FILE SERVICES FOR DNA SEQUENCING

Big Data Service Provider Uses Morro Data to Streamline Operations and Improve Customer Satisfaction

©2010, Ilumina Inc. All rights reserved.

SeqMatic (https://www.seqmatic.com/) is a next generation DNA sequencing service provider based in Fremont, California. SeqMatic was founded by world leading scientists who were part of the revolutionary R&D team at Illumina (a company whose products enable researchers to explore DNA at an entirely new scale). SeqMatic's clients include some of the largest and most prestigious genomics and transcriptomics research facilities in the country, including Stanford and Harvard Universities, the United States Department of Agriculture (USDA), the Food and Drug Administration (FDA), and the Allen Institute for Brain Science.

Compared to the initial Human Genome Project which took over 10 years and cost nearly $3 billion, Next Generation DNA sequencing (NGS) today is incredibly fast and inexpensive, making large-scale whole-genome sequencing accessible and practical for the average researcher. The Illumina HiSeq X Ten System, for example, can churn out 18,000 genome studies a year. This is enabling researchers to use DNA sequencing across many fields and for multiple purposes. As Next Generation Sequencing Machines have high upfront costs, many researchers outsource their studies to DNA sequencing service providers like SeqMatic. Once SeqMatic receives the samples, they run the DNA sequence and send the results files back to the customer for further analysis.

However, the extraordinary value delivered by NGS requires data intensive operations. A single DNA sequencing run can churn out up to 1 TB of raw data, which then needs to be stored, protected, and transferred to their customers. Given the sheer amount of data SeqMatic needs to send to its customers, file transfer costs easily get out of hand. In addition, DNA sequence results can contain dozens of individual files, each of which must be individually checked to ensure that the transfer was successful, making the process painfully time consuming.

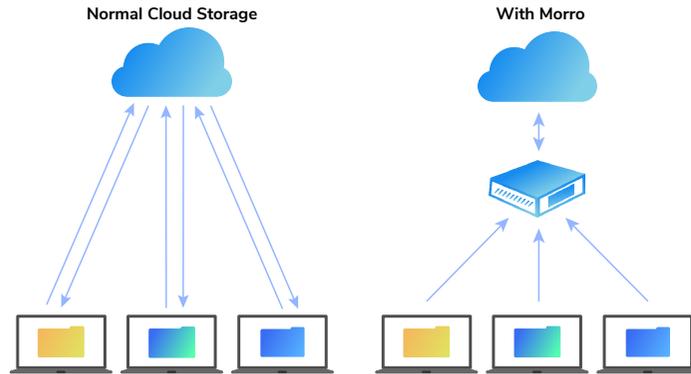The company's additional file handling needs include:

1. Data Scale: DNA sequencers churn out a tremendous amount of unstructured data. Relying on traditional file transfer methods is error prone and time consuming.

2. Data Growth: New Sequencer Machines operate at much higher resolution and speeds, generating four to sixteen times more data than prior generation machines per run, and each run can now be more extensive.

3. Protection: This data is tremendously valuable – especially at pharmaceutical companies. The file transfer must be highly reliable and secure and use a high availability architecture to protect against data loss.

4. Large File Transfer: Each day, SeqMatic simultaneously sends hundreds of thousands of large files to their clients. Currently, they transfer these files via FTP which is time consuming and requires multiple steps.

5. Data Sharing: The test results need to be easily accessible by their customers at any time, so file sharing amongst various teams and across multiple locations is critical. In addition, file locking and version control is required to protect the integrity of all files.

Solving Their Big Data Needs

After assessing their options, SeqMatic selected Morro Data CloudNAS for its ability to meet all of their needs. With the ability to share files across multiple locations at LAN speeds and deliver 99.999999% (eight nines) protection against data loss, CloudNAS provides everything SeqMatic requires in a file transfer solution at a price they can afford. And with its cloud-first storage architecture and Cache & Sync technology, CloudNAS greatly simplifies the transfer and access of all files using a regular broadband connection.

SeqMatic deployed CloudNAS in their environment and immediately began reaping the benefits of its ease of use and LAN speeds. In a matter of seconds, SeqMatic's employees were able to drag and drop files to their Morro CloudNAS with its drive letter interface. Morro then syncs all files with the cloud in the background, and sends clients a notification that their files are ready for download via a share link. Clients with occasional needs can access the files using the Morro Connect app or via the Team Portal web interface. Seqmatic provides clients with more frequent or data intensive needs, their own CacheDrive. By activating a Morro feature called prefetch, files are preemptively downloaded into the destination CacheDrive as soon as available. With prefetch turned on, SeqMatic's clients enjoy immediate access to files on the local LAN storage with no additional effort.

Morro Delivers a Win-Win

With CloudNAS, SeqMatic has reduced its file transfer costs and significantly streamlined its operations while delivering additional value to its customers. This has resulted in lower costs and greater efficiency for SeqMatic, therefore improving ROI. And the faster, easier, more reliable availability of all data has significantly improved customer satisfaction. Now Seqmatic and its clients can focus on the core value without being burdened by time consuming and error prone file transfer. At the same time a copy of the data is available on the cloud for archive and long-term access without consuming local storage space.