

Two-Pronged Challenge: Storing Large Amounts of Data for the Long Term

04-25-2008

Contributed by Elizabeth Millard

- Enterprises must find innovative ways to file archives in light of increasing compliance, productivity and cost-control needs.
- For the enterprises that choose to store data for decades, finding a long-term solution can be challenging, particularly as media changes.

Regulatory mandates like the Health Insurance Portability and Accountability Act (HIPAA) or the Sarbanes-Oxley Act, as well as awareness about potential litigation, have already caused public companies to store data for longer amounts of time and in a secure fashion, but many enterprises are using a strategy that goes above and beyond what's simply required.

Data archiving management as opposed to backup, which creates a temporary copy of a data that's usually overwritten later has become a source of concern at companies that want data stored for decades, not just a few years. But the sheer amount of data being produced and the uncertainty of long-term media could present a two-pronged challenge for the future.

According to a recent Enterprise Strategy Group (ESG) report on file archiving, many companies are grappling with explosive growth in file-based (also called "unstructured") content, such as Web pages, word processing documents, spreadsheets, presentations, multimedia files and other proprietary file formats with industry-specific applications. The research firm notes that enterprises have to move beyond rudimentary approaches to file archives in order to address increasing compliance, productivity, and infrastructure cost-control requirements.

Other ESG research shows that current archive practices are usually based on tape, or enterprises have archived data mixed with mission-critical data on primary storage systems. These two approaches aren't ideal. So, what's an enterprise to do if there's data that needs to be locked down and kept safe until 2020 or beyond?

"It's truly a real problem, and it's going to become even more of one in the next five to seven years, as the drives start becoming outdated or just die from age," says Paul Clifford, founder of technology consultancy Davenport Group.

The situation directly addresses several aspects of a company's technology strategy, such as performance requirements, system scalability, data availability and redundancy, says Bob Picardi, COO at RAID Inc., a managed storage service provider. It will also keep the debate about tape technology going strong, he adds. "There's the mindset of tape being the least expensive alternative for this kind of storage, but as disks get more utilized, tape could become a thing of the past."

Reliability, in particular, is the biggest issue when talking about archiving since it's already such a problem when it comes to backups, notes Chris Cummings, senior director of data protection solutions at NetApp. If a company has trouble making a reliable temporary repository of data, then creating a solid, decades-ready data pool will be even more daunting.

In making sure data and unstructured content can live long, healthy lives in digital suspension, enterprises need to buy devices and software that seem like they can last, but they must also create deeper strategic management around long-term storage, Clifford notes. "You don't want your technical resources determining your policies about information retention," he says. In other words, just because you have the storage capacity, it doesn't mean you should be dumping every bit of digital jet-sam and flotsam into an archive. Clifford adds, "It comes down to understanding the business and talking to other managers about what they need and for how long. Archiving intelligently begins with these conversations, not with buying more technology."

During these talks, non-IT executives should be informed about how archiving works, Clifford says. Many assume that data gets stored like file folders in a company basement, where items can just be pulled out whenever needed.

He notes, "Just because you have data sitting on tapes, doesn't mean it can be retrieved in the same format and instantly. People need to

know this when making their department's archiving decisions."

Also crucial in gauging whether an internal IT department will handle the archiving or whether it'll be outsourced to a storage firm is the strength of existing staff resources, Clifford says. "Generally, IT is putting out fires, so they need to think about whether they have the expertise required to address long-term archiving issues," he says.

Archiving tasks include upgrading systems, migrating information, bringing on new technology, retiring older technology, and staying informed about changing standards, native formats and storage changes, explains NetApp's Cummings. Also, security concerns increase as more data is stored, he says, so an IT department will have to be savvy about multiple layers of encryption and managing access.

Of course, hardware is also important. An enterprise should closely follow the numerous discussions about tape technology, although those are rife with contradictions: Tape is dead! Long live tape! But, in general, conversations with tape and disk vendors could be useful in understanding a company's specific needs, Picardi says.

"There is a great deal going on with combinations," he notes, "such as secondary archiving and different types of disk drives all blended into one."

Although long-term data archiving is creating a greater challenge for enterprises, particularly as energy concerns around power and cooling are increasing, Cummings believes vendors will continue being more innovative about blending media and creating technology built to last.

"Within the next five years, I think we'll see some good gains in terms of dynamic technology focused on this specific issue," he says. "In the meantime, that gives companies time to work out their data management decisions about long-term storage."