This paper focuses on cloud storage industry analysis. Some products of public cloud storage providers are main objects: Google Drive, Dropbox, Amazon Cloud Drive and etc. Based on system dynamics, the study represents a whole map of cloud storage stakeholders. In addition, price and storage space as two main factors are researched to illustrate current situation. Using Vensim simulates future number of users about Google Drive. The result shows that Google Drive has a flexible pricing strategy and Dropbox is a security system. However, most factors are hard to digitize. In the future, user satisfaction test will help system simulation better.
COMPARISON OF CLOUD STORAGE

by
Xin Zhang

A Master’s paper submitted to the faculty of the School of Information and Library Science of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in Information Science.

Chapel Hill, North Carolina
April 2013

Approved by

_______________________________________
Bradley M. Hemminger
# TABLE OF CONTENT

1. Introduction .............................................................................................................. 2  
   1.1 Background ........................................................................................................... 2  
   1.2 Purpose of the Study ............................................................................................ 2  
   1.3 Problem Statement ............................................................................................. 3  

2. Review of Related Literature ...................................................................................... 4  
   2.1 Cloud Computing ................................................................................................. 4  
   2.2 Cloud Service Providers ...................................................................................... 6  
      2.2.1 Google Drive .................................................................................................. 7  
      2.2.2 Dropbox ......................................................................................................... 8  
      2.2.3 Amazon Cloud Drive ...................................................................................... 10  
   2.3 Cloud Storage Prices ............................................................................................ 12  
   2.4 Cloud Storage Space ............................................................................................ 13  
   2.5 System Dynamics ................................................................................................. 13  

3. Methodology & Analytic Techniques ........................................................................ 16  
   3.1 Industry Research .................................................................................................. 16  
   3.2 Build Model .......................................................................................................... 17  
   3.3 Data Collection & Analysis .................................................................................. 18  
   3.4 System Dynamics ................................................................................................ 19  
   3.5 Related Factors Analysis ..................................................................................... 19  

4. Results and Discussion .............................................................................................. 21  
   4.1 Cloud storage industry chain analysis ................................................................. 21  
   4.2 Price and Storage Space ....................................................................................... 24  
   4.3 System Dynamics Simulation .............................................................................. 29  
   4.4 Database and Security ........................................................................................ 30  
   4.5 Limitation and Future Work ............................................................................... 32  

5. Conclusion .................................................................................................................. 34  
   Bibliography .............................................................................................................. 35
1. Introduction

1.1 Background

For private cloud or public cloud, cloud storage is an important part of cloud computing infrastructure. Great cloud storage solution architecture would help users to achieve highly efficient, rational, and automated storage resources. That’s also a measure of the cloud computing infrastructure success mark. However, in today's information era of expansion, the capacity and complexity of the data is larger than at any time in the past. So, faced with increasingly complex data environment to build a high efficiency, scalability and flexible cloud storage solutions users of IT infrastructure must plan accordingly.

Besides the traditional database structured data, currently there is an explosive growth of unstructured data, such as e-mail, documents, images, audio, and video, in addition to mixed data backup and archiving of unstructured and structured data. The reason for this situation is increasingly rich and complex user applications, especially in the popular cloud computing trend, and relying on traditional storage architecture has not proven to be capable of handling the data integration and complexity.

1.2 Purpose of the Study

The purpose of this study is to research Cloud Storage Systems, study Industry offerings, and evaluate how different attributes (price, storage space, database, and etc.) affect customer selections of IT technology options.
1.3 Problem Statement

The problem addressed in the study concerns whether there is a distribution of customer selectivity of Cloud Storage, especially for Public Cloud Storage: such as Google Drive and Dropbox. Their features are about price, storage space, or database type. Research hypothesis is that all the users don’t have preferences based on brands cloud storage options, but instead choose systems primarily based on price, storage space and database internal structure, and they do not consider UI experience, or other dimensions.
2. Review of Related Literature

There is a large body of research on cloud computing and cloud storage. For this paper a brief synopsis of each major topic area will be discussed below. Cloud Computing is the origin of cloud storage so that we need to research for it. Cloud service providers can be divided into two parts: public cloud and private cloud. This study will focus on public cloud storage service. In addition, price and storage space are two main factors to contrast different services. And system dynamic would illustrate the dynamic relationship.

2.1 Cloud Computing

Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network (typically the Internet). The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with a user's data, software and computation.

Cloud computing depends on the sharing of resources and network utility (Peter and Timothy, 2011). The basic concepts of cloud computing can be traced back to the 1950s, when mainframe, in academia and business, through a thin client terminal / terminal computers, often referred to as a "dumb terminal", because they are used for communication, but without internal computing power. In order to use expensive main frame computational power more efficiently, the evolution of the practice that allows multiple users to share physical access to the mainframe from multiple terminals and
share CPU time. This eliminated periods of inactivity on the mainframe and allowed for a greater return on the investment. The practice of sharing CPU time on a mainframe became known in the industry as time-sharing (Christopher, 1959).

In the 1990s, the telecommunications company, formerly focused to provide point-to-point data circuits, began offering Virtual Private Network (VPN) service is comparable to the quality of service, but at a lower cost. Through the exchange of traffic, because they think the right balance between the using of the server, they can make more effective use of the overall bandwidth of the network. They began to use symbols to representing the cloud of services the vendor is responsible for, separate from what the user is responsible, thus defining a cut-off point between the two. Cloud computing extends this boundary to include server and network infrastructure. Fig 1 is the basic workflow about cloud computing:

Cloud computing shares characteristics with:
- Autonomic computing — Computer systems can make self-management.
- Client–server model — Client–server computing refers broadly to any distributed application that distinguishes between service providers (servers) and service requesters (clients).
- Grid computing — "A form of distributed and parallel computing, whereby a 'super and virtual computer' is composed of a cluster of networked, loosely coupled computers acting in concert to perform very large tasks."
- Mainframe computer — Powerful computers used mainly by large organizations for critical applications, typically bulk data processing such as census, industry and consumer statistics, police and secret intelligence services, enterprise resource planning, and financial transaction processing.
- Utility computing — The "packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility, such as electricity."
- Peer-to-peer means distributed architecture without the need for central coordination. Participants are both suppliers and consumers of resources (in contrast to the traditional client–server model).
- Cloud gaming—also known as on-demand gaming—is a way of delivering games to computers. Gaming data is stored in the provider's server, so that gaming is independent of client computers used to play the game.

2.2 Cloud Service Providers

Nowadays, more and more IT Internet service providers (ISP) provide cloud services. Cloud storage has a huge market prospects. For public cloud storage, there are some main providers: Google, Dropbox, Amazon, Apple, and etc.
2.2.1 Google Drive

Google drive, by Google, was released on April 24, 2012, so that users could have shared access to file storage and synchronization services through cloud base storage and file-sharing. A set of productivity applications, collaborative editing documents, spreadsheets, presentations, Google Docs, are now referred to as “Google Drive”. Work began on this much earlier, with rumors first circulating as early as March 2006.

Google Docs is Google's "software as a service" office suite. Create documents, spreadsheets, presentations, and share through the Web interface, import/export functions or by email. Files can be saved to the user's local computer in a variety of formats (ODF, PDF, HTML, RTF, text, Microsoft Office). Files will be automatically saved to Google's servers, in order to prevent the loss of data, and revision history is automatically saved, so the last edit can be viewed (although this is basic implementation, is adjacent to the amendments, there is currently no way to discover and isolate changes in long documents.) You can tag files and archive for the purpose of organization. The latest versions of Firefox, Internet browser, Safari browser and Chrome browser running on Microsoft Windows, Apple OS X, Linux operating system, Chrome operating system all officially support this service.

The System Health database is built upon Bigtable (Chang et al, 2006), a distributed data repository widely used within Google, which itself is built upon the Google File System (GFS) (Ghemawat et al., 2003). Bigtable takes care of all the data layout, compression, and access chores associated with a large data store. It presents the abstraction of a 2-dimensional table of data cells, with different versions over time making up a third dimension. It is a natural fit for keeping track of the values of different variables (columns) for different machines (rows) over time.
The System Health database thus retains a complete time-ordered history of the environment, utilization, error, configuration, and repair events in each machine’s life. Analysis programs run on top of the System Health database, looking at information from individual machines, or mining the data across thousands of machines. Large-scale analysis programs are typically built upon Google’s Mapreduce (Dean and Ghemawat, 2004) framework. Mapreduce automates the mechanisms of large-scale distributed computation (such as work distribution, load balancing, tolerance of failures), allowing the user to focus on specific attributes.

2.2.2 Dropbox

Dropbox cloud storage can do file synchronization, and the client software is a file hosting service. Dropbox users can create a special file local file folder named “dropbox” on their own computer which can then be synchronized across multiple devices, so it seems to be in the same folder (with the same content), regardless of which computer is used to access the content. The files can also be placed in this folder are available through...
a website and mobile application. Dropbox uses a premium business model, and provides users with a free service that is limited in storage space, with fee based accounts providing larger storage capacity. In addition to sharing, DropBox is also positioned as a backup service, so that users can conveniently save a backup copy of their data on the cloud. OPSWAT reported in their December 2011 market share report that Dropbox held 14.14% of the worldwide backup client market, based on number of installations.

Both the Dropbox server and desktop client software are primarily written in Python (Rian, 2011). The desktop client uses GUI toolkits such as wxWidgets and Cocoa. Other notable Python libraries include Twisted, ctypes, and pywin32. Dropbox ships and depends on the librsync binary-delta library (which is written in C). Dropbox declares that all files stored online by Dropbox are encrypted and kept securely on Amazon's Simple Storage Service (S3) in multiple data centers located across the United States. Fig 3 is the client interface of Dropbox:

![Fig 3 Interface of Dropbox](image)
Once you install the free Dropbox app on your Macs and PCs, you place the files you want access to into your Dropbox folder (Snell, 2010). Dropbox has introduced a new service that allows workgroups to share up to 1TB of online capacity, while also allowing IT shops to add or remove users. Dropbox for Teams is priced at $795 annually for five users, with additional seats available for $125 each. The base plan includes 1TB of storage, and each additional seat comes with 200GB (Lucas, 2012). Dropbox was one of the earliest mainstream players in the file-synchronization services area, but it has faced tough competition in recent months from a few other companies, such as SugarSync and CX, that have typically touted their services as being more user-friendly.

2.2.3 Amazon Cloud Drive

Amazon Cloud Drive is a web storage application from Amazon. The user can choose to store their purchased music in the Amazon Cloud Drive website or application via the Amazon MP3 Android (version 2.0 or later) through the Amazon MP3 store purchase. Amazon Cloud Drive account provides users 5 GBs of free storage space, however, the music is not included in the purchase through the Amazon MP3 store storage limits. The music stored in Amazon Cloud Drive, users can choose to download a recognized by Android devices using the Amazon MP3 application, or download it to a recognized computer using the Amazon MP3 download. Fig 4 is the interface of Amazon Cloud Drive.

After the dot-com bubble, Amazon cloud computing development played a key role through modern data centers, the favorite computer network, with as little as 10% of their capacity at any time, just leave the room occasional spikes. Result in significant internal efficiency, so that a small, fast-moving "two-pizza team" (team small enough to feed two pizzas) can be faster and easier to add new features, the new architecture of the
cloud, Amazon launched a new product development efforts to the external customers to provide cloud computing, based on utility computing, Amazon Web services (AWS) was launched in 2006 (CarlBrooks, 2010).

In fact, Amazon dominates the industry because it has a best product. Nasuni, which sells data protection services that work across any type of cloud storage, says it has been testing the 16 largest cloud storage providers (CSPs) since April 2009 to determine the best services for its customers. They published test results in 2011 represented that Amazon S3 had the fewest outages and best uptime, and was the only CSP to post a 0.0 percent error rate in both writing and reading objects during scalability testing.

Last but not least, Amazon is looking to mitigate the pressure on its cloud computing margins with custom servers. It has tied up with Asian manufacturers like Quanta and Foxconn, which became famous for manufacturing Apple’s iPhone and iPad, to manufacture its own servers. It sources its chips directly from Intel and in doing so has followed in the footsteps of Facebook, Google, Microsoft and others going around middle

![Fig 4 interface of Amazon Cloud Drive](image)

In fact, Amazon dominates the industry because it has a best product. Nasuni, which sells data protection services that work across any type of cloud storage, says it has been testing the 16 largest cloud storage providers (CSPs) since April 2009 to determine the best services for its customers. They published test results in 2011 represented that Amazon S3 had the fewest outages and best uptime, and was the only CSP to post a 0.0 percent error rate in both writing and reading objects during scalability testing.

Last but not least, Amazon is looking to mitigate the pressure on its cloud computing margins with custom servers. It has tied up with Asian manufacturers like Quanta and Foxconn, which became famous for manufacturing Apple’s iPhone and iPad, to manufacture its own servers. It sources its chips directly from Intel and in doing so has followed in the footsteps of Facebook, Google, Microsoft and others going around middle
men such as HP and Dell. The move away from standard machines towards cheaper gear will free up cash and should help the company expand its operations more quickly.

2.3 Cloud Storage Prices

Cloud users can eliminate their own storage infrastructure, relying on the cloud only. The migration from an owned infrastructure to a leased one has the immediate benefit of avoiding capital investments in favor of a more flexible expense management based on operational expenses only (Mastroeni and Naldi, 2011). The decision to migrate should consider the risks associated with both alternatives.

Capital investments are one-off in nature but they may lead to savings in the long run (Mastroeni and Naldi, 2011). In addition, switching to the cloud may expose the cloud user to the lock-in phenomenon and price rises (though the decision to continue buying disks is exposed to disk price rises as well) (Mastroeni and Naldi, 2012). Another option is a hybrid one, where the cloud user can also keep its own infrastructure and use the cloud for backup purposes only.

In either case, the cloud storage market is forecast to expand, and the need to evaluate the different offers grows as well. Though a service can be thoroughly described just by the complete set of its features, the most relevant point of differentiation seems to be the price. The risk of commoditization of cloud computing (of which cloud storage may be an ancillary function) has been pointed out by Durkee (Durkee, 2010), as has been the case with the web hosting industry. In fact, at present cloud providers are pushing price as the most attractive leverage to get users choose their platform.

Though we realize that the complexity of a service proposition cannot be captured by price only, in this paper we focus on price. A comparison of the technical merits of commercial cloud platforms has been reported in (Hu et al., 2010), but no economic
analysis has been accomplished so far. However, a comparison of pricing plans for cloud storage is needed to decide whether to migrate or not. In every analysis conducted so far about the opportunity to migrate, one or more pricing plans have been adopted to draw conclusions about that opportunity: in (Walker et al., 2010) and (Mastroeni and Naldi, 2011), just Amazon's prices have been employed, while in those prices has been used along with four other providers (Rackspace, GoGrid, Nirvanix, and EMC Atmos). In addition, analysing the pricing plans proposed by cloud providers helps understand the price structure of the industry and its economies of scale.

2.4 Cloud Storage Space

The service models are divided in Cloud Software as a Service (SaaS), Cloud Platform as a Service (PaaS) and Cloud Infrastructure as a Service (IaaS) (Vaquero et al., 2009). Computing resources like servers and network can be replaced, but the core of most of the organizations is the information, usually stored in data centers. For this reason security and availability are the first issues when companies are deciding to migrate part of their data to the cloud, generally by the internet (Wu et al., 2009).

2.5 System Dynamics

System dynamics is an approach to understanding the behaviour of complex systems over time. It deals with internal feedback loops and time delays that affect the behaviour of the entire system (MIT System Dynamics in Education Project (SDEP)).

Chris mentioned that system dynamics is an efficient approach to analysis Cloud computing (Chris and et al., 2012). This paper investigates this energy-saving claim by reviewing trends in and between four domains of this problem and identifying their key drivers in a business setting. Moreover, Fig 5 represents how three recent reports consider energy consumption and a column for this paper is included in this table as a comparison.
MIT system design and management center illustrates that how a large
corporation successfully builds a new business by a new technology. They said that
creating a private cloud is about $500M for an enterprise, because large corporations
consider hybrid solutions for non-critical application. Fig 6 is the selection view of cloud
computing adoption model by system dynamics:

**Fig 6 Cloud computing adoption model**
H. Wang et al. shows preliminary studies on Amazon EC2 cloud service and on a local cloud computing, which have revealed an interesting interplay between distributed systems and economics related to pricing. They believe that this new angle of looking at distributed systems potentially fosters new insights into cloud computing (H. Wang et al. 2010).

On the paper System Dynamics Modeling for the future IT Development: Applied to education, Health Care, and Smart Work System in Korea, Lee mentioned system dynamics model to simulate the future trends in the field of education, health care, and smart work system. The simulations focus on every successful policy strategy fields. In addition, this study system dynamics modeling provide a unique method. New attempts to build a prototype model applicable to different policy areas been adopted for the new trial. The most critical variables which can be applied to different policy fields are cloud computing, smart devices, and the Internet of Things mechanism; we call these variables C∙S∙I frame. The following Figure 7 shows the frame:

![Fig 7 Three Critical Variables in the Archetype](image)
3. Methodology & Analytic Techniques

3.1 Industry Research

Industry research is the cornerstone to carry out all the consulting business, with long-term follow-up monitoring of specific industries. It could analyze industry demand, supply, operating characteristics, capacity with a wide range of content, industry chain and value chain (McDonald, 2007). Industry research is the analysis of the overall situation and development trend of an industry, including the evolution of life cycle of the industry, the market capacity of the industry, industry of room for growth and profit margins, industry trends, industry, critical success factors, barriers to entry and exit.

In fact, there are four key contents about industry research:

![Fig 8 industry research](image_url)
- Environmental analysis: it works for the most direct impact on enterprises, the role of the external environment.

- Structural Analysis: the structural analysis of the industry is mainly related to the industry's capital structure, market structure, and other content. In general, the main industry barriers to entry and analysis of the level of competition in the industry.

- Market analysis: main content relates to the nature of the industry, market demand, and the requirements of its development and changes, the market capacity of the industry, the industry's distribution channel model, and sales methods.

- Organizational analysis: requirements and reality of the living conditions of the research industry reflect the main content: relevance within the enterprise, industry specialization, level of integration, the level of economies of scale, organizational change status.

3.2 Build Model

Through the industry analysis, a preliminary understanding of the industry has been gotten. The next step is build model for stakeholders and industrial chain. For the cloud storage, stakeholders could be divided into internal and external parts. Even though in most situations, cloud storage providers would play more than one role.

**Stakeholders:**

- Cloud platform service provider
- Cloud resources service provider
- Cloud application service provider
- Cloud platform developer
3.3 **Data Collection & Analysis**

Data collection is any process of preparing and collecting data, for example, as part of a process improvement or similar project. The purpose of data collection is to obtain information to keep on record, to make decisions about important issues, or to pass information on to others. Data are primarily collected to provide information regarding a specific topic (Weller, 1988).

The Basic process of Data Collection is:

1. Pre collection activity — agree on goals, target data, definitions, methods
2. Collection — data collection
3. Present Findings — usually involves some form of sorting analysis and/or presentation.

For cloud storage providers, the qualitative analysis and the quantitative analysis are both necessary. Through system analysis we would get some dimensions data in this environment. In fact, the data need to be quantified includes free space, price of payment options with storage space, maximum single file upload size, desktop and mobile app, and etc.

First of all, list a comprehensive listing containing all of the above dimensions about Google Drive, Microsoft SkyDrive, Dropbox, Apple iCloud, and ASUS WebStorage is
generated for these mainstream public cloud storage providers. Secondly, according to its main points of difference, the study will analysis them step by step. At last, the number of users is also key factor to measure providers and cloud storage system.

3.4 System Dynamics

System Dynamics is an analysis of information feedback system of discipline, but also an understanding of system issues and resolve system problems cross integrated subject. For a system methodology, the system dynamics is the unification of the structure, the function of the method and historical method. It is based on the system theory, absorption of cybernetics, and information theory (Forrester, 1956).

After some basic research about cloud storage industry, a diagram about cloud storage of industrial chain should be established. The related factors should be divided into two pieces: internal and external. The total system should include profit, revenue, user requirement, cloud technology, cloud infrastructure, cloud security capabilities, industrial scale, cost, etc.

Vensim is a software which made by the U.S. Ventana Systems, Inc. It could develop the concept of documentation, simulation, analysis, graphical interface software and optimize dynamic system models. Vensim provides an easy and flexible way to establish a causal cycle (casual loop), inventory (stock) with a flow chart model. I choose it to establish cloud storage model, import related factors, analyze their relationship (+ or -), input data, and do some basic simulation.

3.5 Related Factors Analysis

Factor analysis is a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors.
Exploratory factor analysis (EFA) is used to uncover the underlying structure of a relatively large set of variables. The researcher's a priori assumption is that any indicator may be associated with any factor. This is the most common form of factor analysis. There is no prior theory and one uses factor loadings to intuit the factor structure of the data (Russell, 2002).

Confirmatory factor analysis (CFA) seeks to determine if the number of factors and the loadings of measured (indicator) variables on them conform to what is expected on the basis of pre-established theory. Indicator variables are selected on the basis of prior theory and factor analysis is used to see if they load as predicted on the expected number of factors. The researcher's a priori assumption is that each factor (the number and labels of which may be specified a priori) is associated with a specified subset of indicator variables. A minimum requirement of confirmatory factor analysis is that one hypothesizes beforehand the number of factors in the model, but usually also the researcher will posit expectations about which variables will load on which factors. The researcher seeks to determine, for instance, if measures created to represent a latent variables really belong together (Richard, 2004).

Cloud Storage industry chain would represent related factors. In addition, system dynamics model with some quantitative analysis will find some trend of price, storage space, and users. However, these are not the only reasons cloud service users considered. Security is also another key factor. Nowadays, cloud service is used for governments, military, corporations, financial institutions, hospitals, and private businesses which amass a great deal of confidential information about their employees, customers, products, research and financial status. Most of this information is now collected, processed and stored on electronic computers and transmitted across networks to other computers.
4. Results and Discussion

4.1 Cloud storage industry chain analysis

Industry chain is the concept of Industrial Economics, based on certain technical and economic association with different departments. Based on specific logic chain relations and the temporal and spatial layout relations objective relationship shaped state industrial chain in order to create maximum value of the estate, its essence is reflect the "1 + 1 > 2" value-added effect of this value. This kind of “added” often from the multiplier effect of the industrial chain. It means that the industrial chain effective change will lead to other related industries corresponding occurrence times multiplier effects.

Fig 9 Cloud Storage Industry Structure Diagram
This diagram represents 10 key points of cloud storage industry structure diagram. The core part is cloud storage service provider which is a cloud provider to provide some of the components of cloud computing---Infrastructure as a Service (IaaS), Software as a Service (SaaS), platform as a service (PaaS) ---other enterprises or individuals. Cloud storage providers are sometimes referred to as cloud service providers or CSPs. Platform, resources, and application as three main cloud service providers occupy the business. In other words, a mature cloud storage firm could provide a full suite of integrated services. Stability of the platform, diversification of applications and richness of resource would allow cloud storage providers capture the market opportunities.

In addition, external conditions providers are also necessary for cloud storage service. Platform developers and system integrators have interaction with internal system. Hardware providers and cloud application developers make system emerge. Then the network operates as a bridge communicating with users. Of course, terminal supply should give users the medium.

There are a lot of things to consider when you assess cloud service providers. The cost is usually based in each time you use the utility model, but there are also some changes to consider. The physical location of the server, and may be a factor of the sensitive data.

System dynamics analysis to identify the impact of system changes the key factor. Industry chain is still developing. Stakeholders need to focus on how to enhance the comprehensive capabilities of the cloud storage so that it can grow up as soon as possible. Qualitative analysis is the key factors about the impact cloud computing industry chain ability by the causal diagram. From a macro perspective, a good environment for the development of the industry chain is the key to affect the development of the cloud
storage industry chain. From industry chain, the size of the industry chain, cloud storage investment scale and the popularity of cloud storage can reduce costs through the economies of scale effective. Same with other industrial chain is good economic benefit is the key driving force to promote the development of cloud storage. Moreover, when more and more people want to get value-added in cloud storage, how to create the efficient business model to improve the ability of cloud storage service would be a huge discussion. User-based service and business model would be the focus of cloud storage. Obviously, with improvement of cloud security capabilities, users’ requirement of cloud services is increasing.

Fig 10 Cloud storage causality diagram of industrial chain

Investment scale and popularity to industrial scale are also the positively related impacts. In fact, three factors which are user requirement, comprehensive capacity, and
the cost of the industrial chain have more influence factors than the others. In the other words, how to improve cloud storage product, how to enhance customer satisfaction, and how to increase profits are three problems which cloud storage industry need to solve.

4.2 Price and Storage Space

Price and Storage Space are two most intuitive factors of Cloud Storage, because there is no relationship between network transmission speed and cloud storage providers. Dropbox as the industry leader provides multi-prices. Fig 11 shows the price and storage space from Dropbox 2011.

However, with Google Drive and other cloud service providers improving, Dropbox modified their business charging strategy. Even though it’s also the ruler of the market in this area, based on the board users, Microsoft and Google have also strongly settled the market.

<table>
<thead>
<tr>
<th>Storage</th>
<th>Price</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>50G</td>
<td>$9.99/month</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$99.00/year</td>
<td></td>
</tr>
<tr>
<td>100G</td>
<td>$19.99/month</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$199.00/year</td>
<td></td>
</tr>
<tr>
<td>350G</td>
<td>$785.00/year</td>
<td>team collaboration space; include 5 license and 350G space, if team want to add another license need $125/year; in addition, if team decide to add 100G space, then need pay more $200/year</td>
</tr>
</tbody>
</table>

Fig 11 Dropbox 2011

Fig 12 is the comprehensive chart about 5 main cloud storage providers. Free Space is the first thing that needs to be considered when a new user comes in. Fig 13 is the contrast of Free Space, and it shows that Dropbox provides the minimum free space...
for users, which is only 2 GB. Maximum single file upload size supports users a wider transmission needs. Google Drive can support 10GB single file upload, but Dropbox can only make 300MB. For the Desktop App, Dropbox and ASUS WebStorage get some points here, besides Windows and Mac system, they can represent on Linux system that facilitates a lot of linux users. For the Mobile App, Microsoft SkyDrive and ASUS WebStorage could be installed on Windows Phone. Moreover, Dropbox can support BlackBerry phone system. MS office Support is another factor. Obviously, Microsoft SkyDrive could support all functions. And Google Drive can only support read. Unfortunately, others do not have this function but you download and edit on your own devices. Collaborative work is very important for Google Drive. Based on traditional Google Documents, users can edit documents in different devices at the same time.

<table>
<thead>
<tr>
<th>Cloud Storage</th>
<th>Google Drive</th>
<th>Microsoft SkyDrive</th>
<th>Dropbox</th>
<th>Apple iCloud</th>
<th>ASUS WebStorage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Free Space(GB)</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>25GB</td>
<td>$2.49/month</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>20GB</td>
<td>$29.88/year</td>
<td>$10/year</td>
<td>N/A</td>
<td>$40/year</td>
<td>N/A</td>
</tr>
<tr>
<td>50GB</td>
<td>N/A</td>
<td>$25/year</td>
<td>$9.99/month $99/year</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>100GB</td>
<td>$4.99/month $59.88/year</td>
<td>$50/year</td>
<td>$9.99/month $99/year</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>200GB</td>
<td>$9.99/month</td>
<td>N/A</td>
<td>$19.99/month $199/year</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>400GB</td>
<td>$19.99/month</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>500GB</td>
<td>N/A</td>
<td>N/A</td>
<td>$49.99/month $499.99/year</td>
<td>N/A</td>
<td>$99.99/year</td>
</tr>
<tr>
<td>1TB</td>
<td>$49.99/month</td>
<td>N/A</td>
<td>more than 5 users and $795/year</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2TB</td>
<td>$99.99/month</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4TB</td>
<td>$199.99/month</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>8TB</td>
<td>$399.99/month</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>16TB</td>
<td>$799.99/month</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Maximum file upload</td>
<td>10GB</td>
<td>2GB</td>
<td>300MB (Paying)</td>
<td>250MB (Paying)</td>
<td>1GB (Paying)</td>
</tr>
</tbody>
</table>
In order to satisfy users from different groups, Cloud Storage providers customize lots of packages. Fig 12 represents those packages on payment options. In fact, they provide in different spaces so that it not easy to horizontal comparison all of them. I choose 100GB and 200GB to contrast (Fig 14 and Fig 15). Google Drive and Microsoft.
SkyDrive has superiority in price. Google Drive is cheaper than Dropbox $39.12/year which is 65% for 100GB.

![Fig 14 Price for 100GB](image)

And the number is $79.12/year which is also 65% for 200GB.

![Fig 15 Price for 200GB](image)

Fig 16 is the number of user online for Dropbox. It was set up on 2007, and the users increased to 0.1 million when it was on Sep 2008. On Oct 2009, the number of user
turned into 3 million. Additionally, on May 2010, users came to 5 million; and on Apr 2011, users came to 25 million. Currently, Dropbox has more than 100 million users.

For Google Drive, they have about 40 million users when it released one month on July 2012. In fact, based on Gmail, Youtube, Google + and other products, this number would increase as an amazing number (Fig 17).
4.3 System Dynamics Simulation

The total system represents as Fig 10 using Vensim. User requirement, comprehensive capacity, and the cost of the industrial chain are three key points of system. However, it is not easy to quantify all of the factors. For a specific cloud storage provider, Fig 18 describes a brief model of Google Drive Users.

![Fig 18 brief model of Google Drive Users](image)

New Users = Google Drive Users * Increasing Rate.

Inactive Accounts = Google Drive Users * Inactive Rate

Google Drive Users (2012) = 40 million

Increasing Rate = 150% (Based on Dropbox increasing rate from 2011-2012)

Inactive Rate = 3% (Based on IDC research 2012 cloud computing)
Fig 19 Google Drive Users Simulation

Fig 19 is the simulation of Google Drive Users. In fact, simple consideration we could forecast that this year Google Drive Users will reach 100 million at least. The reason is not only this simulation but also from Fig 17- Google have more than 0.4 billion Gmail Users.

4.4 Database and Security

Based on Cloud storage causality diagram of industrial chain (Fig 10), there are some other related factors. Database and Security are two significant ones. As Cloud Storage providers, the selection of Database and Searching language would decide backstage supporter.

For Google Drive, Google Cloud SQL is a MySQL database that lives in Google's cloud. It all the features and functionality of MySQL, the following is a list of some of the extra features and some features not supported. Google Cloud SQL is easy to be used,
do not need to install any software or maintenance, small to medium-sized application ideal choice. Google Cloud SQL is currently available for Google App Engine applications that are written in Java or Python. It can also be accessed from a command-line tool, and other admin and reporting tools are available. In addition, Google Drive SQL can support 16GB RAM and 100GB storage. The interaction between Google APIs Console and Google Drive SQL could allow interactive web access to your database. Synchronous or asynchronous geographic replication and import or export databases using mysqldump.

However, there are also some restrictions: the individual instances size up to 100GB, users could not defined their own functions, MySQL replication is not supported. Moreover, some of MySQL statements and functions could not be used here:

- LOAD DATA INFILE
- SELECT ... INTO OUTFILE/DUMPFILE
- INSTALL/UNINSTALL PLUGIN ...
- CREATE FUNCTION ...
- LOAD_FILE()
- SHA2()

So Google Drive saves documents on their own server. But Dropbox is not. Once a file has been added to your Dropbox, the file is then synchronized to the Dropbox secure online server. Dropbox online storage file encryption and security use the Amazon Simple Storage Service (S3) in multiple data centers located across the United States.

Amazon S3 storage service is an Internet-facing. The service is designed for developers to reduce the size of network calculation difficulty. Amazon S3 provides a
concise Web interface, and through it, users can at any time in any location on the Web access storage and retrieval of data of any size. This service allows all developers can access the same with high scalability, reliability, security and cost-effective infrastructure, Amazon utilizes their own infrastructure to run its global network of websites. The service aims to allow maximization of benefits, and to pass this advantage to developers.

For Dropbox, their synchronization uses SSL to transfer data stored and encryption by AES-256. Dropbox can use the HTTPS protocol through SSL encryption terminal node data safely upload/download to Amazon S3. Amazon S3 also provides a variety of options for static data encryption. If Dropbox want to manage their own users’ encryption keys, Dropbox can use Amazon S3 Encryption to encrypt data, and then upload it to Amazon S3. Or, if Dropbox want to make the Amazon S3 for you to manage encryption keys, you can use Amazon S3 server-side encryption (SSE). With Amazon S3 SSE, Dropbox need to write data element, the addition of a request header can upload data encryption. It can retrieve data automatically when it decrypts.

In the other hand, when Google Drive had a network failure in 2013/3/18, some users not being able to access their service. The affected users use Twitter to express their dissatisfaction. Unfortunately, security is the Google Drive currently lacking. The biggest problem is that the certification review. To access Google Drive one need only be logged-in to the user's Google account. Therefore, if you have Gmail or Google services open in a web-browser window, your files are accessible to anyone having this access.

4.5 Limitation and Future Work

In fact, Fig 10 is the total system map with system dynamics. However, most of the factors are hard to quantify. Furthermore, as a new system model, cloud storage providers
do not have much more relative information. Google Drive just released last year so that it is not easy to measure its future value. Obviously, the development of the Internet is rapid. Maybe some other products would replace cloud storage.

For the future work, this study will more focus on user satisfaction measurement. I make questionnaire to survey user selecting of price, storage space, and user’s selectivity. The price and storage space have 10 pieces. In addition, the user’s selectivity has 5 pieces: Strongly Agree, Agree, Neither Agree Not Disagree, Disagree, and Strongly Disagree.

Some other questions are also involved in my survey:

1. Do you think Price is the most important thing when using cloud storage?
2. Do you think Storage space is the most important thing when using cloud storage?
3. Will you upgrade your Google Drive account more than 5GB? (would be charge)
4. Will you upgrade your Dropbox account more than 2GB? (would be charge)

With those satisfaction measurements, more specific quantitative data could be taken. System Dynamics model would be simulation with more data. And that is an efficient approach to analyze this area.
5. Conclusion

This study sought to investigate different cloud storage providers. Reviews of Cloud Computing made a lot of background about this area. Basic information about Dropbox, Google Drive, Amazon Cloud and etc. were investigated. With some research into the IT industry, system dynamics can be an efficient approach to analyze cloud storage industrial chain. Moreover, stakeholders map could represent all participants. Data Collection is about price, storage, free space, max single file upload size, Desktop App, Mobile App, and MS office support.

In addition, a brief model of Google Drive Users is calculated forecasting their users in 2013. The result of prediction is more than 0.1 billion users. Security and Database Language are also two key points in system model. This study mentions that Google Drive have advantage in price, storage space, and collaboration of documents. However, security problems are currently a drawback with Google Drive. Dropbox would be a best one in this area.

For the limitations of this study, most of the factors in system model are hard to quantify. Moreover, as a new system model, cloud storage providers do not have much more relative information. This study would bring some user tests for some relative questions using questionnaire. Different dimensions from 1 to 5 could be import into system model. Additionally, that will bring a more actual result.
Bibliography


WATERS, JOHN K. (2011, April) MOVE IT OR LOSE IT Cloud-Based Data Storage. Education Digest, Vol. 76 Issue 8, p28-34, 7p.


Dropbox Overhauls User Interface. PC Magazine Online


Meeting report from the IP over ATM working group of the IETF. *CH: Switch. Retrieved August 22, 2010.*


