BBC News and Digital Archive Transformation

Customer Challenge

The BBC is the creator and custodian responsible for a vast collection of digital assets documenting the history of an entire nation, stored in a pair of on-premises data tape-based archive systems.

Facing large impending CapEx refreshes to sustain these aging and proprietary archive systems, combined with concerns regarding future viability of the archive software solution itself, BBC decided to investigate their options. These concerns combined with the immense scale of data involved, almost 30 petabytes of unique content, presented BBC a daunting challenge. With ever increasing daily global production demands on the archive, it was also imperative that archive transformation planning ensured zero-impact on sustaining operations and addressed several complex surrounding workflows.

BBC sought a partner skilled in similar massive-scale archive transformation and migration initiatives to help assess what technical, operational, and commercial options exist.

Data Analysis and Empowerment

The BBC engaged Cloudfirst.io to perform a Legacy Archive Audit and Modeling Service across their existing on-prem legacy archive infrastructures located in England and Scotland.

The Cloudfirst data sciences team analyzed all technical and performance aspects of the BBC archive from inception through to present day, across all surrounding workflows. By modeling real-world archive growth, utilization, and performance, Cloudfirst extracted hundreds of crucial KPIs empowering BBC to make data-driven archive transformation decisions. Cloudfirst's insightful analysis resulted in a fifty-page report, comprised mainly of data visualizations, and defined a 12-month migration timeline for 30 petabytes of digital assets stored on thousands of T10000 and LTO8 data-tapes in the legacy archive system.

During the analysis phase, Cloudfirst worked closely with the Amazon Web Services (AWS) account team to produce various five-year total cost of ownership (TCO) models based on BBC's real-world data. These models compared each potential approach against the baseline of remaining "status quo" with their existing on-premises legacy archive infrastructures. Scenarios examined a cloud-based option for disaster recovery and business continuance, but also demonstrated a viable path to "active archive" in the cloud. By assessing archive utilization over close to 10 years of history, Cloudfirst was able to develop various reference architectures to target specific Amazon S3 storage classes for individual BBC workflows to ensure the most cost-effective path to cloud. Leveraging these analysis results, TCO models, and migration timeline, BBC was able to construct an internal technical and financial business case to justify a move to AWS.
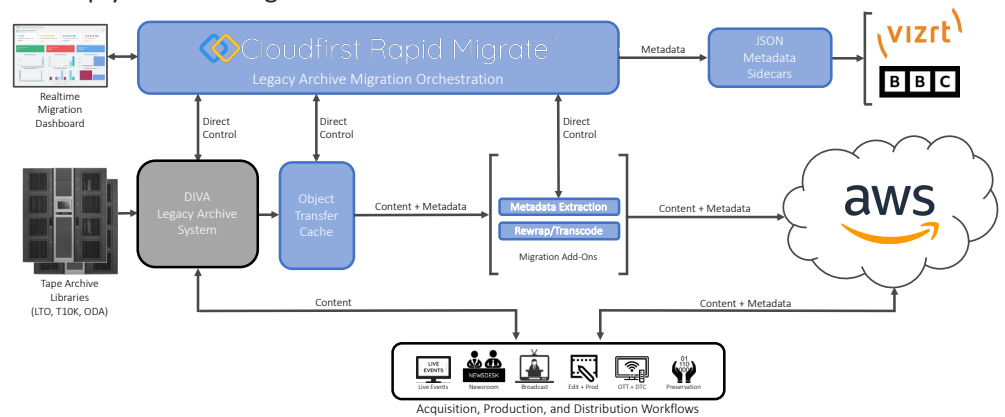
About BBC

The British Broadcasting Corporation (BBC) is a British public service broadcaster headquartered at Broadcasting House in London.

Originally established in 1922 as the British Broadcasting Company, it evolved into its current state with its current name on New Year's Day 1927. The oldest and largest local and global broadcaster by stature and by number of employees, the BBC employs over 21,000 staff in total, of whom approximately 17,900 are in public-sector broadcasting.

Archive Migration as a Service

Cloudfirst was selected by BBC to handle the entire legacy archive lift-and-shift to Amazon S3 with their Archive Migration as a Service (AMaaS) managed migration service. This AMaaS service follow a SaaS-based pricing model depending on the scale of unique content to be migrated from the legacy archive, with all costs exposed upfront. BBC was also able to take advantage of the AWS Migration Acceleration Program (MAP) to help offset migration costs. Although analysis showed a 12-month migration timeline, Cloudfirst was able to accelerate this and complete migration one month ahead of schedule, averaging 3PB per month.

One of the advantages of Cloudfirst's AMaaS managed migration service is assets never leave the client's own control. Existing production workflows continue uninterrupted while migration transparently progresses in the background. There is no need for upgrades or expansion as legacy archive resources required for migration are simply used for migration when available.

During migration all proprietary wrappers and interfaces are removed, decoupling assets from the legacy archive platform forever. Once migration is successfully complete, clients are left with unencumbered pure media assets in their own cloud accounts. This enables future cloud-centric archive workflows and clients are free to decommission their legacy archive platforms at any point thereafter.

As custodians of a nation's history, preservation is a key area of focus for BBC. Cloudfirst was the first migration partner to implement Amazon S3 Checksum/Fixity support (released mid-2022) in their AMaaS offering to meet BBC demands. Cloudfirst's unique implementation ensures bit-for-bit legacy archive migrations. With their unique workflow implementation, Cloudfirst is the only migration partner truly focused on end-to-end preservation grade migration services.

At the heart of Cloudfirst's AMaaS managed migration services is their Rapid Migrate migration orchestration platform. Rapid migrate dynamically senses and adjusts to day-to-day production demands ensuring zero operational impact during migration. The Rapid Migrate real-time dashboard provides a "single pane of glass" for clients to track progress and performance, while Cloudfirst experts manage the complexities of the migration operation behind the scenes.

To further reduce costs, BBC moved 24x7 operational and technical support for these legacy archive systems to Cloudfirst. Their 24x7 global support team was able to ensure smooth operations under increasing demands during migration. Cloudfirst also assumed a proactive operational role to help BBC ensure that existing production workflows continue to run optimally. Cloudfirst performed daily assessments of the BBC archive environment throughout the migration to ensure all resource and systems were performing to their full capabilities.

Metadata is Key

Cloudfirst believes metadata intelligence is imperative to future viability and accessibility to these valuable digital archives. Rapid Migrate retrieves core metadata from the legacy archive platform, extracts important technical metadata from the content itself (bitrate, codec, frame accurate timecodes, and more), and carries ancillary MAM/asset metadata along with the assets during migration. This invaluable metadata is protected and preserved throughout the AMaaS process and can be used to dynamically route individual assets to specific Amazon S3 storage classes on-the-fly during migration.

Metadata collected and mined during migration is used to immutably tag migrated objects in Amazon S3. This allows BBC operators to quickly locate content directly without having to use complex user-facing tools (MAM, etc.) and build metadata-driven cloud workflows, fully decoupled from existing on-premises tools. The entire collection of metadata, with pointers to the new locations of each digital object, is written out to JSON sidecar files on an asset-by-asset basis enabling metadata lakes, automation, or other cloud workflow construction.

Client Benefits

- Empowered data-driven digital archive transformation strategies
- Cloudfirst managed end-to-end migration for 30PB of valuable content
- All proprietary wrappers and interfaces removed during migration
- Zero interruption to sustaining production workflows during migration
- Metadata intelligence enables future media accessibility
- Unique preservation-grade migration ensures long-term asset protection
- Helped eliminated third-party service providers significantly reducing costs
- Halted continued capital investment in aging and proprietary tape archives
- Cloudfirst 24x7 legacy archive support improved SLAs and reduced costs

Services Involved

- Cloudfirst Legacy Archive Audit and Modeling Service
- Cloudfirst Archive Migration as a Service (AMaaS)
- Cloudfirst 24x7 Legacy Archive Proactive Support Services
- Amazon S3 Storage (S3, GIR, and GDA)
- Amazon Direct Connect
- Legacy Archive Environments
  - o Telestream DIVArchive
  - o Oracle StorageTek SL8500

cloudfirst.io