

集群计算环境下的计费系统设计

宋文豪 文军* 姚远哲

电子科技大学信息与软件工程学院, 四川 611731

摘要: 集群计算能够为各种应用提供可动态扩展的计算能力、存储空间和信息服务, 具有动态可伸缩性、容错性和可用性, 然而这些需要大量的节点提供支持, 只有为参与服务的节点提供相应的补偿机制, 才能提高其参与服务支持的积极性。本文正是对应上面的需求, 提出了一种集群服务计费系统。计费系统包括资源采集、计量和计费, 资源采集器周期性地采集资源用量信息, 传输给计费模块, 然后计算出对应的资源消费, 其中资源的费率能够根据当前基础设施的负载情况以及服务质量动态地变化。该计费系统的实现能为客户提供资源消耗的计量与计费, 并能更好地促进计算服务的发展。

关键词: 集群计算; 计费系统; 资源管理; 系统负载; 服务质量

一、前言

集群计算^[1]是一组相互独立的、通过高速网络互联的计算机, 它们构成了一个组, 并以单一系统的模式加以管理, 集群配置是用于提高可用性和可缩放性。通过集群技术, 可以在付出较低成本的情况下获得在性能、可靠性、灵活性方面的相对较高的收益。随着集群计算技术的发展以及普及, 企业与个人对于集群服务的需求逐渐增大, 因此集群服务的提供开始走上商业化的道路, 多租户和按需采购模式使得拥有强大的计量和计费机制变得至关重要。因此服务提供商在服务商业化的进程中能否取胜的关键便在于是否具有科学的计量, 计费标准。现有的模式有基于网格的, 以及基于云服务模式的IaaS, PaaS以及SaaS^[2]。目前国内外典型的商业IaaS云服务有亚马逊Amazon Web Service (AWS)^[3], 微软的Windows Azure^[4]以及国内的盛大云等。以AWS为例, 其主要服务包含Amazon EC2以及Amazon S3, 分别对计算资源以及存储资源计费。其中EC2只需要为自己使用的EC2实例付费, 通过对用户使用的CPU、硬盘、内存、网络流量等计费因子实施计量并结合固定费率得到服务费用。其他几种云服务的计费策略与AWS的计费策略也存在一定的相似性。由此可见, 在传统的商用IaaS计费模型中, 用户按照一个预设的固定费率实施付费, 收费标准单一、不合理, 没有充分考虑定价策略的业务相关性以及用户服务质量、集群负载情况。

基于网格系统中的各个节点的计费系统中, 用户需要在各自群内节点上进行计费, 在这种情况下, 针对网格的特殊性出现了很多非传统意义上的计费方案, 如虚拟用户技术和临时计费系统^[5], 其中比较有代表性的有网格银行所提供了一种价格协商的方法, 讨论了各种记录中数据格式统一的问题并提出了组件间交互的多种协议。通过订立合同来保证计费的无欺诈, 并通过来收集资源消费的信息产生资源消费记录, 从而实现计费的功能; 多组织网格计费系统通过技术收集资源使用过程中的基本静态信息和任务执行过程中的动态信息, 其中包括数目、执行时间、存储容量、网络流量等等, 从而计算获得用户应该支付的费用^[6]; 多组织网格计费系统通过技术收集资源使用过程中的基本静态信息和任务执行过程中的动态信息, 其中包括数目、执行时间、存储容量、网络流量等等, 从而计算获得用户应该支付的费用。

面向用户的计算过程是一个动态, 实时的过程, 每个特定时间段对资源的消耗情况是动态变化的, 需要一个动态费率来对用户实施实时计费。显然, 现有的计费模型无法满足此需求。因此需要开发一款可实时变换费率的弹性计费模型。在此计费模型中, 用户的资源费率会基于当前基础设施集群的负载情况动态变化, 用户在当前时段的资源耗费便可根据此动态费率计算出来。与此同时, 服务提供商也可根据统计到的系统负载情况对基础设施群进行更好地管理。

本文提出的计费系统不仅能够统计常规的物理资源消耗情况, 还会根据当前的系统负载以及服务质量来动态地调整资源费率。整个计费系统划分为资源采集模块以及资源计费模块, 资源采集模块被部署到各个虚拟机上, 资源采集器周期性地采集资源用量信息, 并传输给计费模块, 计费模块根据计费策略库, 计算出当前时间段各种资源的费用。

文章后面的组织结构如下: 第一节对此领域的其他研究成果进行了分析比对, 以突显此系统的独特性所在。第二节详细地阐述了此计费系统的架构, 资源采集器及计费模型设计情况。第三节为大家展示了此系统各功能模块的具体

* 通讯作者: 文军, 1966年10月, 男, 汉族, 现就职于电子科技大学信息与软件工程学院, 副教授, 博士研究生。研究方向: 大数据处理。

实现方式。最后总结，并对深入研究做出展望。

二、相关工作

在这一节中，我们调查分析了近几年针对云计算和集群环境下计费模型的研究成果，以展示该领域的研究现状。并将这些研究成果与本文所提出的计费模型作比较，突显此模型的独特性所在。

相较于最初按月租赁基础设施的计费模式，胡军^[7]等人所提出的计费系统引入了按量计费的思想，整个计费系统被划分为两大功能模块，数据采集模块被部署于各个虚拟机上，并周期性的获取虚拟机的资源用量信息，最后通过计费模块对用户实施按量计费。但是该系统存在一个很大的问题，在计费的时候，计费模块仅着眼于用户使用虚拟机的总时间，通过使用虚拟机的时间来计算用户的费用，虽然这的确是一种按量计费，但是太过于笼统，客户并不能明确自己消耗了哪些资源，以及这些资源的消耗情况。本文所提出的计费模型将进程的CPU使用量，进程读取/写入磁盘量，网络上传/下载量等资源用量信息作为计费因子，并分别对它们的耗量实施计费。

当前针对浮动利率的服务计费系统的研究并不多，肖哲彬^[8]提出了一种动态的计量计费模型，可依据集群当前资源稀缺度动态调整资源价格。资源的稀缺程度在一定程度上反映了当前的供求关系，因此通过动态调整稀缺资源的价格可以平衡资源的供需关系。在动态调价的基础上，此系统设计并实现了基于浮动利率的平台计量计费系统，并在日任务量千万级的条件下对任务实例的资源用量进行分析计算，给出用户账单。相较于本文，首先本文的计费系统应用于集群平台下，其次，针对费率调整的问题，本文着眼于当前基础设施集群的负载情况对用户服务质量的影响，而并不是旨在寻求供求关系的平衡。谈及系统负载，Narayan^[9]所提出的智能计费模型可基于历史数据，使用自回归整合移动平均（ARIMA）统计模型预测平台设施上的负载，以此来估算下一个时间段的资源价格，用户便可根据此价格数据决定是否继续使用该服务。

通过以上分析可发现，集群平台下的计费模型大多还是以用户为中心，充分考虑用户的实际资源用量，让用户费用尽可能地贴近合理范畴，不管是采用何种技术，研究者们最终都希望将用户权益放在首位，实现一个真正意义上的按需、按量计费系统。

三、模型设计

整个计费系统从结构上分为两大功能模块：资源采集模块，资源计费模块。资源信息采集模块收集各个云实例上的资源用量，并将其发送给计费模块；计费模块主要统计所接受到的计量数据，并按照服务商设计的计费策略，对特定用户实施分类按量计费、存档操作。

四、资源采集器模块

在资源采集模块中，采用资源监控工具，实时收集部署在各个虚拟机应用进程，在数据采集模块中调用其内部API，可以获取到主机ID，进程组ID，进程所有者的用户ID，进程的CPU使用量，进程读取磁盘量，进程写入磁盘量，网络上传量，下载量等资源用量信息了。

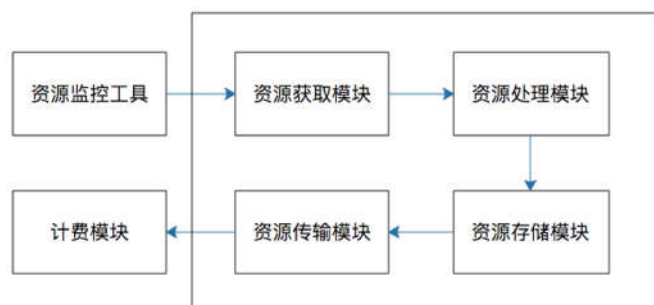


图1 资源采集功能模块图

如图1所示，采集模块主要分为四大功能模块：资源获取模块，资源处理模块，资源存储模块，资源传输模块。其中，资源获取模块调用资源监控API来获取计费因子的原数据，并且记录获取到该数据的时间。资源处理模块对获取到的数据进行处理，直接计算出每个虚拟机各种资源消耗量，经传输模块发送给计费模块，供其处理计费，将处理后的数据存入数据库中。

五、资源计费模块

资源计费模块获取来自资源采集模块的各种计费元素后，结合相应的计费策略，计算出特定用户的账单信息。其功能组件如图2所示，主要功能模块为：数据接收模块，数据存储模块，计费策略模块，计费生成模块。

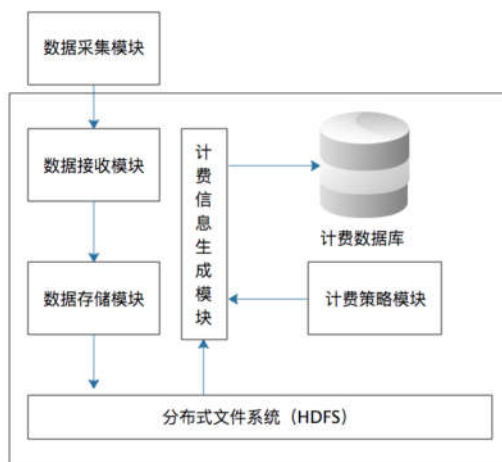


图2 计费功能模块图

资源计费模块获取到的用量信息包含：主机ID，进程所有者的用户ID，进程CPU使用量，进程读取/写入磁盘量，网络上传/下载量。资源计费模块会收到大量的来自不同虚拟机的资源用量信息，使用MapReduce对这些资源进行处理^[10, 11, 12]，将进程所有者的用户ID以及各类资源耗量信息提取出来，并分别作为MapReduce^[13]中的<K.V>对输入，经过map和reduce操作，得到某用户当前时间段内的各种资源用量的信息，其具体处理流程如图3所示。

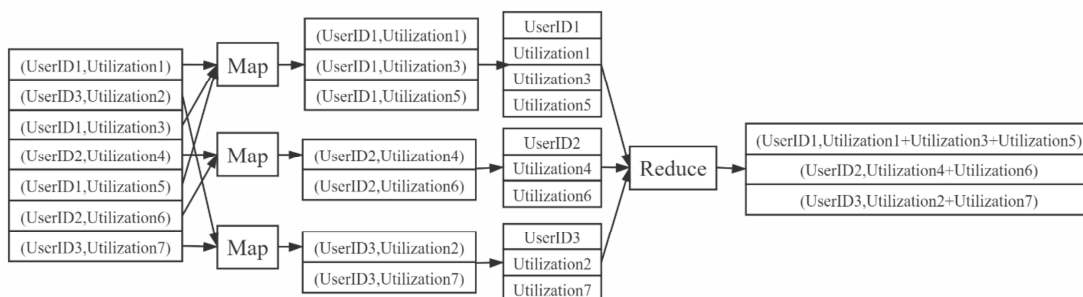


图3 资源用量数据MapReduce处理流程图

针对任何一类资源用量，服务提供商都可以给出一个基础费率C_{base}，根据当前的系统状态S，存在一个β S变化率，该变化率可影响当前的费率，其关系式如下所示：

$$C_{current} = C_{base} \cdot \beta_s \tag{1}$$

其中C_{current}代表当前的实际费率，β本身其实代表一种变化率，该变化率是由服务提供商设置，随着系统负载^[14]而变化。比如当下载/上传速度小于某个值的时候就调整β的值，因为此时的用户服务质量变低，其收费标准也应该随之变低。

取得了资源用量及计价信息后，把两个数据相乘便可以计算出在当前时间段里的费用信息，其计算公式如下所示：

$$Bill = C_{current} \cdot u_{interval} \tag{2}$$

其中u_{interval}代表当前时间段内某种计费因子的资源用量。

将其累加到之前的资源花费中便得到了某计费因子的总费用Bill_{sum}，计算公式如下所示：

$$Bill_{sum} = \sum_{t=1}^n Bill_t \tag{3}$$

某个客户使用资源的总费用是累计各个资源用量的总费用。

六、系统实现

本计费系统由多个数据采集器、多个数据汇聚器、一个负载均衡器和一个计费服务器所组成，其中，数据采集器，分别设置并运行在于该集群系统中的各个宿主机，负责采集运行于宿主机的各虚拟机的实时状态信息，并将采集的各虚拟机实时状态信息通过数据汇聚器上传到计费服务器；数据汇聚器，负责汇聚各数据采集器采集的实时状态信息，并在负载均衡器的控制下，将汇聚的虚拟机实时状态信息传输给计费服务器；负载均衡器，用于收集各个数据汇聚器和数据采集器的负载信息，依据预先设定的负载均衡策略平衡各个数据汇聚器的负载，避免单个数据汇聚器出现

热点而成为该系统的性能瓶颈；计费服务器，负责将数据采集器上传的虚拟机实时状态信息存入分布式数据库，同时依据预设的计费算法利用并行计算组件对存储于分布式数据库中的数据进行处理，以获取计费结果，再用计费组件将该计费结果展示给系统用户或管理员。

其中资源采集模块的核心是资源监控^[15, 16]，一切费用的产生都是源于资源用量变化，在此系统中，资源监控工具被部署在各个虚拟机上，实时地监控虚拟机的资源用量情况，资源用量信息的获取流程如算法1。当有了各虚拟机的资源用量信息，可以根据公式（1）与公式（2）分别计算出当前时间段的资源费率与资源费用，并将资源费用累加到之前总的资源费用中，最后将最新的费用信息存储到计费数据库中。具体的计费流程如算法2中描述。

七、资源监控

算法1:

```

begin Monitoring the resource utilization
while VM is Active do
    foreach instance in VM do
        resourceUtilization ← getResourceUtilization (VM)
    end
    Store (resourceUtilization)
end
end

```

算法1描述从虚拟机获取资源用量的流程，当接受到getResourceUtilization调用函数时，就根据调用请求返回指定的资源用量信息，返回的信息通过Store函数存储于本地文件中。

八、资源计费

算法2:

```

Data: resourceUtilization,VM,time
begin Calculate the Bill
foreach timeInterval in time do
    BillAmount←getStoredBillamount(VM,timeInterval)
    foreach VM in Infrastructure do
        resourceUtilization← getStroedutilization(VM)
        currentPrice←convertBaseprice(timeInterval)
        intervalBill←resourceUtilization* currentPrice
    end
    BillAmount ← BillAmount+ intervalBill
    Store (BillAmount)
end
end

```

算法2描述资源计费流程，计费程序通过getStroedutilization函数的调用获取由虚拟机发送过来的用量信息，调用convertBaseprice函数对当前系统的负载情况以及服务质量来对基础费率进行调整，结合资源用量信息以及实时费率计算出当前时间段的资源费用信息，然后累加到数据库中去。

九、结论

本文设计实施的集群服务计费系统成功地实现了用户资源用量地按量计费，并且根据基础设施的负载情况及服务质量动态地调节即时费率，实现科学合理的资源计费，同时，通过订阅方式的计费方法，可以根据需要，完成集群内特定节点的计费任务。它的部署能够记录加入的节点系统所提供的服务，并以此提供一定的补偿，从而能够吸引更多的服务提供者加入集群服务。

参考文献:

- [1]Michael Armbrust,Armando Fox,Rean Griffith,etc.,A view of cloud computing.Commun.ACM 53, 2010(4),50-58.
- [2]袁玉宇,胡文博,基于IaaS云计算平台的弹性计费模型,中兴通讯技术, 2012,18(06):34-37.
- [3]Cloud A E C,Amazon web services,Retrieved November, 2011,9.

- [4]Chappell D,Introducing the windows azure platform,David Chappell & Associates White Paper, 2010.
- [5]McGinnis,Thigpen,etc,Accounting and accountability or distributed and grid system,2nd IEEE/ACM International Symposium in Cluster Computing and Grid. 2002(5).
- [6]Alexander Barmouta,etc,A grid accounting services architecture for distributed systems sharing and integration, 2006.2.
- [7]胡军,崔毅东,OpenStack环境下计费系统的研究,北京:中国科技论文在线.
- [8]肖哲彬,基于浮动利率的PaaS平台计量计费系统设计与实现,哈尔滨工业大学, 2017.
- [9]Narayan A,Rao S,Ranjan G,et al,Smart metering of cloud services, 2012 IEEE International.IEEE, 2012:1-7.
- [10]Borthakur D,HDFS architecture guide,Hadoop Apache Project, 2008,53.
- [11]Karun A K,Chitharanjan K,A review on hadoop—HDFS infrastructure extensions, 2013 IEEE Conference on.IEEE, 2013:132-137.
- [12]Wu J,Ping L,Ge X,et al,Cloud storage as the infrastructure of cloud computing, 2010 International Conference on.IEEE, 2010:380-383.
- [13]Dean J,Ghemawat S,MapReduce:simplified data processing on large clusters,ACM, 2008.
- [14]Kim K H,Youn H S,Kang Y C,Short-term load forecasting for special days in anomalous load conditions using neural networks and fuzzy inference method,IEEE Transactions on Power Systems, 2000,15(2):559-565.
- [15]OmPrakash D,Waheed M A,A Virtual Machine Resource Management Method with Millisecond Precision, 2017.
- [16]Chen J,System and method for monitoring and managing logistics resource information:U.S. Patent Application 09/973,305, 2003-4-1