Research on building elastic cloud computing based on Windows Azure platform

Wei GuangHui^{1,a}

¹ Chongqing College of Electronic Engineering, Chongqing, 401331 ^{a.}290768260@qq.com

Keywords: Large data, elastic cloud, rule research

Abstract. In this paper, in view of the big data era enterprise cloud computing applications, analysis the Windows Azure platform to build elastic cloud computing based on, including: Windows Azure platform, automatic telescopic application block and restriction rules, reaction rules and automatic retractable life cycle, and which relates to the primary XML code editing and analysis, in order to achieve enterprise application function construction of elastic cloud computing.

Introduction

From 2009 onwards, big data has become a popular word in the Internet technology industry. Big data and cloud computing are linked together, because real-time large-scale data set analysis requires a distributed processing framework to distribute work to tens, hundreds or even tens of thousands of computers. However, cloud computing is a business behavior, according to the enterprise occupied Cloud Calculation of the amount of resources to implement cost calculation[1]. Therefore, in the era of big data companies building cloud computing services must be faced with the problem is to obtain a balance between cloud computing performance and cost, this point in pay attention to high yield, high efficiency of modern enterprises is particularly important. This requires all enterprises in the construction and use of cloud computing services, you need to use a flexible cloud computing capabilities[2].

Windows Azure platform

Windows Azure platform is the operating system on the Microsoft cloud platform, running in Microsoft's data center. Windows Azure platform uses virtualization to abstract the hardware resources, the enterprise released to Windows Azure application running on VM (virtual machine). Different enterprises released to Windows Azure applications may be different VMS on the same physical host, these different VM need to share network I / O and CPU core and disk space, these different VM is based on the same physical hardware structure of a layer of abstraction with transplantation and scalability benefits, business users can be in VM can be realized with scalability of elastic cloud computing function, so as to realize the balance between the cloud computing performance and cost.

VM can migrate and distribute among different physical hosts, through virtualization technology, multi tenant and demand integration, so as to achieve high data center utilization, so as to save the cost of enterprise users. Virtualization technology includes vertical virtualization and level virtualization technology. Vertical virtualization refers to a particular virtual machine has the amount of resources, namely Windows Azure computing resources occupied by the environment

enterprise users customize their own virtual machine role; and level virtualization is directly increase the number of virtual machine instances, the virtual machine instance contains Web and worker are two kinds of roles.

Implementation of elastic cloud computing in Windows Azure platform

When the enterprise users will be applied to the cloud computing application to Windows Azure operation, is in the Windows Azure cloud platform to apply for a virtual machine, in the virtual machine contained in the release of the Web and Worker two types of role. Enterprise users can specify the number of examples for each role, the number of examples show that the larger the computing capacity, that is, the more resources to occupy Windows Azure cloud platform, and thus the cost of the enterprise to pay more. Therefore, enterprise users in the application of the enterprise will be released to the Windows Azure, you can indicate the size and quantity of each role, and according to the dynamic changes in the role of examples.

The Autoscaling Application Block. The Autoscaling Application Block is Windows Azure Component of enterprise library integration package, Enterprise users can use it to build a flexible cloud computing. Application of the flexible block to play a role through the rules of enterprise user defined, through these rules, can control the application in the virtual machine load changes, when the reaction occurs, and how to react. In Windows Azure, These rules is divided into two categories, restricting rules, such as the number of characters in the virtual machine make the limitation of the maximum and minimum; reaction rules, such as the collected from the application of the counter or measure to change number of role instances.

Restriction rule. The restriction rules can specify the effective time area of the rule according to the schedule, and set the number of different roles (including Web and Worker roles) in different time zones. The role of the number of maximum and minimum value, the minimum value to ensure that the application to meet the quality of service agreement, the maximum value is used to control the operating costs of Windows Azure applications.

In enterprise cloud computing applications, the load mode is predictable. For example, to engage in the one or two industry, the largest load often occurs in the normal working hours, while engaged in the third industry enterprises, the maximum load will occur at other times more[3]. Therefore, different enterprises according to their own application needs, through the time period to customize the rules.

The need for Web roles and Worker roles in different time periods, respectively, defines the number of different instances. Suppose now a web role roleA, the default minimum examples for Rmin, the results for Rmax (the default priority takes a low value of 1), the work of section starting time Tstart and termination time tend, in section at least examples for Rwmin. Results for Rwmax (the priority from the higher values of 90). In the case of not including the reaction rules, as shown in figure 1:



Figure 1 RoleA simple restriction rule

Based on the above simple restrictions on the RoleA rules, by the enterprise IT professionals to create automatic scaling rules, the XML code is as follows:

<rules

xmlns=http://schemas.microsoft.com/practices/2012/entlib/autoscaling/rules enabled="true"> <ConstrainrRules>

<rule name="default" description="windows azure active" enabled="true"rank="1" > <actions>

<range min=" Rmin"max=" Rmax"target="RoleA"/>

</actions>

</rule>

<rule name="working" description="active at working times" enabled="true" rank="90"> <actions>

<range min=" Rwmin"max=" Rwmax"target="RoleA"/>

</actions>

<timetable startTime="Tstart"duration="Tend"><daily/>

</timetable></rule></constaraintRules><reactiveRules/><operands/></rules>

Need to pay attention to, because in Windows Azure cloud platform charging mode is according to the clock hours to billing[4], so in the above code, work section starting time Tstart best is the whole point, and termination time tend the best is the whole point of reducing X minutes (X minutes for contraction operation will take the cooling time, x value determined by the need for cooling the number of characters, usually set to 10 minutes). Otherwise, a clock hour will be according to the clock hour enters billing. The XML code into the Windows Azure cloud computing server XML document file in order to allow the application of automatic block access.

For a normal use of the enterprise cloud computing, will use a number of different rules, an annual limit of rules may have different maximum value and minimum value, according to the automatic telescopic block conflict adjustment algorithm was stretching:

I .if the current number of instances is lower than the minimum value of all the restrictions, the number of instances will rise to the minimum of the smallest of all the rules;

II if the current number of instances is greater than the maximum value of all the restrictions, the number of instances will be reduced to the maximum value that is the smallest of all;

I. other conditions remain unchanged.

Response rule. Rules that limit is the maximum and minimum limits on the number of characters, and reaction rule is adjusted within the range of limited licensing rules, the upper and lower bounds of the reaction rules formulated by the rules, the reaction rules for difficult to advance the projected demand sudden increase or decrease. The reaction rules is based on automatic telescopic application block surveillance data point values to adjust the number of characters, the count includes: counter or metrics, aggregate functions and calculate aggregate values used in the time period. Reaction rules use expressions to specify stretching conditions, this expression can be a simple numerical and threshold comparison, can also be complex expressions, including inequality of combination is composed of a plurality of operands.

Web role in a period of time to reject the number of requests and Web CPU role of the asp.net utilization as an example, to illustrate the application of the reaction rules, write the following XML code section:

<when><all>

.

<greater operand="PublicWeb_AspNetrequestRejected_Avg_20m"than"5"/>

```
<greater operand="PublicWeb_CPU_Avg_20m"than"85">
</all></when>
......
<when><all>
<lessoEqual operand="PublicWeb_AspNetrequestRejected_Avg_20m"than"1"/>
<lessOrEqual operand="PublicWeb_CPU_Avg_20m"than"35">
</all></when>
.....
```

Reaction rules are usually in pairs, a rule for extended operation (to increase the number of characters) and another rule for tightening operation (to reduce the number of characters), such as XML code segment above, the front part is extended operation: when 20 minutes to the asp.net and the number of requests for more than 5, and web role of CPU average utilization rate of more than 85%, increase a role instance; the back part is tightening operation, to meet the conditions of reduction of role instances, of course, or its scope is in rules that limit the determined scope instances within[5].

Conditions for the use of automatic telescopic applications. Enterprise use of cloud computing, is not the time to use the automatic expansion of the application block. Such as the relatively simple application or user with less number of applications, or intend to will Windows Azure application of part of the cost of fixed, as well as to what do not have horizontal scalability (not by adding examples to improve the performance of the application, you should not use automatic telescopic application block. Therefore, automatic telescopic application block should use in such several conditions: one is the automatic response in a period of time, web and worker roles of load with time change is large, the second is time to additional resources can be predicted, which can forecast to what time will be the sudden demand, complex applications, and is in need of application program running cost of the enterprise management. In these situations, we can use the example of the role of expansion and contraction to achieve the balance between the performance and cost of cloud computing.

Conclusion

In the era of big data, the human society is increasingly dependent on the computing and storage capabilities provided by the data center. For traditional data center, the resources waste depends on the application design and mode of operation, many applications will be in accordance with its runtime spend resources to apply for resources and leading to a large number of additional allocations of resources can not be effectively recycled. The construction of flexible cloud computing can avoid this phenomenon, improve resource utilization, reduce the cost of enterprise investment in cloud computing.

References

[1] Youping Tang, Shuguang Zhou. Research on Application of Case Method [J]. Journal of Sichuan College of Education, 2010, 26(1):27-30

[2] Zhenghui Li. Critical Process and Enlightenment of American Case Method [J]. Law Review of Nanjing University, 2012 Autumn Volume, 337-358.

[3] Bolan Zhong, Zhifeng Cui. Eliminate Mistaken Ideas of Case Teaching [J]. Vocational Education Research, 2008, 5:126.