Backblaze

Kanopy

Blockbuster Success: Scaling a Streaming Platform with Reliability and Redundancy

Use Cases
Primary Storage / Distribution

Keywords
Content Delivery / On-Demand Streaming
CacheFly

25,000+
Videos

Situation

Kanopy, a video streaming platform, serves more than 4,000 libraries and 45 million library patrons worldwide. They used an all-in-one video processing, storage, and delivery provider to manage 25,000+ videos, but reliability became an issue after missing files led to service disruptions. They needed cost-effective origin storage and safe backups they could count on.

3x
Redundancy
At The
Same Cost

<3
Day Migration

Solution

After moving to Backblaze B2 Cloud Storage in less than three days via our Cloud to Cloud Migration service, Kanopy unlocked a partner network that allowed them to use best-of-breed providers. Now, films are transferred from Kanopy's video processing infrastructure directly to Backblaze B2 where they're then served to users through the CacheFly content delivery network (CDN).

Result

As a key part of their new tech stack, Backblaze B2 gives Kanopy the ability to scale efficiently and rapidly accelerate content onboarding. Backblaze B2's reliability means Kanopy can provide a seamless customer experience. Taken together, Backblaze allows Kanopy to deepen their value proposition by providing thoughtful films their millions of users can't find elsewhere.
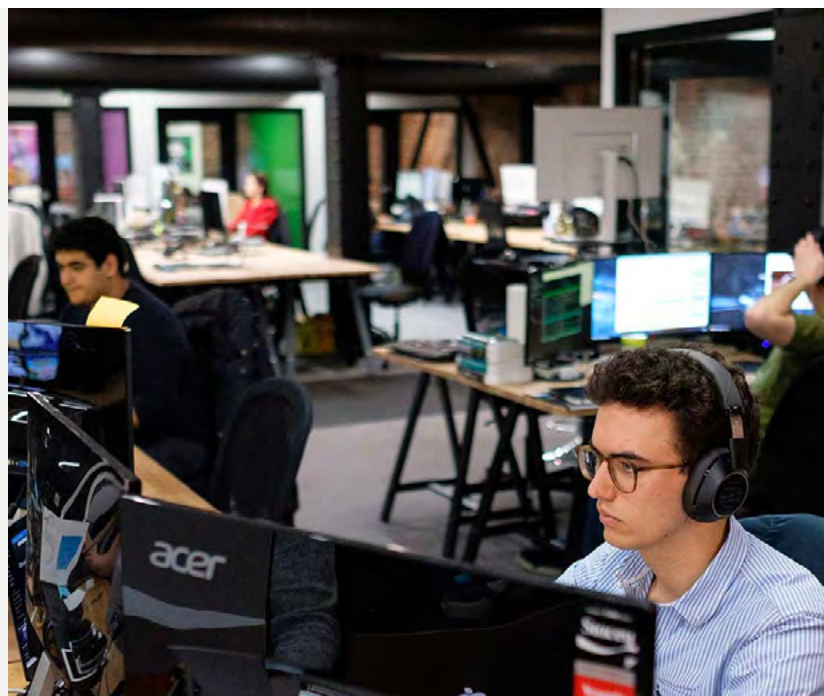
kanopy
THOUGHTFUL ENTERTAINMENT

Founded in 2008, Kanopy is a video streaming platform providing academic institutions and public libraries with essential films that foster learning and conversation. Today, they serve more than 4,000 library customers, providing free, on-demand access to video content for 45 million patrons worldwide. Their mission is to ensure that everyone has access to enriching films that bridge cultural boundaries, spark discussion, and expand worldviews.

Houston, We Have a Reliability Problem

Dave Barney, Chief Technology Officer at Kanopy, often describes the company as "Netflix for libraries." It's an easy comparison, until you realize Kanopy has 25,000+ titles under management, many thousands more than Netflix. That extraordinary depth is their key value proposition. Where Netflix entertains the masses with content that appeals to the biggest audiences, Kanopy serves the professor who needs a hard to find 1960s title for their existentialism in French New Wave cinema class, or the film buff looking for a long-lost favorite.

Rather than compete with Netflix, Kanopy pursued a blue ocean market in academic and, later, public libraries. The libraries pay for the service, offering patrons free access to films that can't be found anywhere else in the world. "On the public library side, we regularly hear from users

saying, 'I can't believe you have this film' or 'I haven't seen this film in 30 years,'" Barney noted. Unlike Netflix, when a title's popularity dips, Kanopy doesn't immediately retire it. Instead, they've ramped up content onboarding, adding the equivalent of Netflix's entire library, if not more, every year, with plans to continue expanding.

That growth made the reliability and cost of their storage solution all the more important. "We don't have deep pockets like Netflix, even though we have to keep a lot more data spinning. With all that content, reliable and inexpensive storage are two very important things from an operations and cost-efficiency standpoint," Barney explained. When customers started complaining about reliability issues, the development team knew that they had to make a change to their legacy storage and distribution infrastructure if they wanted their expansion plans to succeed.

> ❝ **We're rapidly accelerating the amount of content we onboard. Had reliability not become an issue, cost containment very quickly would have. Backblaze helped us attain both.**
>
> **Dave Barney,** Chief Technology Officer, Kanopy

What We've Got Here Is a Failure to Back Up

Founded in 2008 in Australia, Kanopy's early delivery model paralleled Netflix—they shipped videos on DVDs. As technology evolved, they relocated their headquarters to San Francisco and transitioned to streaming video, engaging an all-in-one video processing, storage, and content delivery network (CDN) provider to manage their media.

"It was a trusted provider, so nothing should have gone wrong," Pierre-Antoine Tible, Lead Video Software Engineer, said. "With time, we discovered things can go wrong."

In order for a film to be streamed without delays or buffering, it must first be transcoded, or broken up into smaller, compressed files known as "chunks." A feature-length film may translate to thousands of five to 10-second chunks, and losing just one can cause playback issues that disrupt the customer viewing experience.

"We weren't able to recover the files [our legacy provider] lost, so it was disabling HD for the films," Tible acknowledged. "On top of that, if we passively checked the CDN, it said the file was there. We had no way to discover missing files unless a user complained." Once they discovered a missing file, finding the original source file and transcoding it again took time and resources. That's if they could find it at all—one thing their legacy CDN didn't provide was backups. If the file couldn't be located in their primary storage, it was gone.

"It was very bad in terms of engineering, provisioning, and user experience," Tible reported. For a video streaming service of Kanopy's scale, a poor user experience was not acceptable. Nor was operating without a solid backup plan. They resolved to look for better cloud partners to increase reliability and redundancy.

> **❝ We don't have deep pockets like Netflix, even though we have to keep a lot more data spinning. With all that content, reliable and inexpensive storage are two very important things from an operations and cost-efficiency standpoint.**
>
> **Dave Barney,** Chief Technology Officer, Kanopy

The Beginning of a Beautiful Friendship: Kanopy + Backblaze

First, Kanopy had to find a way to move their data out of their legacy provider and make it S3 compatible in order to structure their tech stack the way they wanted. Their legacy CDN used its own storage type, so Kanopy's development team wrote a script themselves to move the data to AWS, where they planned to set up their video processing infrastructure.

While AWS works well for compute resources like transcoding and encryption, Kanopy's AWS bills would skyrocket if they used Amazon S3 for origin storage. They stream out 15,000+ unique titles each month, or roughly half of their total collection, which would incur massive egress fees through Amazon S3. It was not an option for storing production data long term, but making their data S3 compatible allowed them to consider other S3-compatible providers.

Prioritizing reliability above all else, Kanopy tested a few solutions for origin storage, including Backblaze B2 Cloud Storage and IBM. IBM also offered S3 compatible APIs, so the data would have been easy to move, but using IBM for storage meant implementing a CDN they didn't have experience with.

After running a proof of concept, Backblaze proved more reliable and gave them the ability to use their preferred CDN at the time to continue delivering content around the globe. Kanopy completed the migration of production data from Amazon S3 to Backblaze B2 using Cloud to Cloud Migration, moving 150TB in less than three days and discovering a few final missing files in the process. They've since switched to CacheFly CDN. Since Backblaze works smoothly across different cloud services, they were able to make the CDN change with no problem.
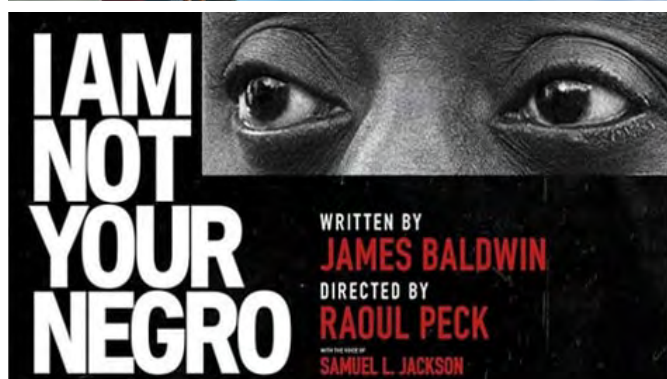
The Partner Network
That Dreams Are Made Of

Kanopy's vendor-agnostic, multi-cloud tech stack provides them the flexibility to structure the most cost-effective workflows. They use IBM Aspera to receive videos from contract suppliers like Paramount or HBO onto their video processing infrastructure on AWS. There, they transcode and encrypt the files for digital rights management purposes, generating upwards of 10,000 chunks per film in some cases. They send the transcoded chunks from AWS to their Backblaze B2 origin storage using a data transfer service and save an archive copy on Amazon Deep Glacier, giving them much needed redundancy in addition to their origin store copy on Backblaze B2.

Data stored on Backblaze B2 automatically feeds into CacheFly CDN. This allows content to live in Backblaze B2, yet be delivered with low latency and high speed.

The Backblaze partner ecosystem allowed Kanopy's development team to structure their tech stack using integrated, best-of-breed providers. "Backblaze's partners were important," Tible confirmed. Barney added, "The Backblaze partners were fortunate wins."

Here's Looking at the
Next 25,000 Movies, Kid

Though reliability was a main driver in Kanopy's efforts to overhaul their tech stack, looking forward, Barney sees their new system enabling Kanopy to scale even further. Backblaze offers simple pricing that allows Kanopy to achieve redundancy below their expected budget. "We're rapidly accelerating the amount of content we onboard. Had reliability not become an issue, cost containment very quickly would have. Backblaze helped us attain both," Barney attested. "Backblaze was a natural fit for us because it was more reliable and less expensive, enabling us to get ahead of where the business is going." Since Backblaze offers free egress between many partners and cloud services, including CacheFly CDN, Kanopy also pays $0 in egress fees to transfer data.

Adding even more titles deepens their value proposition and enables them to support independent film. As independent cinemas continue to close across the world, their service fills that void and gives independent filmmakers a channel for reaching consumers. How many titles will Kanopy add in the future? 10,000? 25,000? Every movie ever made? The great thing about their service is that you can find out. You just need to renew your library card.

About Backblaze

The Backblaze B2 Storage Cloud is purpose-built for ease. It offers always-hot, S3 compatible object storage that supports your workflows via third-party software integrations, APIs, CLI, and web UI. And it's priced for easy affordability at rates a fraction of other cloud providers. Businesses in more than 175 countries use the platform to host content, build and run applications, manage media, back up and archive data, and protect and recover from ransomware.

backblaze.com

Backblaze