National Library of Scotland & Scality

Government

The National Library of Scotland houses one of Scotland's national cultural collections, and as a Legal Deposit Library it is entitled to claim a copy of everything published in the UK. The growing collections occupy 120 miles of shelving, with thousands of new items being added every single week. The library has a set a goal for one third of the collections to be in digital format by 2025. With that much data to hold, protect, and make accessible they needed a new approach. They found a partner in Scality.

Scalability

With already 31 million items in its collection, and fi ve thousand additions every week, growth is an absolute reality. As they continue to digitize the existing collection while constantly adding more new items as well, easy growth is an absolute requirement.

Low TCO

Conservative curation compels the National Library of Scotland to maintain three copies: one in each of their data centers, so ease of management and the ability to scale without replacement, coupled with software licensing that doesn't penalize them for choosing to replicate the data, can mean big savings.

Lower Stress

No LUNS means less maintenance for the National Library of Scotland, and no more lengthy tape backups from which they wondered if they would ever be able to restore. All that, and they can lose a data center and everything is fine, even available throughout.

With a history dating back to 1682, the National Library of Scotland was formally founded in 1925 by an Act of Parliament. At that time the collection was approximately 750,000 items. Now just shy of 100 years later, the collection holds more than 31 million items of massive value and importance. Its collection

Mass Digitalization

A collection that broad and vast requires a lot of care. In 2015 the National Library of Scotland published an ambitious strategy, including the aim to have a third of the collections in digital format by 2025. This is achieved by the acquisition of born-digital

"Tape is cheap, but it's labor-intensive and leaves some uncertainty, could we restore from the tapes if we had to?"

Stuart Lewis, Associate Director, National Library of Scotland

includes books, manuscripts, maps, music (both print and recorded), photographs, posters, postcards, newspapers, magazines, electronic journals, videos, the UK web archive and more. The holdings include such vastly differing and highly-valued items as one of the only 21 complete Gutenberg Bibles known to exist; the last letter written by Mary Queen of Scots just before her execution in 1587; first edition books, even comic books, including Harry Potter and the Sorcerer's Stone from 1997 and The Beano Comic from 1938; the set for The Cheviot, the Stag, and the Black, Black Oil, which is in the form of a book that measures three by four meters; programs from Football games; politicians' pre-election promises in the form of campaign fliers and more.

published materials, along with a new mass digitization program. Last year the Library digitized over 200,000 items, most of which are available for free public consumption on the website. "This throughput means we are generating a lot of data which needs to be digitally preserved and stored", said Stuart Lewis, Associate Director of Digital, National Library of Scotland. "We don't just store high resolution images, but also derivative data sets (from OCR, for example)." The Library has a lot of data on magnetic media that they have to convert as well, as Lewis told us, "We're transferring a lot of films from magnetic media, VHS, for example. We've only got about ten years to get that done before it may be impossible to buy players or spare parts, and we won't be able to rescue them."

Business Benefits

Licensing Model

Especially because replication is their preferred resiliency scheme, the National Library of Scotland appreciates Scality's licensing model. With Scality RING, customers purchase a storage software license based on the amount of data they want to protect—the original data. Not for the amount of data stored (which is doubled by replication). "It's like we get 2 for 1," said Lewis.

Scalability

The storage can grow as they achieve their digitization goals. All they have to do is add more capacity, linearly. There's no rip & replace, and no technology refresh: add drives to servers, add servers when the current ones are at capacity.

Standard S3 Interface

The National Library of Scotland made a decision to standardize on S3 so that they would have a single interface for cloud and on-prem object storage. This decision was sufficiently important to them to motivate a re-engineering of their home-grown asset management system to support S3. A robust, standard S3 interface was key to their decision to use RING. for EDF. This, their first foray into Software-Defined Storage, won the team an innovation award.

Replication for Preservation

Their former preservation storage involved one copy on SAN, backedup on tape. As the digitization initiative got underway, they had to re-think their preservation storage strategy, so put it out to bid. They decided to keep three copies in different geographic locations, preserving their precious data, and wanted to use multiple technologies for safety with checksums to ensure data integrity twice each year. Eliminating tape, ensuring streamlined growth and enabling simple on-prem and cloud storage synergy with S3 emerged as key goals. "Tape is cheap, but it's labor-intensive and leaves some uncertainty, could we restore from the tapes if we had to?" said Lewis. "And, it's hard to read every year to ensure that the data is OK."

National Library of Scotland made the move to Scality RING after getting proposals from "all of the expected vendors." They chose Scality RING as a foundation to their preservation storage strategy. In addition to providing the resilient S3-compatible storage they wanted, Scality further advanced the Library's data integrity assurance model, "Scality developed a Bitrot checker for digital preservation that streamlines our integrity checking process," said Lewis.

"We had been running a biannual full check of all files (read, compute checksum, compare, fix if required); now this runs in the background, and it's period configurable. We can download an audit report of results. It all works very well – stable and resilient."

The test came one day when the air conditioning failed in one of the data centers, and many of the servers shut down due to the heat. "Scality kept it all together with the remaining data center, and re-synced the replication as soon as the downed data center came back up," added Lewis.

Confidence Means looking at More Use Cases

Now that Scality has shown its strength, the National Library of Scotland is looking at adding use cases. "We're looking at moving other backups, replacing tape with disk-to-disk backup; moving the non-print collections and even offering it as a service to other organizations, given its multitenancy capabilities," added Lewis.

"Working with Scality has been brilliant"

Stuart Lewis, Associate Director, National Library of Scotland

Solution Components
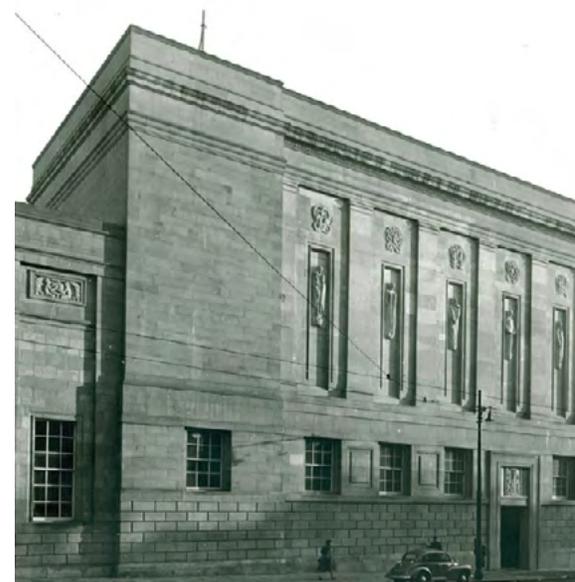
Software-Defined Storage

- Scality RING with S3 Object
- Storage Interface Cross-Region Replication 2 RINGS: Edinburgh and Glasgow data centers and a third copy
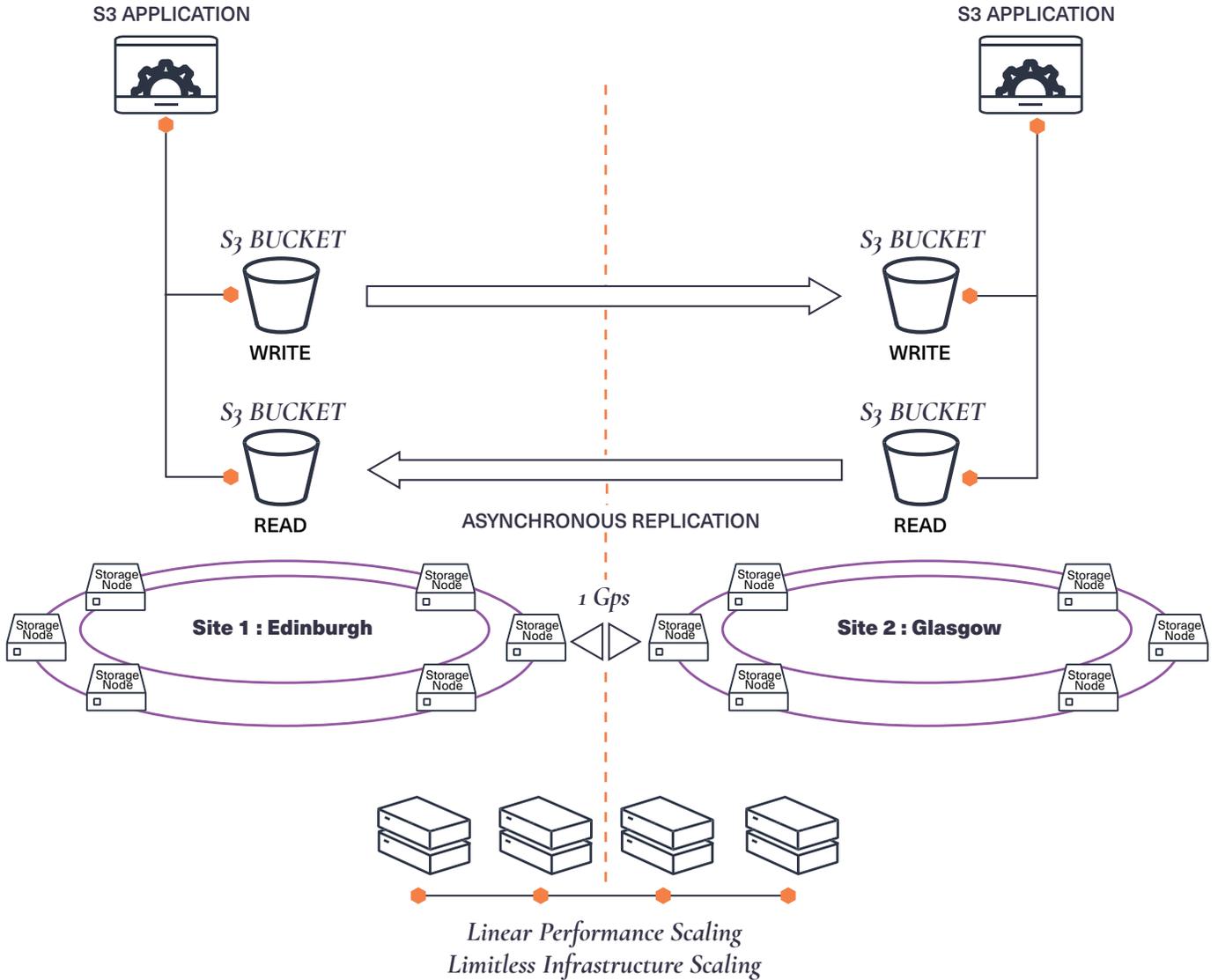- in AWS Glacier Deep Archive

Hardware

- 6 HPE Apollo 4510 Servers per site

Third Party Applications

- In-house developed asset management system, upgraded to work natively with S3

S3 APPLICATION

S3 APPLICATION

S3 BUCKET

WRITE

S3 BUCKET

WRITE

S3 BUCKET

READ

ASYNCHRONOUS REPLICATION

S3 BUCKET

READ

Storage Node

Storage Node

Site 1 : Edinburgh

Storage Node

1 Gps

Storage Node

Storage Node

Site 2 : Glasgow

Storage Node

Storage Node

Storage Node

Storage Node

Storage Node

Storage Node

Storage Node

Linear Performance Scaling
Limitless Infrastructure Scaling

About Scality Scality builds the most powerful storage tools to make data easy to protect, search and manage anytime, on any cloud. We give customers the autonomy and agility necessary to be competitive in a data-driven economy. Recognized as a leader in distributed file and object storage by Gartner and IDC, we help you to be ready for the challenges of the fourth industrial revolution.

Let us show you how.

Follow us on Twitter **@scality** and **@zenko** Connect with us on **www.scality.com**
San Francisco. Paris. Washington, D.C. Boston. Tokyo. Dusseldorf. London.

🐦 in f 📷

SCALITY