SERVICE PROVIDERS' GREATEST NEW GROWTH OPPORTUNITY:

Building Storage-as-a-Service Businesses

Today, with the availability of off-the-shelf cloud storage platforms, service providers of all sizes can now build cloud storage services with very little risk. Storage-as-a-Service (STaaS) can now be easily added to a service provider's offerings with minimal cost, little deployment time, and significant ROI. As the enterprise market adopts the cloud model and more applications (and thus use cases) are created, STaaS stands poised to be a massive revenue generator. In addition, this same cloud storage system can reduce operating costs for the service provider by utilizing the same infrastructure for secondary storage for cloud compute systems.

This whitepaper provides an overview of the market size for STaaS as well as describes the technologies and ecosystem which make STaaS viable. The paper goes into detail about design best practices, using the Cloudian Community Edition—which allows a 10TB system to be deployed with no software license fees—as an example of the value and promise of STaaS.

ACCORDING TO 451 RESEARCH, STORAGE AS A SERVICE REPRESENTS A LARGE AND RAPIDLY GROWING MARKET WITH A CAGR OF 47% WITH A TOTAL MARKET OF NEARLY $6B IN 2015.

STaaS Market Size

According to 451 Research, Storage as a Service represents a large and rapidly growing market with a CAGR of 47% with a total market of nearly $6B in 2015. Of this $6B, online backup and archiving represents 20% (a 50% – 50% split between backup and archiving). 80% of the $6B, or $4.8B in revenues, is driven by dedicated, capacity-based, raw storage services as part of Platform-as-a-Service (PaaS) or Infrastructure-as-a-Service (IaaS) where the storage can be purchased on a stand-alone basis[1].

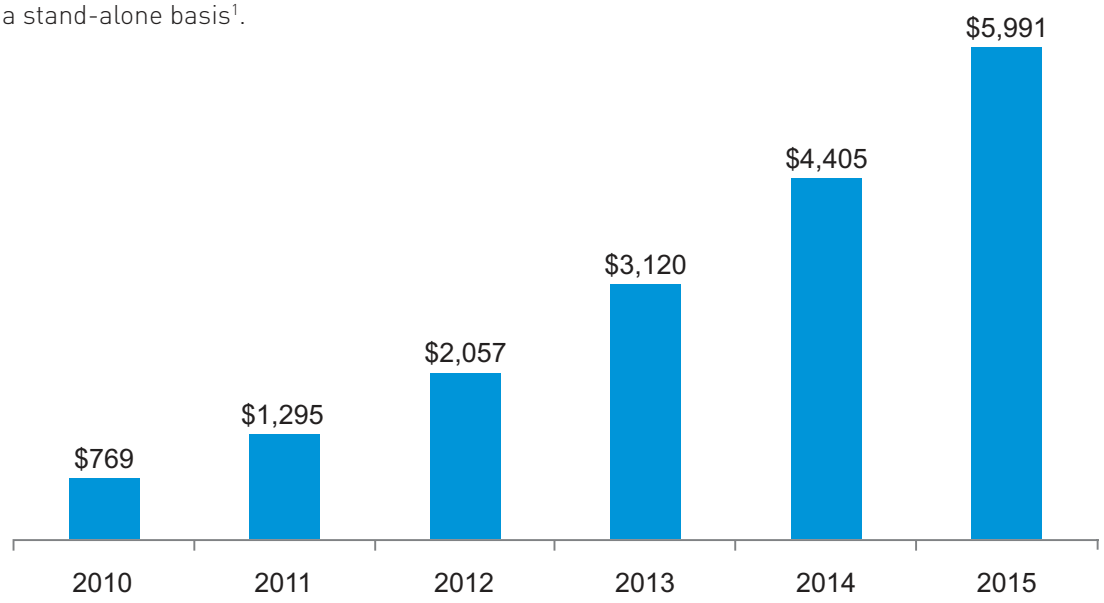


FIGURE 1:
Storage-as-a-Service Total Revenue, source 451 Research

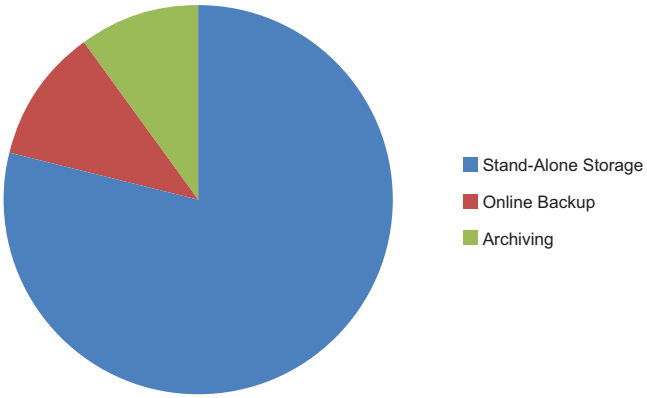


FIGURE 2:
2015 Storage-as-a-Service Applications, source 451 Research

Cloud Storage is a Necessity

Why is this market growing so rapidly? The core value proposition and growth driver of cloud computing is a shift away from a capex-centric, advanced-reservation model to a fully automated, opex-driven, on-demand model. With storage being a critical component of the cloud, this movement is also changing the way storage is procured and provisioned. Storage represents a portion of the underlying cloud infrastructure and an enabling technology, and a true cloud offering cannot exist without a variant of cloud storage[1].

THE CLOUD MOVEMENT IS DRIVING INNOVATION IN STORAGE ARCHITECTURES

The cloud movement is driving innovation in storage architectures because it requires levels of scale, reliability, performance, functionality and cost that, at multiple petabytes of capacity, cannot be delivered by traditional storage system designs[1]. Object storage technology provides the basis to meet the needs of a cloud storage system.

Why Object-Based Storage?

Physically, object-based storage arrays are comprised of stackable, self-contained disk units like any other SAN. Unlike traditional storage arrays, object-based systems are accessed via HTTP and the clusters are built with commodity hardware. That access method, plus their ability to scale to petabytes of data, makes object-based systems a good choice for public and private cloud storage solutions. An entire object storage cluster of disparate nodes can be easily combined to become an online, scalable file repository[2].

To retrieve data, the object storage software reads metadata and object ID numbers that are associated with data within the object storage environment. When retrieving data, the associated object ID number and metadata is read and the data is made available via the object storage software. This eliminates the need to delve into deep file structures, and intelligent caching can speed the process. The metadata also enables storage administrators to apply preservation, retention and deletion policies to their data[2].

Object storage is a good fit for use cases where static, hugely scalable archival storage is required and the use of SAN arrays designed to host virtual servers or highly dynamic application systems would be cost-prohibitive. It is also a viable alternative for data archiving to tape or optical disk; that media is slow, non-scalable and difficult to retrieve[2].

According to the Forrester report File Storage Costs Less in the Cloud Than In-House, the cost of storing 100 TB of data internally is close to $1 million per year. In contrast, for an enterprise to store the same 100 TB of data in the cloud, the cost is estimated to be $250,000 annually[5]. This clearly demonstrates the value for enterprises to move to STaaS. For the service provider, building the STaaS for this 100TB can now be done with significant profit margins using off-the-shelf software and commodity hardware.

The Value of Being S3 Compatible

In addition to choosing an object storage-based solution, it is important to closely consider which API you use. Amazon's AWS S3 has become the de facto standard API for cloud storage, and for good reason. As of Q4 2011, Amazon has stored approximately 760 billion objects, which is up from 40 billion in 2010. This amazing growth has energized the S3 ecosystem and as of 2012 there are over 350 tools, applications and SaaS using the S3 API. Within the ecosystem there are applications for backup, archiving, project management, media management and many more. There are tools in the ecosystem for file/bucket management, secondary storage tools, API integration tools, and others.

Supporting the S3 API is an important decision for service providers and enterprises. Not only does it afford service providers and enterprises all of the tools, applications, and gateways found within the ecosystem, it also provides a ready-made market. Although Amazon does not break out its AWS revenues, it does list those revenues in an "other" category which is at a $2B run rate for 2012, the majority estimated to be for AWS. Additionally, utilizing the widely adopted S3 API enables the very important hybrid storage model expected to be used by a large number of enterprises regardless if they initially utilize a 'cloud-in' or 'cloud-out' approach to cloud adoption.

SUPPORTING THE S3 API IS AN IMPORTANT DECISION FOR SERVICE PROVIDERS AND ENTERPRISES.

Today, in the Amazon world, you build and launch new applications directly onto cloud IaaS. Then, as you get to scale and a significant amount of steady-state capacity, you pull workloads back into your own data center (cloud-in), where you have Amazon-API-compatible infrastructure. Because you have a common API and set of tools across both, where to place your workloads is largely a matter of economics. You can develop and test internally or externally, though if you intend to run production on AWS, you have to take its availability and performance characteristics into account when you do your application architecture. You might also adopt this strategy for disaster recovery[3].

Alternatively, you move your data center (with its legacy applications) into the modern era with virtualization, and then you build a private cloud on top of that virtualized infrastructure; to get additional capacity, business agility, and so forth, you add external cloud IaaS (cloud-out), and this requires a compatible provider[3].

What enables this hybrid model is the common API. But the cloud-in approach has proven difficult without an off-the-shelf S3 compatible storage platform that can be quickly, easily, and inexpensively deployed. Without the availability of Cloudian, or other S3 compatible object stores, enterprises would have to expend significant engineering and IT resources in order to develop in-house solutions or utilize open source software. The cloud-out approach going forward is going to largely depend on many service providers offering STaaS with S3 compatibility allowing enterprises choices in STaaS providers.

Not All S3 Compatibility is Equal

Most object storage systems, by definition, comply with a basic set of APIs. They either support REST or SOAP, or both. Objects are stored in buckets with metadata associated with each object. Object storage systems provide for a distributed storage environment and enhance reliability through replication. On top of this basic object storage layer is another API that provides for the user functionality of Read, Write, Delete, etc. In the S3 API this basic functionality is covered with Get, Put, Delete, etc. At this point alone, a system could claim S3 compatibility. Indeed, while most cloud storage solutions offer some level of S3 API compatibility, this generally only addresses the most basic API calls to create, list and delete buckets and objects. The common approach is to use an API translator from the S3 syntax into the vendor's native API set.

Cloudian is the only cloud storage solution that is natively tracking AWS S3 in terms of API fidelity and feature coverage. Cloudian natively supports the advanced S3 API functions that are so important to both enabling robust functionality and ensuring hybrid model compatibility. AWS is a rapidly innovating and fast moving target, and Cloudian is committed to keeping pace with enhancements as they become available. In addition, Cloudian has provided extended functionality for Multi-tenancy, Multi-data centers, Virtual buckets, flexible management GUI, and REST API for system management (refer to figure 3).

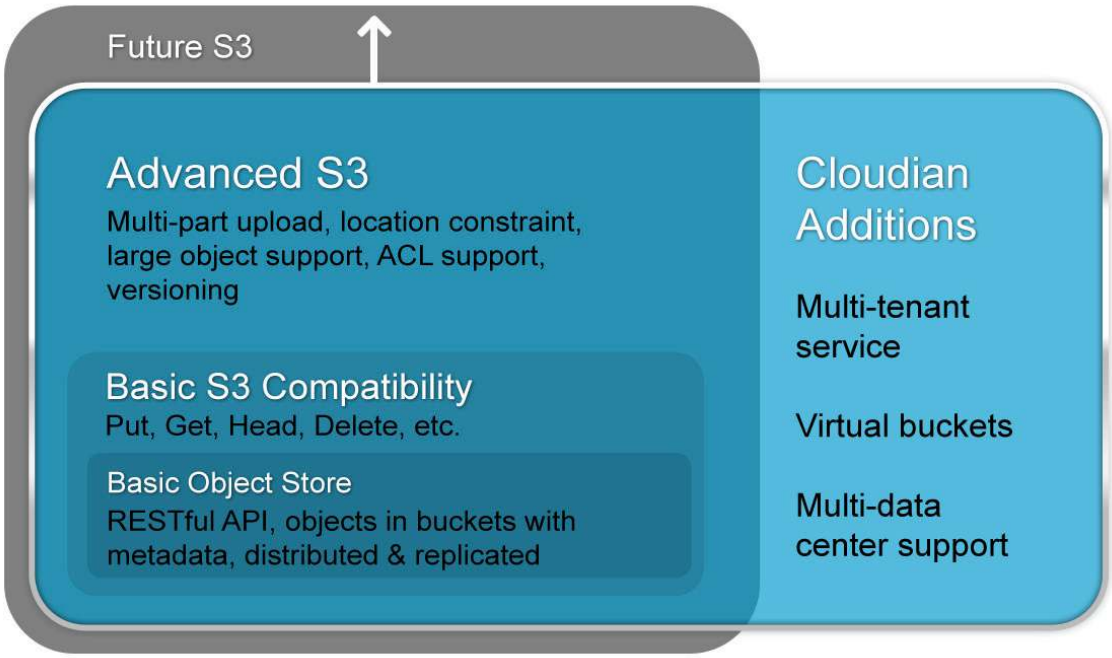


FIGURE 3:
S3 Compatibility Stack

Gateways

Cloud storage is well suited for a variety of storage workloads where the data is unstructured. In addition to backup and archive data, files themselves are also considered unstructured data and are a good candidate to take advantage of the benefits of cloud storage services. Enterprises are used to working with files and file systems and in order to bridge between the file system and the object system used by cloud storage, cloud gateways have emerged.[5]

Cloud gateways are an emerging type of technology that act as brokers to major cloud services, allowing enterprises to not have to do API integration with the service provider themselves. A cloud storage gateway is a network appliance or server, which resides at the enterprise premises and translates cloud storage APIs (REST) to block-based storage protocols such as iSCSI or Fibre Channel or file-based interfaces such as NFS or CIFS. Some cloud storage gateways also include additional storage features such as backup and recovery, caching, compression, encryption, duplication and provisioning[4].

Cloudian is actively partnering with many of the cloud on-ramp providers. Each offer their own feature set and advantages to the service provider and enterprise. Gateways are also priced in different ways, ether as a hardware/software solution with an upfront cost or as a service billed via the service provider.

Some of the cloud gateways that Cloudian suggests are:

- Citrix
- Riverbed
- Twinstrata
- StorSimple
- Panzura
- Fobas

Cloudian® Use Cases

We will now illustrate the power and ROI of moving to an object storage solution via the example of the Cloudian platform. There are two primary use cases for the Cloudian object storage platform (refer to figure 4):

1. Storage as a Service
2. Secondary Storage for Cloud computing platforms

Enterprises and Service Providers can utilize the Cloudian Platform for applications such as back up, device synchronization, collaboration, archiving, CDN storage, and low cost secondary tier storage in a scale out environment. In addition there is a second use case that can lower the cost of

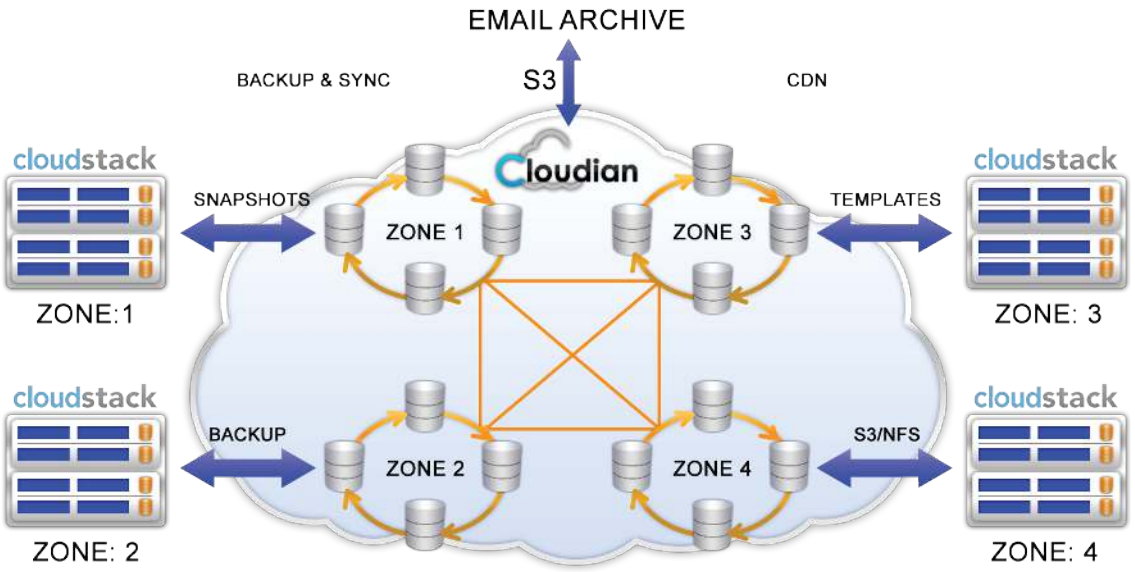


FIGURE 4:
Using Cloudian as STaaS with CloudStack

storage for both Service Providers and Enterprises who are utilizing a cloud computing platform. Cloudian object storage can be used as the secondary storage subsystem for these platforms. Not only does it substantially lower the cost of the subsystem, but it also allows for templates and snapshots to be accessed across availability zones, significantly improving the overall reliability of the cloud computing environment.

Cloudian System Design Overview

Cloudian employs a fully distributed and replicated peer-to-peer architecture with no single point of failure. It easily scales horizontally using commodity hardware so deployments can cost effectively start with a few commodity servers in a single datacenter and then scale up as usage increases to thousands of servers distributed across multiple datacenters managing hundreds of petabytes of data.

Cloudian's distributed architecture with automatic replication and recovery services make it highly resilient to network and node failures without data loss. Similarly, when scaling the storage datacenter cluster or performing maintenance, changes in node availability are automatically detected and existing data rebalanced. (refer to figure 5)

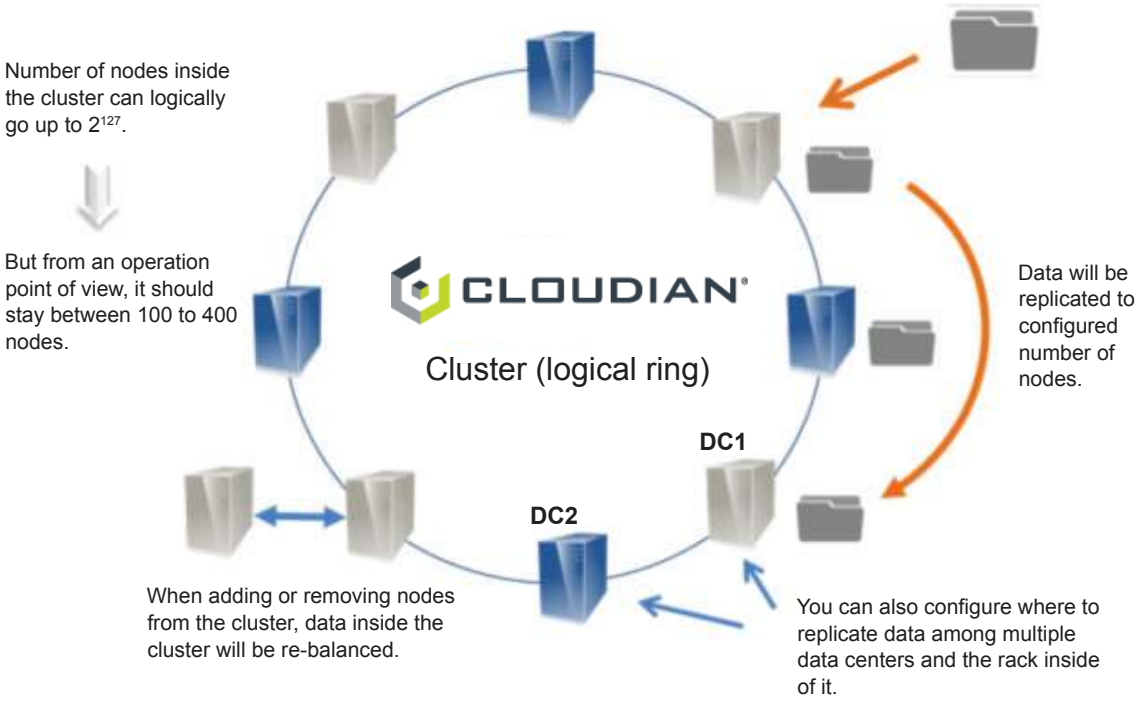


FIGURE 5:
The Cloudian Ring scales to an unlimited number of nodes

The Cloudian Platform consists of (refer to figure 6):

- Storage Nodes: Heterogeneous commodity servers running Linux, which can be of any size and number of SATA hard drives with no theoretic maximum capacity per node.
- S3 API layer: Provides the S3 compatible interface to all other S3 tools and applications.
- Administrative Services: Provides services such as provisioning, user and group management, billing, and usage controls.
- Administrative API: Provides interface to third party administrative functions such as billing and monitoring.
- User Interface (Cloudian Management Console): Complete web-based user interface for users and administrators.

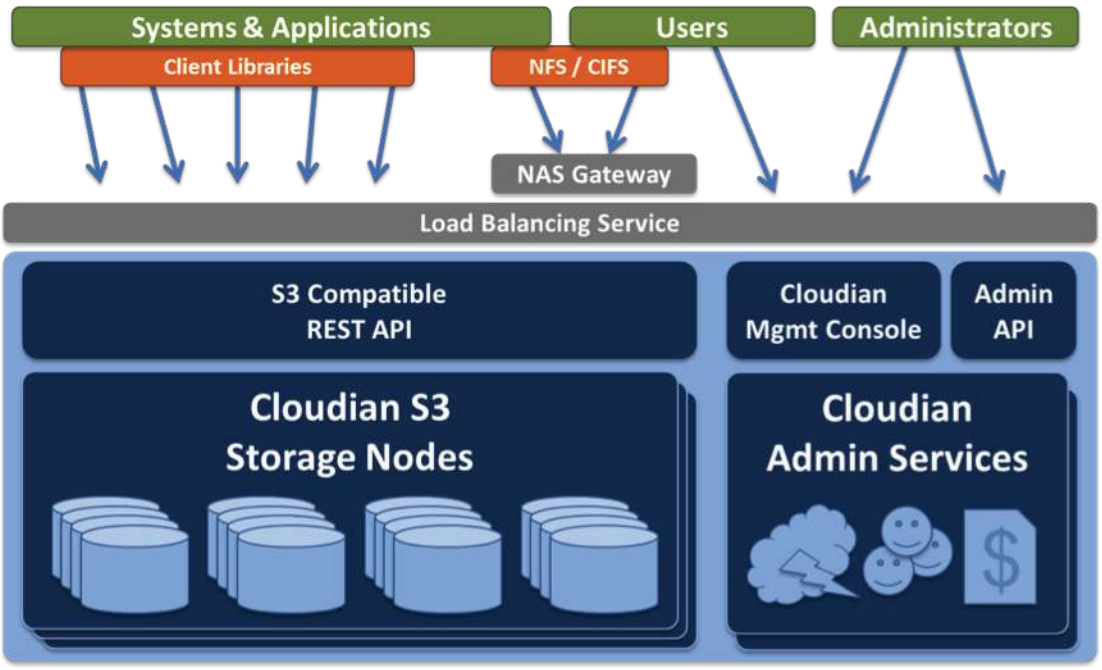


FIGURE 6:
The Cloudian Functional Components

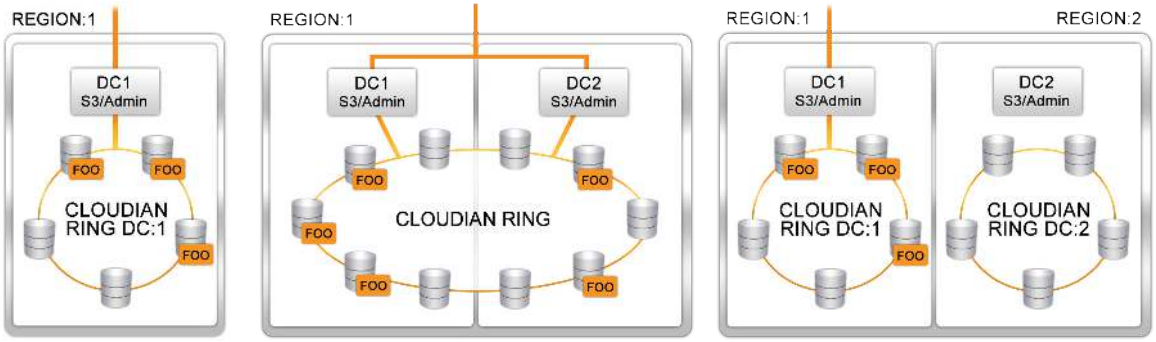
When you are building your system there are many configuration and architecture choices that are available (refer to figures 7, 8, 9). These include:

- Number of nodes
- Storage per node
- Number of datacenters
- Number of regions (rings)
- Number of replica

FIGURE 7:
Single DC, 3 replica

FIGURE 8:
Single Region, 2 DC, 5 replica

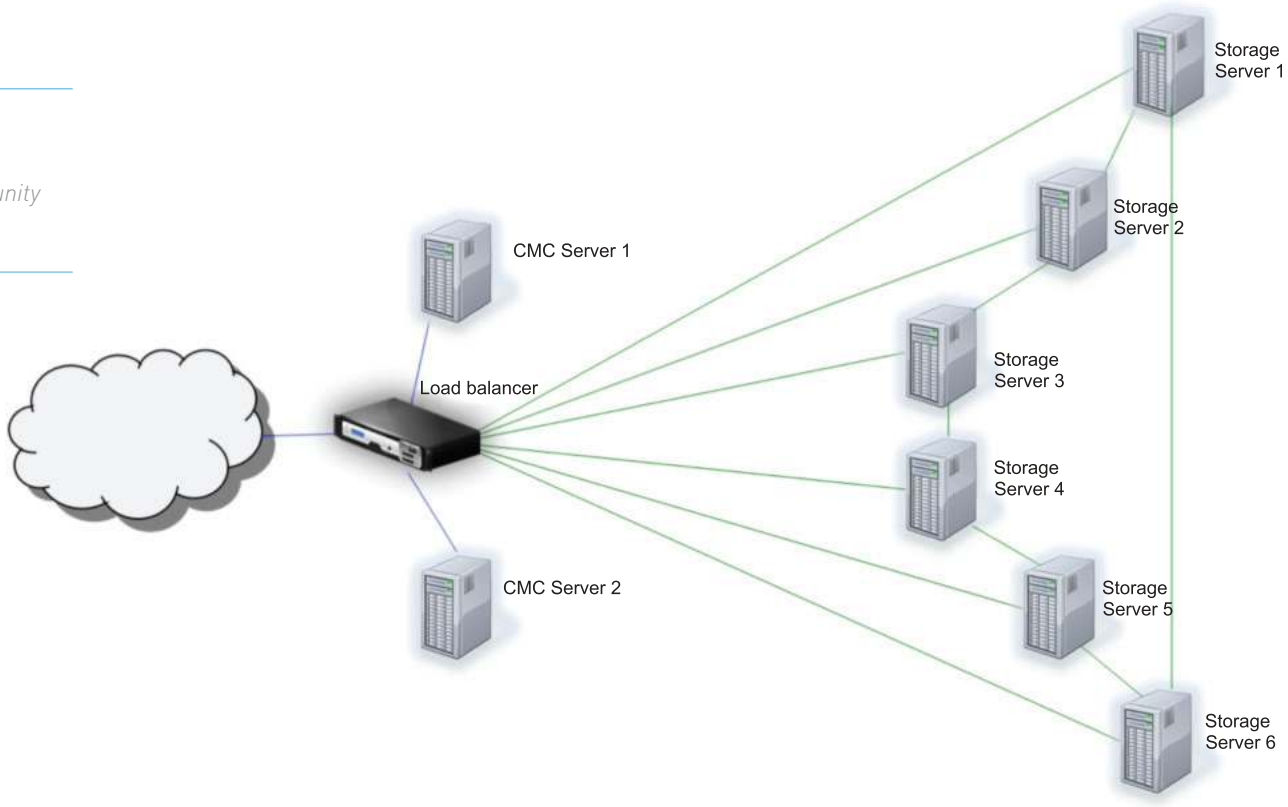
FIGURE 9:
2 Regions, 2 DC, 3 replica



System Design

The Cloudian system is very simple from an architectural point of view. Figure 10 shows a typical 6 node design where each node can provide from 2TB to 135TB of storage. Since the Cloudian Community Edition supports 100TB with no license fees, each node would support 16TB. Other than the commodity servers to act as storage nodes, you only need a load balancer to distribute the 2 main types of HTPP traffic: the S3 API traffic and the User Interface (Cloudian Management Console, CMC) traffic. Cloudian also recommends 2 network segments in a production environment: one for internal inter-node traffic, and the other for actual service traffic.

FIGURE 10:
A Typical 10TB Cloudian Community Edition System



The User Interface traffic is the same as any regular webserver traffic. A load balancer that supports sessions is highly recommended so that 2 or more CMC nodes can be used to provide redundancy (see figure 10).

The S3 API traffic is what provides direct access to the user's data, and it is distributed to all the storage nodes in the system. Cloudian's peer-to-peer architecture allows for any node to handle the client's request, providing built-in redundancy.

As an example, figure 10 above illustrates how a 6-node Cloudian system can be setup. The web-based User Interface is setup on 2 separate nodes for clarity's sake. The User Interface can also be combined with all the storage nodes to further reduce costs. The load balancer only needs to distribute the User Interface traffic to the CMC nodes and the S3 traffic to any storage node. Scaling out the system then only involves adding new storage nodes (commodity servers) and adding it to the load balancer rotation.

Conclusion: STaaS Hold Promise of ROI, Improved Services

As this paper demonstrates, STaaS holds great promise for both new revenue streams for Service Providers and better storage solutions for their Enterprise Customers. Given the shocking fact that cost of storing 100 TB of data internally is close to $1 million per year, the marketplace stands poised for service providers to offer cost-effective alternations via powerful but easy to deploy object storage solutions such as the Cloudian Storage Platform. Service Providers can get started today by deploying the Cloudian Community Edition, which allows a 10TB system to be deployed with no software license fees. To get started, visit http://www.cloudian.com/get-started.html.

Works Cited

[1] Cloud Computing, Sector Focus: Storage-as-a-Service, 451 Research Market Monitor, March 2012.
[2] Object-based storage promises scalability and flexibility for static, unstructured data needs. Martin Taylor. April 2012.
[3] Lydia Leong, Research VP in the Technology and Service Provider group at Gartner
[4] Wikipedia
[5] Andrew Reichman, File Storage Costs Less in The Cloud Than In-House, Forrester Research, August 25, 2011

About Cloudian

Cloudian, Inc. is a Foster City, Calif.-based software company specializing in cloud storage software. The main product is Cloudian®, an Amazon S3-compliant cloud object storage platform that enables service providers and enterprises to build reliable, affordable and scalable cloud storage solutions. Cloudian is actively partnering with the leading cloud computing environments including Citrix Cloud Platform and OpenStack, cloud on-ramp providers, and the vast ecosystem of tools and applications that is afforded through true S3 compatibility. Cloudian's customers include Vodafone, Nextel, NTT, SoftBank, Nifty, and LunaCloud. The company has additional offices in China and Japan. For more information or to try Cloudian today, please visit www.cloudian.com.

CLOUDIAN
Cloudian, Inc. • Metro Center Tower • 950 Tower Lane, Suite 300 • Foster City, CA 94404
Tel: 1.650.227.2380 • Email: info@cloudian.com • **www.cloudian.com**