

Advancing cancer research with fast transfer of large Proteomic data

AT A GLANCE

Industry

Life Sciences

Products

Aspera Connect Server
Aspera SDK

Challenge

ESAC needed a fast and secure transfer solution that could handle large mass spectrometry data for integration into their data portals.

Solution

ESAC selected Aspera Connect Server to transfer large files from research centers to the Data Coordination Center.

Results

- With one Aspera server, ESAC can meet the needs of two use cases – the DCC with a private portal and a public portal.
- Aspera reduces program costs by eliminating the need for costly global shipments in addition to allowing unlimited free browser-based clients.
- The Aspera platform was seamlessly integrated directly into the NCI portals creating a great end-user experience.

ESAC, Inc. provides research data management, bioinformatics, and healthcare information technology solutions for government, commercial, and academic clients.

ESAC was awarded a 5-year contract with the National Cancer Institute's (NCI) Office of Cancer Clinical Proteomics Research to manage the launch of a new Clinical Proteomic Tumor Analysis Consortium (CPTAC) Data Coordinating Center (DCC). ESAC was tasked with creating one data center and two data portals: one a secure and private repository for members of the consortium to upload and download shared data; and the other a separate public portal that would provide any member of the scientific community with access to data.

ESAC relies on Aspera's high-speed transfer software to move large proteomic data sets to and from the data repositories that support each portal with the efficiency, security, and reliability that scientific researchers can depend on.

CHALLENGE

While setting up the data center and accompanying web portals, ESAC encountered one major challenge: because global proteomic profiling is done using mass spectrometry, file sizes are often exceedingly large, plus the quantity of data accumulates rapidly. A typical experiment is composed of separate mass spectrometry data sets for each tissue sample and may reach several dozen GB.

Previously, researchers had shipped physical hard drives to share data amongst their workgroups. However, this proved to be slow and expensive. Another transfer tool they had used resulted in data loss because of the solution's inability to package and send all data reliably.

ESAC sought a file transfer solution that they could integrate directly into the CPTAC web portals and that would provide fast transfer speeds, security, and data integrity verification while significantly improving the overall end-user experience.

SOLUTION

ESAC selected the Aspera Connect Server to handle all transfers in and out of the data portals.

With Aspera seamlessly integrated within the data portals, ESAC was able to configure the portal used by the general public to only allow downloads, while designating greater flexibility and two-way transfer capabilities for the private portal available to the consortium team members.

Data is generated at the Proteome Characterization Centers – there are five in total that were awarded, and each center comprises multiple research sites. Researchers simply login to the secure web portal and use the Aspera Connect Web-browser Plugin to upload data into the DCC private repository, where other individuals

BENEFITS

Maximum speed

Aspera maximizes available bandwidth, so scientific researchers can upload and download data to the secure web portal at maximum speed, regardless of file size, transfer distance, or network conditions.

Strong security

With Aspera, ESAC was able to build a secure web portal that protects researchers' data with built-in security features that include SSH authentication, encryption in transit and at rest, and data integrity verification for each transmitted block.

Ease of integration

ESAC used Aspera's comprehensive SDK for an easy integration with the web portals, providing users with a smooth and seamless experience.

Monitoring and control

Aspera Connect Server provides complete visibility over the transfer environment, and ESAC tracks transfer performance with daily logging and reporting. Verifications for each transmitted block, bandwidth while remaining fair to all other network traffic.

ABOUT ESAC

ESAC, Inc. provides research data management, bioinformatics, and healthcare information technology solutions for government, commercial and academic clients. Headquartered in Rockville, MD, our mission is to play a vital role in helping clinicians, scientists and researchers across the globe improve all aspects of human health.

This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN261201100106C.

within the consortium can then download and view the data. Once a workgroup determines the data is ready for the public, it is indexed with metadata by the DCC and placed on the public data portal.

The public portal opens the door to validation of scientific observations, investigation of novel proteins and proteoforms, and shared access to cancer research in the scientific community. The portal offers data at different analysis stages, so investigators with different backgrounds can download the data level that is best suited for their research program.

RESULTS

ESAC, Inc. was able to reduce data delivery times from days with physical shipping methods to hours or even minutes with Aspera. Plus, Aspera reduces program costs by allowing for unlimited free browser-based clients with the Aspera Connect Server, in addition to eliminating costly global shipping and handling of hard drives.

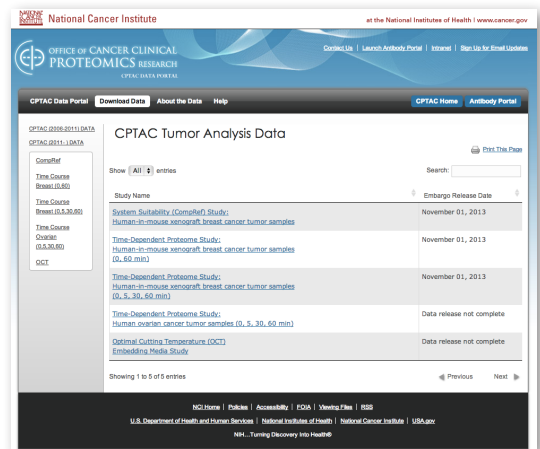
“Having a system where there is high integrity of file transfer and verification of each packet was a really big deal with this community.”

 Karen A. Ketchum
Director at ESAC, Inc.

“Having a system where there is a high integrity of file transfer and verification of each packet was a really big deal with this community,” said Karen A. Ketchum, PhD, Director at ESAC, Inc.

Concurrently, high-speed downloads from the public repository via Aspera benefit the scientific community by making data readily available for researchers to utilize.

“With one Aspera server, we can serve two use cases. That was key,” added Ketchum.



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.