The Road Into the Cloud

The Need for a New Business Model for Enterprise Datacenter Infrastructure Providers

Insights in Engineering Leadership White Paper

Abstract

Hyper-Scalars currently use cheap commodity hardware and their own homegrown software to provide their services in the Cloud. This paper addresses the issue current enterprise datacenter infrastructure providers face with their existing business model which focuses heavily on providing hardware and software features for data safety and disaster recovery directly to companies. New business models will have to be adopted by these providers in order to continue to flourish in the age of the Cloud and its associated services.

Number: 2014.4.10
Date: April 10, 2014

Authors
Michael Baikie / *Director, SW/FW Test Engineering, SanDisk*
Robin Huber / *Sr. Director of Engineering, Software Development, NetApp*
Kaladhar Voruganti / *Sr. Technical Director, CTO office, NetApp*
Fanny Wong / *Sr. Director of Product Engineering, SDDC Division, VMware*

1. Introduction

The emergence of hyper-scalars like Amazon AWS, Microsoft Azure, and Google App Engine has a potential to be disruptive towards the existing business and delivery models of traditional IT infra-structure companies such as Oracle, SAP, VMWare, NetApp, EMC, IBM, Cisco, HP and Dell. The fundamental reason is that the hyper-scalars develop most of the hardware and software in house and purchase very little hardware and software from the traditional infra-structure companies. Additionally, the hyper-scalars are competing with the traditional infra-structure companies by offering platform as a service (PaaS), infra-structure as a service (IaaS) and software as a service (SaaS) to the existing customers of the infra-structure companies.

In section 2, we first analyze how the current customers of the infra-structure companies are reacting to the emergence of the hyper-scalars. In section 3, we provide an analysis on what potential business models and technology delivery options a traditional infra-structure company can pursue in order to deal with this changing landscape.
This white paper focuses its analysis to only infra-structure companies that are providing traditional **storage area network (SAN) and network attached storage (NAS)** solutions (e.g., EMC, NetApp, Hitachi, IBM and HP).

In this report, by a hyper-scalar we mean service providers who build their infra-structure using commodity hardware and a combination of home-grown and open source software. They hire a lot of PhDs/specialists to home grow their infra-structure. The hyper-scalars operate using a low margin model, and thus, do not like to offer high margins to the infra-structure providers. Furthermore, unlike the traditional network storage companies, the hyper-scalars are not building storage solutions that can satisfy every workload type, but instead they are targeting their solutions for some well-defined solution sweet spots. For example, Amazon AWS offers only 4 types of storage solutions (Dynamo DB, EBS, S3, Glacier) that they hope will satisfy the needs of 80% of the customers (80-20 rule).

2. Changing Landscape

In order to understand the changing landscape for the traditional infra-structure companies, we have classified their current set of customers into three buckets (as shown in Fig. A-1). The three buckets are: 1) Small and Medium Businesses (SMBs) that have less than 100 employees 2) Enterprise Customers that have 100+ employees and 3) Service Providers who are providing hosting services to other enterprises. We now take a single customer use case from each of these three buckets and then present an analysis on both how and why they are reacting to the emergence of hyper-scalars.

2.1 SMBs and Hyper-Scalars (Yelp Case Study)

Yelp provides a service where customers write reviews about their experiences about a particular business, and these reviews are subsequently publically available to others. In order to provide this service, Yelp needed to have an infra-structure to both host the review comments, and also for running analytics software that analyzes and generates trends based on the reviews. By creating their IT infra-structure on Amazon AWS, Yelp was able to save 55K up front on capital infra-structure costs (1). Furthermore, as their data analytics jobs required more computing power, they were able to dynamically obtain more Amazon EC2 resources. Yelp was able to get their infra-structure up and running in a matter of few days instead of months. Finally, Yelp developers were able to focus on building their applications rather than on integrating various pieces of software.

Small and medium businesses are using public clouds (specifically hyper-scalars) because this allows them to not make large capital investments upfront in hardware and software. Furthermore, the public cloud model allows them to *elastically* change resource consumption as their user demand fluctuates. The hyper-scalars are also providing an integrated set of application development APIs that provide messaging, database, storage, content distribution, security etc services that allows developers to quickly develop and deploy their applications. Thus, resource purchasing decisions are being increasingly made in these organizations by developers and not data center professionals. At a recent venture capital funding competition in Cochin India, out of 49 proposals, ninety percent of them were planning to not have dedicated data center professionals but instead were going to just have developers build the required applications using Amazon AWS (2).

2.2 Enterprises and Hyper-Scalars (Netflix Case Study)

Originally, Netflix had their online video service hosted on their internal data centers. However, they quickly realized that they were not able to accurately predict the growth rates in the number of their customers. Thus, they were not able to build data centers fast enough to keep up with their customer growth rates and decided to move their IT infra-structure on to Amazon AWS (3). Depending upon the value of the data, Netflix decides whether to keep it in house or host in in Amazon AWS. For example, Netflix hosts the customer credit card data in house, while it hosts the video data as well as user preference meta-data in Amazon AWS. Netflix had to solve some important technical challenges in order to port their software to make it run on Amazon AWS (4).

Depending upon the value of the data, enterprises selectively decide which data needs to be hosted in house and which data can reside in the public cloud (in hyper-scalars). Enterprises also choose the cloud model, for elasticity reasons, when the demand for a particular service is unpredictable. However, when the demand stabilizes, they prefer to move the service to a private data center for cost reasons. For example, Zynga game provider moved its older game titles to internal clouds due to predictable usage models, whereas, its new games are hosted in the public cloud. UC Berkeley study has shown that beyond a certain size, it is cheaper for enterprises to have their own private clouds rather than use a public cloud because resource costs can be amortized at larger scale (5). For cost reasons, companies

have also begun to use public clouds for 1) storing archival data 2) as a standby disaster recovery site, and 3) also for cloud bursting when their local data centers cannot handle the sudden increased demand for more computing power. In conclusion, different enterprises will be at different points on the spectrum with respect to the percentage of their IT infrastructure that can be moved to hyper-scalars.

2.3 Service Providers and Hyper-Scalars (T-Systems Case Study)

T-Systems is a subsidiary of Deutsche-Telekom. It provides application hosting services for many major enterprises in Europe like Philips, Shell, Volkswagen, Daimler, etc (6). Unlike the Hyper-scalars which for the most part develop their IT infra-structure in house, T-Systems purchases it hardware and software from traditional IT infra-structure vendors. T-Systems is targeting hosting services for enterprises that want five 9s availability and predictable performance. Currently, hyper-scalars are not able to offer five 9s availability to their customers, and the sweet spot for the hyper-scalar design center is web applications that are written by individual developers and small companies. One could argue that service providers like T-Systems could leverage hyper-scalars and provide value-add enterprise level services on top of it. Currently, there is reluctance for service providers like T-Systems to pursue this route because they will get locked in to a particular hyper-scalar's APIs, and they also lose control over pricing decisions.

Increasingly, the traditional service providers like T-Systems are coming under intense pressure to reduce their margins due to competition from hyper-scalars. Thus, most service providers (who are not hyper-scalars) are exploring the use of open source alternatives because they too do not want to pay high margins to the traditional infra-structure vendors. However, unlike in the case of hyper-scalars, enterprises still have to deploy, manage and support the solutions that are built out of open source offerings. Currently, the existing open source offerings are not that easy to deploy and they do not scale, and thus, a lot of effort is required to deploy, manage and support these offerings. Service companies like Cloudera, Hortonworks, Redhat etc are beginning to emerge that are doing for these open source offerings what RedHat and IBM did for Linux. Thus, over time, a credible alternative will emerge for enterprises to manage their private clouds using open source alternatives. This trend could adversely impact both traditional infra-structure vendors as well as the hyper-scalars.

3. What should the Storage Infrastructure Providers do?

It is very evident that the hyper-scalars with their Infrastructure as a Service (IaaS) capabilities pose a serious threat to conventional datacenter infrastructure suppliers (including storage companies). This is further accentuated by the fact that IaaS is one of the fastest growing segments of the cloud services market as shown in Fig A-2.

According to Gigaom Research, the current worldwide cloud market is growing by 126.5 percent year over year, driven by 119 percent growth in SaaS and 122 percent growth in IaaS (7). Furthermore, it has been shown, given today's costs, that one hour of cloud

services (providing a virtual server, storage, and processing) is already less costly than what many datacenters are currently paying to power their internal infrastructure for one hour (8).

The cloud represents the biggest technology and market transition since the transition from mainframe to client-server computing. It fundamentally changes how IT services are being delivered and consumed, as well as how the underlying infrastructure is designed and deployed. Specifically for the incumbent enterprise storage infrastructure vendors, it is a paradigm shift from being a *storage vendor* to a *storage service provider*, for both on-premise and public cloud.

Therefore, a new business model is needed for storage providers in order to continue to succeed in this new and evolving market. While, at first glance, it looks like that Hyper-scalars are going to totally replace the on-premise storage, and therefore, put all the storage infrastructure providers out of business, a second look reveals that this isn't so. Hyper-Scalars cater to a very specific customer, one that can tolerate a longer period of downtime (Hyper-scalars currently only promise 99.9% uptime which translates into about 9 hours of downtime a year) and at least for the time being, a customer that isn't too worried about the security of their data in the public cloud. Traditional storage providers on the other hand have already built up the architecture, hardware, and accompanying software to be able to promise a customer 99.999% of uptime (which translates to only about 5 min per year of downtime) and in a private cloud environment can continue to provide the data security that customers have been accustomed to over the years. Enterprise data (e.g., large banks, the stock market, crucial customer data containing private information or financial transactions, etc.) continues to demand high security, the minimum amount of downtime possible, and the assurance that it will be available almost instantaneously. Therefore, with some adjustments, the infrastructure storage providers actually still have great opportunities ahead of them and there appear to be four main business models for storage providers to succeed in this new world of Cloud IaaS:

1. *Become a Cloud Software Provider*: Become a pure software provider and sell the RAID stack along with storage specific value added features for block and file directly to the cloud providers which will require heavy integration with either their proprietary software (e.g., AWS) or open source software such as OpenStack
2. *Become a Storage Service Provider for the Private Cloud:* Partner with enterprise companies and help them build out their own private cloud with software and hardware that's being provided by the storage companies in a very similar fashion as it is today in the existing business model but with higher software integration and cloud awareness
3. *Become a Storage Service Provider for the Hybrid Cloud:* Again, partner with enterprise companies to build out their private cloud and also provide software to seamlessly move and access data in a virtual private cloud or a Hyper-scalar's public cloud
4. *Become a hyper-scalar:* Become a hyper-scalar and provide the same service as Amazon Web Services (AWS), Rackspace, Microsoft Windows Azure, and Google Compute Engine but with enterprise class storage which is backed with 20+ years of storage experience and code and feature hardening. Partner with service providers and help them build out their own public cloud

We will now look at these four models in more detail and map them against the projected market size.

3.1 Become a Cloud Software Provider

This business model, which would be to only provide storage software to the hyper-scalars, virtual private clouds, and companies with private clouds, is certainly sustainable in the future. The projected market for this opportunity is difficult to assess but it could easily be around $3B, or roughly 15% of the current SAN and NAS market of $21B (Fig. A-3).

This would be a big shift for companies like EMC and NetApp who tend to lead their sales pitch with hardware before getting into software features. The sales force of these types of companies would have to re-learn their sales habits.

A software only model provides great margin and this would provide these companies with a business model that many of their shareholders will appreciate from a profit margin perspective. However, the overall revenue stream for these companies would dramatically decline from today's values. The assumption with this model is that hyper-scalars will continue to build out their hardware with commodity off-the-shelf products and that the companies building out their own private cloud would do something very similar to keep cost down and benefit from the economies of scale on the hardware. This then means that the storage providers provide value with their block and file storage software in terms of:

1) overall features
2) integration into the most prevalent architectures such as AWS and OpenStack
3) seamless disaster recovery
4) software reliability/quality that has been built up over more than two decades

3.2 Become a Storage Service Provider for the Private Cloud

This business model is very much in line with what storage providers do today except that it would now call for further cloud awareness in the software, and higher levels of software integration at the solution level. Specifically, the storage companies would provide a service of setting up a private cloud for the enterprise on premise whereby they can continue to sell their own storage products as they do today, along with also bringing in expertise and consulting on the rest of the solution (compute, etc).

This will require the storage providers to fully integrate their existing software with the entire solution stack such that it gives a seamless experience for users of the internal cloud. So, in this business model they would be able to easily sustain their current revenues while adding growth in the areas of providing the cloud build-out service and other value add software that streamlines the customer's cloud experience within their own company. The market opportunity in this model currently is $21B (Fig. A-3) and is actually projected to grow to $40B based on forecasted world-wide storage needs (9).

For enterprise customers the private cloud is their most likely approach in the next 10+ years given their desire to not have to share physical resources with other companies (they need guaranteed service levels), their absolute need for high reliability and availability (require 99.999%+ of uptime) and their fear of data security in the public cloud. It also gives them the flexibility to expand (burst) into a public cloud on demand with non-critical data, and thus,

keeps total cost of ownership (TCO) under control. Given current pricing from the public cloud vendors it is also evident that any enterprise with large amounts of data (which is almost certain at the enterprise level) isn't saving nearly as much money as it appears on the surface – see table A-1 for the cost structure of 40TB of data stored in the public cloud. At higher data levels there is an absolute cutover where it is cheaper to have the storage on-premise in a private cloud than residing on hyper-scalar storage.

3.3 Become a Storage Service Provider for the Hybrid Cloud

This business model is an extension of the private cloud model. Enterprise companies will want to and/or have to move some data into a more publically available cloud but will continue to have fears of data security and lack of high data availability which will prevent them from moving their data into a hyper-scalar. Therefore, virtual private clouds will become very important to the enterprise. A virtual private cloud is basically a public cloud that is run by other enterprise companies that are providing cloud services. However, they have built this cloud model on enterprise class hardware and software, and therefore, do not have some of the reliability limitations as the hyper-scalars. Storage software integration with OpenStack and Cloud Operating Systems will be very important to succeed in this space.

Given this approach, the storage infrastructure providers can adapt their business model to further enhance their software and provide seamless software extensions/solutions for movement of data between the private cloud and the virtual private cloud and even the public cloud of hyper-scalars (for data that the enterprise companies are willing to store there).

The expansion into a more public cloud opens the market by another $5B above the private cloud, and therefore, the total addressable market for storage infrastructure providers is projected to be around $45B (9).

3.4 Become a Hyper-Scalar

This business model has several downsides for the large storage providers that can actually afford to do so (e.g., NetApp, EMC, HP), in a sense that they would have to take on some of the now more established cloud providers like Amazon who already run on very thin margins and have built the entire infrastructure to support IaaS (which spans way beyond just storage). Amazon's AWS is reported to be adding 1,000 servers a day, seven days a week, in order to keep up with compute demand and provide the promised level of service to its customers. This is an astounding number of resources which any new entrant into this space would have to build up to, which clearly is a very large expense. Thus, one of the major challenges faced in building a hyper-scalar like site is the high cost of build-out. This causes this solution to have a high bar of entry, and is a major deterrent to many of the current infrastructure companies (see Fig. A-4).

An additional challenge in this space would be the need to partner with other infrastructure providers that provide CPU/Server or network capabilities and would therefore cause cost sharing challenges. Furthermore, the "white box" Open Compute Platforms and Open

Storage groups are deriving common off the shelf standards for adding infrastructure to the data centers. Thus, the margins on infra-structure hardware will be extremely low.

Based on these large hurdles it is not recommended that a storage infrastructure provider become a standard hyper-scalar unless that storage vendor has really deep pockets. However, there is a variant of this business model that can be more easily adopted. That model establishes a partnership with service providers such as Verizon, Comcast, etc. in the build-out of their own public cloud. In that sense they would become mini hyper-scalars with a very specific customer base, namely their existing customers. In this environment the storage infrastructure providers can sell their hardware and software (service providers are not inclined to gain expertise in an area that has nothing to do with their core business) which has been adapted to be cloud aware and also provide consulting services on the overall cloud solution based on their years of experience with data center build-outs. The business opportunity for the storage vendors in this space is incremental to that of the hybrid-cloud and would allow the company that partners first and partners well with service providers to grab a bigger share of the projected $45B overall storage market.

4. Summary

For storage infrastructure companies to prevail in the coming decade it will be necessary for them to develop their data management solutions to be “cloud aware”, operate with Cloud Operating Systems, account for multi-tenancy and implement cloud protocols such as the RESTful interface.

If storage infrastructure providers adopt a combination of the business models outlined in this paper, and thus, continue to add their value in delivery of data security and reliability, they will prevail in the private, hybrid and the service providers’ public cloud infrastructure. The business models required to be adopted in order to serve the various market spaces/customer categories are listed in Fig. A-5.

Positioning in cloud environments will be achieved through strong partnerships in order to solve the customer’s challenges in maintaining service level objectives (SLOs) that they are currently used to (99.999% uptime) and providing assurances on their data security concerns.

In order to maintain dominance in the storage space, the large incumbents (EMC, NetApp, HP, Hitachi) will have to look for companies on the rise in areas of high density storage HW solutions to SW startups that are adding unique value to the OpenStack platforms that are in development today. Acquisitions based on some of these startups should be closely considered by the storage infrastructure providers in order to enhance their storage SW with a more integrated and feature rich cloud experience. While smaller storage players will continue to occupy various niche markets we do not believe that they will be able to provide the storage reliability and data security that is required by the customers now and in the future.

In summary, the key success factors for making the technology and business model transition from storage vendor to storage cloud service provider are:

1. **Transition from hardware centric to software driven**. The key differentiators for delivering storage services are agility, efficiency, and cost. Storage hardware needs to be abstracted, pooled, managed and provisioned through automated software based policies for all types of applications. For many storage vendors, this represents new development models, processes, and skill sets to build software with a service delivery mind set
.
2. **Have a hybrid cloud plan** that leverages the vendor's existing strong enterprise data center storage and data management footprint. The stickiness of data and the underlying data services and data protection technology, as well as compatibility with the install base are cornerstones to develop a successful bridge from on-premise to the cloud. Customers that have on-premise data centers or private clouds will want their public cloud design to be as similar as possible leveraging what they already use and trust. High availability/reliability and data security will remain a key driver for storage build-outs in the hybrid-cloud.

3. **Design for the cloud and new classes of applications.** Openness, extensibility, and scale are key design points. Have open APIs that are industry standard and provide support that bridges multiple storage vendors and storage types (block, file, and object). A flexible design for both scale up and scale out for web-scale applications will also be critical.

4. **Partnerships**. Storage vendors need to partner with other vendors, partners, systems integrators, service providers, hyper-scalars, and customers in order to be able to build out the private cloud, the virtual private cloud and the mini-hyper-scalar environments of service providers while providing feature rich cloud-aware storage software that easily and seamlessly allows the data to move between the various clouds.

The road into the cloud requires changes on the part of the storage infrastructure providers, especially on the software front, but it is a prosperous road ahead and has an expanding market space and great and strong revenue growth potential for the companies that adapt and start implementing a combination of the business models described in this paper.

Appendix

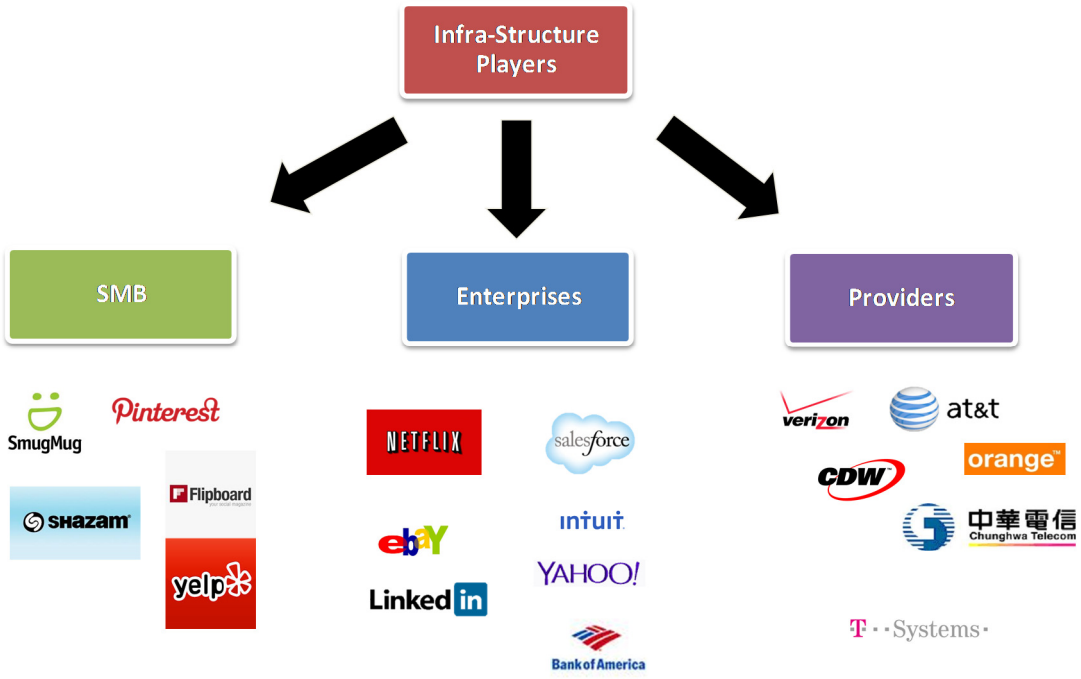

Figure A-1: Example of customer landscape for infrastructure providers

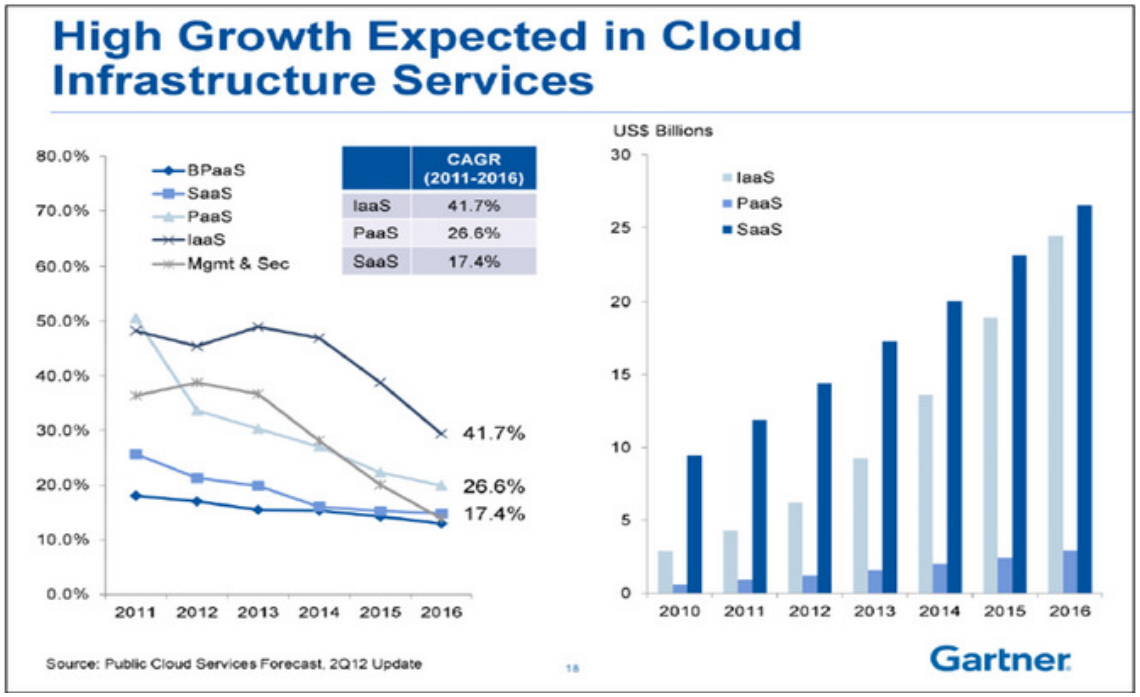

Figure A-2: Infrastructure as a Service growth potential

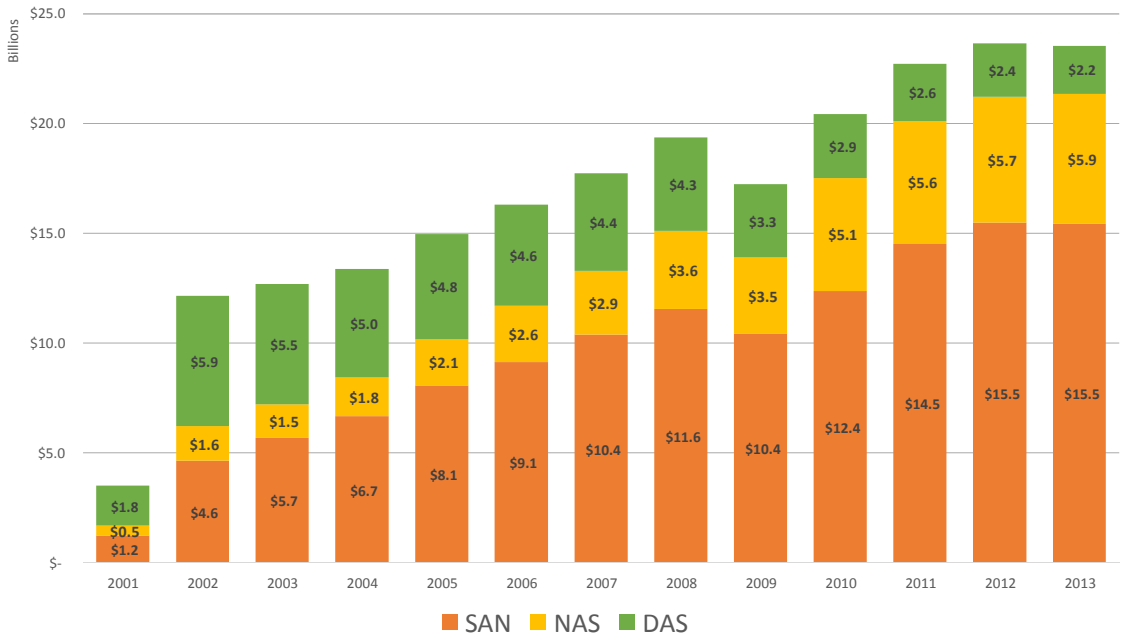

Figure A-3: Storage Market from 2001-2013 (Source: IDC Tracker, idctracker.com)

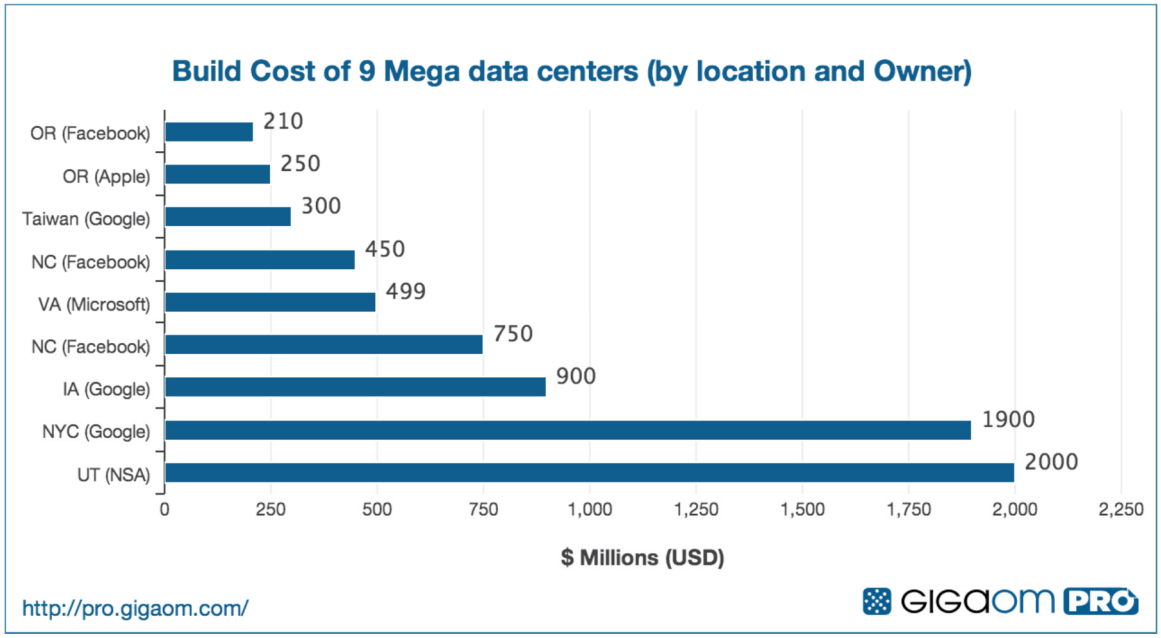

Figure A-4: Build cost of mega data centers (Hyper-Scalar like)

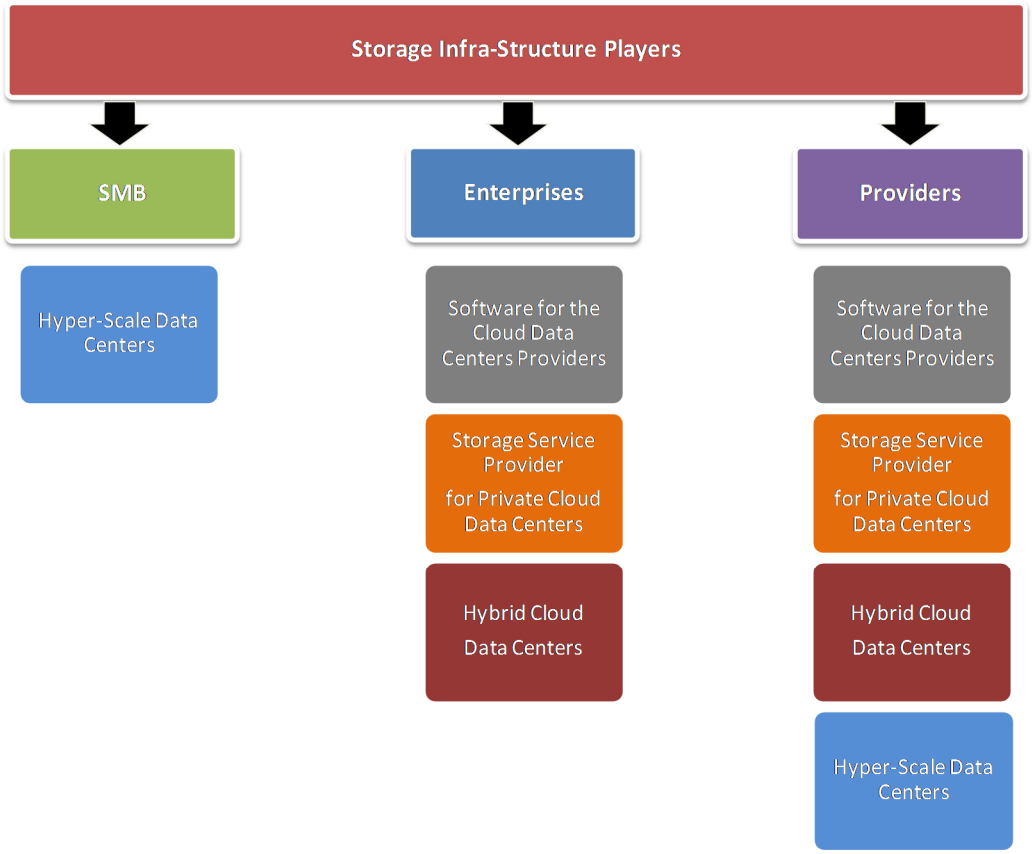

Figure A-5: Business models to be applied against the infrastructure customer/market space

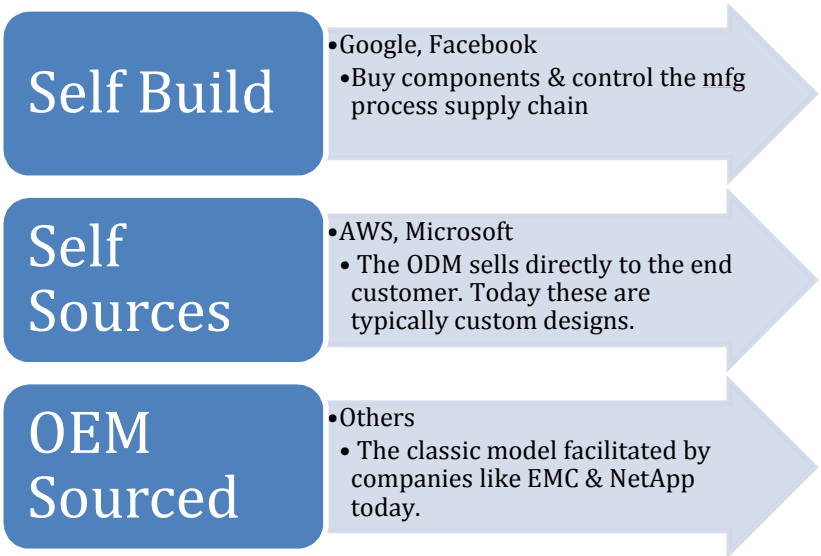

Figure A-6: Supply Chain Models

Figure A-7: Possible business models for storage infrastructure providers

Table A-1: Cost to store and retrieve 40TB of data from public cloud vendors

	Monthly Backup Cost			Cost per 40TB dataset restore		
	AWS (1)	Azure (2)	Google (3)	AWS (1)	Azure (2)	Google (3)
Storage at Rest	$3,302.00	$2,340.00	$2,845.70	N/A	N/A	N/A
Transfer bandwidth	$0.00	$0.00	$0.00	$3,900.10	$3,900.00	$3,510.09
Protocol	$2.00	$0.01	$0.40	$24.14	$0.01	$20.00
Total Monthly Storage Cost	$3,304.00	$2,340.01	$2,845.74	$3,924.24	$3,900.01	$3,530.09
40TB dataset with 3.3% change delta. Read workload 40% 4KB GETs (for metadata) and 60% 1MB GETs (for data). Write workload of 100% 1MB PUTs. Pricing effective as of 03/13/14						
Sources:						
1 - http://aws.amazon.com/s3/pricing/ 2 - http://www.windowsazure.com/en-us/pricing/details/storage/pricing-changes/ 3 - https://developers.google.com/storage/pricing						

Bibliography

1. **Amazon.** AWS Case Study: Yelp. *Amazon.* [Online] http://aws.amazon.com/solutions/case-studies/yelp/.
2. Amrita Cloud Security Alliance Conference. [Online] https://cloudsecurityalliance.org/media/news/csa-startup-pitchfest-at-secureit.
3. **Amazon.** AWS Case Study: Netflix. *Amazon.* [Online] http://aws.amazon.com/solutions/case-studies/netflix/.
4. **Cockroft, Adrian.** Adrian Cockroft AWS Architect. [Online] http://www.slideshare.net/adrianco.
5. **Armbrust, Michael, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia.** UC Berkeley Above the Clouds Report. [Online] http://radlab.cs.berkeley.edu/publication/285.
6. **T-Systems.** T-Systems. [Online] http://www.t-systems.com/.
7. **Lawrence, Andy and Ascierto, Rhonda.** The datacenter of the future, Part1: innovation accelerates. *451 Research.* [Online] March 14, 2014. [Cited: March 14, 2014.] https://451research.com/report-short?entityId=80473.
8. **MSV, Janakiram.** *A checklist for stacking up IaaS providers.* Gigaom Research. San Francisco : Giga Omni Media, 2014.
9. **Kidd, Jay.** *Chief Technology Officer.* Sunnyvale, April 1, 2014.

National rankings consistently place UC Berkeley's undergraduate and graduate programs among the worl's best. Berkeley is home to top scholars in every discipline, accomplished writers and musicians, star athletes, and stellar scientists—all drawn to this public university by its rich opportunities for groundbreaking research, innovative thinking and creativity, and service to society.