing FTP also face IM file transfer.

Physical Media

Organizations may record data to a DVD, hard drive, or flash storage, then physically ship it to the recipient. This process can be useful when there are huge data sets or when there is poor connectivity between the two parties, but it has some obvious drawbacks:

Post-production houses exchange terabytes of data by shipping physical media. This is __labor-intensive__ and __expensive__.

- Preparation Time: This process requires human interaction, from recording the data to printing out mailing labels and shipping the media.
- Cost: Depending on the urgency of the files, costs for using a courier service can easily add up.
- Delivery Time: Since physical delivery is involved, the process can take 1 to 5 business days. This may be an unwelcome delay and makes time-sensitive delivery an unreliable proposition.

Overcoming File Transfer Issues with Acceleration

__Overcoming__ TCP/IP __bottlenecks__ allows digital delivery to become viable

Most companies realize that transferring files online is more cost-effective than shipping physical media. By overcoming TCP/IP bottlenecks, digital delivery becomes a realistic option. File transfer acceleration is possible through two complementary but separate methods. Generally speaking, these are optimizing throughput and reducing data being sent.

Optimizing throughput with UDP

__UDP__ takes latency out of the equation because it does not require acknowledgment

User Datagram Protocol (UDP) squeezes performance from an IP network by not implementing all features of TCP. UDP is a "connectionless" protocol that does not depend on sequenced acknowledgements and retransmissions. Although this implies potential for incomplete transfers, vendors are free to build their own reliability into their applications.

Since UDP allows packets to be received in a different order than they were sent it does not stall if packets are dropped. Lost packets can be requested with a command channel: data flows over UDP while error-correcting commands are periodically exchanged over a TCP/IP stream. Without the same acknowledgment requirements as TCP, latency is effectively eliminated.

Optimizing TCP with Multiple Streams

Multiple streams means __less "dead air"__ during a transfer, but is resource intensive

With standard FTP, only one data channel is used. This channel becomes idle during acknowledgement and other housekeeping. By increasing the number of data streams, there is less "dead air," as one or more streams is likely to be sending data at any given time.

This method of transferring files is effective when network degradation is not extreme. However, the process is extremely processor intensive due to the large number of concurrent threads that must be running.

Reducing Data with Compression

Data compression is a general term for a group of technologies that encode large files in order to shrink them down in size. Fewer bytes to send will naturally result in a faster transfer, whether the network is the internet, an intranet, or a local area network (LAN).

*Smaller file sizes require **less time** to transfer*

Data compression involves applying an algorithm to data that stores repetitive bits as a "shorthand". The receiving end uses an appropriate decoder to restore these short forms to their original state, resulting in an identical copy of the original file.

Some files can be compressed more than others. Text-based files, such as documents, reports, database files are highly compressible. On the other hand, many media files such as MP3 or JPG are already compressed and cannot be compressed further.

Minimizing data with Delta Transfers

*Transferring a file delta can **reduce data** sent by a staggering amount*

A file may already exist on both sides of the transfer, but has been modified at the source. When a difference is detected, an algorithm can calculate the differences and store them in discrete files. These small delta files are transferred instead of resending the entire file, then applied as a "patch," updating the destination file in order to make an identical copy. The benefit is clear: imagine being able to send a 4MB delta instead of a 2TB, for example. One popular Unix application that supports delta transfers is rsync.

The Cost of Slow File Transfers

Slow transfers incur significant "hard" and "soft" ***costs***

Most enterprises do not realize the cost of slow file transfers. The most significant hard costs is bandwidth; soft costs include lost hours in terms of employee productivity and project completion. Fixing the problem of slow transfers yields significant returns both on hard and soft costs.

The number crunching

Acceleration technologies provide return on investment (ROI) by maximizing existing infrastructures. Spending money on more bandwidth makes little sense if current bandwidth is under-utilized. Based on the following chart, consider some scenarios in which acceleration provides cost or productivity benefits:

Monthly Internet Costs						
Bandwidth (mbps)	1.5	3	5	10	50	100
Cost (month)	$75	$250	$500	$750	$1000	$1300
Pricing based on GoWebMan and Rogers Cable						

Improving file transfer returns ***significant time savings*** *to the organization*

Scenario 1—reclaim lost time: A post-production house in New Zealand transfers files to and from its LA office. Their 50mbps link costs $1000/mo, and is theoretically very fast. However, the average latency of pan-Pacific transfers is 200ms, with an average packet loss of 1%. Under these conditions, FTP will only transfer at 490kbps, a mere 1% of the large pipe.

With UDP-based acceleration, transfer speed will reach near link speed, or 100X faster than FTP. Art that formerly took a hundred minutes now takes one. Without spending any more money, the company has regained ninety-nine valuable minutes that can be applied to productivity and project completion.

Scenario 2—save IT expense: A company transfers files between North America and Europe over a 10 mbps connection. Due to latency and packet loss, the effective throughput is 613kbps or 6% of bandwidth. The internet service provider suggests the company purchase a faster link.

Maximizing *bandwidth is* ***less expensive*** *than buying a faster link*

Instead of adding to their monthly operating expense, the company deploys acceleration technology. By using their bandwidth to its fullest, the company experiences faster transfers and has a virtual savings of $670/mo compared to the upgrade option.

Summary of Scenarios				
Scenario	**Line Speed**	**Latency / Loss**	**Previous speed**	**Acceleration Advantage**
1 – reclaim lost time	50 Mbps	290ms RTT 1% packet loss	490 Kbps (1%)	Increase transfer speed by 100X, reclaiming man hours and improving turnaround
2 – save IT expense	10 Mbps	160ms RTT 0% packet loss	613 Kbps (6%)	Increase transfer speed by 15X, giving company higher performance and avoid $670/mo expense.

Optimizing bandwidth generates both virtual and actual returns on investment made in implementing a file acceleration solution.

Factors to Consider for your File Transfer Solution

The optimum file transfer solution should be able to connect the various file storage systems of the enterprise regardless of network topology, physical location or operating system. Many commercial vendors provide a file transfer solution—some offer a comprehensive platform, while others provide a subscription service solution. There are universal considerations to be made for each option:

- **Acceleration:** How important are accelerated file transfers to your enterprise? Do any of your processes involve large file transfers that may be reduced?
- **Automation and Scheduling:** Will you benefit from automating routine file transfers? Would automatic scheduling help your business complete projects and improve its bottom line?
- **Web-Based Interface:** Should end-users have access via web browser, or is a client-side application preferred?
- **Ease of deployment:** How many servers and clients do you need to install and where? Is installation feasible for remote locations?
- **Security and Compliance:** What are the major security issues that your organization faces? Are there any security mandates you must meet, and does the proposed solution meet them?
- **Tracking and Reporting:** How much tracking functionally do you need? What type of reports would you like to generate, and does the solution support this?
- **After-sales support and service:** Does the vendor provide reasonable support and service for the product?

The Ideal Solution

Scalable
The solution should scale in a cost-effective way. The solution should be easily expandable

Reliable
Once a file transfer is initiated, users should expect delivery. The system should recover from errors.

Simple
The system must have a gentle learning curve for end-users

Secure
The solution should be able to meet existing or potential security mandates

Learn more about our technology at ***www.filecatalyst.com***

About FileCatalyst

FileCatalyst software solutions are developed by Unlimi-Tech Software, Inc. Founded in 2000, Unlimi-Tech Software is a privately held corporation based in Ottawa, Canada, operating with a global reseller network. Unlimi-Tech products are focused on solving file transfer challenges in diverse environments, from end user desktops to sophisticated WAN and satellite-based multi-cast systems.

Learn more about Unlimi-Tech Software and FileCatalyst:

www.filecatalyst.com
info@filecatalyst.com

tf +1.877.327.9387
t +1.613.667.2439
f +1.613.667.2439