



# Enabling Cloud Storage for the Enterprise

## At a Glance

Cloud Storage is another facet to the “cloud” ecosystem that is emerging in the marketplace. This paper covers:

- Benefits of enterprise cloud storage solutions
- Inhibitors to widespread adoption in the enterprise
- Why new tools will be needed to overcome the inhibitors
- Emulex’s thought leadership on an interesting new set of tools on the horizon

As companies rapidly develop new technologies and practices to deliver on the promise of “the cloud”, the infrastructure surrounding the cloud architecture needs to quickly evolve to provide options for service and support. Many of the gaps in this infrastructure that have inhibited storage-as-a-service concepts from gaining traction are now being filled by more robust, creative solutions from companies like Emulex. By targeting these gaps in the cloud infrastructure, Emulex is uniquely positioned to provide tangible capital and operational cost savings by integrating technology that allows seamless transition to a cloud storage model.

Solutions targeting cloud infrastructure seem to be appearing on every technology company’s roadmap, even to the point of old technology being repositioned to take advantage of the buzz created around cloud. However, despite offerings from companies as disparate as, say, VMware compared to Iron Mountain, they all have one thing in common: they are focused on delivering dynamic scaling and pay-as-you-go acquisition.



Figure 1 Cloud storage ecosystem.

\* Venture-backed start-up company

## Enabling Cloud Storage for the Enterprise

Dynamic scaling is the cloud architecture promise of elastic storage capacity on-demand—use what you need only when you need it. Today's storage paradigm is centered on the premise of over-provisioning. Data pools tend to grow at a fairly constant and predictable rate. However, storage hardware cannot be purchased in small increments. When a company begins to run low on storage capacity, they must buy new arrays—a significant expense—that sits largely unused for a period of time, perhaps years, as the data grows to fill the new hardware. This cycle repeats itself over and over, with large chunks of the company's capital being budgeted periodically to increase the available storage pool. A comprehensive analysis is provided by Michael Armbrust, et al in *Above the Clouds: A Berkeley View of Cloud Computing* (<http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html>).

Conversely, cloud storage provides a variable cost, pay-as-you-go acquisition model that scales dynamically right along with the data pool. When utilizing cloud storage, there is no longer any need to purchase storage in large chunks. Companies pay only for what they need when they need it, taking advantage of the expected price declines in storage capacity costs over time. Furthermore, contracts with storage service providers (SSPs) are typically viewed as operating expense (OpEx) versus capital expense (CapEx). Converting large pools of archived storage from CapEx to OpEx can have significant positive benefits to companies' bottom line. The cost savings equation is one of the primary benefits of cloud storage.

### Sounds Great, But...

The buzz created by cloud storage vendors and venture capitalists is slowly gaining momentum as certain benefits of the strategy become clearer: pay-as-you-go pricing, OPEX versus CAPEX, reduction of hardware, etc. However, skeptics point to the big picture of total cost of ownership (TCO), which includes aspects of which enterprise customers are more wary. Data security, availability and questions of ownership are all things that keep the CIO awake at night, making him or her reluctant to place his or her company's data in the hands of a third-party vendor. Not to mention perhaps the biggest challenge, integration of cloud storage into their existing infrastructure. Each of these issues is worthy of further discussion.

### Data Security

When data is kept within the confines of a data center, there are recognized methods for ensuring that it is kept safe. When moving data outside of the data center, as is the case with public cloud storage, security concerns become a top priority.

A key to improving both security and availability issues is the incorporation of an information dispersal algorithm (IDA). The class of IDA algorithms was first explored by Michael Rabin for security and fault tolerance of data distribution. The IDA (after encryption) splits data into a configurable number of slices that include error correcting codes, with those slices dispersing to a number of storage endpoints. Individual slices are not useful by themselves to exploit the original data. IDA also tolerates loss or corruption of slices, allowing reconstruction of the original data from a predefined number of slices.

IDA combined with current encryption technology provides additional security for data stored in the cloud. Because data objects are sliced up and dispersed using IDA, even a security breach at a cloud SSP where the encryption keys are compromised is less likely to yield useful data to the intruder. To make use of the data in the cloud, a hacker or SSP employee would have to also gain access to a quorum of the data slices stored elsewhere. Assuming that the data is dispersed across multiple SSPs that are different companies, this is unlikely. IDA works seamlessly with database encryption, host-based storage security, or self encrypting disk drives providing an additional layer of security. Additionally, this increased level of data protection for cloud environments is achieved without additional security management overhead.

### Availability

Even the most sophisticated data centers in the world experience outages, and this holds true for cloud SSPs as well. Unfortunately for the industry, several outages of public cloud vendors have been publicized and somewhat sensationalized lately. That's the nature of being a public provider of a service, whereas a data center outage, within the walls of any Fortune 1000 company, is unlikely to be publicized in the same way.

Nevertheless, outages do happen, and Murphy's Law ensures that they will happen at the worst possible time. Remember, the data most likely to be stored in the cloud is cold archive data, so the likelihood of it needing to be retrieved urgently is low. But rest assured that there will be cases where an archived object will need to be retrieved at the exact moment that the cloud SSP is unavailable for some reason.

Again, this is where IDA provides an interesting solution. Using IDA, the data objects will have been sliced and dispersed to multiple cloud vendors, and an outage at one SSP is unlikely to also affect the others, so the object can be reconstructed using the data stored at the SSPs that remain active and the user need not worry about—or even know—that one SSP was down. Dispersing data across multiple SSPs is the best—and possibly only—way to achieve a level of service from cloud storage that approaches the five 9s that data center CIOs expect. (A side benefit here is that now there is no pressure on SSPs to offer the five 9 SLA, which will keep costs of the service lower.)

# Enabling Cloud Storage for the Enterprise

## Data Ownership

One final argument for IDA. With data objects dispersed across multiple cloud SSPs, no single SSP has a complete data set. If the relationship between the user and the SSP ends for whatever reason, the CIO need not worry that the SSP followed through on their commitment to delete the data because it is incomplete and unusable to anyone.

## Integration into the Data Center

Integrating the new cloud storage paradigm into the proven, hardened and trusted IT infrastructure may be the most difficult challenge out of all of those mentioned here. While there is simply no way to gain the advantages of cloud storage without some additional work (no free lunch), there must be a method for adding a cloud storage tier with minimal disruption to existing IT processes and procedures.

Many would say that simply adding an agent to an existing server, providing access to the cloud, is enough. True, this does provide access for the data center to cloud storage, but it also introduces new worries, such as dealing with cloud APIs, and being locked into a single cloud SSP. This method may work well for the small business customer, but will not be attractive to the enterprise.

An ideal solution is a cloud storage on-ramp, independent from the cloud SSPs, that seamlessly integrates with existing data center infrastructure, providing access to multiple cloud SSPs. This solution has the added benefit of shielding the IT administrator from all of the various, non-standard cloud APIs and transport protocols that are foreign concepts in most data centers today. Ideally, this cloud on-ramp solution will provide a block-level interface using Fibre Channel (FC) or SAS protocol and present itself as simply another storage target. With this type of cloud storage on-ramp the enterprise customer truly begins to feel comfortable with a cloud storage tier.

## E3S – Enabling Elastic Storage

Emulex is developing a solution that will allow enterprise IT architecture to bridge the gap to the elastic storage pools of cloud storage providers seamlessly. Emulex's Elastic Enterprise Storage (E3S) provides the tools to fulfill enterprise requirements for cloud storage regardless of the cloud SSP chosen:

- Customized, seamless integration with existing storage platforms
- Transparent bridge between block domain and cloud SSPs
- Improved protection and availability of data in the cloud
- Robust metrics on cloud SSP performance

In its simplest form, E3S is a block storage (FC or SAS) on-ramp to the cloud, converting standard block-level data to cloud objects (extents) and sending the objects to the cloud for archival. Because E3S is a block device, it can be integrated into existing storage architectures as a target device for virtualization appliances (for example, see Figure 2), virtual tape libraries, de-duplication appliances or any number of other block initiators without significant changes to existing firmware or processes.

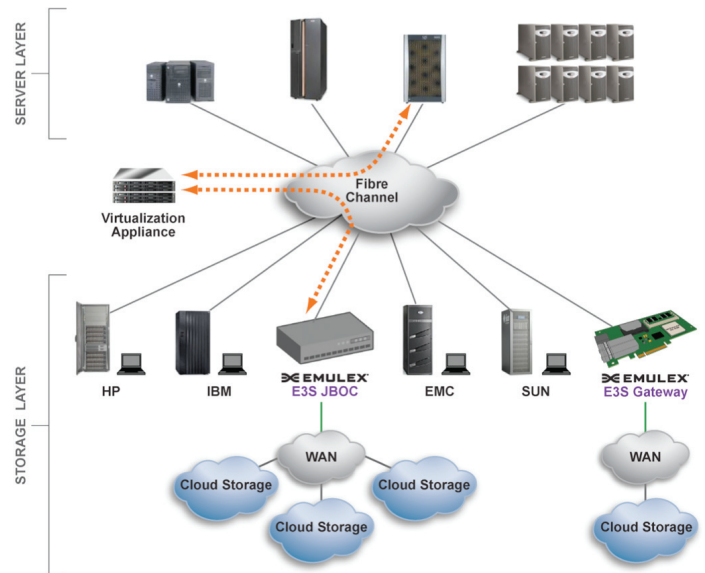


Figure 2 E3S in the enterprise—virtualization target.

The block-storage-to-cloud-storage on-ramp can provide various levels of service and functionality depending upon the customers' intended use and desired level of security, availability and performance. The simplest version of E3S is a gateway that will provide a bridge between block data and a single cloud SSP. This application lends itself well to cold archive data that has a low likelihood of being needed quickly and has a low impact if the SSP is unavailable for a period of time for any reason.

# Enabling Cloud Storage for the Enterprise



Figure 3 HA Cloud JBOC Enclosure.

An enhanced version of E3S would be implemented as an E3S JBOC™ (Just-a-Bunch-of-Clouds). The E3S JBOC would have dedicated compute horsepower to provide a more robust feature set, including dedicated cache, greater processing power and the ability to disperse data across multiple cloud SSPs, thereby greatly improving both availability and security of the data. Using IDA, along with additional data integrity options, E3S JBOC can guarantee data availability in cases where a cloud SSP is out of service or has experienced data corruption or loss. These features will be discussed in greater detail in a future paper.

Simply bridging block storage to cloud SSPs is only part of the value of the E3S toolset. CIOs are not keen on the idea of training their personnel on new, foreign protocols such as ReST and SOAP or learning to write to the cloud APIs simply to gain the ability to archive data in the cloud. A cloud on-ramp device must abstract these APIs and protocols in a way that makes them completely transparent to the IT administrator. Couple this with the opportunity to create competitive bids for cloud storage provider services, and E3S becomes a big win.

E3S is designed to support the APIs of multiple cloud SSPs and their transport protocols so that the enterprise IT administrator will never have to deal with them. Furthermore, complex operations surrounding encryption, compression, data integrity and IDA are simple configuration options controlled through the E3S user interface. Emulex Enterprise Elastic Storage designs make enterprise integration of cloud storage as simple as installing a familiar JBOD.

## Why Emulex?

Emulex has a long history of success in abstracting complex storage issues that goes back before 2004 with the introduction of FC-SATA protocol bridges. Emulex's innovative solutions enable the use of inexpensive SATA disk in enterprise storage arrays.

In an interesting parallel to bridging the enterprise to the cloud storage tier, the FC-SATA bridging technology was much more than simply bridging physical layers. The ultimate solution allowed storage OEMs to use any SATA disk available, while the Emulex bridge presented the storage system with the personality of the OEM's favorite FC HDD. The complexities of various SATA disk personalities were abstracted and hidden behind the bridge, allowing storage OEMs to greatly improve the time to market for lower cost enterprise storage solutions.

Today's issues with using cloud storage in the enterprise are similar but on a larger scale. Customers need a solution that makes the complex issues of cloud storage disappear, allowing seamless integration into their existing storage infrastructure. If this were not the case, the CIO would need to waste too much OPEX making the cloud work, which would ultimately defeat the purpose of using the cloud in the first place.

## Cloud Storage and the Enterprise Customer

The debate continues in earnest regarding the exact role of elastic storage in the enterprise. Will customers utilize public cloud providers, build elastic data stores in-house via private clouds or will there be a combination of both approaches, sometimes called a hybrid cloud? Regardless of how these questions are ultimately answered, one thing seems clear: there are attractive business drivers to adoption of elastic storage and enterprise customers are anxious for the storage industry to dissolve the technological barriers that are preventing them from taking full advantage of what these technologies have to offer. By providing a seamless bridge between existing IT infrastructure and cloud storage, as well as eliminating much of the complexity associated with new cloud storage protocols, Emulex provides a robust solution to many of the issues facing enterprise customers as they seek to take advantage of this new storage paradigm.

This paper has touched on the high-level attributes of Emulex Elastic Enterprise Storage. In future papers, we will explore in more depth how E3S overcomes the barriers to enterprise adoption of cloud storage, use cases for the technology and more.



**World Headquarters** 3333 Susan Street, Costa Mesa, CA 92626 +1 714 662 5600  
**Wokingham, UK** +44 (0) 118 977 2929 | **Munich, Germany** +49 (0) 89 97007 177  
**Paris, France** +33 (0) 158 580 022 | **Beijing, China** +86 10 68499547  
**Tokyo, Japan** +81 3 5322 1348

### Connect with Emulex

[twitter.com/emulex](https://twitter.com/emulex)
[friendfeed.com/emulex](https://www.facebook.com/emulex)
[bit.ly/emulexlinks](https://www.linkedin.com/company/emulex)
[bit.ly/emulexfb](https://www.youtube.com/channel/UC...)

[www.emulex.com](http://www.emulex.com)

©2009 Emulex, Inc. All rights reserved. This document refers to various companies and products by their trade names. In most, if not all cases, their respective companies claim these designations as trademarks or registered trademarks. This information is provided for reference only. Although this information is believed to be accurate and reliable at the time of publication, Emulex assumes no responsibility for errors or omissions. Emulex reserves the right to make changes or corrections without notice. This report is the property of Emulex and may not be duplicated without permission from the Company.