MUSEUM

A PERPETUAL EDGE ON PROTECTION

How CSI is helping a national museum to preserve millions of specimens in its collections and make the information available to a wider audience.

Our client: is a non-departmental public body sponsored by the UK Department for Digital, Culture, Media and Sport. The museum employs almost 1,000 staff and runs a series of educational and public engagement programmes to tackle scientific questions of wide interest.
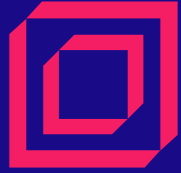
SCIENTIFIC RESEARCH DASHBOARD

80m
It is estimated that there are over 80 million specimens of natural and man-made interest in the museum.

2%
Due to physical restrictions, fewer than 2% of the available specimens are available to the public.

70%
Innovative storage and archiving solutions reduces data centre footprint by up to 70%.

CONTEXT

With hundreds of millions of specimens and archives across the globe, museum collections hold much of the information needed to tackle fundamental scientific and social challenges - from conserving the biodiversity to finding new ways to combat disease.

However, most of this information is only available to a handful of scientists. Our client wants to unlock this vast source of data to a wider audience.

GOAL

To improve the performance, stability, cost-efficiency and security of data storage, archive and recovery – and consequently to support the ongoing programme to digitise and release data about the 80 million items in the museum's collection.

CHALLENGE

Having installed the original deployment, CSI has supported a large IBM Spectrum Scale (GPFS) unstructured data store at the museum for the last ten years. Backup and archiving is provided by an extensive tape library.

With the progress of the museum's digitisation programme, storage capacity is under pressure and the requirement is expected to increase ten-fold in the next five years.

The museum also needs to make the stored data available on demand for collaboration or public consumption. Accessing large batches of archived data from tape results in contention and is slow and inconvenient which impedes sharing with external users.

A new solution was required to protect the digital data and to take advantage of emerging technology such as cloud storage.

SOLUTION

CSI worked with the museum to design new high-performance, scalable and resilient storage infrastructure that simplifies the addition of capacity and increase in performance as usage demands change. The solution stores data on tiers of disk (rather than tape) and combines the benefits of IBM's Elastic Storage Server (ESS) and Cloud Object Storage. Initial capacity is double that of the existing system and can grow to 10PB in the next five years.

With "always on" data checking and rebuilds, ESS is designed to eliminate missed writes, bit rot, and other potential corruptions of data.

Data written to the mirrored ESS system is migrated to cloud object storage in three geo-dispersed sites – in contrast to the existing system which provides only a single copy on tape.

OUTCOMES

The new storage and archive solution eliminates the need to manage a tape library and reduces the data centre footprint by 70%.

Data retrieval from cloud object storage is significantly faster and enables easy external access to the museum's digital assets.

With ESS versioning and block-level checksums, data is more secure and cloud resiliency means that data is available even in the event of a data centre outage.

The new solution also supports immutable storage providing a further level of data protection.

CSI has also introduced Spectrum Discover metadata management software that will give the museum enhanced visibility into its data.

"IBM SPECTRUM SCALE HAD ALREADY PROVEN ITS VALUE. THE CONTINUED DEPLOYMENT IN THE FORM OF IBM ESS APPLIANCES WITH TIERING TO IBM CLOUD OBJECT STORAGE REPRESENTS A MODERNISATION OF THE CLIENT'S INFRASTRUCTURE. THIS IS A PLATFORM FOR SUSTAINABLE FUTURE GROWTH AT PETABYTE SCALE."

Daniel Banche
Technical Architect
CSI

Tectrade
A CSI GROUP COMPANY