Case Study

Getty Images

STORAGE > HIGH SPEED STORAGE

Digital Forensics Company Modernizes Data Management

ellwood Evidence sticks with open source as it searches for a solution to handle its growing storage and management requirements.

Karen D. Schwartz | Jun 21, 2021

The “e” in ellwood Evidence is lowercase by design, but that’s about the only element that’s small in any way about the Toronto-based digital forensics company. The company is big in every way that counts—ambition, technology know-how and business savvy.

which began asking it a lot of questions related to incident response and e-discovery. Today, ellwood Evidence focuses almost solely on digita

Because of the company's deep roots in managed services, CISO William Ellwood and his team have a deep knowledge and respect for technology. Over the years, they have become expert at using open source tools and have built their foundation on the OpenZFS file system and the open source Linux-based Debian operating system. To handle data management and storage, for example, the team used ZFS-based file stores running on Debian-run servers with replication between sites. ZFS uses snapshots to track changes in the file system.

Over time, as ellwood Evidence became more successful, it started to outgrow its approach to data collection, management and storage. While the technologies the company built around open source components were solid, they couldn't handle the storage and management requirements of 120TB of data, distributed among about 300 million files. Data types run the gamut, from JPEG images and .txt files, to data from smartphones, security systems, photos and social media services, to boxes of paper documents. And because of industry requirements, ellwood Evidence is obligated to preserve the integrity and security of all data it handles—something that has become increasingly complex.

"We found that as data volumes grew, we were spending more time managing the data flows and making sure the systems we had strung together were doing their job effectively," Ellwood said. A lot of that time—up to 10 hours per week—was spent manually managing its own ZFS environment.

Finally, in 2019, it became apparent that the company had outgrown its homegrown approach. Ellwood then spearheaded the search for a solution that would centralize and organize its files, and allow it to scale out as well as up to preserve data availability. At the same time, open source technology still had a place in ellwood Evidence, and Ellwood wanted to make sure that any solution the

data gravity. As the company's main systems administrator—another hat Ellwood wears—he spent a lot of time thinking about the ideal location for data to meet the operational, availability and durability requirements of the data itself, and the applications that use the data.

"You have to think about it in two ways: the business workflow/business unit logical location of your files, and the physical file locations," he explained. "That has impact on disaster recovery and security obligations and chain of custody, which for us is a big deal."

Other important factors included the ability to scale easily across sites, and handle both large file sizes and large numbers of files.

Ellwood ended up choosing Hammerspace, hybrid cloud storage software that works in conjunction with a company's storage repositories. With a maximum 16-exabyte file size and 256 quadrillion zettabytes of storage, space would no longer be an issue.

And because Hammerspace manages metadata separately from data, it was a particularly good fit for a digital forensics company. That meant it would have the flexibility Ellwood needed to manage and protect the company's diverse and growing data. It would allow the IT staff to configure the file shares the way they needed and set their own policies. If one set of data required 11 nines of durability, nine nines of availability and specific performance metrics, that could be easily achieved. If the company needed to specify that files that match specific criteria need 10,000 IOPS, no problem. While Ellwood calls this flexibility, Hammerspace calls it a "storageless" model.

Seamless Switch

After testing the solution for three months in late 2020, ellwood Evidence went live

"We just spun up an NFS server that exposed our existing ZFS data sets. From there, Hammerspace assimilated and inventoried the data, and br Hammerspace file system," Ellwood explained. "We can plug and play essentially any of the open source solutions we decide to deploy. I can add them in as a new storage solution to Hammerspace, and it just starts being used."

The Hammerspace technology has adapted well to ellwood Evidence's workflow. The first step is data collection, where the company converts data from smartphones, laptops, servers, hard drives and other form factors to a digital form called a forensic image. Forensic images are essentially verbatim direct copies from devices and documents that include not only the text, but deleted files and metadata. Once in forensic form, the data is pushed to a virtual staging server and then to Hammerspace in the company's data center, where all machines are virtual.

Since making the switch, Ellwood has seen some positive changes. It has definitely reduced infrastructure complexity, he said. "It has allowed me to conduct midday outages on particular servers much more effectively. If I have my objectives set up correctly, I don't have to worry about what systems will be affected if I take something down."

Meeting governance, risk and compliance requirements also is easier. "We just say what availability we want on a particular set of files, what kind of durability, which sites we want to have access to these files, and then it just happens," he said.

The solution's flexibility also has given Ellwood room to experiment through A/B testing to evaluate different types of storage systems and configurations without fear of making a mistake. "I can very easily decide this is my test database, this is one of my production systems, and today I want to move that storage back end to an entirely new server. If that doesn't work, I can just push the data back, and the client won't even notice," he explained.

explained, the digital forensics company had to replicate data to a regional target at another site. This resulted in a lot of manual processes to ensure t[illegible] lost and that a valid snapshot of the data existed before spinning up a cold site to a warm site. The new system is essentially active-active, which means Hammerspace synchronizes changes within sites within 15 seconds. If there is a conflict, it is tagged and made available.

Over time, Ellwood hopes to squeeze more value out of Hammerspace. He's already planning to push out NVMe PC4 storage systems, which should work seamlessly with the technology. "We'll deploy the server and then tell Hammerspace it's an NFS file. That's it," he said. "If it doesn't work, we decommission that storage volume and Hammerspace moves the files off of that storage server."

Ellwood also plans to take more advantage of the ability to set up rules that apply to metadata and build them into the file system. For example, ellwood Evidence uses FileCloud—essentially a web-based option to Dropbox—that allows metadata to be added to files. Ellwood plans to experiment with setting rules that selectively mark items as immutable. That means that if something is classified as a piece of evidence, that data, once written, can't be changed.

"We have a great store of plain text and documents scattered around across matters and projects. I'd love to be able to manifest arbitrary views on my file system," he said. "For example, because it's a metadata database that gives you mediated access to your files, I would like to be able to say, 'Show me any file that matches this criterion in this folder structure, maybe by attribute.' Then you could say, 'Wherever these files are, I want to see everything from 2017 by file category.'"

TAGS: CLOUD STORAGE

IBM Boosts Protection of All-Flash Storage Range
JUL 20, 2021

A Guide to Edge Sto
JUN 09, 2021

New Storage Trends Promise to Help Enterprises Handle a Data Avalanche
APR 08, 2021

Flash Storage: A Guide to Best Practices
FEB 17, 2021