GBH Adapted to the Pandemic with the Help of ExpeDat™ and HyperStore™

Just before the global pandemic hit, award-winning public television and radio producer GBH of Boston had embarked on a massive media archive project. Data Expedition, Inc. and Cloudian helped adapt it into a more efficient remote work setup.

> **❝ Out of the box, it worked completely.❞**
>
> **– Shane Miner,**
> Chief Technology Officer, GBH

Overview

For more than 65 years, GBH, a public television and radio service based in Boston and the largest producer of content for PBS, has created award-winning programming viewed and heard around the world. That's a lot of media, much of it produced before digitization, so the station began a project through a National Endowment for the Humanities (NEH) challenge grant and capital campaign, to digitize and preserve much of that media so it could be more accessible to more people.

GBH's goal was to create a system that could store both archival media as it was digitized and new media as it was produced. That meant not just storing the data, but also supporting the kinds of high throughput access needed for day-to-day editing workflows. Archivists, producers, editors, and distributors all needed fast and secure access to the object storage.

Cloudian®, specializing in S3-compatible object storage, and Data Expedition, Inc.® (DEI®) specializing in high-performance data transport, were the partners to build this ambitious system. When the pandemic hit, bringing new remote workflow challenges, the flexibility of these systems proved essential to keeping the work flowing.

Challenges

GBH has a large Archive and team of archivists managing lots of physical assets, both analog and digital. With a grant from the NEH, they began working on digitizing the most at-risk assets to ensure preservation and access. The goal is to digitize and store 3 petabytes (PB) of data over the three- to five-year life of the project.

At the same time, existing digital assets are being moved into centralized storage and ongoing workflows are being realigned to ensure that media is archived from the start and at every step of production. This means ensuring

> *The Cloudian box is the storage unit and the ExpeDat software sits in front of that, interfacing with the S3 API."*
>
> *– Seth Noble,*
> CEO,
> Data Expedition, Inc.

secure, real-time access to the central storage for editors and producers wherever the productions may be taking place.

Then the COVID-19 pandemic brought a new mandate: every workflow was now remote. Shane Miner, GBH's Chief Technology Officer, explained, "When we got into the COVID situation, and had to send people home, there was concern about how we would edit and produce the shows . . . will we need to send people home with their edit stations and hard drives? That didn't fit with the model that we wanted to have, which was that all the footage would come into a centralized place, and everyone would work on it from there. This new situation was the exact opposite of that, and we had to figure out some way to make this work."

Solution

In the months before the pandemic, the archivists began automating the transfer of digitized content to Cloudian's HyperStore using Data Expedition's ExpeDat command-lines tools. Deploying the ExpeDat server in GBH's network DMZ created a secure gateway between the internet and their internal network. ExpeDat provided accelerated data transfer through the outer DMZ firewall, then translated the data stream into S3 multipart through the internal firewall to the HyperStore archive.

"We have a pretty good internet connection at GBH – a gig – so we were able to get the files as fast as they could digitize it. It's a three- to five-year project they are undertaking – about 3 petabytes (PB) worth of data. Right now, about 30 terabytes (TB) of data per month – picking up speed as they go," Shane said.

When COVID-19 suddenly forced editors to work from home, production teams needed a creative work around that could handle the massive increase in network access while still ensuring ease and security. The first step was to reconfigure the existing edit stations as virtual desktops outfitted with Avid and Adobe Premiere products. This gave editors access to GBH's centralized storage, while minimizing the amount of data going back and forth to their homes.

"When you are using one of those, it's just like your computer in the building with access to object storage and high-speed editing storage," Shane said.

But they still needed a friendly way for producers and external contributors to get the raw media files in and for editors and post-production teams to access the full resolution products. ExpeDat Desktop provided a familiar graphical user interface. Shane added that they needed ExpeDat to provide a friendly way for producers and external people to get the media files in, along with a nice GUI (graphical user interface). Since most people are familiar with an FTP or file transfer protocol-type program, it was straightforward for the editors to use from their home offices.

"We could configure all the permissions on our end and hand the editors a little app they could download, log in, and then start pushing footage up,"

Shane explained. "Then we set up some workflows on the back end, so they are able to access our object storage." At just a few megabytes and with no need for installers or admin privileges, the ExpeDat clients make it especially easy to get started.

The GBH Archives will benefit from this set up, too. As the media files are digitized, the vendor can drop digitized files directly into object store, alleviating the need to shuttle hard drives or LTO (Linear Tape-Open) tapes back and forth. The GBH archivists also add metadata, tags and transcodes into various formats for accessibility.

DEI CEO Seth Noble noted that this is a great example of best practices for both performance and security. "The Cloudian box is the storage unit and the ExpeDat software sits in front of that, interfacing with the S3 API. The DMZ with its external and internal firewalls provides layers of protection as does the protocol switching."

Added Shane, "We had a lot of concerns security-wise about exposing our S3 endpoint directly. This way we are layering it behind those layers of protection and authentication in ExpeDat so we make sure only people who should be connecting are connecting."

Setup was easy. "Out of the box, it worked completely," Shane said.

There are other accelerated transfer tools as well, but they are usually priced by traffic or other means that are less advantageous to the M&E industry because video files are huge. A one-hour show could be up to 30 terabytes, which is problematic in those other business models.
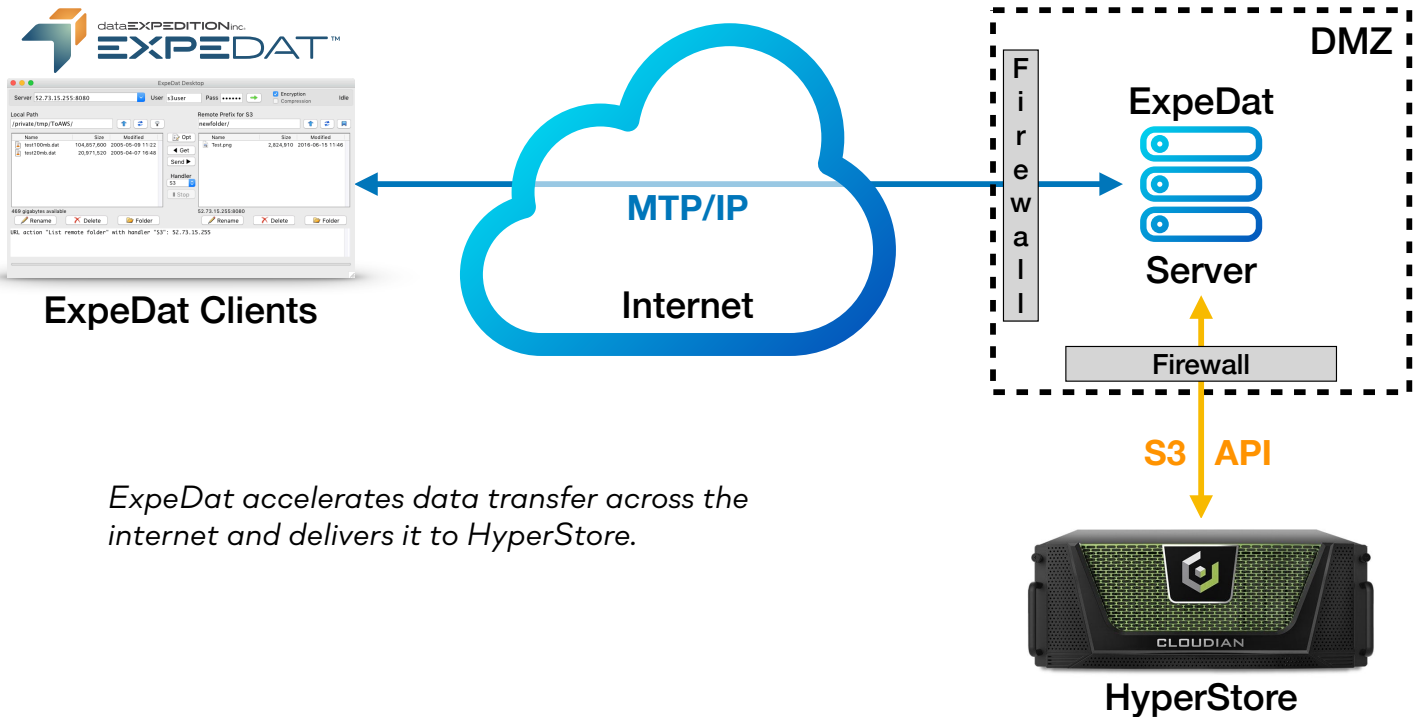
ExpeDat accelerates data transfer across the internet and delivers it to HyperStore.

Results

GBH now has fast, reliable, and efficient data transport and storage that is easily accessible to archivists, editors, and contributors. As for throughput, Shane tested at the DMZ and reached their internal limit of 6 gigabits per second (Gbps). "We put a cap on the external interface, so it does not interfere with other traffic we have," he added. "We originate national streams for TV and radio stations and give priority to that traffic." These speeds give plenty of room for expanding bandwidth in the years to come.

Despite the challenge posed by remote editing teams because of COVID-19, there is a silver lining – GBH may continue to utilize this workflow as an alternative to bringing people onsite.

"This isn't just a COVID work around," Shane said. "This is the way we want to work, and that's important."