DATA TRANSFER MODEL FOR CLOUD COMPUTING

Rambabu Vatti, Manish Karani*

Electronics and Telecommunication Department, Vishwakarma Institute of technology, Pune, India

ARTICLE INFO

Corresponding Author:
Manish Karani,
Electronics and
Telecommunication Department,
Vishwakarma Institute of
technology, Pune, India
mdkarani@gmail.com.

KeyWords: Cloud Computing,
Challenges, 4G, Availability of
service, Model, Data Criticality.

INTRODUCTION

Cloud computing, a new kind of computing model which has the potential to transform a large part of the IT industry, makes software more attractive as a service and redefines the way IT hardware is designed and purchased. Large capital investments in hardware or the human expense incurred is no longer a requirement for developers with innovative ideas for new Internet services. The Cloud Computing can also be termed as dynamic computing because it provides resources when required (dynamically).

Cloud Computing includes software applications provided as services over the Internet and the hardware and systems software in the datacenters that provide those services made available in a pay-as-you-go manner. The services provided are referred to as Software as a Service (SaaS). The datacenter hardware and software is what we will call a Cloud. The service being sold is Utility Computing. Thus, Cloud Computing is the sum of SaaS and Utility Computing, but does not include Private Clouds. People can be users or providers of SaaS, or users or providers of Utility Computing.

Given below is the equation which indicates the trade-off between cost of cloud and cost of maintaining a datacenter as mentioned in [2].

\[
\frac{\text{UserHours}_{\text{cloud}} \times (\text{revenue} - \text{Cost}_{\text{cloud}})}{\text{UserHours}_{\text{datacenter}} \times (\text{revenue} - \frac{\text{Cost}_{\text{datacenter}}}{\text{Utilization}})} \geq 1
\]

Four aspects form the hardware point of view are new in Cloud Computing.

a) The illusion of infinite computing resources available on demand, thereby eliminating the need for Cloud Computing users to plan far ahead for provisioning. [2],

b) The elimination of an up-front commitment by Cloud users, thereby allowing companies to start small and increase hardware resources only when there is an increase in their needs [2],

c) The ability to pay for use of computing resources on a short-term basis as needed (e.g., processors by the hour and storage by the day) and release them as needed, thereby rewarding conservation by letting machines and storage go when they are no longer useful [2],

d) The elimination of the need of storage space and manpower for maintaining the hardware.
TYPES OF CLOUDS

Cloud computing can have three types of clouds: Public, Private and Hybrid Clouds.

A. Private cloud

It is a proprietary architecture subscribed by an organization, accessed by users within the organization; general public cannot access hosted services from the private cloud. This is protected by the firewall.

B. Public cloud: It is not proprietary of any organization; the services provided in these clouds can be accessed by any organization [4].

C. Hybrid cloud

In hybrid cloud, the services are offered to the limited and well defined number of parties [4].

Fig. 2 Hybrid Cloud model

Style of Cloud Computing:

SAAS (Software as a service)

A cloud provider hosts and provides access to a software application over a network, mostly Internet.

Table 1. Weight factors for cloud adoption calculation [5]

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of IT resources</td>
<td></td>
</tr>
<tr>
<td>Number of servers</td>
<td></td>
</tr>
<tr>
<td>Annual revenue from IT offerings</td>
<td></td>
</tr>
<tr>
<td>Workload variability</td>
<td></td>
</tr>
<tr>
<td>Average usage</td>
<td></td>
</tr>
<tr>
<td>Type of services</td>
<td></td>
</tr>
<tr>
<td>Amount of data handling</td>
<td></td>
</tr>
<tr>
<td>Sensitivity of data</td>
<td></td>
</tr>
<tr>
<td>Criticality of work done</td>
<td></td>
</tr>
</tbody>
</table>

The cloud user/SaaS provider hosts and manages software in their own data center and makes it available to multiple leaseholders and users over internet. From the user’s point of view, this can save some cost on servers and software. From the provider’s point of view, they can also save cost as they only need to maintain one program.

PAAS (Platform as a service)

Platform as a service provides development environment as a service. This model is primarily used by software developers who focus on the development cycle and monetizing of new applications.

IAAS (Infrastructure as a Service)

Infrastructure as a Service enables us to install and execute the software. Here, users can gain access to virtualized server. IaaS targets operating systems, hardware, CPUs and embedded systems, networks and storage. This enables a homogenous virtualized environment where specific software will be installed and executed. Amazon provides infrastructure as a service [1].

CRITERIA FOR ADOPTION OF CLOUD:

Criteria which should be considered while adopting the cloud are mentioned below

1. Size of IT resources
2. Utilization of existing resources
3. Sensitivity of data
4. Availability of services
5. Data lock-ins and data transfers
6. Performance unpredictability
7. Software licensing
8. Legal terms and conditions.
Calculations: [5]
Largeness value:
L= NoS *CNoS + NoC * CNoC + AR * CAR (1)
Average Usage value:
AU=ToS*CToS (or) ToP*CToP+(4-SCB)*CSCB
Peak Usage value:
PU=DoP*CDoP+PbA*CPbA (3)
Value of Workload Variability:
WV=PU*CPU+AU*CAU+ADH *CADH (4)
Value of Data Sensitivity:
DS = SoD (5)
Value of Criticality:
C = CWD (6)
Suitability index = L*CL + WV*CWV+DS*CDS*ADH+C × CC
× (65-L)
Where,
NoS = Number of Servers
CNoS = Credit of Number of Servers
NoC = Number of Countries it is Spread Across
CNoC = Credit of Number of Countries it is Spread Across
AR = Annual Revenue
CAR = Credit of Annual Revenue
SCB = Size of Customer Base
CSCB = Credit of Size of Customer Base
ToS = Type of Service
CToS = Credit of Type of Service
ToP = Type of Project
CToP = Credit of Type of Project
DoP = Duration of Peak
CDoP = Credit of Duration of Peak
PbA = Peak by Average
CPbA = Credit of Peak by Average
CPU = Credit of Peak Usage
CAU = Credit of Average Usage
ADH = Amount of Data Handling
CADH = Credit of Amount of Data Handling
CWV = Credit of Work Variability
CC = Credit of Criticality.

DATA TRANSFER MODEL:

Fig. 3 Data transfer Model for Cloud computing

Model Description:
1) The approach is intermediate to the concept of cloud and present private datacenters.
2) Instead of giving all the hardware and software responsibilities to a cloud provider, a company will have a base station geographically proximate to the branches of the company and/or nearest in distance to the cloud providers center. This will decrease the distance of data transfer and hence, will reduce the cost of data transfer.
3) As companies presently have data centers of their own, the model makes use of the already built infrastructure, however all the data centers of the company would not be utilized. The model would use only 3-5 data centers of the company around the world.
4) The above diagram shows the model for a company spread globally with 3 base stations and global coverage of the base stations.
5) The hardware and systems software maintenance, data updating work and management in the base station, would be done by the cloud provider under the supervision of the client company. Hence the critical and confidential data would not have to be stored in the cloud provider’s data center.
6) The base station will act as an interface between the cloud provider and other branches of the company, the base station can be a working branch of the company.
7) The base station will have certain hardware storage and all data transfers and software/platform services between the cloud provider and branch centers of the company will take place through the base station using 4G technology.
8) 4G systems will prove far cheaper than 3G, since they can be built atop existing networks and won't require operators to completely retool and won't require carriers to purchase costly extra spectrum. In addition to being a lot more cost efficient, 4G is spectrally efficient, so carriers can do more with less. [6].
9) As all the data transfers and software/platform services will take place through the base station, the number of centers around the world the cloud has to serve will reduce and the cloud will only have to serve the base stations of companies. The amount of data and software/platform services will increase proportionally. But it is better than handling more number of centers with distributed data and service requests. This will reduce the bottleneck situations and help to give equal priority of service to all clients.
10) The hardware storage in the base station will function as a cache memory, frequently used data will be stored in the base station, and after the work and processing on the data is complete and has to be stored for a long time or is not frequently accessed, it will be shifted to the cloud for storage. This will help reduce the cost of data transfers.
11) The company would have to subscribe to the cloud provider only for the servers and extra storage space required at peak instances.
12) At one point of time, transferring data to the Amazon cloud was so expensive that, FeDexing disks to the cloud center was much cheaper than transferring through the internet, even though sending disks by courier services and uploading the data on the internet was time consuming. [2]
13) Another issue faced by the cloud is of data sensitivity and criticality, the data can be breached at two points, first during transfer of the data through the internet and the other threat is when it is stored in the cloud. The first breach is already taken care of by using 4G as it is and IP based network, also firewalls and data encryption can be introduced. For the second breach the model is helpful. Some experts in the field are apprehensive of storing their firm’s confidential and critical data in the cloud. The base station can store the sensitive and critical data which the company does not want to upload on the cloud.
14) As the number of centers demanding services is grouped together and one representative service center is demanding service from the cloud, the obstable of performance unpredictability which is related to I/O resource sharing and its virtualization in the cloud can be overcome.

Limitations of the model:
1) The base stations electricity cost and miscellaneous maintenance would have to be by the company.
2) The global implementation of 4G will take approximately 2-3 years from date.

Advantages of the model:
1) Use of already built hardware.
2) Provides data confidentiality and security.
3) Extra storage space and servers can be provided by the cloud, thereby decreasing the cost of adopting and using the cloud services the cloud further.

AVAILABILITY OF SERVICE:
Another challenge faced by the Cloud is availability of service. Organizations have an apprehension about whether cloud services will have adequate availability. Outages in AWS, AppEngine, and Gmail are given below.

<table>
<thead>
<tr>
<th>Service and Outage</th>
<th>Duration</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3 outage: authentication service overload leading to unavailability</td>
<td>2 hours</td>
<td>2/15/08</td>
</tr>
<tr>
<td>S3 outage: Single bit error leading to gossip protocol blow up</td>
<td>6-8 hour</td>
<td>7/20/08</td>
</tr>
<tr>
<td>AppEngine partial outage: programming error</td>
<td>5 hours</td>
<td>6/17/08</td>
</tr>
<tr>
<td>Gmail: site unavailable due to outage in contacts system</td>
<td>1.5 hours</td>
<td>8/11/08</td>
</tr>
</tbody>
</table>

A proposed solution for this is using multiple cloud providers. The high-availability computing community has long followed the mantra “no single source of failure,” yet the management of a Cloud Computing service by a single company is in fact a single point of failure. Even if the company has multiple datacenters in different geographic regions using different network providers, it may have common software infrastructure and accounting systems, or the company may even go out of business. Large customers will be reluctant to migrate to Cloud Computing without a business-continuity strategy for such situations.

The advantages of using multiple cloud providers are listed below,
1) The company will have an option while paying for software’s and software platforms, the company can select the cloud provider providing software’s at a lower price.
2) If one of the clouds is out of services, the company can still survive by working in shifts till the cloud has continued its services, as shown below.

Fig. A company’s work load divided in shifts during outage of cloud

RESULTS
Economics of cloud computing shows that hype which exists around it will soon disappear [4], considering the above model and solutions we see that the mathematical equations discussed above for criteria of adoption of cloud is simplified using 4G.

Suitability index = L*CL + WV*CWV+DS*CDS*ADH + CC × (65-L)

The terms involving criticality and sensitivity in the above equation will become negligible, making criteria for cloud adoption simpler.

Consider the equation

UserHours\text{cloud}* (revenue - Cost\text{cloud}) = UserHours\text{datacenter} * (revenue - (Cost\text{datacenter} / Utilization))

Now, Cost\text{cloud}= Storage space + servers
We assume, 20% data is critical stored in data centers

Cost\text{cloud} = 0.8*storage space + servers

Hence, cost of cloud decreases and makes it more feasible for adoption of cloud.

CONCLUSION AND FUTURE WORK
Hence, we conclude that 4G can be a good option for overcoming obstacles faced by the Cloud. The model described above deals with data sensitivity and criticality and confidentiality and reduces cost of data transfers. Opting for multiple clouds is an option which can solve the problem of availability of service; it also gives the company to choose while buying software and platforms, it also helps during outages of cloud services.

Future work in this direction would involve designing protocol for data transfer involving interoperability between cloud providers, so that during unfavorable events the cloud users and in turn the SaaS users are not affected.

REFERENCES


[4]. Xu Wang, Software School of Xiamen University, Xiamen, China E-mail: wangxu047@gmail.com. Beizhan Wang Software School of Xiamen University Xiamen, China E-mail: wangbeizhan@gmail.com. Jing Huang, Department of Land and Resources of Jiangxi Province Nanchang, China E-mail: huang_jing@126.com (2011) IEEE. “Cloud computing and its key techniques”.

---

<table>
<thead>
<tr>
<th>Table 2: Outages in AWS, AppEngine and Gmail [2]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service and Outage</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>S3 outage: authentication service overload leading to unavailability</td>
</tr>
<tr>
<td>S3 outage: Single bit error leading to gossip protocol blow up</td>
</tr>
<tr>
<td>AppEngine partial outage: programming error</td>
</tr>
<tr>
<td>Gmail: site unavailable due to outage in contacts system</td>
</tr>
</tbody>
</table>

---
Karani et al/ Data Transfer Model for Cloud Computing

[5]. Sameer Rajan Apurva Jairath, Govt.of India, MCIT, DIT Department. of C.S. & Engg., National Informatics Centre (NIC) G.G.T.S., Naharlagun, Arunachal Pradesh (India) Jabalpur M.P. (India), e-mail: sameer.rajan@nic.in, e-mail: apurvajairathit@gmail.com (2011) International Conference on Communication Systems and Network Technologies. “Cloud Computing: The Fifth generation of Computing”.

[6]. Berislav Biocic, Drasco Tomic, Dario Ogrizović*, *Hewlett-Packard Croatia, Zagreb, Croatia, **Center for advanced computing and modeling / Faculty of Maritime Studies, Rijeka, Croatia (MIPRO 2011). “Economics of the Cloud Computing”.