1-1-2013

A high performance peer to cloud and peer model augmented with hierarchical secure communications

Zhe Sun
University of Wollongong, zs789@uow.edu.au

Jun Shen
University of Wollongong, jshen@uow.edu.au

A high performance peer to cloud and peer model augmented with hierarchical secure communications

Abstract

This paper presents a secure storage model named Peer to Cloud and Peer (P2CP). P2CP uses the cloud storage system as a backbone storage system. However, when data transmission occurs, the data nodes, cloud user, and the non-cloud user are involved to complete the transaction all together. The users, typically the client peers, can communicate with each other directly, thus bypassing servers on the cloud. Similarly, cloud servers can communicate with each other in a P2P mode. We also introduce a "hierarchy security" method to guarantee the data security in the proposed P2CP storage model. A key feature of our P2CP is that it has three data transmission tunnels: the cloud-user data transmission tunnel, the clients' data transmission tunnel, and the common data transmission tunnel. Assuming that the P2CP model follows the Poisson process or Little's law, we not only mathematically prove that the speed of P2CP is generally better than that of the pure peer-to-peer (P2P) model, the peer to server and peer (P2SP) model or the pure cloud model, but also testify the results through simulations. Beyond security, we also investigate the performance of another characteristic of usability of the data storage, namely availability, where P2CP is more robust to the failures of peers or servers in the cloud environment.

Keywords

secure, peer, performance, cloud, model, high, hierarchical, augmented, communications

Disciplines

Engineering | Science and Technology Studies

Publication Details

Sun, Z. & Shen, J. (2013). A high performance peer to cloud and peer model augmented with hierarchical secure communications. Journal of Systems and Software, 86 (7), 1790-1796.

A ~~Secure~~ High Performance Peer to Cloud and Peer Model Augmented with Hierarchical Secure Communications

Zhe SUN [a, b, *], *Jun SHEN*[a],
[a] School of Information Systems and Technology, University of Wollongong, Wollongong, NSW, Australia
[b] Information Management Center, Huaneng Shandong Shidao Bay Nuclear Power Co., Ltd, Shandong, China
zs789 @ uowmail.edu.au, jshen @ uow.edu.au
[*] *Corresponding author: Tel: +86 631 7357661, Fax: +86 631 7357666*

Abstract:
This paper presents a secure storage model named Peer to Cloud and Peer (P2CP). P2CP uses the cloud storage system as a backbone storage system. However, when data transmission occurs, the data nodes, cloud user, and the non-cloud user are involved to complete the transaction all together. The users, typically the client peers, can communicate with each other directly, thus bypassing servers on the cloud. Similarly, cloud servers can communicate with each other in a P2P mode. We also ~~set up~~introduce a "hierarchy security" method to guarantee the data security in the proposed P2CP storage model. A key feature of our P2CP is that it has three data transmission tunnels: the cloud-user data transmission tunnel, the clients' data transmission tunnel, and the common data transmission tunnel. ~~P2CP uses the cloud storage system as a common storage system. When data transmission occurs, the data nodes, cloud user, and the non-cloud user are all together involved to complete the transaction.~~ ~~A~~Assuming that the P2CP model follows the Poisson process or Little's law, we not only mathematically prove that the speed ~~-~~of P2CP is generally better than that of the pure Peer to Peer (P2P) model, the Peer to Server, Peer (P2SP) model or the pure cloud model~~-~~, but also testify the results through simulations. Beyond security, we also investigate the performance of another characteristic of usability of the data storage, namely availability, where P2CP is more robust to the failures of peers or servers in the cloud environment.

Keywords
Cloud, Performance, Security, P2P.

1. Introduction

Cloud computing is an area of information systems that is undergoing rapid development. Many large corporations, ~~e.g.like~~such as Google, Amazon and Microsoft, have recently been focusing on developing and releasing a number of related storage products such as Google file system (GFS) (Sanjay et al. 2003), Amazon elastic compute cloud (EC2), Azure, etc. All of these use cloud distributed storage models based. These models are appealing as they can lead to a significant decrease in the utilization rate of bandwidth. During download session, the current alternative file sharing models based on Peer to Peer (P2P) communication all suffer high utilization rates of bandwidth and lower availability.

In this paper, with regard to original contributions, we have analysed several existing distribution storage models and designed a hybrid model, P2CP, which exploits the P2P protocol to enhance the data transmission performance and at the same time it uses a cloud storage system to provide continuous availability. It can be considered as a P2P storage system ~~combined~~integrated with a cloud storage system. ~~For our purpose, w~~We assume that the P2CP model follows the Poisson process or Little's law and mathematically prove that the speed and availability of P2CP is generally superior to that of the pure P2P model, ~~-~~the Peer to Server and Peer (P2SP) model ~~(Merkur, 2002, Sun 2009)~~ (Merkur, 2002, Sun, 2009) or the pure Cloud model. Furthermore, we propose "hierarchy security" to guarantee the security of P2CP storage model which include three security layers that are internal security layer, contracted security layer and functional security layer.

The paper is organized as follows: Section 2 introduces related works and background on distributed storage models;~~,~~ moreover, it ~~; and~~ introduces ~~a study of~~ the existing distribution storage models.~~;~~ Section 3 proposes and introduces the security hierarchy in the new P2CP storage model; Section 4 details performance evaluation of the P2CP model by mathematical modelling; Section 5 discusses data ~~availability~~ usability and its implications on security of P2CP.~~;~~ Section 6 is the conclusion.

2. Related Work

This section overviews the three main models for distributed storage, P2P, P2SP and cloud storage models. It highlights key applications, strengths and weaknesses of each of the three models.

In a pure P2P storage model, each peer is equal. Peers act as both clients and servers. In the P2P storage model, there is no master server to manage the network, metadata, and data. Especially in a commercial machine environment, each peer is mutable, which makes the whole network ~~unstable~~volatile. In a P2P storage model, users get data from each other.~~,~~ ~~w~~When a user joins to the network, ~~and then;~~ they become either ~~to~~ a server or a seed. The advantage of the P2P storage model is that it efficiently exploits the network bandwidth, but sometimes, the server or seed that contains the particular resource does not exist in the network, so the file sharing process has to stop. ~~So, t~~The disadvantage of the P2P storage model is that it is hard to offer persistent data availability. Typical P2P applications are Gnutella ~~before~~ prior to version 0.4 (Kirk, 2003), Freenet (Clarke, 2000), Sorrento (Hong et al. 2004), etc. The architecture of the pure P2P storage model is shown in Figure 1.

The Eliot file system (Stein et al. 2002) is a reliable mutable file system based on P2P block storage. The system exploits a metadata service in an auxiliary replicated database separated and generalized to isolate all mutation and client states. It consists of the following four components: an un-trusted, immutable and~~,~~ reliable P2P block storage substrate known as the Charles block service; a trusted and~~,~~ replicated database, known as the metadata service (MS), ~~storing~~ which stores mutable nodes, directories, symlinks, and superblocks; a set of file system clients; and zero, one~~,~~ or more cache servers intended to improve performance, but ~~are~~ not necessary for correctness. FS2You (Sun et al. 2009) is another large-scale online storage system. It also has

four main components: directory server, tracking server, replication servers and peers. With the peers' assistance, it makes semi-persistent files available and reduces the server bandwidth cost.

Serverless Network Storage (SNS) is a persistent P2P network storage application. It has four layers, ~~which are~~namely: operation logic; a file information protocol (FIP) that exploits XML-formatted messages to maintain files and disk information; a proposed security layer; and a serverless layer, which is responsible for routine network state information (Ye et al. 2003).

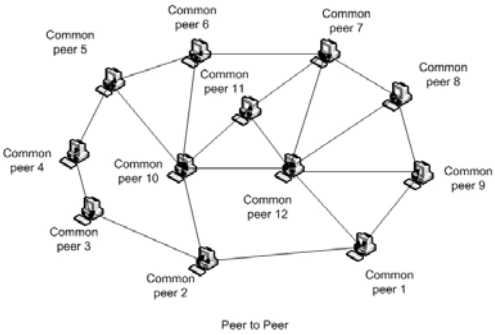

Figure 1: P2P Network model.

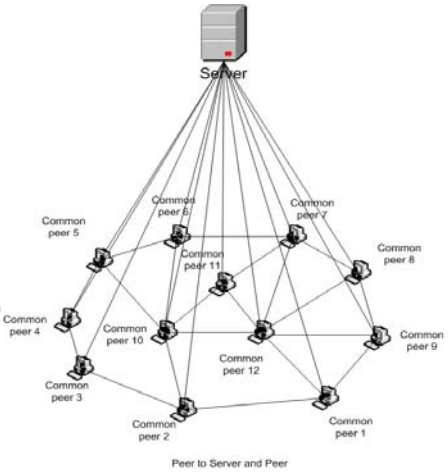

Figure 2: P2SP Network model.

To solve the problem of non-persistent availability in pure P2P storage model (Figure 2), a hybrid P2P model emerged that is Peer to Server and Peer (P2SP). In this storage model, peers are distributed into the client group or the server group. The client group responses handle the data transmission, and the server group acts as a master server to coordinate the P2P structure. However, the workload of the master servers is very heavy,~~ and~~ furthermore, without the server group, the P2P network does not work. ~~As~~Some ~~W~~well-known P2SP applications are eMule (Merkur, 2002), BitTorrent (Cohen, 2001) and FS2You (Sun , 2009). ~~-~~For example, FS2You is a large-scale online storage system. With the peers' assistance, it makes semi-persistent files available and reduces the server bandwidth cost. When the clients are going to download data, ~~firstly, ~~they first download data from the server, ~~and then,~~then they exchange data with ~~each~~ other clients. If the other peers are not available, the client will download all the data from the server. No matter the server is a cluster or a distributed server system, the client will connect to one single physical ~~one single~~ server which is in the cluster or distributed server system.

In (Fang et al. 2009), differences between the pure P2P network and the P2SP network are analysed. The work assumes ~~is~~ that the peer arrival rate and departure rates follow the Poisson process or Little's law,~~.~~ which will be explained and used ~~T~~in ~~hese two assumptions will be explained in~~ the section 4 of this paper as well. ~~Finally, t~~They proved that P2SP has higher performance than P2P based on two assumptions.

Cloud computing consists of both applications and hardware delivered to users as services via the Internet. With the rapid development of cloud computing, more and more cloud services have emerged, such as SaaS (software as a service), PaaS (platform as a service) and IaaS (infrastructure as a service). The concept of cloud storage is derived from cloud computing. It refers to a storage device accessed over the Internet via Web service application program interfaces (API). A traditional storage cloud system is a high performance, scalable, reliable, and available file distribution system (a typical architecture is shown in Figure 3). There are many existing cloud storage systems, for example, Amazon S3 (Amazon, 2006), the Google file system (Sanjay et al. 2003), HDFS (Borthakur, 2007), etc. These systems consist of master nodes and multiple chunk servers. Data is accessed by multiple clients and all files in the system are divided into fixed–size chunks. The master node maintains all file system metadata. At start-up or whenever a chunk server joins the cluster, the master node registers each chunk server with their chunks of information. Clients never read and write file-data through the master, but request from the master ~~a chunk server~~ to contact a chunk server.

The work in (Feng et al. 2010) analyses several existing cloud storage platforms such as Simple Storage Service, Secure Data Connector, and Azure Storage Service, with their focus on the problem of security. The work identifies the problem of repudiation and proposes a non-repudiation protocol suitable for cloud computing environments by using third authorities certified (TAC) and secret key sharing (SKS) techniques. The key problem is that clients get data from the individual data nodes, but the clients do ~~renovate~~ not have any communication among themselves. In this paper, we not only consider renovating storage mechanisms of P2P or P2SP-like systems, we also consider security guarantees in our new P2CP. We will focus on how to establish secure communication tunnels among participants in a hierarchical model. However, we will not discuss the basic security frameworks implementable in the P2CP as we have many mature infrastructures to exploit.

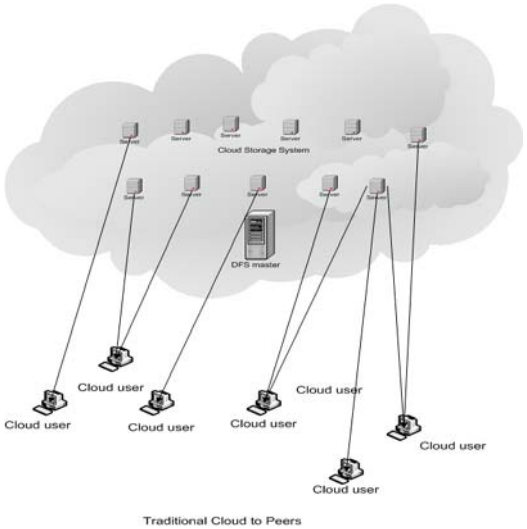

Figure 3: Traditional cloud network model.

3. ~~A Secure Cloud~~P2CP Model

We propose a storage cloud model, which is the peer to cloud and peer (P2CP) model (Sun, 2011). This means that cloud users can download data from the storage cloud and exchange data with other peers at the same time, regardless of whether the other peers are cloud users or not. There are three data transmission tunnels in this storage model. The first is the cloud-user data transmission tunnel. The cloud-user data transmission tunnel is responsible for data transactions ~~between~~ among the cloud storage system and the cloud users. The second is the clients' data transmission tunnel. The clients' data transmission tunnel is responsible for data transactions ~~between~~ among individual cloud users. The third is the common data transmission tunnel. The common data transmission tunnel is responsible for data transactions between cloud users and non-cloud users.

Figure 4 is an example to show how a P2CP cloud model works. In Figure 4, we can see that cloud user 2 is downloading data from data node 1, which is in the cloud, via cloud-user data transmission tunnel~~, and~~. In the meantime ~~at the same time~~, cloud user 2 is exchanging data with cloud user 1, cloud user 3, via clients' data transmission tunnel, ~~and~~ while common peers 2, 5, and 6 via common data transmission tunnel. By exploiting multi data transmission tunnels, cloud users can achieve a high download speed. On the other hand, P2CP model avoids extremely high workload for cloud servers ~~as~~ when the number of cloud users increases. Non-cloud users could not access resources directly in the cloud. When the same resources existed in both ~~of~~ P2P network and~~-~~ the cloud, non-cloud users could exchange the data with cloud users to achieve a high download performance. ~~When the resources are committed to other transmitting activities, non-cloud users may still get access to resources in the cloud which are not in common with the P2P networks.~~

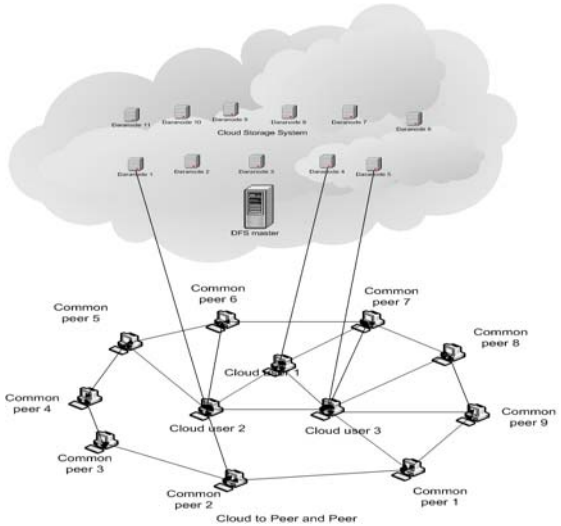

Figure 4: P2CP Network model.

In the pure P2P storage model, peers are divided into seeds, which are denoted by S, and leeches, which are denoted by L. Initially, seeds have the whole file, and leeches do not have any block of the file, but as time passes, leeches obtain blocks and exchange blocks with other peers. When the leeches get the whole file, they may leave the network, or stay in the network as seeds. In the P2SP network storage model, the difference is that it has a server group. Normally, in the Cloud storage model, there are three replicas of the file existing in different data nodes, and each data node keeps different amounts of blocks of the file. In the P2CP storage model, the storage cloud replaces the role of the server in the P2SP model. Compared to work mentioned in Section 2, our model addresses load balance issues via separating peers and cloud servers. Other existing models such as Groove (Ozzie, 2005), as known as comparable to Microsoft SharePoint (Chou, 2006), and Tahoe (O'Hearn, 2008) ~~-~~tended to balance loads between peers and cloud serves in different ways. However, in our P2CP model, peers may communicate directly and flexibly between each other without tight dependence on servers, though some advanced features such as backing up, caching, security and versioning of data may still be elevated or mitigated to servers just because peers' storage and computing capacities are supposedly inferior to those cloud servers.

P2CP storage model contains cloud storage model and P2P storage model, thus, the advantages and disadvantages of P2CP are similar with cloud storage model and P2P storage model. Both of P2P storage model and cloud storage model involve security issues, so does P2CP model. The work of (Diaz-Verdejo 2009) claims that the four main security topics in P2P environment are copyright protection, trustworthiness, privacy and prevention and protection against attacks, and intrusions. Furthermore, the concept of data life cycle has been proposed in the work of (Xiaojun, 2010), ~~it~~ which also mentioned the security requirements of data in cloud. ~~Herein~~ Next we concentrate on the security of P2CP storage model and data security.

To solve the security problems in P2CP storage model, we use "hierarchy security" method ~~which refers~~similar to (Hiroyuki, 2010). In ~~their~~ Hiroyuki's work, they proposed two trust layers, which were internal trust layer and contracted trust layer. In P2CP storage model, we propose three security layers, which are internal security layer, contracted security layer and functional security layer. Internal security layer and contracted security layer are establishing security statements, where the security level will be guaranteed ~~as~~ by statements, or security service agreements. Security statements are offered by cloud services provider which declares the quality of services that providers offer. The functional security layer is a compilation of the augmented security functions. The security functions could be developed by cloud providers ~~but also they could be developed~~ or by third party companies. Internal security is the highest security level, which applies to the cloud and cloud to users. Contracted security is the secondary security level, which applies~~y~~ to cloud-users. Functional security layer applies to certain parts of cloud users and all non-cloud users.

We separate the P2CP model into three components, which are cloud, peers and cloud users. Figure 5 shows the security layers and relationship of data transmission tunnels. The data transmissions between different components, belongs to different security layers. As mentioned previously, three data transmission tunnels existed in the P2CP storage model. The cloud-user data transmission tunnel belongs to internal security layer. This is~~:~~ because in this transmission tunnel, cloud-users only communicate with cloud, the data security is guaranteed by the cloud

services provider. In clients' data transmission tunnel, cloud-users communicate with each other. ~~,~~ T~~t~~he data security is guaranteed by the cloud-users; furthermore, cloud services provider will double check the data after the transmission at the same time. ~~Thus~~In other words, the clients' data transmission tunnel belongs to contracted security layer. In the common data transmission tunnel, cloud-users communicate with non-cloud-users, the data security is guaranteed by these users themselves.~~, and~~The cloud services provider will not check the data after the transmission,~~;~~ thus ~~it need a~~ extra security functions are required to guarantee the data security. ~~So,~~Consequently, common data transmission tunnel belongs to functional security layer.

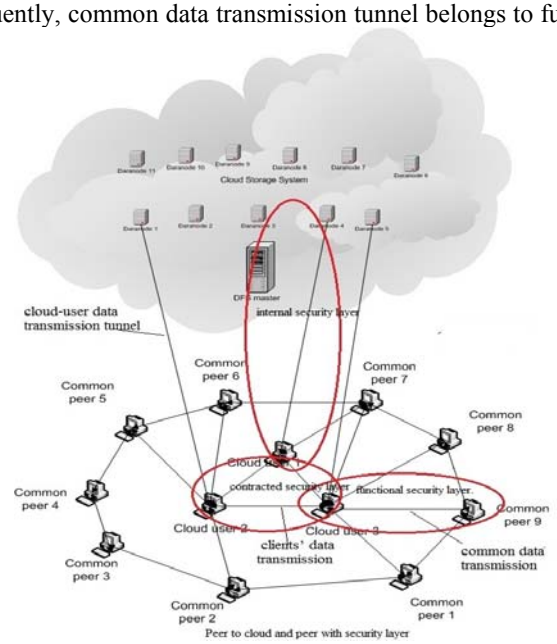

Figure 5: Security layers of P2CP Network model.

4. Performance ~~Comparisons and~~ Evaluation

In this section, we evaluate our P2CP storage model against the three storage models described in Section 2: the pure P2P model, the cloud model and the P2SP model. For the network storage models, the two most important parameters for the performance are average downloading time and usability (which we will discuss in section 5). In this section~~part~~, we will compare average downloading times of the above models by a mathematical model. We assume the following parameters:

Seed: each seed upload bandwidth is U_s; the number of seeds is N_s.

Peer: each seed upload bandwidth is U_p; the number of peers is N_p.

Server: for each server, average upload bandwidth is U_{se}; the number of servers is N_{se}.

The average number of peers and seeds is N.

Cloud: for each data node upload, bandwidth is U_c; the number of data nodes is N_c.

F is the size of the file.

T is the average downloading time. T_c is ~~C~~cloud downloading time; T_{p2p} is P2P downloading time; T_{p2sp} is P2SP downloading time; T_{p2cp} is P2CP downloading time.

t is the current time of data transmission happening.

O is usability.

U is the average upload bandwidth of peers and seeds.

λ: arrival rate of peers arrive at the network.

μ: departure rate of peers leave the network.

λ: must be greater than μ, otherwise, P2P network will not exist.

Formatted: Justified

4.1 Comparison based on Poisson Process

The Poisson distribution is very useful for modelling purposes in many practical applications. It has been empirically found to well approximate many circumstances arising in stochastic processes(Adan, 2001). For our purpose, we assume that peers arrive and leave nearly according to a Poisson process. This assumption is consistent with literature (Fang et al. 2009). The numbers of peers and seeds existing in the pure P2P network modelled on M/G/∞ queue. We assume that two peers constitute the smallest pure P2P network; the smallest P2SP network includes one server and one peer; the smallest cloud includes one master node and one data node; and the smallest P2CP network includes one smallest cloud and one peer. We can get the number of peers and seeds that exist in the pure P2P network with time goes:

$$N = (\lambda - \mu)t \tag{1}$$

If a peer costs T time to download a file with size F in the P2P network, according to Poisson process, we could get the formulas as follow:

$$\sum_{k=1}^{N_s} \int_0^T U_s dt = F \tag{2}$$

$$\sum_{k=1}^{N_s} [(\lambda - \mu)U_s] \frac{1}{2} T^2 = F \tag{3}$$

If it costs a peer time T to download a file with size F in the P2SP network, we could get the formulas as follow:~~we get:~~

$$\sum_{k=1}^{N_s} \int_0^T U_s dt + T \sum_{k=1}^{N_{se}} U_{se} = F \tag{4}$$

$$\sum_{k=1}^{N_s} [(\lambda - \mu)U_s] \frac{1}{2} T^2 + T \sum_{k=1}^{N_{se}} U_{se} = F \tag{5}$$

If it costs a peer time T to download a file with size F in the P2CP network, we could get the formulas as follow:~~we get:~~

$$\sum_{k=1}^{N_s} \int_0^T U_s dt + T \sum_{k=1}^{N_c} U_c = F \tag{6}$$

$$\sum_{k=1}^{N_s} [(\lambda - \mu)U_s] \frac{1}{2} T^2 + T \sum_{k=1}^{N_c} U_c = F \tag{7}$$

If it costs a peer time T to download a file with size F in the cloud system,

we could get the formulas as follow:~~we get:~~

$$N_c * U_c * T = F \tag{8}$$

The relationship between the number of cloud host servers and relative throughput is volatile and according the normal Cloud storage system configuration, we get (9) for convenience of computation, supposedly one data piece has 3 duplicates over the Cloud.

$$N_c \geq 3N_{se} \tag{9}$$

We assume that:

$$A = \sum_{k=1}^{N_{se}} (\lambda - \mu) U \tag{10}$$

$$B = \sum_{k=1}^{N_{se}} U_{se} \tag{11}$$

$$C = \sum_{k=1}^{N_c} U_c = 3B = 3\sum_{k=1}^{N_{se}} U_{se} \tag{12}$$

Then, we can deduct and get:

$$T_c = \frac{2F}{6U_{se} * N_{se}} \tag{13}$$

$$T_{p2p} = \frac{2F}{\sqrt{2FA}} \tag{14}$$

$$T_{p2sp} = \frac{2F}{\sqrt{B^2 + 2FA} + B} \tag{15}$$

$$T_{p2cp} = \frac{2F}{\sqrt{C^2 + 2FA} + C} = \frac{2F}{\sqrt{9B^2 + 2FA} + 3B} \tag{16}$$

We easily get the result as follow:

$$T_{p2p} > T_{p2sp} > T_{p2cp}$$

For our comparative purposes, we assume that the size of the file is 100,000 KB, the upload bandwidth of the peers and seeds are 20 KB/s, the upload bandwidth of a server is 100 KB/s, the arrival rate of peers is 2 peers/s and the departure rate is 1 peer/s. When the peers' and servers' arrival rate is lower than departure rate, the number of peers and seeds will go to zero, and then the P2P network will extinct.~~not existing.~~ Figure 6 clearly shows that the alternative with minimal cost of download time is P2CP. The maximum download time is found with P2P. P2SP falls in the middle when there are not too many peers. The difference in download time is quite obvious. When more peers join the network, download time decreases.

Formatted: Justified

Formatted: Justified, No widow/orphan control, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Formatted: Centered

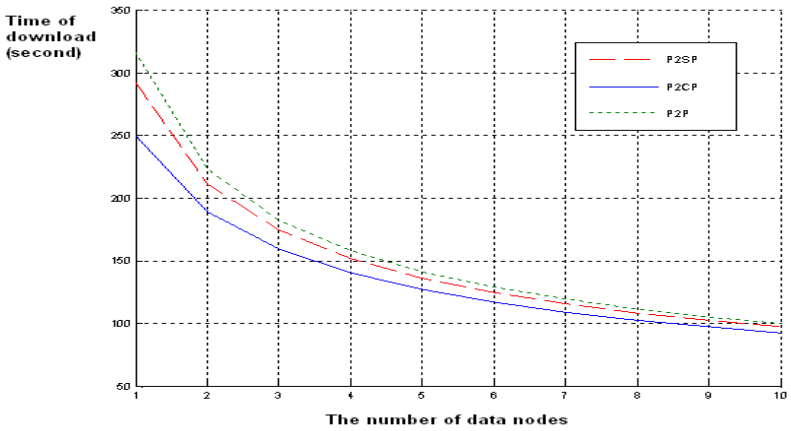

Figure 6: Time of download.

This clearly shows that the alternative with minimal cost of download time is P2CP. The maximum cost of download time is found with P2P. P2SP is in the middle. When there are not too many peers, the difference in download time is obvious. When more peers join the network, the cost of download time becomes less and less.

With the growth of upload bandwidth for the peers, we have another test. Assume that the size of a file is 100000 KB, the upload bandwidths for peers and seeds are 20 KB/s, 40 KB/s, or 60 KB/s, while the upload bandwidth for the server is 100 KB/s. The arrival rate of peers is 2 peers/s, and the leaving rate is 1 peer/s. Figure 7 shows that when there is an increase of upload bandwidth for the peers, the download time inversely decreases. At the same time, differences in download time between P2P, P2SP, and P2CP are also reduced. Pure cloud storage model performance is not shown in Figure 6, because the result changes significantly. In some instances it outperforms the P2P and the P2SP models depending on the chunk distribution in the cloud storage system, but it never outperforms our P2CP storage model.

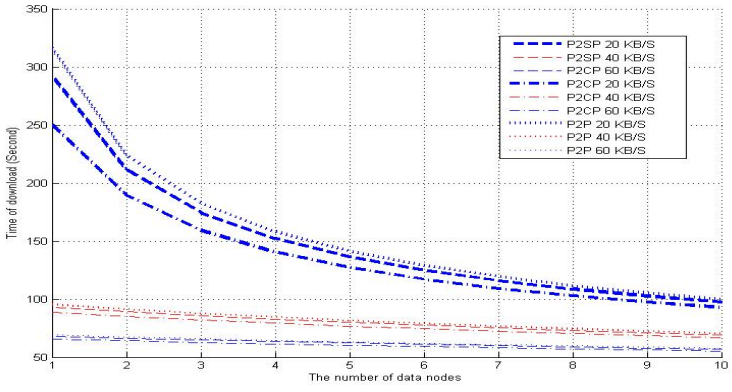

Figure 7: Comparing download time.

4.2 Comparison based on Little's Law

~~It is difficult to prove that~~Sometimes the peers' and seeds' arrival and departure rates are ~~accurate~~ not following~~according to~~ the Poisson process. Therefore, ~~herein~~ we ~~instead~~ herein use Little's law instead to relate L (number of peers), W (sojourn time), and λ (average number of users) (Adan, 2001) as $L = \lambda W$. Based on Little's law, we can get (due to space limitation, we skipped maths deductions):

$$N = (\lambda - \mu)T \tag{17}$$

$$T_{p2sp} = \frac{2F}{NU + 2B} \tag{18}$$

$$T_{p2p} = \frac{2F}{NU} \tag{19}$$

$$T_{p2cp} = \frac{2F}{NU + 6B} \tag{20}$$

$$T_c = \frac{2F}{6B} \tag{21}$$

According the (18), (19) and (20), we get:

$$T_{p2p} > T_{p2sp} > T_{p2cp}$$

Thus, minimum download time is still only possible with P2CP, then P2SP, and lastly with P2P (Sun, 2011).

5. ~~Discussion of Data~~Usability of Data Storage ~~Usability~~~~Availability~~

For a storage service, usability, which refers to download speed, system availability ~~, speed~~ and data security, ~~speed~~ ~~are~~ demands high priority considerations. In the previous section, we ~~proved~~ demonstrated that the data transmission speed of P2CP is superior. In this section, we compare and discuss the availability of P2CP in comparison with other models from the point of view of the whole network and of shared resources.

According the work of (Sands, 2009), common hardware failures of peers are happening more often in the network or clusters. It also suggests that certain servers are sometimes more prone to fail in a networked environment, though their availability is always much higher than single peers. ~~According the work of (Sands, 2009), we know that common hardware failures are often happened in the clustered, e.g.for example, certain servers are expected prone to fail sometimes in a networked environment.~~ Similarly, ~~I~~in our comparative evaluation, we assume the failure rate of each peer is 1%, and the failure rate of each server is 0.1%. In the following discussions, ~~W~~we assume that two peers constitute the smallest pure P2P network; the smallest P2SP network includes one server and one peer; the smallest cloud includes one master node and one data node; and the smallest P2CP network includes one smallest cloud and one peer.

~~From the point of view of whole network availability, based on the above, we observe the following:~~ In the P2P network, even if only 1 peer existed in the P2P network, when the user connects to the peer, the P2P network can still be set up. Thus the maximum failure rate of the

Formatted: Font: 16 pt, Bold, (Asian) Chinese (PRC), (Other) English (Australia)

Formatted: Font: (Default) Times New Roman, 12 pt

Field Code Changed

Formatted

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Widow/Orphan control, Tab stops: 1.27 cm, Left

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt

pure P2P network is 1%. ~~To~~ As for the failure rate of the P2SP network, failure for one ~~machine~~ node will not lead to a whole breakdown of the P2SP network. If the server is shutdown, the network becomes a P2P network; on the contrary, if the peer is offline, the network becomes a client and server based. Only when both~~,~~ the server and peer~~,~~ break down at the same time, the whole network will shutdown. Thus the maximum failure rate of the P2SP network is 1%*0.1% =0.001%. ~~To~~ Regarding the failure rate of the cloud network~~, according the features of cloud, we know that no level of~~, neither master nor data nodes shutdown will lead to the whole cloud network being fully disabled; except ~~until~~ when both master node and data node are broken at the same time. ~~So, t~~The maximum failure rate of the cloud network 1%*0.1% = 0.001%. P2CP network could run without peers, even the failures happened to master node or data node.~~,~~ U~~u~~ntil all peers ~~is~~ quit ~~gone~~ and both master node and all data nodes are broken, the whole P2CP network will not shutdown. So, the failure rate of P2CP is 0.1%*1%*1%=0.00001%. Thus, in the worst network situation, the most stable network storage model is P2CP.

From the point of view of a particular shared resource, ~~we know that~~ the storage services follow the long tail law (Anderson, 2006). This means that the particular resource may be very popular at the beginning but the demand eventually falls significantly for a long time. In the P2P storage model, initially the particular resource is frequently downloaded and uploaded in the network, so users can access the particular resource easily. However, when the particular resource is no longer popular, and the peers who hold the information for the particular resource ~~leave~~quit, ~~the P2P network is still there, but~~ the resource ~~is~~ will ~~not available~~be absent though the P2P network is still present.

Both the cloud storage model and the P2SP storage model tried to solve this problem. They use a series of servers or a single server to record the particular resource to guarantee the availability, ~~but~~ probably with different transmission efficiency. The transmission efficiency of the P2SP storage model is only improved~~,~~ when the particular resource is popular and the transmission efficiency is high;~~,~~ but when the particular resource is unpopular, the transmission efficiency ~~is~~ will be lower. Interestingly, ~~T~~the cloud storage model gets the opposite result. Therefore, ~~O~~only the P2CP storage model achieves the best result to balance the popularity and efficiency. Regardless of whether the particular resource is in fashion, the availability and speed are ~~very good~~better than other storage models.~~-~~ From the evaluation results of Section 4.1, we can clearly see that in whatever the situation, the cost in time for P2CP is the lowest, and the usability is highest.

From the point of view of security in the life cycle of data~~ life cycle~~, data creation phase, data storage phase, data archival~~e~~ phase and data destruction phase should be classified to internal security layer, because ~~of~~ all of these phases occur in the cloud. Only data use and sharing phase~~,~~ involves all security layers, because cloud-users would have opportunity and obligation to balance the transmission speed and data security. For example, when cloud-users need high transmission speed, they can use all transmission tunnels in different security layers. Otherwise, if cloud users just need ~~high~~ better data security, they could just exploit transmission tunnels in the internal security layers and the contracted security layers to achieve this goal. Moreover, cloud-users could only employ cloud-user data transmission tunnel in internal security layer to achieve extreme data security. ~~Subsequently~~Hence, security over the data life cycle in the proposed P2CP storage model can be guaranteed at different levels.

6. Conclusion

In summary, we have introduced a secure cloud storage system model to enhance data transmission performance and provide persistent ~~availability~~usability, which has been named P2CP. ~~The conclusion of~~ According to our comparative studies ~~presented in this paper,~~ based on statistical modelling, P2CP not only enhances the utilization rate of bandwidth that exists in cloud storage systems, but also ~~may be a solution of~~solves the problem of persistent availability in the P2P network model. We ~~prove using a mathematical model~~demonstrate that the utilization rate of bandwidth and the persistent availability of the P2CP model ~~should be~~are better than ~~for~~ the pure P2P model, the P2SP model, or the cloud model. ~~Moreover,~~

In addition, we ~~set up~~utilise a "hierarchy security" ~~method~~ architecture, which includes three security layers, namely ~~that are~~ internal security layer, contracted security layer and functional security layer, to guarantee the data security in P2CP storage model, where the cloud-users have opportunity to balance the transmission speed and data security. In future ~~work~~, we will ~~pursue~~ conduct an empirical evaluation based on ~~building the~~ the prototype P2CP storage system and test the robustness of security and performance of computation as well as overhead of~~and~~ communication.

References

Adan, I. and J. Resing (2001) Queueing Theory, pp. 111-113 Eindhoven, The Netherlands.

Amazon (2006) "Amazon S3," http://aws.amazon.com/s3/ (07/12, 2010).

Anderson, C. (2006) The Long Tail, in Why the Future of Business Is Selling Less of More, 2 edition, vol. 24, pp. 170–177: Hyperion.

Borthakur, D. (2007) "The Hadoop Distributed File System: Architecture and Design", The Apache Software.

Chou, Y. (2006) "Get into the Groove: Solutions for Secure and Dynamic Collaboration," TechNet Magazine., http://technet.microsoft.com/en-us/magazine/2006.10.intothegroove.aspx (22.March, 2011).

Clarke., I. (2000) "Freenet," http://freenetproject.org/ (07/12, 2010).

Cohen, B. (2001) "BitTorrent," http://www.bittorrent.com/btusers/what-is-bittorrent (07/12, 2010).

Díaz-Verdejo, J., García-Teodoro, P. and Maciá-Fernández, G. (2009) "Environmental security in P2P networks" Proceedings of the First International Conference on Advances in P2P Systems, pp 138-143

Fang, L., L. Peng, Y. Jie, and L. Zhenming. (2009) "Contrastive Analysis of P2SP Network and P2P Network. Wireless Communications", WiCom '09. 5th International Conference on Networking and Mobile Computing, pp. 1-5.

Feng, J., Y. Chen, W.-S. Ku, and P. Liu. (2010) "Analysis of Integrity Vulnerabilities and a Non-repudiation Protocol for Cloud Data Storage Platforms", 2010 39th International Conference on Parallel Processing Workshops (ICPPW), pp. 251-258.

Hiroyuki, S., A. Kanai, and S. Tanimoto. (2010) "A Cloud Trust Model in a Security Aware Cloud", 10th Annual International Symposium on Applications and the Internet, pp121-124.

Hong, T., A. Gulbeden, Z. Jingyu, W. Strathearn et al. (2004) "A Self-Organizing Storage Cluster for Parallel Data-Intensive Applications". Proceedings of the ACM/IEEE SC2004 Conference Supercomputing, pp. 52.

Kirk, P. (2003) "Gnutella," http://rfc-gnutella.sourceforge.net/).

Merkur (2002) "eMule," http://www.emule-project.net/home/perl/general.cgi?l=1 (07/12, 2010).

O'Hearn, Z. W. and B. Warner (2008) "Tahoe: the least-authority filesystem", Proceedings of the 4th ACM international workshop on Storage security and survivability, pp. 21-26. Alexandria, Virginia, USA: ACM.

Ozzie, R. (2005) "Microsoft, Groove Networks to Combine Forces to Create Anytime, Anywhere Collaboration," http://www.microsoft.com/presspass/features/2005/mar05/03-10GrooveQA.mspx (22 March, 2011).

Sands, A. and V. Tseng (2009) "An analysis of reported laptop failures from malfunctions and accidental damage", SquareTrade.

Sanjay, G., G. Howard, and L. Shun-Tak (2003) "The Google file system", in Proceedings of the nineteenth ACM symposium on Operating systems principles, pp. 29-43. Bolton Landing, NY, USA: ACM.

Stein, C. A., M. J. Tucker, and M. I. Seltzer. (2002) "Building a reliable mutable file system on peer-to-peer storage. Proceedings. 21st IEEE Symposium on Reliable Distributed Systems", pp. 324-329.

Sun, Z., Shen, J. and Beydoun, G. (2011) "P2CP: A New Cloud Storage Model to Enhance Performance of Cloud Services," in Proceedings of International Conference on Information Resources Management, Conf-IRM, 2011

Sun;, Y., F. Liu;, B. Li;, and B. Li; (2009) "FS2You: Peer-Assisted Semi-Persistent Online Storage at a Large Scale", in INFOCOM 2009, IEEE, pp. 8. Rio de Janeiro

Xiaojun Y.; Q. Wen;, (2010), "A View about Cloud Data Security from Data Life Cycle " 2010 International Conference on Computational Intelligence and Software Engineering (CiSE) pp: 1 - 4

Ye, W., A. I. Khan, and E. A. Kendall. (2003) "Distributed network file storage for a serverless (P2P) network". The 11th IEEE International Conference on Networks,, 2003, pp. 343-347.