

The natural selection of storage

The Natural History Museum deals with an explosion of specimen data

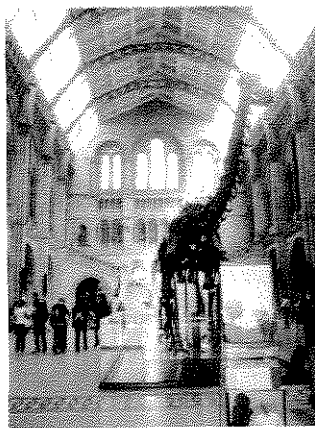
London's Natural History Museum has a physical storage challenge like no other. More than 28 million specimens in its zoology collection; roughly the same number of insects and other entomology specimens; 9 million in the palaeontology department; and over 5 million in the botany collection. And you can add to that almost 800,000 books and 20,000 photographic prints.

The electronic data held on those items has historically been relatively small but as researchers have adopted digital devices as essential tools – ultra-high-resolution cameras, electronic microscopes, scanners and video and sound recorders – demand for storage at the museum has soared dramatically.

Until recently, that was putting a huge strain on the museum's storage systems and associated processes. Although it had a basic storage area network (SAN) fabric based on an IBM FAScT500 disk array and an IBM TotalStorage 3583 Tape Library, the available capacity was proving completely inadequate: a maximum of 4.2 terabytes (TB) of disk storage and 6TB of tape.

"The SAN was absolutely maxed out," says Gavin Malarky, senior infrastructure analyst at the museum. Not only was the ICT unit acutely aware of this, the users were too. They were being told that, if they wanted more space, they had to delete existing files. Unwilling to do so, for two or three years researchers simply went out and bought low-cost external hard drives and CD writers.

The obvious dangers here were that locally held data was being stored on unreliable media and not being backed up. "Digital data needs to be treated as an asset every bit as valuable as specimens and books," says Paul Richards, head of



ICT at the museum. "Because of digital acquisition, we didn't know what the actual requirement was – it was hidden once people stopped coming to us"

TOWARDS A PETABYTE

To be cost effective, any replacement system also needed to offer different – and easily managed – tiers of storage. Much of the data the museum holds is accessed infrequently after an initial period of interest following initial publication. So, although there was a requirement for high-cost online disk capacity, there was an equally important need for near-line and offline storage.

The solution, implemented with the help of storage systems integration partner Tectrade, came in the form of three IBM disk subsystems – a DS4300 based on high-quality, fast disks and two DS4100 models that use lower-spec. and lower-cost SATA disks. That has enabled the museum

to scale its data under management on disk to more than 50 terabytes.

It has also upgraded its tape library to an IBM TS3554 system capable of holding 350TBs on high-speed LTO3 tapes. Currently, the system holds over 50TB on tape but it was chosen for its ease of expandability. "We can see the tape scale to a petabyte," says Malarky.

As important as these subsystems were, a critical overlay tool has added even greater value: IBM's storage virtualisation appliance, the SAN Volume Controller (SVC).

Scalability was key, says Malarky, and the SVC virtualises the storage environment, creating all devices as a single pool of storage. The museum simply plugs more disc systems into the SAN when required. "The benefit is that, if we want to add another 10TBs, we can get an IBM another DS4300 or buy disks elsewhere, sit them behind the SVC and, at the front end, it is all the same – simply a storage resource to the server," he says.

That frees skilled staff from repetitive storage management tasks. "The ease of management that virtualisation gives us made our lives a lot easier – flexibility and expandability within a single management console," says Malarky.

Updating to faster LTO3 tape drives has cut the back-up window dramatically. "We don't talk about archiving data any more, we just move it from online to near-line," explains Malarky. Back-up times have been cut by over 80%.

The demand curve for storage will only get steeper, says Richards, as the museum starts holding more researchers' videos and perhaps high-definition television used within exhibitions. "To meet those demands, we needed systems that would fit with what we might want to do further down the line." ☺

The natural
selection
of storage

