Aspera's *fasp* Track for High-Speed Data Delivery

Data transfer protocol facilitating global data access and collaboration.

Michelle Munson, co-founder of Aspera

LEAH FASTEN

BY KEVIN DAVIES

Software engineers Michelle Munson and Serban Simu , the co-founders of Aspera in Emeryville, California, both worked in application-level networking since leaving graduate school, and were exposed to the problem of transporting data over wide area networks (WANs) early in their careers.

"We'd worked on related areas, particularly in transferring digital media content, and knew there was an unsolved problem," says Munson. That problem boiled down to: Why doesn't the Transmission Control Protocol (TCP) work well for moving bulk data over WANs? And what were the alternatives?

"We didn't originally set out to make a transport, as we'd assumed there'd be open-source technologies for reliable transfer," says Munson. Indeed there are, but Aspera, a bootstrapped company with roots in Munson's garage, argues that the performance of its commercial software outstrips the open-source alternatives (see, "We Can't Fix the Internet"). Munson claims that the typical increases in speed experienced life sciences companies based on network capacities and bottlenecks range from 10-fold to 100-fold.

When Munson and Simu investigated the alternatives for high-speed data transport, they found that none of the available transport approaches held up. TCP is a reliable transport protocol that powers FTP, HTTP, CIFS and NFS, SCP and RSYNC, among others. But given the fundamental problems of TCP over networks with high round-trip time and packet loss, which severely limits the speed of large data transfer over WANs, Munson

and Simu set out to engineer a new protocol that did not have any artificial bottlenecks under WAN conditions.

The pair was able to forego any external investment or venture capital because of early customers that allowed the company to grow software around the technology. The first two types of companies to test our technology and put us over the edge were affiliated with the Department of Defense [DOD] and media/entertainment." (The DoD connection was somewhat serendipitous, and came from Munson knowing a then small contractor in DOD intelligence that was having difficulty transporting data over networks.)

About two years ago, the genomics /life sciences community discovered Aspera, becoming the firm's third key vertical market—particularly in the field of next-generation sequencing (NGS).

Each vertical has its own issues, but Munson says the problems confronting the intelligence community—the collection and dissemination of unstructured data such as video surveillance and high-resolution imagery—are not fundamentally different from life sciences: both groups need to share and exchange large amounts of data over global Internet networks in rapid time.

Munson and Simu wrote the first version of the *fasp* protocol and remain intimately involved in technical product development although, with a twinge of regret, Munson says her coding days are behind her. It's largely the impetus of key communities such as life sciences and

digital media (not to mention others that have ever increasing quantities of data to share) that is pushing research and development around the transport, she says.

Following Protocol

Aspera's *fasp* is a communications protocol that aims to satisfy the burning need posed by two fundamental problems surrounding the movement of file-based data from storage A to B across a network—reliability and bandwidth. "There's a fundamental efficiency problem, and then there's a congestion or bandwidth control problem, because the user doesn't know the bandwidth or the other traffic, making it unsafe to blast traffic over the network," says Munson.

"If you use a TCP protocol, you'd experience a severe bottleneck due to congestion control as it affects speed due to round trip delay and packet loss," says Munson. "That's the baseline."

Munson says there have been many attempts to build various types of simple "data blasters"—reliable transport alternatives to TCP over IP or the User Datagram Protocol (UDP)—in which the data traveling over an unreliable IP channel has reliability implemented in a protocol above IP. But there's a big drawback: "From the controls perspective—i.e. how

such blasters re-send dropped packets over the IP network—it is extremely inefficient," says Munson. "They generate heavy duplicate transmission of the data and tend to overrun the network bandwidth."

"This was shown in the literature and is what we confirmed when testing many open-source solutions, and this ultimately led us to create *fasp*."

Aspera elected to implement the *fasp* software protocol specifically as an App protocol rather than in the network stack as a driver. "We chose to do that to make it available so end users on their computers could use it without having admin rights to install and run," says Munson. "That was important: It allowed our technology to be used in a very simple way and users to start experimenting with this."

Aspera has been deployed by the European Bioinformatics Institute (EBI), the Broad Institute, and other companies and academic groups, including the University of Washington, University of Maryland, and Memorial-Sloan Kettering in New York. Among the firm's highest profile successes is work for the National Center for Biotechnology Information (NCBI) at the NIH as part of the 1000 Genomes Project. "The one that exposed our software to the community (and caused us to come to Bio-IT World Expo in 2009) was NCBI," says Munson.

The 1000 Genomes Project requires transferring and exchanging data from institute to institute, across continents. Users accessing 1000 Genomes data visit one of the four public websites that disseminate more than 7 Terabases of genomic data. They can browse and download data with FTP and/or the Aspera protocol—using the Aspera Connect free web browser plug-in.

"In those cases, the dimensions of improvement over FTP go up with more bandwidth and more difficult networks/distance," says Munson. Whereas the speed of FTP is theoretically fixed based on round-trip time and packet loss, *fasp* fills the available bandwidth. "The difference is the bottleneck speed of FTP and bandwidth capacity," says Munson. "*fasp* does not compress the data and achieves its speed up in the transport efficiency." For example, from the US to Australia,

the FTP bottleneck speeds is 1 Megabit/second or less. With Aspera on a 100 Mbps link with all bandwidth available, however, it's virtually 100 Megabits/sec.
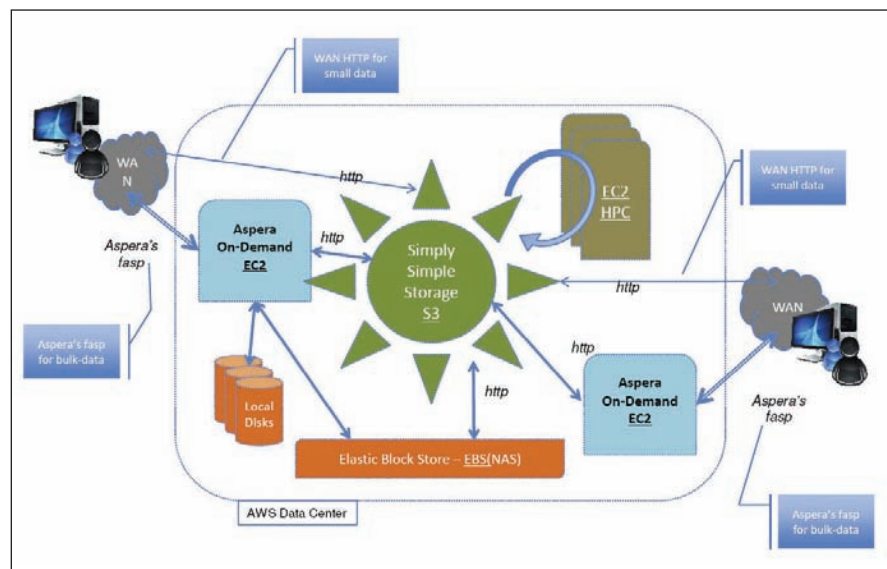
Munson says BGI Shenzhen, the high-capacity Chinese genome research center, will soon become another hub on this data transfer pathway. BGI's location makes an optimal data transfer solution essential, because there is typically a 200-400 millisecond round-trip time and high packet loss into China. "As you go into China, especially mainland China, the wide area networking problem is unbelievable," says Munson.

For BGI, "on those types of networks,

Aspera will allow users to transfer file data over the WAN and write directly to S3.

large data transfer is not only inefficient, it's often impossible," says Munson, because of the distance and packet loss. "There's a massive opportunity and capability to process the data on these locations, but the problem in moving data to and from these locations becomes paramount. For an economy of scale, shipping disks won't work."

Another niche that Munson expects to fill is where two research or medical institutes are sharing data between each other. Formerly they might have used Unix SCP or R-sync (open source). But Aspera can be used like a Unix utility, while transferring using the *fasp* protocol, which allows easy automation of data transfer between institutes.

User Needs

To take advantage of *fasp*, users require no special networking, hardware, or fiber channels. "We run over standard IP networks," says Munson. "The user experiences the *fasp* software as a file transfer application or an embedded transport in someone else's application."

For the most part, the *fasp* protocol doesn't vary from vertical to vertical, life sciences to digital media. But there are some special software adaptations in life sciences, says Munson. "It is the same transport core, but what is emphasized and refined by LS users has to do with upper end speeds. We have an adap-

tive rate control that adjusts the rate of transfer to match the available network bandwidth and disk throughput." That is especially important on 1-10 Gbps networks, where the network capacity often outstrips the file system or disk I/O speed.

This is an important issue in life sciences. Network bandwidth is quite large, and there is access to Internet 2, so the transfer bottleneck using fasp becomes shared access to the disk system as data goes in and out. Aspera's rate control has both disk-based and network-based adaptation components. "We released the disk-space component during the time we've been working with LS community," says Munson.

Recently, Aspera was used in its first

single transfer session of more than 10 Terabytes of data.

"Theoretically *fasp* can transfer data of any size, but in practice, single transfer sessions between institutes have gone up from hundreds of gigabytes in our first year, to now as high as 12 Terabytes at a time in a single transfer," she says. "We made some architectural changes in the way our software was implemented to accommodate that. We have no limits today in our session sizes."

Aspera is also working closely with the data storage community to establish benchmarks for the movement of ultra large file sets over 10 Gbps networks and beyond—including firms such as EMC, HP, NetApp, Isilon, BlueArc, and Panassas.

Cloud Traffic

Aspera's early success in facilitating the transfer and movement of huge datasets begs the question of whether it can assist users in leveraging the Cloud?

"Absolutely, and transfer of data to and from the Cloud is one of the most pressing challenges," says Munson. Aspera's On Demand product enables data transfer to Amazon Web Services (AWS) at speeds (up to a current practical limit) of several hundred Mbps. But Munson says, "We are coming up against technology limitations the way the Cloud is currently deployed, in terms of directly reading and writing large file data to persistent storage." That said, Amazon is moving very rapidly, she says, and improvements are on the way.

"In the near future, users will be able to transfer file data over the WAN and write directly into S3 within the Aspera application, and at high speed." •

'We Can't Fix the Internet'

"Technology is a balancing act between access and cost," Bhavik Vyas, Aspera's director of technology sales, told attendees at the second Bio-IT World Europe conference in Germany in October. From next-gen sequencing to medical imaging and media, managing data is about size, backups, and reliability.

The key problems are: 1) Collaboration requires the Internet; 2) Data transfer is slow over the Internet via public or private WANs; and 3) Fast networks typically have not only slow transport but also very slow storage.

"We can't fix the Internet—no-one can—but we can tackle (2) and (3)," said Vyas.

Vyas lays the blame with slow data transfer over the Internet on TCP for lost productivity and inefficiency. TCP has well-known bottlenecks—high round-trip times (RTT) and packet loss rates, especially on high-bandwidth WANs. "The further you are from your data, the slower TCP will go," he said. And while there are open-source protocols to help avoid congestion, they typically have "high inefficiency and catastrophic effects on packet loss," said Vyas. For example, the RTT from London to New York has a latency time of 60 milliseconds. As TCP performance wanes with distance, the rate can be calculated. Combining distance with loss leads to terrible performance.

Performance Options

One of the options to improve performance is to explore the use of commercial or academic high-speed TCP variants, such as CUBIC, BIC, Reno, FAST TCP, and H-TCP. These can reduce congestions and increase throughput, but on heavily congested WANs, modifying TCP becomes difficult, because it has to be deployed across all workstations—and packet loss can ensue, and the accelerated speed is still impaired on lossy networks. Another problem is that while TCP ensures no data loss, everything is sent in sequence. This results in a stop or slow down for every lost packet. Aspera argues that sequential delivery isn't necessary, and network capacity should be utilized.

A second option is the use of UDP-based transport applications. Many open-source and commercial technologies use UDP for moving data reliably and quickly (in a connection-less way), checking reliability and re-sending data if necessary. But as Vyas pointed out, "If the cost of the improvement is you send 10x more data than you receive (in duplicate retransmission and bandwidth overdrive), then the cost benefit isn't really realized."

In other words, an architecture that facilitates data blasting creates its own problems, limiting the return on investment on multi-Gigabit Ethernet (GbE) and 10-GbE networks.

Developed nine years ago by Yunhong Gu, UDT is a UDP Data Transfer application protocol that is faster than TCP, but Vyas argues there are performance issues in some typical WANs, such that the network can appear 'full' but with un-needed data. A company called VeryCloud provides commercial service for UDT. A new reliable transport option is Aspera's *fasp* protocol, enabling access to and management of data. Vyas calls it purpose-built, reliable, with a theoretically infinite transfer speed and zero receiving cost.

Comparison Notes

Vyas presented metrics comparing the speed and bandwidth cost of *fasp* compared to TCP (Reno TCP and FAST TCP) and UDT. Aspera's *fasp* achieves a throughput of about 90-93 percent depending on round-trip time (anywhere from 20-1000 ms and packet loss of just 5-10%). UDT, by contrast, has low throughput and efficiency (less than 50%) over networks with high round-trip time and packet loss.

Another issue is the "last foot" of the data transport pipeline—storage at the end user. Vyas cited benchmark studies with storage vendor EMC and its Celerra product in which they obtained 3 Gbit/second large data transfer rates over worst-case global WANs with round-trip time of 300 milliseconds, and packet loss rates of 5%. "You can get these speeds if you want to," said Vyas. **K.D.**