Case Study: Texas Tech University System improves storage infrastructure TCA and TCO by adopting Caringo Swarm

TEXAS TECH UNIVERSITY

I had the chance to talk with Ryan Curry (Assistant Managing Director) and Lance King (Server Administrator III) from Technology Operations and Systems Management (TOSM) at Texas Tech University System (Texas Tech). They are in the process of migrating their primary file storage environment (250 TB) from a traditional storage solution to a Caringo Swarm object storage infrastructure. This case study is another brilliant example of how easy it can be to migrate to a modern software-defined platform based on an object store backend, while obtaining benefits both in terms of TCO (Total Cost of Ownership) and service improvement.

The Scenario

The TOSM IT team serves more than 50,000 end users distributed in 200 departments. Students, researchers, professors and others share the same infrastructure and the use cases are numerous, ranging from small (0.5 TB) repositories for departmental file shares to rich media archives larger than 60TB. The total capacity of active storage is more than 250TB, which has to be added to 250TB of backup and archival data for the complete picture.

> *"Overall, I think we will be moving more and more towards object storage since it is simpler to manage and to scale."*
>
> *Lance King, Texas Tech*

The previous infrastructure was built out of a frontend file virtualization platform and multiple arrays, with different characteristics, in the backend. The diverse array types were necessary for achieving an acceptable mix of $/GB and performance, while the frontend was in charge of sharing data volumes and managing an automated tiering mechanism—a complex infrastructure that led to a good compromise of the TCA (total cost of acquisition) but required considerable effort in data and system management.

> *With the introduction of Caringo Swarm and its ecosystem components like FileFly, Texas Tech has finally found a way to get the best TCA while, in the meantime, improving TCO and making the infrastructure ready for the future.*

With the introduction of Caringo Swarm and its ecosystem components, like FileFly, Texas Tech has finally found a way to get the best TCA while, in the meantime, greatly improving the TCO and preparing the infrastructure for future developments in terms of scalability, new applications and innovative services for the end users.

Why Object Storage?

To find the best solution for its needs, Texas Tech went through a thorough 6-month evaluation process that included traditional, software-defined and open source vendors. Eventually, they chose Caringo Swarm, a software-defined object storage solution that allows Texas Tech to obtain the most flexibility (given that they can take advantage of commodity x86-based storage servers) without compromising on other aspects, like data protection, data services, ease of use and performance. In fact, the TCO is improved tremendously thanks to the ease of use and overall robustness of the platform.

> *"Maintaining backups and archives is very expensive. With erasure coding, replication and other data protection features, we can move away from traditional backup. Object storage can take care of itself. Resiliency and flexibility is the real benefit for us."*
>
> *Ryan Curry, Texas Tech*

At the same time, open source solutions were discarded because of their potential complexity in system management and tuning, while traditional storage systems were not cost competitive and lacked the flexibility and agility needed by a challenging environment like the one at Texas Tech.

The migration process, which essentially consists of large file system copies, is coming along smoothly and System Administrators have not had to deal with any issues. The advanced data protection scheme implemented by the object storage backend, coupled with seamless integration offered by Caringo FileFly, makes this end-to-end solution transparent to the end users and traditional backup processes have become unnecessary. This translates into major time savings when it comes to system administration, freeing up resources for other activities.

As always happens with object storage systems, Texas Tech is planning to use Caringo Swarm to perform more than just file services. Future projects include utilizing it as a backup target, and possibly as a storage target for developers creating next-generation applications. Since Swarm can also run in some of the public cloud architectures, future expansion in that direction is possible as well. This will further simplify the infrastructure and improve services for end users.

Why Caringo Swarm?

At the beginning, Caringo Swarm was chosen for the ease of use and the affordable price point. But over time, Texas Tech has found Caringo's support service very responsive. Now that the new storage infrastructure is in production, outages cannot be tolerated and 24x7x365 support is essential for achieving the best service levels for end users.

> *"Sys Admin efforts are very minimal compared to the effort required for the previous platform."*
>
> *Lance King, Texas Tech*

Caringo has also demonstrated that Swarm can start small and grow seamlessly over time. The smallest configuration (3 nodes) can quickly grow up to multi-petabyte configurations by adding additional nodes, without performance becoming an issue. With Swarm, Texas Tech is free to choose the hardware that best fits its needs and the architecture of the front-end interface FileFly, which has several configuration possibilities, decouples performance from capacity. In fact, it can cache most of the writes locally and sync with the object storage backend at a later time. This mechanism is particularly helpful with distributed environments like Texas Tech and in all cases where network latency can vary from site to site.

Caringo's ecosystem goes beyond traditional object storage solutions, which are usually accessed only via APIs, and can also provide file services as well as indexing and smart search features. This approach opens a vast number of possibilities and allows consolidation of unstructured data in a single large repository (eliminating silos) which can be accessed by end users from anywhere and with any device.

Bottom Line

Today, with the right technology, object storage can be adopted by organizations of all sizes without undermining scalability over time. Caringo does exactly that, it has the right solution for allowing IT organizations to start with a small initial investment and grow as data and applications require. The software-defined approach is another important benefit that allows customers to purchase hardware and software separately, which is not only a technical advantage but also a financial one as a result of the different depreciation cycles.

Caringo has been building a very impressive ecosystem around Swarm. Advanced access methods are now available alongside S3 APIs while a very comprehensive and simple web-based UI makes it easy to manage.

Caringo has been building a very impressive ecosystem around Swarm. Advanced access methods are now available alongside S3 APIs while a comprehensive and simple web-based UI makes it easy to manage, enabling the shift from tens of terabytes to petabytes under management a single Sys Admin. Caringo offers a set of end-to-end integrated solutions that cover a large number of use cases ranging from IoT and Big Data to next-generation and traditional file services.

Texas Tech has embraced object storage and Caringo, having found a broad solution to solve their day-to-day storage problems while at the same time being prepared for future challenges.

Think Global, Act Local!

Why Juku?

Jukus are Japanese specialized cram schools and our philosophy is the same. Not to replace the traditional information channels, but to help decision makers in their IT environments, to inform and to discuss the technological side that we know better: IT infrastructure virtualization, cloud computing and storage.

Unlike the past, today those who live in the IT environment need to be aware of their surroundings: things are changing rapidly and there is a need to be constantly updated, to learn to adapt quickly and to support important decisions—but how? Through our support, our ideas, the result of our daily global interaction on the web and social networking with vendors, analysts, bloggers, journalists and consultants. But our work doesn't stop there—the comparison and the search are global, but the sharing and application of our ideas must be local and that is where our daily experience, with companies rooted in local areas, becomes essential in providing an honest and productive vision. That's why we have chosen: "think global, act local" as a payoff for Juku.

Author

Enrico Signoretti is an analyst, trusted advisor and passionate blogger (not necessarily in that order). He has been immersed in IT environments for over 20 years. His career began with Assembler in the second half of the 80′s before moving on to UNIX platforms until now when he joined the "Cloudland". During these years his job has changed from highly technical roles to management and customer relationship management. In 2012 he founded Juku consulting SRL, a new consultancy and advisory firm deeply focused on supporting end users, vendors and third parties in the development of their IT infrastructure strategies. He keeps a vigilant eye on how the market evolves and is constantly on the lookout for new ideas and innovative solutions. You can find Enrico's social profiles here: http://about.me/esignoretti