THE CLIPPER GROUP
Navigator™
Navigating Information Technology Horizons℠
Published Since 1993 Report #TCG2009047 October 21, 2009

Zetta Builds Primary Cloud Storage to Enterprise Specifications

Analyst: Anne MacFarland

Management Summary

There are many situations where strength, resilience, and security must be *built in*, not *bolted on afterward*, such as bridges and skyscrapers. The basic architecture is determinative – future additions will be built around its constraints. **Cloud IT infrastructure, and particularly storage clouds, have a similar need for well-thought-out design.** Unlike the temporary infrastructures offered by compute clouds (think particularly of the test and development clouds that have seen early success), **the value offered by service compute clouds is one of persistence.** In the case of enterprise storage clouds, what they store is of high value to both its owners and to potential predators. Commercial use of cloud services mandates an approach of secure multi-tenancy**.**

Therefore, storage clouds are defined not as much by their vastness as by the mesh architectures and control systems that give them their essential character. As a bridge for a multi-state highway (with associated utility conduits) is different from a bridge built for a two-lane back-country road, so storage clouds built for the enterprise information will be different from storage clouds built for consumer use, specific scope, or transitory information. Nonetheless, enterprise cloud storage is undeniably attractive.

- More organizations, particularly large organizations, are physically distributed and need an elastic, usage-priced infrastructure to meet their dynamic information use needs without affecting other IT systems.
- Many organizations want to avoid the costs of building, staffing and maintaining a local infrastructure – for certain operations.
- Information, used in the right place at the right time, optimizes most businesses.

Zetta, a storage cloud vendor based in Sunnyvale, CA, has approached cloud storage leveraging its distributed computing heritage and storage expertise to meet high expectations. Excluded from their focus, at least for now, are real-time, low-latency operational data (such as data marts and database/inventory clusters), where nanosecond response time is needed. Zetta's focus is on the primary, operational data of an organization's file-oriented applications. Zetta offers a range of protection options to fit the activity, use, and risk profile of the stored information. While backup and offline data recovery are accomplished by the Zetta cloud, Zetta focuses on delivering a more complete and business critical functionality.

Early beta customers range from legal firms and social sites to energy, architectural engineering, and media and entertainment organizations. Their demands have focused on scalable primary file services, active archives, and bursts of storage requirements – none of which are easy to meet with on-premises storage.

Zetta's architecture supports these requirements and more. For more details, please read on.

IN THIS ISSUE

Architecting for Vastness

Much has been made about the increasing bulk of enterprise information. Much has been made about its wider use. The consequences of both are implacable. Every incident of loss of customer data is intolerable.

Inadequacy, tolerable at a smaller scale, will show up as large-scale catastrophes that cannot be hidden**.** Anomalies in data must be addressed proactively because, in the scale of vast repositories, they become significant. Security must be comprehensive. Authenticity of information must be assured even during the inevitable failure and rebuild of infrastructure elements.

The high rate of data growth is expected to continue – something that will require a different approach to management as well as to architecture. Yet enterprise clients still want a detailed view of their cloud storage operations – this view reveals trends in business operations and the costs of those operations. A simple dashboard with red green and yellow dots will not do.

Architecting for secure multi-tenancy ups the ante for each layer of the architecture. As with physical buildings, it is prudent to add door attendants, security checks, not only at the ground floor, but also in the elevator vestibules and at the entrance to individual offices.

Zetta's Storage Cloud Architecture

Zetta's beta customers' demands are instructive. Many saw Zetta as the better alternative to building new storage infrastructure, but they had specific requirements. They needed native file system access. They had multi-threaded applications that required an inherently non-blocking architecture. Even traditional write-intensive operations would generate consequent reads that had to be well supported. Their data was growing daily. 10GbE connectivity was needed to support day-to-day use of cloud resources.

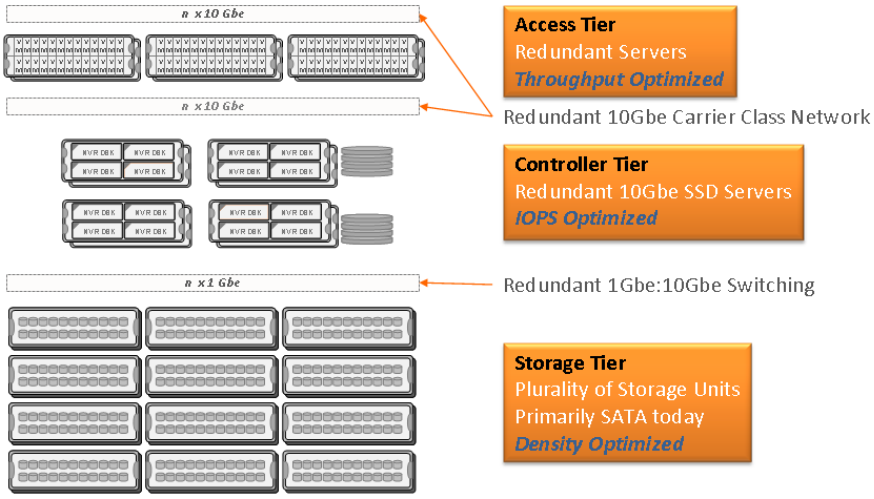
To meet these important requirements, Zetta has chosen to use an *N+3* paradigm (a scale-out paradigm accessorized with three spares) for all three layers of its architecture, reinsuring the resilience of the infrastructure against multiple failures. The top layer is of virtual machines. The middle layer is of solid-state drives meshed with 10 GbE Connectivity. The bottom layer is SATA. (See Exhibit 1, below.) But, it is the functionality built into these layers that provides Zetta's differentiation. Zetta has considered both the security needed for multi-tenancy, and the availability support needed for data on a vast scale.

What Zetta offers its customers is vastly scalable file-based storage. "There are no architectural limitations to expanding to block-based storage and high I/O," says CEO Jeff Truehaft, "but Zetta has chosen to focus initially where enterprise storage growth is most problematic, and that is file systems."

Protocol Translation Layer

Zetta's protocol translation layer is optimized for throughput. It contains nodes that

Exhibit 1 — Zetta's Tiered Architecture



Source: Zetta

perform the tasks of protocol translation, volume virtualization, inline data verification, and quality of service management.

When storing (writing) data, the Zetta cloud generates a write receipt that is a secure SHA[1] hash. The receipt is created as the write enters the *Zetta Virtual Volume*, and is persisted perpetually to verify the validity of an ultimate read back of that file, as well as optionally returned to the customer for verification that what was sent by the customer was what was received by Zetta. Along the way, it is used to detect network errors and silent corruption, and to enforce quality of service. Zetta uses *Xen* virtual machines[2] to allot a VM for each mount point. This gives each customer separation at the access tier – a separation and security that are inherited through underlying levels.

Zetta is *POSIX* compatible, so no applications need be changed to use the Zetta cloud. *CIFS*, *NFS*, *Web DAV*, and other protocols are supported natively, avoiding the need for a protocol translation appliance[3] that might be another point of failure. POSIX compatibility also implies a strong consistency model – required by most in-use enterprise applications today, and a rarity among existing cloud storage providers.

Zetta File System Management Layer

Zetta's middle layer is IOPS optimized, using SSDs, and VRAM to minimize latency and media memory for high throughput. It consists of redundant 10GbE-connected SSD servers. The functionality of this layer includes volume management, data protection management, and quality of service continuity.

In this layer, Zetta captures file metadata and stores it separately from the data in a very fast, SSD-enabled data store. This data store facilitates file anomaly detection and file recreation from its constituent stored chunks.[4]

The middle layer also handles chunk/extent management and chunk/extent hash management. Since security is handled at the protocol (in motion) and data management (at rest) tiers, the middle tier can use hash codes to check for accidental changes in data, and use Zetta's strong encoding algorithm to repair errors using spare system cycles.

Data "Chunk" Management Layer

Zetta's bottom tier is density optimized. This layer handles traditional storage management functions, data validation, and encryption-at-rest of the chunks and extents stored. This granular level of encryption is key to scalable and secure multi-tenancy that still can leverage striping and redundancy to support a variety of qualities of essential service factors (e.g., availability). Think of it as an apartment house with security at entry, private elevators, and a key to your personal spaces.

Each storage node is a full computer, not a dumb disk shelf, enabling the encryption-at-rest and data validation functions. The physical storage is SATA.

Zetta uses its own RAID 6 N+3 system of striping and extending. The striping is across at least eleven nodes with three parity stripes, so up to three full node failures are survivable. This approach also allows the system to avoid rebuilding an entire file. Sensing corruption at the chunk level is key to a fast, local, and less-expensive chunk rebuild. Rebuild and repair inevitably impact performance, so they should be as local as possible.

This resilient, tiered architecture allows Zetta to insure that data, wherever written, is valid well before the read event. It proactively addresses bit-errors and silent corruption – problems that are manageable in small, local arrays, but are significant issues in large clouds. The virtualization that is inherent in Zetta's architecture also allows files to be instantly provisioned and quickly grown.

Allaying Cloud Nightmares with Security, Availability and Integrity

Security pervades the Zetta storage experience end-to-end. For transport to the cloud, Zetta supports protocol-enabled wire-line encryption. Some protocols support this; others do not. To ensure transit security where the protocols available do not support wire-line encryption, Zetta offers support for virtual private circuits with a cross-connect for redundancy or VPNs with SSH.

Zetta has developed its own PKI that uses

---

[1] SHA stands for *Secure Hash Algorithms*. Zetta feels that CRC (Cyclical Redundancy Checking) is not good enough for multi-tenancy situations. Cyclic redundancy checks are not secure, but are used often to check for accidental changes in data in transmission.

[2] Zetta intends to support other virtual machines in the future.

[3] Such an appliance could also introduce locking and file management that do not scale.

[4] In the future, this architecture could also support findability actions, like crawling and indexing, which could be accomplished without hitting the stored data.

standard key management. It is a performance-oriented sub-key system that supports volume rebuilds. For integrity at the chunk level (bottom layer), security must be very granular. This also ensures that encryption at rest will not affect performance, since retrieval is parallelized.

Data integrity and availability are similarly assured at every level of the cloud architecture – integrity by the write receipt, the availability by the N+3 paradigm, the non-blocking architecture of all layers, and the granularity of rebuilds. This is an architecture built to address risk, not to ignore it.

Zetta's cloud solution is a storage architecture built to handle enterprise data loads and enterprise service requirements. It can address the consequences of success as well as assuage worries about insufficiency.

System Management Portal

Zetta's beta customers have said they want snapshots, native file systems access, and geo-replication for disaster recovery. They want a complete detailed view of their data and the events that occur – but they want to set policies for expected actions and let the cloud "make it so." They want to leverage the cloud-based instant automated rebuild when needed – but also want to know when such incidents happen, how long the rebuild will take, and when it is complete. After all, while it may house data belonging to second tier (business critical, as opposed to mission critical) applications, it still is primary data that is important to the enterprise. Zetta's management interface supports this level of detail.

There is also a need for disaster recovery remoteness. In its beta stage, remoteness is limited to the Bay Area of California. Soon after launch, Zetta will add sites in southern California, Virginia, and New Jersey to its Santa Clara original site. Two additional sites in central United States will be added in 2011. With these expanding cloud sites, Zetta can meet this need now for customers in the US and will meet it for all customers in 2010.

Pricing

The Zetta pricing model is for one base price that will include all features, customer service, and customer support. This price will be $.25/GB/month for 1TB. There are discounts for larger footprint volumes, and for longer-term contracts. Zetta will allow month-to-month usage, with the monthly bills based on net average capacity. These structures meet the needs of small organizations and those with seasonal needs, as well as those looking for a longer-term solution for their CapEx pains.

The cost of connectivity to cloud storage is another concern of many cloud customers. Early experience with the efficiency of Zetta's architecture has led to a guarantee that connectivity costs will never be more than 20% of the total cost of Zetta Cloud service, even with adverse use cases, such as high monthly data turnover.

Conclusion

Zetta throws a whole lot of hardware at the cloud storage challenge – but it uses this hardware well to support enterprise-style availability and secure multi-tenancy that are basic requirements for many organizations. What makes their solution affordable is how they have architected their storage cloud for both resilience and very high utilization. If these characteristics suit your organization, Zetta is worth a closer look.

About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

Anne MacFarland **is Director of Information Solutions for The Clipper Group.** Ms. MacFarland specializes in strategic business solutions offered by enterprise systems, software, and storage vendors, in trends in enterprise systems and networks, and in explaining these trends and the underlying technologies in simple business terms. She joined The Clipper Group after a long career in library systems, business archives, consulting, research, and freelance writing. Ms. MacFarland earned a Bachelor of Arts degree from Cornell University, where she was a College Scholar, and a Masters of Library Science from Southern Connecticut State University.

- ***Reach Anne MacFarland via e-mail at Anne.MacFarland@clipper.com or at 781-235-0085 Ext. 128. (Please dial "128" when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.