Billing System Design of Cloud Services

Benzhi Sui¹, Qiuquan Li², Jun Wen¹* and Yuanzhe Yao¹

¹School of Information & Software Engineering, University of Electronic Science and Technology of China, Chengdu, P. R. China
²Shanxi railway engineering vocational and technical college, P. R. China
*Corresponding author

Abstract—Cloud computing can provide scalable computing power, storage space, and information services for a variety of applications. It has dynamic scalability, fault-tolerance and availability, which require a large number of nodes to provide support. Providing corresponding compensation mechanisms for the participating nodes can increase their enthusiasm for participating in service support. In order to provide customers with the use of cloud resources measurement and costs to better promote the development of cloud services. This paper proposes a cloud service billing system. The resource collector periodically collects resource usage information, and transmitted to the billing server. The billing will calculate the corresponding resource consumption based on these data, and store these records in the database to provide user management. The resource rates vary dynamically based on the current infrastructure load and service quality. After test, this system only has partial errors in CPU billing due to the long statistical interval (1 second). The accuracy of the rest reached 95%.

Keywords—cloud computing; cost; cloud billing; pricing; billing models

I. INTRODUCTION

Cloud computing is often hailed as the future of enterprise and enterprise technologies. It can reduce computer charge, reduce software charge, and improve performance. It increased user flexibility in planning, owning, purchasing, and using cloud computing systems. The attractive prospect of cloud computing is carefree IT—honed to enterprise needs and can quickly respond to changing business environment. Billing is the missing link but it is the key to monetizing cloud services. A reasonable billing system can allow service providers to provide better services and can allow consumer consume reasonably.

For the dynamic pay-per-use model in the cloud environment based on real-time pricing of the cloud service, this paper proposes a smart billing model that provides for dynamic pricing of the cloud service based on the load condition of the cloud. The resource collector periodically collects resource usage information, and transmitted to the billing module. The billing module will calculate the corresponding resource consumption based on these data. And the resource rates will vary dynamically based on the current infrastructure load and service quality.

The rest of the paper is organized as follows. Section II introduce related work. Section III descripts our IaaS online billing system framework which includes design ideas, billing model and pricing, collect and dispatch resource information system. Performance results and some conclusions are presented in Sections IV and V respectively.

II. RELATED WORK

[1] defined the billing and the billing process. Billing management is the core component of IaaS in cloud computing. [2] argues that the service should be provided through fixed billing price, which is expensive for short-term use. In this approach, consumers are billed the same amount every month, regardless of actual usage. [3] proposed an approach which estimates cost but did not compare the calculated cost with other cloud computing providers. [4] has introduced a way to describe the advantages of cloud computing in the paid use of an enterprise, which means accurate usage measurement and billing with a large number of users. According to this paper [5], flexibility should be a key feature of the billing system.

[6] proposed the design and implementation of Aquarium, an extensible billing service software. This approach provides portability but does not achieve the desired performance. [7] described three pricing models used by cloud service, which are used to pay for the resources used, namely tiered pricing, per-unit pricing and subscription based pricing.

III. IAAAS ONLINE BILLING SYSTEM FRAMEWORK

A. Design Ideas

From Figure 1, tasks are received on the cloud server and distributed to each node by the task distribution mechanism. Collect the resource consumption of each node to the billing server and calculate the cost of the task on the billing server. Among them, the hardware of different nodes is different, the price is different, the price will be set separately for calculate.

![FIGURE 1. DESIGN IDEAS](image-url)
IV. BILLING MODEL AND PRICING

Billing and pricing are two important aspects of the intelligent measurement model.

**Pricing:** Pricing of a cloud service can be applied based on multiple considerations. Current service providers, whose cloud instances are priced primarily based on duration of use and configuration. Another popular option is to charge the consumers a certain amount of time.

According to [8], intelligent measurement proposes pricing based on the dynamic operational cost of running the service. In this pricing model, the base cost of running the service is specified by the service provider, \( C_{\text{base}} \). Pricing rules which define the pricing overhead for running the service under various load conditions are specified by the service provider. This pricing overhead is given by \( \beta(l, t) \), where \( l \) is the load at time instance of operation \( t \). The current price, \( P_t \), for the interval at a given load condition is given as,

\[
P_t = C_{\text{base}} \times \beta(l, t)
\]  

(1)

where \( P_t \) is operational price at time \( t \), \( C_{\text{base}} \) is base operational cost, \( \beta(l, t) \) is pricing overhead for running the service under load \( l \) at time \( t \).

**Billing:** Dynamic billing on the cloud is a function of the instantaneous load on the cloud, as well as the price information obtained according to the configuration specified by the service provider. Billing calculations involve determining the overall load of the cloud over a recent history interval and obtaining the weighted sum of the load based on the entities and the corresponding pricing information as described above. The total load \( L_t \) on the cloud infrastructure at time \( t \) is the sum of the load on each node as

\[
L_t = \sum_{i=1}^{n} l_i
\]  

(2)

where \( L_t \) is total load on the cloud at time \( t \), \( l_i \) is load index of the individual cloud component. The load obtained in (2) is mapped to a corresponding pricing value at a given interval of time \( t \). The bill amount is computed as a summation of the product of instantaneous pricing obtained in (1) and the utilization of the consumer, \( U_t \). The total bill amount is obtained as,

\[
\text{Bill} = \sum_{i=1}^{n} P_t \times U_i
\]  

(3)

Where \( P_t \) is operational price at time \( t \), \( U_i \) is resource utilization of the consumer at time \( t \).

**Billing of CPU and Memory:** The CPU utilization sampled in the unit time and calculated it according to the dynamic price, and then added it up, because of the different CPU utilization in each period. In the memory section, it calculates the maximum amount of memory the program uses on the server.

\[
B_{\text{CPU}} = \sum_{i=1}^{n} P_{\text{CPU}} \times U_i \Delta t
\]  

(4)

\[
B_{\text{mem}} = M_{\text{mem}} \times P_{\text{mem}}
\]  

(5)

\( P_{\text{CPU}} \) is CPU price at time \( t \). \( U_i \) is resource utilization of the consumer at time \( t \). \( \Delta t \) is sampling interval. \( M_{\text{mem}} \) is the maximum memory consumption. \( P_{\text{mem}} \) is price of different memory consumption levels.

**Billing of Storage and Network Traffic:** The storage section, respectively, counts the total amount of read and write, and calculates. In the network part, the traffic flow of the upstream and downstream is counted and calculated according to different bandwidth.

\[
B_S = S_r \times P_{sr} + S_w \times P_{sw}
\]  

(6)

\[
B_{\text{net}} = N_r \times P_{nr} + N_s \times P_{ns} + P_{bw}
\]  

(7)

\( S_r \) is the total of storage read. \( S_w \) is the total of storage write. \( P_{sr} \) is price of storage reads. \( P_{sw} \) is price of storage write. \( N_r \) is the total of receive traffic. \( N_s \) is the total of send traffic. \( P_{nr} \) is price of receive traffic. \( P_{ns} \) is price of send traffic. \( P_{bw} \) is price of bandwidth.

V. COLLECT AND DISPATCH RESOURCE INFORMATION.

We use a self-developed resource information collection system and deploy it on each node of the cloud system to count the resource consumption. The resource information for each process on the node is sampled every \( \Delta t \). And publish it through the Publish–subscribe pattern. On the Metering server, subscribe to the node that needs to be charged and collect this resource information.

Information collected and sent includes: Host Id, IP, User ID, Process Group ID, Process ID, Process start time, Percent of CPU at \( \Delta t \), Peak memory size at \( \Delta t \), total amount of read, total amount of write, total amount of network traffic receive, total amount of network traffic send, Command of the process and so on.
VI. PERFORMANCE AND RESULTS

The test section sets the price of each part to 1. A test program that runs for 10 minutes was established to collect the CPU running time and maximum memory amount of the test program through top. Through the self-writing test program, control network send and receive, store read and write file size, measure network send and receive total amount, store read and write total, compare it with charge system calculate the cost that get. After a hundred tests, the average is calculated as follows:

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Top</th>
<th>IOBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>60179.65</td>
<td>59873</td>
</tr>
<tr>
<td>Memory</td>
<td>4103</td>
<td>4103</td>
</tr>
<tr>
<td>Net Read Total</td>
<td>32110</td>
<td>32110</td>
</tr>
<tr>
<td>Net Write Total</td>
<td>34110</td>
<td>34110</td>
</tr>
<tr>
<td>Disk Read Total</td>
<td>26323</td>
<td>26323</td>
</tr>
<tr>
<td>Disk Write Total</td>
<td>18902</td>
<td>18902</td>
</tr>
</tbody>
</table>

As you can see from the above table, this system only has partial errors in CPU billing due to the long statistical interval (1 second). The accuracy of the rest reached 95%.

VII. CONCLUSION

This paper presented a novel billing model for a cloud service. The cloud service billing system was implemented based on the user resource in the cloud computing environment, and can dynamically adjust the billing rate at a specific time according to the load condition of the infrastructure. With the help of scientific and reasonable resource billing system, more and more nodes will take part in the cloud computing services, computing resources can be efficiently applied.

ACKNOWLEDGEMENT

This research was financially supported by National High Performance Computing Environment Operation Management Support Platform, No 2016YFB0201403.

REFERENCES


