UNIVERSITY OF MARYLAND SCHOOL OF MEDICINE
INSTITUTE FOR GENOME SCIENCES
BIOTECH

aspera
an IBM® company

Reliable Data Delivery at the Institute for Genome Sciences of the University of Maryland

AT A GLANCE

Industry
Life Sciences

Products

Aspera Connect Server

Challenge

Collecting and distributing terabytes of genomic sequencing data to researchers globally.

Results

- High-speed genomic sequencing data transfers and end-user access
- Full bandwidth utilization with extraordinary bandwidth control
- Reduced operational costs
- Rapid deployment
- Easy user rollout

The Institute for Genome Sciences (IGS) is a research center within the University of Maryland School of Medicine (UMSOM). The Institute uses the powerful tools of genomics and bioinformatics to understand, study, and generate data for the international scientific community.

CHALLENGE

IGS manages the collection, storing and distributing of terabytes of genomic research data globally. Due to increasing file sizes from an average 5GB to 20GB, data movement over FTP faced growing reliability problems, especially with international long distance transfers. FTP was delivering under 20% network utilization and frequent connection failures were resulting in excessive re-transmission.

IGS then converted to physically shipping hard drives, which quickly became constrained by time and personnel resources. Many data sets exceeded disk capacity, which resulted in further increasing the manual workload. It often took two people over two hours to prepare and pack the disks—another 24 hours for overnight delivery—and another four hours to unpack and mount the disks; and if data was being distributed to multiple researchers at different sites for collaboration, additional personnel and time was required.

The researchers were losing valuable research time because of the slow data delivery and poor reliability of data transfers.

"Before Aspera, we were only utilizing a fraction of our bandwidth, transferring at 40-50Mbps. Now that we've deployed Aspera software, we consistently transfer at our maximum allocated rate of 300 Mbps, fully utilizing the bandwidth, while protecting other critical traffic."

Victor Felix
Senior Software Engineer at IGS

SOLUTION

The IGS team began evaluating software-based alternatives to FTP and shipping hard drives. IGS chose the Aspera Connect Server after successful trials.

Critical requirements for the solution included the ability to rapidly deploy across a widely distributed network; and allowing researchers to immediately use the system without extensive training. The configuration is centered on the Aspera Connect Server, established at the IGS Data Analysis and Coordination Center (DACC) as a hub interconnecting multiple databases from other genomic research centers. DACC is responsible for collecting sequencing data, refining and coalescing it, and presenting it back out on the Web so the entire research community can make use of it.

aspera
an IBM® company

BENEFITS

Ease of use
Intuitive web-based platform is easy to configure, manage, and maintain. Users simply install a web-browser plug-in compatible with popular web browsers.

Maximum speed
Enables large data set transfers over any network at maximum speed, source disk to destination disk, regardless of network conditions or distance.

Extraordinary bandwidth control
Provides precise rate control (pre-set and on-the-fly) for guaranteed transfer times. Adaptive rate control fully utilizes available bandwidth while remaining fair to all other network traffic.

Complete security
Includes built-in security using open standard cryptography for user authentication, data encryption and data integrity verification.

Software only solution
Uses standard, unmodified IP networking and is implemented in software as an application protocol. Requires no changes to the operating system and no new appliances.

Robust
Automatically resumes partial transfers and retries failed transfers.

Flexible open architecture
Supports interoperable file and directory transfers between all major operating systems and provides a complete, modern software API to build upon.

Built upon Aspera's flagship Enterprise Server software, the Aspera Connect Server is the ideal solution when large files and data sets are exchanged with a large number of locations and users over the wide area network. Featuring patented, high-speed FASP™ transport at its core, Connect Server transfers large sequencing data from Aspera servers and clients at other facilities with unparalleled performance and reliability, regardless of distance, and with complete security. Its web-based directory listing then allows researchers to easily browse the data they need to access and an install-on-demand plug-in enables high-speed uploads and downloads transparently from within the user's browser.

RESULTS

The IGS team compared the performance of FTP with that of Aspera software based on FASP technology. "Before Aspera, we were only utilizing a fraction of our bandwidth, transferring at 40-50Mbps," said Victor Felix, a senior software engineer at IGS. "Now that we've deployed Aspera software, we consistently transfer at our target rate of 300Mbps. That is a significant improvement over FTP — not even counting the reliability improvement — which is incredible."

"Instead of waiting 24 hours shipping disks or worrying if FTP will fail, Aspera enables us to know precisely when the data will get to its destination."

Victor Felix
Senior Software Engineer at IGS

Other benefits, adds Felix, are reduced operational costs, reduced capital equipment, and faster turnaround to researchers. "Using Aspera, we don't have to handle and ship drives. We can move all the data online. We have gone from spending 4 hours packing disks to 5 minutes using Aspera. Instead of waiting 24 hours shipping disks or worrying if FTP will fail, Aspera enables us to know precisely when the data will get to its destination.

ABOUT IGS

The scientific discoveries that are being made at IGS are helping to unravel the mysteries of biological systems and to improve healthcare for people around the world.

About Aspera

Aspera, an IBM Company, is the creator of next-generation transport technologies that move the world's data at maximum speed regardless of file size, transfer distance and network conditions. Based on its patented, Emmy® award-winning FASP® protocol, Aspera software fully utilizes existing infrastructures to deliver the fastest, most predictable file-transfer experience. Aspera's core technology delivers unprecedented control over bandwidth, complete security and uncompromising reliability. Organizations across a variety of industries on six continents rely on Aspera software for the business-critical transport of their digital assets.

Learn more at www.asperasoft.com and follow us on Twitter @asperasoft for more information.