

云计算下基于容错 QoS 的虚拟机资源分配策略

黄 飞, 汪 芸

(东南大学 计算机科学与工程学院 计算机网络与信息集成教育部重点实验室, 江苏 南京 211189)

摘 要: 为了提高云环境下虚拟机的可靠性及执行用户请求的成功率, 采用了基于 cloudsim 的开源框架构建了云计算平台, 提出了基于容错 QoS、随机分配及按序分配三种资源分配策略, 并对提出的三种策略进行了对比. 实验结果表明, 基于容错 QoS 的虚拟机资源分配策略有性能上的优势.

关键词: 云计算; 容错 QoS; 资源分配; cloudsim

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2013)03-0136-04

Virtual Machine Resource Allocation Strategies Based on Fault-tolerant QoS in Cloud Computing

BEN Fei, WANG Yun

(School of Computer Science and Engineering, Southeast University,
Key Lab of Computer and Information Integration, MOE, Nanjing 211189, China)

Abstract: In order to improve the reliability of virtual machines and the success rate of user's requests in cloud environment, this paper proposes to build the cloud computing platform based on the open source framework cloudsim. Three strategies including fault-tolerant QoS, randomly allocation and sequentially allocation are presented and compared. Empirical results show that the resource allocation policy of fault-tolerant QoS outperforms the other policies.

Key words: cloud computing; fault-tolerant QoS; resource allocation; cloudsim

1 引言

云计算是分布式计算、效用计算、虚拟化技术、Web 服务、网格计算等技术的融合和发展^[1], 其目标是用户通过网络能够在任何时间、任何地点最大限度地使用虚拟资源池, 处理大规模计算问题. 虚拟化可以说是云计算的核心技术之一, 通过虚拟化, 可以动态地组织异构的计算资源, 隔离具体的硬件体系结构和多样化的软件系统平台, 灵活构建满足不同应用需求的计算环境^[2].

云计算平台在收到云计算用户的任务请求后会根据用户 QoS 创建相应的虚拟机, 用以执行用户的请求, 一个虚拟机可以同时服务多个请求. 文献[3]采用虚拟化技术, 将资源归类形成多级资源池, 结

合不同策略(共享策略、私有策略、收回策略和借还策略)完成资源的调度. 该方法可有效减少资源空闲, 提高资源利用率, 但没有从用户花费的角度进行考虑, 也没有从资源的动态性方面进行考虑, 资源提供的过程中缺乏灵活性. 文献[4]使用随机整数规划优化资源提供方法, 采用预订和即付即用两种方式给用户提供服务. 该策略能提供动态的资源分配方法, 考虑用户各个阶段的资源花费, 能达到最小化客户资源花费的目的, 但在不确定资源的预测精确性上还有待提高.

文献[5]在本地资源管理系统 HPC4U 和 AssessGrid 中引入了 SLA 感知的概念, 采用运行时间估计的资源管理策略, 提供协作的监测协议, 支持故障容错, 实现 QoS 管理. 文献[6]对基于 QoS 的资

收稿日期: 2012-06-27; 修回日期: 2012-07-21

基金项目: 国家“八六三”高科技计划项目(2011AA040502)

源管理系统进行了研究,提出了 QoS 层次模型、QoS 参数的映射和转换等在资源管理系统中的重要作用。

上述工作都从不同方面研究了云计算下的资源利用策略,但没有从用户角度考虑容错问题,即假设物理主机不出错,但实际情况并不如此,大多数的云计算平台搭建在廉价服务器的基础之上^[1-2],因此云计算中的物理主机集群很容易发生个别主机失效的错误。针对此问题,本文提出了三种虚拟机资源分配策略:基于容错 QoS、随机分配和按序分配的策略,其中基于容错 QoS 的分配策略具有较好的性能。本文提出的三种资源分配策略均具有较好的容错性,对容错调度^[7]也具有一定的意义。

2 容错 QoS

容错 QoS 分为上下两层:用户层和虚拟层,用户层的容错 QoS 从用户的角度出发考虑用户对容错的需求,因此用户层 QoS 应该表达直观,易于理解并且可区分。虚拟层的容错 QoS 与平台关系比较密切,其表示方式与具体的资源和采用的容错技术相关,同时,用户层容错 QoS 和虚拟层容错 QoS 之间存在某种映射关系,通过映射,可以将用户层的容错 QoS 转化为虚拟层的容错 QoS,为云计算平台虚拟机的创建准备前提条件。

云计算服务模式分为三层,从上到下依次为软件即服务(SaaS)、平台即服务(PaaS)和基础设施即服务(IaaS),虚拟化技术大多被采用在 PaaS 层,比如调用 IaaS 层接口创建符合一定需求的虚拟机,与此相对应,用户层的容错 QoS 体现在 SaaS 层而虚拟层的容错 QoS 对应 PaaS 层,容错 QoS 的层次化表示与云计算的服务模式相吻合。

2.1 容错 QoS 的表示

容错 QoS 在不同的层次有不同的表示方法,如图 1 所示。在用户层,考虑到用户不可能提出具体的容错参数,比如需要运行多长时间或者需要几个节点作为副本,因此用户层的 QoS 用 UserFTQoS 参数表示,该参数具有高、中、低,三个等级值,这样比较直观容易理解,也符实际需求,而在虚拟层的 QoS 则需要细化用户层 QoS 三个容错级别,用采取容错措施所需要的节点的个数来反映,也就是说采用多少副本节点来执行此任务,用 VirtualFTQoS 参数表示,该参数值为一整型。

2.2 容错 QoS 的映射

容错服务质量在不同层之间存在一种映射关

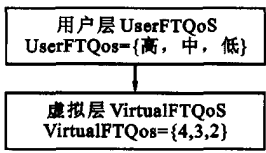


图 1 容错 QoS 层次映射

系,映射规则可以有很多种,本文给出的映射规则如下:

$$\begin{cases} \text{VirtualFTQoS}=4, & \text{if UserFTQoS=高} \\ \text{VirtualFTQoS}=3, & \text{if UserFTQoS=中} \\ \text{VirtualFTQoS}=2, & \text{if UserFTQoS=低} \end{cases}$$

在这种映射关系下,用户层的高级容错服务质量意味着系统能够屏蔽主机崩溃错误和软件执行错误两类,对应 4 个节点中有 3 个用来执行用户请求,1 个作为备份;中级容错服务质量意味着系统能够屏蔽主机崩溃错误或者软件执行错误中的一类,具体操作时可依据实际情况而定,对应 3 个节点可以采用复制技术或者选举的方式;低级容错服务质量只可以屏蔽主机崩溃错误,对应 2 个节点 1 个用来执行用户请求,1 个作为备份。由于实施高级容错 QoS 的任务时,可适用的资源较多,可以屏蔽更多类型的错误,此时这些任务比带有较低容错 QoS 的任务具有较大的执行成功概率。

3 虚拟机资源分配策略

虚拟机的构建有必要考虑容错性,在进行资源分配时,应尽可能地将此虚拟机构建在几个物理主机上,这样就可以采用复制的方法进行容错。

为了权衡虚拟机的容错性和传输代价问题,本文提出了网络距离的概念。网络距离是数据中心内部任意两个物理主机在网络拓扑上的距离,反映了两个节点是否“彼此邻近”和它们之间带宽的大小。理论上网络距离不能太大,因为距离过大则传输数据的时间代价变大,但网络距离又不能太小,距离太小,主机同时失效的概率就会增大,因此,在实际分配资源时要综合考虑这两个因素。

3.1 网络距离

典型的云计算中的服务器集群网络拓扑如图 2 所示。服务器集群由数据中心组成,数据中心内部由服务器机架组成,一个机架内的服务器之间传输速度很快,而不同机架内的服务器节点传输可能慢一些,数据中心之间的节点的传输可能更慢。

本文将两节点间的带宽以及是否在同一物理位置作为衡量网络距离的标准,即对于以下四个场景,可用带宽依次递减的方法表示网络距离:

- (1)同一节点中的进程;
- (2)同一机架上的不同节点;
- (3)同一数据中心不同机架上的节点;
- (4)不同数据中心中的节点.

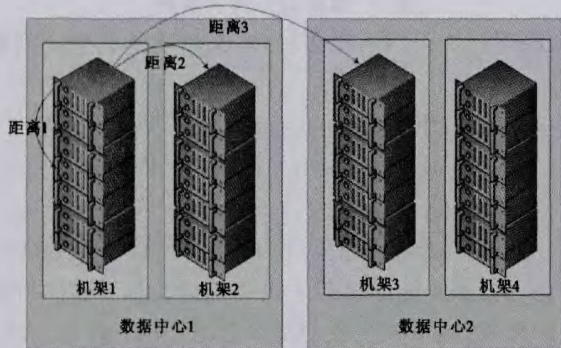


图2 服务器集群拓扑结构及网络距离的表示

任意一个节点可表示为 $d1/r1/n1$,即在数据中心1($d1$)机架1($r1$)中的节点($n1$).因此,在上述四个场景中的距离可描述为:

- (1) $distance(/d1/r1/n1,/d1/r1/n1) = 0$ (同一节点中的进程);
- (2) $distance(/d1/r1/n1,/d1/r1/n2) = 1$ (同一机架上的不同节点);
- (3) $distance(/d1/r1/n1,/d1/r2/n3) = 2$ (同一数据中心不同机架上的节点);
- (4) $distance(/d1/r1/n1,/d2/r3/n4) = 3$ (不同数据中心中的节点).

3.2 基于容错 QoS 的资源分配策略(FTQoS)

用户的任务请求动态地到达云计算平台,对于队列中的每一个任务请求,FTQoS 算法为其选择合适的节点来创建虚拟机并在其上执行此请求,其算法伪代码如图3所示.

FTQoS 算法首先计算出满足任务请求容错 QoS 需求的副本节点数量 num (算法第 1-3 行),如果所需的节点数等于 2,则首先随机选择一个机架,选出此机架上局部队列最短的那个节点作为第一个节点,然后从距离此节点距离为 2 的节点中选出局部队列最短的那个节点作为第二个节点(算法第 4-8 行),如果所需的副本节点数等于 3,则按照前面的方法选择第一个节点,与前面不同的是接下来选择距离此节点距离为 2 的节点中选出局部队列最短的两个节点(算法第 9-13 行),如果所需的副本节点数等于 4,选择前 3 个节点的方法和前面一样,第四个节点从距离第一个节点为 2 的节点中选择局部队列最短的那个节点,但此节点与前两个节点不在同一

个机架上,如果数据中心只有两个机架则在同一机架上选择(算法第 14-17 行).最后是任务的分发、虚拟机的创建和任务的执行(算法第 18-23 行).

```

1. for each requesti , find the primary and backup nodes,
   do
2. find = FALSE;
3. calculate the needed nodes num;
4. randomly choose one rack;
5. and select the least loaded node in the rack;
6. if num==2
7.   select one more least node from the nodes whose
     distance is 2;
8.   if success then find = TRUE;
9.   else
10.    select two more least node from the nodes whose
       distance is 2;
11.    if success&&num==3
12.      find = TRUE;
13.    else
14.      select one more least loaded node form the
        nodes whose distance is 2;
15.      if success then find = TRUE;
16.    end if
17.  end if
18.  if find == TRUE then
19.    create the vitural machine;
20.    send the job to the primary and backup nodes;
21.  else
22.    reject the requesti;
23.  end if
24. end for

```

图3 FTQoS 算法的伪代码

3.3 随机分配策略和按序分配策略

随机分配策略和按序分配策略比前一种策略简单,随机分配策略(FTRandom)在选择资源时采取的是随机的方法,例如用户容错 QoS 为中,按照前面所述的映射规则应该为其选择 3 个物理主机,在随机分配的策略下就随机地选择 3 个物理主机并在其上创建虚拟机.按序分配策略(FTSimple)与此类似,不同的地方在于选择物理主机的时候按主机的序号选择,云计算中的资源在加入云中的时候会有一个资源 id,FTSimple 算法即是按照此序号按顺序为虚拟机分配资源.

4 实验验证和结果分析

本文对 cloudsim 云计算仿真平台^[8]进行了扩展,在原有虚拟机资源分配策略的基础上,主要增加了 VmAllocationPolicyFTQoS 等几个资源分配策略类.实验环境为 30 个物理主机、1 个数据中心还

有数量不等的虚拟机,通过使用本文提出的三种资源分配策略将这些虚拟机创建在 30 个物理主机上,用以进行实验验证,性能方面的比较。

4.1 平均距离

平均距离计算的是创建一个虚拟机的各个主机间的平均距离,表现的是此虚拟机的容错性能和通信代价。如图 4 所示。

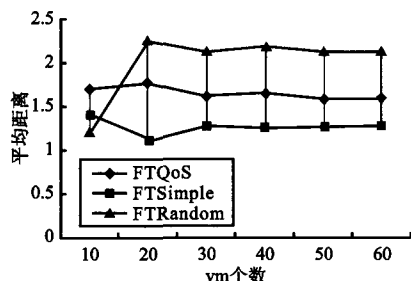


图4 平均距离

从图 4 可以看出 FTSimple 的平均距离最小,可以想象其通信代价会很小,但其容错性不好,因为创建它的主角平均距离短同时失效的概率大。而 FTRandom 的平均距离最大,其通信代价大但容错性好,相对而言,FTQoS 介于它们之间,表现出一种比较折中的性能。

4.2 虚拟机失败率

横坐标是创建虚拟机的个数,纵坐标是失败的虚拟机所占的比例。失败模型为同一机架上的主机失效相互之间不独立,有一个概率上的相关性,不同机架上的主机失效相互独立。如图 5 所示。

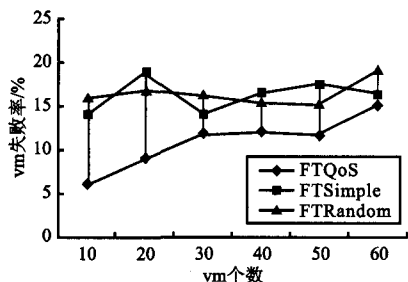


图5 虚拟机失败率

从图 5 可以看出 FTQoS 的失败率最小,其他策略都比它高,并且随着创建虚拟机数量的增加,三种策略的失败率都随之增加,但 FTQoS 仍然表现出良好的性能。

5 结束语

本文提出了三种云计算环境下的虚拟机资源分配策略,三种策略均考虑了容错 QoS,将容错性作为一个可区分的指标,并通过实验比较得出 FTQoS 策略与其他两种策略 FTRandom 和 FTSimple 相

比,在容错性、负载均衡性还有通信代价方面都具有较好的性能。三种虚拟机资源分配策略都有助于提高云计算平台的稳定性和用户请求执行的成功率,对推动云计算的发展具有一定的意义。

下一步要做的工作是在此基础上的用户请求调度,包括主机级、虚拟机级、用户级的任务调度,还有容错 QoS 的重协商,即当不能满足用户提出的容错要求时,可以通过降低用户的容错 QoS 进行调度。

参考文献:

- [1] 刘鹏. 云计算[M]. (2 版). 北京:电子工业出版社, 2010.
- [2] 曾龙海,张博峰,张丽华,等. 基于云计算平台的虚拟集群技术研究[J]. 微电子学与计算机, 2010, 27(8): 32-35.
- [3] Buyya R, Abramson D, Giddy J, et al. Economic models for resource management and scheduling in grid computing[J]. Concurrency and Computing, 2002 (14): 1507-1542.
- [4] Neumann D, Stoesser J, Weinhardt C. Bridging the adoption gap-developing a roadmap for trading in grids[J]. Electronic Markets, 2008, 18(1): 65-74.
- [5] Nimis J, Anandas ivam A, Borissov N, et al. SOR-MA- business cases for an open grid market: concept and implementation[C]// Proceedings of the 5th international workshop on Grid Economics and Business Models(GECON 08). [s. l.]: IEEE, 2008: 173-184.
- [6] 高宏卿,邢颖. 基于经济学的云资源管理模型研[J]. 计算机工程与设计, 2010, 31(19): 4139-4212.
- [7] Zheng Q, Veeravalli B, Tham C K. On the design of fault-tolerant scheduling strategies using primary backup approach for computational grids with low replication costs[J]. IEEE Computers, 2009, 58(3): 380-393.
- [8] Calheiros R N, Ranjan R, Beloglazov A, et al. CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms[J]. Software: Practice and Experience, 2011, 4(1): 23-50.

作者简介:

贲 飞 男, (1987-), 硕士研究生. 研究方向为分布式计算。

汪 芸 女, (1967-), 教授, 博士生导师. 研究方向为分布式计算、传感器网络和容错算法。