

云计算环境下基于主动预测的节点部署模型研究

马自堂 陈鹏 李兆兴

(解放军信息工程大学三院 郑州 450004)

摘要 随着云计算的不断普及,越来越多的用户选择将自身的业务迁移至云计算系统。用户的使用习惯与社会日常的运行规律也伴随着大量用户涌入云计算系统,如每早8点集中地向云计算系统申请资源节点,这给系统带来了一种可预期的资源冲击。针对上述问题,提出了一种基于主动预测模式的资源部署模型。该模型首先根据预测模块中Holt-Winters季节指数平滑模型的算法周期长度来预测下一个时间周期的任务请求量,通过设计的主动预测算法判断是否应对当前的任务请求量做出响应并得出其具体数量、位置等参数指标,以实现对用户使用规律的主动应对。使用CloudSim进行仿真实验,系统地评判模型的性能。实验结果表明,AF-HW模型在应对可预测的海量并发任务请求时可有效地提升单点及整体的响应速率,使用户得到更好的体验。

关键词 云计算,预测,资源部署,Holt-Winters模型,整体最优

中图分类号 TP302.7 文献标识码 A DOI 10.11896/j.issn.1002-137X.2015.9.026

Research on Resource Deployment Model Based on Active Prediction in Cloud Computing

MA Zi-tang CHEN Peng LI Zhao-xing

(The Third Institute, PLA Information Engineering University, Zhengzhou 450004, China)

Abstract With the growing popularity of cloud computing, more and more users choose to migrate their business to the cloud computing system. Users' usage habits and social routine working laws swarm into the cloud computing system along with the influx of large numbers of users, such as applying intensively to cloud computer system for resource nodes as early as 8:00, which leads into a kind of predictable resources conflict. In view of the problems above, a resource deployment model based on active prediction was proposed. Firstly, the task request amounts of next cycle are predicted according to the algorithm cycle length of Holt-Winters seasonal exponential smoothing model in prediction model, and determination of whether to make response to the current task request amounts or not and the specific amount, location and other parameter indicators should be made according to the active prediction algorithm designed to achieve active response capabilities to users' usage patterns. The simulation experiment was conducted using CloudSim, and the performance of model proposed was judged systematically. Experimental results show that AF-HW model can effectively enhance the single-point and overall response rate when responding to predictable massive and concurrent task requests, so that users can get a better experience.

Keywords Cloud computing, Predictive, Resource allocation, Holt-Winters model, Systems optimize

随着云计算系统在应用层面上的不断发展,用户的认可度和参与度不断上升。可以预计在未来的某一时刻,人们已经不再拥有个人电脑,取而代之的是一种简洁的终端接入设备,设备的另一头是能力强大的云计算系统^[1]。人们利用云计算系统完成日常的办公、娱乐等活动,而不必考虑其具体的运维、升级等繁杂的工作。云计算出色的性能表现和灵活的商业服务模式使其不断走向成功。

当用户源源不断地将自身的业务迁向云计算系统时,云计算系统需要应对的不仅仅是增加的系统资源使用量,还要应对一种看不见的潜在挑战即用户的使用习惯(社会活动运行规律)^[2]。举一个例子解释这种潜在的挑战:目前社会的绝大多数企业和个人在某一个特定的时间段如早上8点左右上

班,到达工作岗位的第一件事,就是打开办公室的个人终端系统,向云计算系统申请一个合适规格的虚拟机,开始一日的工作。当用户的使用达到一定规模以后,这种使用方式就会对云计算系统的性能表现带来冲击,我们不能将希望寄托在用户改变这种使用习惯上,而应将关注的目光聚焦在如何应对此类给云计算运营商带来“幸福的烦恼”的问题上来^[3]。VMware Distributed Resource Scheduler(DRS)是一种通过获取虚拟机环境中的资源信息来计算平衡和分配容量的工具。VMware DRS可以跨资源池地不间断地监控资源利用率,并根据业务需要和不断变化的优先级的预定义规则,在多台虚拟机之间智能地分配可用资源。但是该策略只有在任务请求到达后,系统根据资源节点的使用情况来执行操作,这种迁移

到稿日期:2014-09-28 返修日期:2015-01-20 本文受武器装备预研重点基金资助项目(9140A15060311JB5201)资助。

马自堂(1962-),教授,主要研究方向为信息安全、密码系统工程;陈鹏(1988-),硕士生,主要研究方向为云计算、系统优化, E-mail: hoops@whu.edu.cn(通信作者);李兆兴(1990-),男,硕士生,主要研究方向为云计算、大数据、数据挖掘。

式的资源调度方式具有一定的滞后性。美国高通公司研制了 StreamBoost 技术^[4], 该技术主要应用于千兆位级 Wi-Fi 路由器, 用于解决多设备接入对网络性能体验造成的影响。该技术的核心思想可以归纳为系统可主动识别不同任务请求的类型并为其分级, 以便进行相应的服务, 其可在不改变原有网络带宽的情况下大幅提升用户的使用体验。同时该产品还加入了云服务功能模块, 这种服务可不断增强路由器的智能, 在新应用和设备添加到网络中时, 可以识别和优化管理这些应用和设备。

本文提出一种虚拟机预制机制来应对云计算系统所面临的大量用户具有周期特性的并发任务请求挑战。利用 Holt-Winters 季节指数平滑模型预测方法, 设计了一种可提前应对可预测的大量并发任务请求的模型, 以摆脱现有虚拟机部署单一被动应对的局面^[5]。其通过化被动为主动的方式提高了云计算系统的响应速度和用户体验。

1 相关工作及本文研究思路

1.1 被动响应策略

目前解决任务请求响应问题的策略主要集中在 3 个方面, 首先是对传统虚拟机部署技术的改进, 如虚拟机模板技术^[6], 其主要思想是在虚拟机模板中预装操作系统与常用或指定的应用软件, 并预配置了虚拟设备, 系统通过复制模板完成虚拟机部署, 但当模板文件较大时, 复制文件的时长对系统响应请求的性能影响较大; 其次是对镜像文件传输过程的优化, 镜像流传输技术^[7,8]的主要依据是在用户启动虚拟器件时的对象为操作系统、中间件、应用软件等部分, 而其在整个虚拟器件文件当中只占有一小部分, 应用镜像流技术使系统仅仅在下载部分虚拟器件时就可启动虚拟机; 最后从利用带宽效率的角度研究, 协同部署技术^[9]是将虚拟机镜像在多个目标物理机之间的网络中传输, 可在不造成网络通讯瓶颈的基础上加快部署资源节点的效率。以上 3 类技术在层次上属于递进关系, 即按照时间顺序依次进行, 但又可将其统一归纳成被动的应对任务请求模式, 因为只有任务请求到达后, 系统方执行相应的部署资源节点等操作。

1.2 主动响应策略

借鉴 StreamBoost 技术思想与成熟的商业营销管理模式^[10], 本文提出了一种主动应对任务请求的响应策略, 其可提前感知任务请求量的变化, 对其适时地做出动态的调整。在商业营销量与库存量的比例策略中, 会随着不同的时间段、消费环境做出不同的策略部署。当销售旺季来临或特殊节假日迎来短时的客流高峰时, 加大库存量同时提高导购效率, 使所有客户得到服务。淡季时, 会减少库存以优化购物环境, 增加导购的服务质量, 从而提高客户的满意度。这种差异化的服务模式是以整体的最优化为目的, 而不局限于取得局部的最优解。借鉴此种思路, 当云计算系统面临可预见的使用高峰时, 系统应具备提前感知、预制计算节点的能力。

1.3 AF-HW 模型

利用 Holt-Winters 季节指数平滑模型^[11], 本文提出了一种基于 Holt-Winters 的主动预测模型 (Active Forecast Model Based on Holt Winters, AF-HW)。其实现架构如图 1 所示。

AF-HW 模型在架构上分为两个不同的时间周期, 架构的左半部分实现对任务请求的实时响应, 是对目前主流响应

任务请求的一种概括; 右半部分为主动预测下一时间段任务请求量的框架, 遵循既定算法中规定的时间周期。首先利用 Holt-Winters Model 根据记录的历史数据预测下一时间段的任务请求量, 并通过 AF (Active Forecast) 算法 (在第 2 节介绍) 判断云计算系统是否需要预制计算节点, 当判断为是时, 通过调用系统的部署算法库来判断计算节点需要部署的数量、位置等相关细节, 并在部署完成以后将相关的地址、数量等信息推送到部署计算节点模块, 完成计算节点的预制工作。此外, 历史数据日志中还要记录当前预制计算节点对于下一时间段任务请求的匹配情况, 使 AF 算法可根据当前环境做出动态的调节。

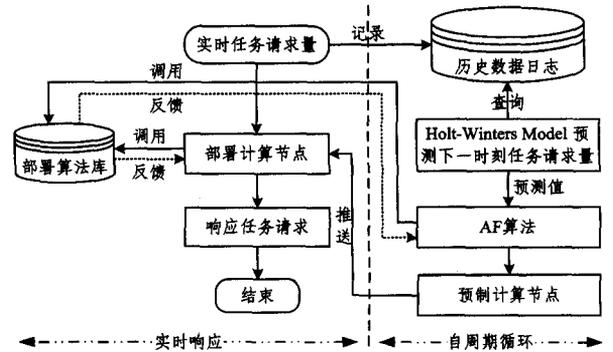


图 1 AF-HW 模型实现框架

通过引入 Holt-Winters Model 和 AF 算法使云计算系统具备主动预测任务请求的能力, 为云计算系统提前预制计算节点, 加快系统响应性能做出了量化的指标。

2 任务请求量预测算法设计

针对具有周期特征的用户使用习惯方式及其数量呈现出的分布情况, 选取 Holt-Winters 季节指数平滑模型作为对任务请求量的跟踪预测模型。因 AF-HW 模型针对周期性的、满足季节规律的短时大量需求, 并将云计算系统的发展趋势及 Gartner 公司提出的技术成熟度曲线^[12]作为参考, 可将目前云计算系统所要应对的任务请求挑战归纳为一种具有线性趋势和加法时序变化的序列。因此, 选取 Holt-Winters 加法季节指数平滑模型。下面简单给出该模型对下一周期任务请求量的具体计算方法。

首先给出模型输入变量 x_t 的平滑指数 \hat{x}_t 的表达式:

$$\hat{x}_{t+k} = m_t + n_t k + S_{t+k}, t = s, s+1, \dots, T \quad (1)$$

其中, m_t 为截距, 代表当前时刻的任务请求量; n 代表斜率, 即任务请求量的变化率; S 为系统的季节因子, 选取天 (24h) 作为一个季节周期; s 为周期的长度, 选取 $s = \frac{S}{96} = 0.25h$ 。以下

为平滑指数表达式中 3 个参数的递推公式:

$$\begin{cases} m_t = \alpha(x_t - S_{t-s}) + (1-\alpha)(m_{t-1} + n_{t-1}) \\ n_t = \beta(m_t - m_{t-1}) + (1-\beta)n_{t-1} \\ S_t = \gamma \frac{x_t}{m_t} + (1-\gamma)S_{t-s} \end{cases} \quad (2)$$

最后得出预测值:

$$\hat{x}_{T+k} = m_T + n_T k + S_{T+k-s} \quad (3)$$

其中, S_{T+k-s} 为最后一个样本数据的季节因子。

接下来给出模型的初始值计算公式。

a) 计算前两个周期的平均任务请求增量:

$$\bar{V}_1 = \frac{1}{S} \sum_{i=1}^S x_i, \bar{V}_2 = \frac{1}{S} \sum_{i=S+1}^{2S} x_i \quad (4)$$

b) 计算初始截距值与初始斜率:

$$n_{2S+1} = \frac{1}{S} (\bar{V}_2 - \bar{V}_1), m_{2S+1} = \bar{V}_2 + \frac{S-1}{2} n_{2S+1} \quad (5)$$

c) 计算前两个周期的季节因子:

$$S_{1t} = x_t - \bar{V}_1 + \left(\frac{S+1}{2} - k\right) n_{2S+1} \quad (6)$$

$$S_{2t} = x_{t+S} - \bar{V}_2 + \left(\frac{S+1}{2} - k\right) n_{2S+1} \quad (7)$$

d) 计算前两个周期的平均季节因子:

$$S_t'' = \frac{1}{2} (S_{1t} + S_{2t}) \quad (8)$$

e) 将季节因子做正态化处理:

$$S_t' = \sum_{i=1}^S S_t'', S_t = \frac{S}{S'} S_t' \quad (9)$$

对未来任务请求量的定量化描述是下一步设计预制计算节点的必要前提。Holt-Winters 加法季节指数平滑模型可有效预测云计算系统任务请求量的规律性变化,并可减小突发任务请求量(噪音)的干扰,对任务请求量的跟随性强,为云计算系统应对大量并发的、满足季节性规律的任务请求量提供了参数依据,为提高云计算系统响应性能打下了基础。

3 AF 算法设计

针对代入 Holt-Winters Model 所得到的任务请求预测值,AF 算法应满足以下需求:

- 1) 判断当前预测值是否达到系统提前预制计算节点的要求;
- 2) 当系统预制计算节点后评估其预制性能及改进策略。具体算法如表 1 所列。

表 1 Active Forecast algorithm

Input
f_x , the forecast value in x time quantum.
t, forecasting cycle.
n_t , the number of resource allocation under one forecasting cycle.
n_t', n_t within SLA.
N_t , actual usages in t time quantum.
The algorithm of Active Forecasting Model
1. $\alpha \leftarrow$ users defined.
2. $T_0 \leftarrow$ initial moments.
3. $f_{x+1}, f_{x+2}, \dots, f_{x+\alpha-1} \leftarrow$ assume that we have known the values at T_0 .
4. IF ($\frac{f_{x+\alpha}}{n_t} \geq \alpha$) THEN//initial value of α is five.
5. {set $\sum_{n=1}^{\alpha} f_{x+n} = n_t, \alpha++$,
6. END IF }
7. GoTo 16;
8. IF ($\alpha n_t' - \sum_{n=1}^{\alpha-1} f_{x+n} \leq f_{x+\alpha} \leq \alpha n_t$)
9. set $f_{x+1}, f_{x+2}, \dots, f_{x+\alpha-1}$ to n_t' ,
set $f_{x+\alpha} = f_{x+\alpha} - (\alpha-1)n_t' // f_x'$; the modify value about f_x ;
10. END IF;
11. ELSE
12. set $f_{x+\alpha}$ to n_t' ,
13. FOR ($f_{x+n} = f_{x+n} + f_{x+\alpha} - n_t', (n=\alpha-1); f_{x+n} > n_t'; n--$) $f_{x+n} = n_t'$.
14. $\alpha--$;
15. END ELSE
16. END

从表 1 可以看出,算法通过提前判断下 α 个时间段的任务请求量来判断系统是否需要提前预制计算节点,通过具体的比较算法进一步确定部署的时间段及数量。系统可通过 α 取值的自调节性来实现对系统任务请求量波峰波谷的自适应调节,使系统可充分应对高峰时期的任务请求,又可以在系统使用量减少时缩短预制周期,以减少系统能耗。

4 性能测试及分析

4.1 实验设计与评价指标

为了验证 AF-HW 模型的性能表现,采用云计算仿真软件 CloudSim3.0^[13] 对其进行模拟运行实验。通过对 VmAllocationPolicy、DatacenterBroker 类的修改,分别添加 AF 算法、计算节点部署算法,以实现模型的功能。模拟实验参数设置如下。

(1) 任务请求量数据采集:为了更逼真地模拟云计算系统日常的使用情况,本文任务请求量数据采用 BaiDu 统计 3.0 版^[14] 中对百度推广等系列相关网站的实时响应数据(见图 2) 为单日每小时系统的页面浏览量(Page View, PV)和平均访问时长。为了突出 AF-HW 模型解决问题的主要矛盾,实验中假设每一个任务请求在相同时长的任务指令长度(MI)、单个虚拟机的执行速度(MIPS)相同。这样,可将系统的 PV、平均访问时长等同于系统的任务请求量个数与每个任务的指令长度。为了与预测算法的周期因子同步,需要对采集的数据进行预处理,将 1h 时间长度的监测值切分为 0.25h 以获得一组监测数据。如 B 点的值为 $V_b = \frac{x_b}{60} (V_{b前} + V_{b后})$, 其中 x_b 为 B 点钟值, $V_{b前}, V_{b后}$ 为 B 点相邻整数点的值。图 2 中 Forecast Value 为应用 Holt-Winters 季节指数预测模型得出的每 0.25h 系统的任务请求输入量,可以看出,在拟合 2013-10-05 单日请求量时,模型可有效追踪未来请求量的变化,较好地反映了发展趋势。

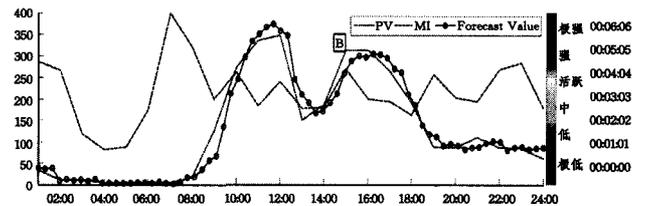


图 2 2013-10-05 单日负载量及预测值

(2) 任务属性信息:任务指令长度[8000, 12000]MI, 虚拟机执行速度[250, 400]MIPs, 虚拟机性能利用率为 50%, 存储量(int ram)[1024, 2048]MB, 带宽(bw)[1000]。0.25h 内系统可有效部署节点数 n_t [300, 600], 满足 SLA 节点数 $n_t' = 0.65n_t = [195, 325]$ 。

实验中,将 AF-HW 模型分别采用 Greedy 分配策略^[15] 与 Ordinal 分配策略^[16], 并与单纯采用上述两种算法的模型进行性能比较。系统利用 2013-09-06 至 2013-10-04 的数据预测 2013-10-05 每 0.25h 的值,并输入 2013-10-05 的实际值作为输入量。为比较模型在不同环境、压力下的性能表现,实验设计了不同的虚拟机执行速度与单位时间内有效部署计算节点数的组合方式,建立如式(10)所示的一个类似 Key/Value 模式的矩阵:

$$C=[N \ V]=\begin{bmatrix} n_{i1} & v_{11} & v_{12} & \cdots & v_{1j} \\ n_{i2} & v_{21} & v_{22} & \cdots & v_{2j} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ n_{im} & v_{m1} & v_{m2} & \cdots & v_{mj} \end{bmatrix} \times 50\% \quad (10)$$

其中,矩阵中的每一行代表一种实验环境, n_{i1} 表示单位时长可有效部署计算节点的数量, v_{ij} 代表一个节点的计算速度。

其性能评价指标如下:

(1)Allocationspan,反映单节点对请求的响应性能,是衡量是否满足与用户签订的SLA(service level agreement)的重要指标^[17]。其值取决于各节点完成请求的最大时间。计算公式如下:

$$Allocationspan=\max \{CT_Vm_i\} \quad (11)$$

其中, CT_Vm_i 表示各节点完成部署的时长。

(2)Finishedtime,该指标反映系统整体的性能与灵敏度,表示系统对当前时间段最后一个请求做出响应的时刻,是对模型整体性能表现的度量。计算公式如下:

$$Finishedtime=RT_Vm_i_last \quad (12)$$

两种评价指标分别从用户的QoS体验与系统整体的性能表现出发,在说明二者重要性的同时,从两个不同维度评价模型在应对大量并发任务请求时的性能表现,提供判断AF-HW模型的量化指标。

4.2 AF-HW模型性能评估实验

为了验证模型提出的预测模式与AF算法的设计在实际应用中的表现,实验中设计了5种不同强度的环境来模拟云计算系统所处的真实环境。由于虚拟机执行速度、单位时间可有效部署计算节点数这两个参数对系统性能体验有相互促进的效果,因此实验参数设置如表2所列。

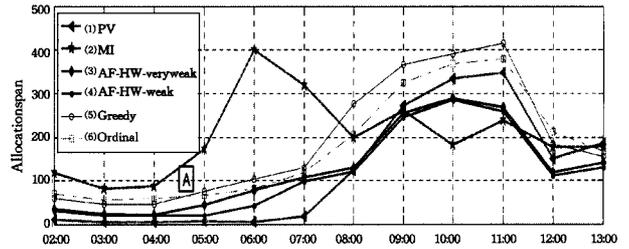
表2 5种不同负载环境下的参数组合

No.	n_{ij}	v_{ij}	intensity
1	300	[250~280]	极低
2	300	[280~330]	低
3	400	[330~350]	中
4	500	[350~400]	强
5	600	[350~400]	极强

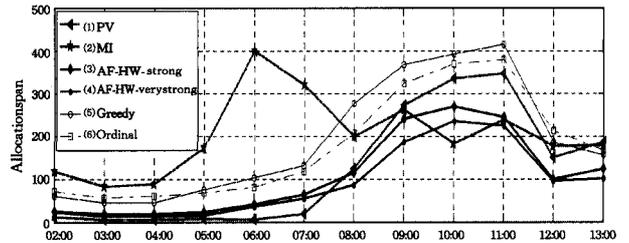
在前两种参数设置中,固定 n_{i1} 值,观察 v_{ij} 值对系统的影响;在后两种参数设置里,固定 v_{ij} 值,改变 n_{i1} 的取值,以便观察 n_{i1} 对系统响应性能带来的影响。而第三种参数设置作为基准,方便两组数据的比较。下面分别对两种评价指标进行实验。

(1)图3(a)、(b)反映了AF-HW模型在4组(1、2、4、5)不同环境强度下的两两Allocationspan指标对比结果。从图3可以观察到,不论在何种强度的环境下,AF-HW模型均要优于不添加预制功能的Greedy、Ordinal分配策略,说明模型可在不同工作负荷下优化Allocationspan指标,提高用户的QoS体验。图中A点为Greedy策略与Ordinal策略性能表现的拐点,原因是随着PV、MI值的增大,系统底层资源使用量波动情况加剧,Greedy算法的预测结果出现失真等情况,不能很好地跟随资源的变化情况,造成性能表现变差。从图3(a)可以观察到,在06:00时刻,AF-HW模型在极低和低强度环境下的性能差距达到最大值,此时云计算系统输入指标MI达到极大值,而在PV值达到极大值或PV的斜率达到最大的时候(09:00、11:00时刻),两曲线差的绝对值并没有明显放大,该特征说明系统提升 v_{ij} 指标在应对MI指标时有良好的作用。

好的作用。在图3(b)中,曲线(3)、(4)的绝对差值出现在09:00:00时刻,为PV斜率的极大值点,而在06:00时刻,绝对差值没有明显放大,说明系统 n_{i1} 指标对PV值敏感,提升系统单位时间有效部署节点的效率可应对访问量激增的问题。



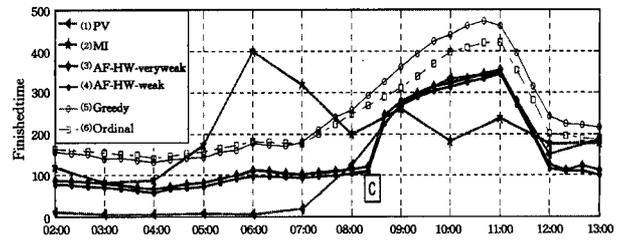
(a)在低负载和极低负载强度下的AF-HW曲线



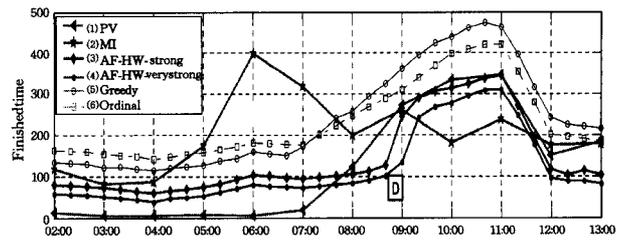
(b)在高负载和极高负载强度下的AF-HW曲线

图3 不同负载环境下的Allocationspan指标

(2)由于无法获取系统实时的PV指标,同时也为了简化实验复杂度,在对Finishedtime指标进行测评时,认为任务请求在单位时长内由初始时刻集中释放。实验结果如图4所示。



(a)在低负载和极低负载强度下的AF-HW曲线



(b)在高负载和极高负载强度下的AF-HW曲线

图4 不同负载环境下的Finishedtime

首先观察Greedy与Ordinal策略曲线在07:00时刻之前,Finishedtime指标波动的范围很小,在这之后,开始随着PV值的增大而迅速增加,且Greedy策略的变化率要大于Ordinal,这说明Greedy策略处于强负载环境中时,仍追求局部的最优解而忽略了整体的利益,导致其在两种指标的评判下均劣于Ordinal策略,这同时也从侧面反映出当系统底层资源环境处于剧烈变化时,部署策略倾向于随机有利于提升整体的性能表现。在图4(a)中观察到,曲线(3)(4)的拟合程度很高,说明 v_{ij} 的变化对Finishedtime指标的影响很小。在图中C点处曲线(3)(4)出现阶跃式提升,其变化时间节点滞

后于 Greedy、Ordinal 策略,通过分析可以认为是由于曲线(3)(4)带有预制计算节点的功能,其对负载的波动有一定的预判、吸收作用,因此当 PV 指标的斜率开始攀升时,策略(3)(4)仍可保持较好的适应能力,Finishedtime 指标几乎没有受到干扰。但 C 点之后,由于绝对数量的增大,负载量开始逼近系统所能承受的范围,导致预制量没有紧紧跟随任务量的增长,系统整体的响应时间被拉长。在图 4(b)中,曲线(3)(4)的绝对差值变大且拐点 D 出现的位置发生后移,这是由于增大系统单位时间有效部署计算节点的能力,可在应对绝对任务数的增大问题方面有积极的作用,可有效提升系统应对大任务量环境的能力。综合分析图 3、图 4 可以得出结论:AF-HW 模型对 Allocationspan、Finishedtime 指标的贡献均优于优化前的系统,该模型可有效地提升系统单点与整体的性能表现。

结束语 为了使云计算系统能有效地应对符合一定规律的海量并发任务请求给系统带来的冲击,在给用户提供良好 QoS 体验的同时满足与用户签订的 SLA,实现云计算出色的性能表现,本文提出一种基于 Holt-Winters 季节指数平滑模型预测的计算节点部署策略模型。模型通过加载系统历史请求数据预测下个单位时间段的任务请求量分布情况,并根据设计的 AF 算法判断系统是否应为本时段的任务请求数据量做出响应,及相应的部署数量、位置。AF-HW 模型还可自适应地调节系统实时响应请求的情况,充分利用冗余预制的计算节点,并可有效地根据请求量的分布状态自扩大或缩减预制周期,达到对请求变化的动态跟踪及降低系统能耗的目的。为了验证该模型的性能,设计了性能评估实验,针对模型的两个主要模块分别进行了实验。证明了 Holt-Winters 加法季节指数平滑模型在跟踪系统单日的任务请求量时拟合程度较高,说明其预测的准确及有效性;在对单个计算节点、系统整体响应性能指标评判时,设计了 5 种不同的环境强度进行模拟,证明在实际应用中相对于不添加预制功能的模型,该模型在响应性能方面有明显的优势,较适合类似于云计算等大规模的动态环境。

参 考 文 献

- [1] Ioannis A M, Helen D K. Evaluation of gang scheduling performance and cost in a cloud computing system [J]. *Journal of Supercomputing*, 2012, 59(2): 975-992
- [2] Anand R, Jeffrey D U. Mining of massive datasets[M]. London: Cambridge University Press, 2011
- [3] Jerri L, Joe T, Mary E T. Google Analytics, 3rd Edition[M]. New Jersey: Wiley, 2009
- [4] Qualcomm. The Wi-Fi evolution-an integral part of the wireless landscape[R]. USA: Qualcomm Incorporated, 2013
- [5] Shiraz M, Abolfazli S, Sanaei Z, et al. A study on virtual machine deployment for application outsourcing in mobile cloud computing[J]. *The Journal of Supercomputing*, 2013, 63(3): 946-964
- [6] 陈彬. 分布环境下虚拟机按需部署关键技术研究[D]. 长沙:国防科学技术大学, 2010
Chen Bin. Research on key technologies of on-demand deployment of virtual machines in distributed environments [D]. Changsha: National University of Defense Technology, 2010
- [7] Samer A K, Dinesh S, Prasenjit S, et al. VMFlock: Virtual Machine Co-Migration for the Cloud[C]//Proceedings of the 20th international symposium on High performance distributed computing (HPDC'11). New York: ACM, 2011: 159-170
- [8] 罗晶. 嵌入式虚拟机系统镜像存储的研究[D]. 武汉:华中科技大学, 2012
Luo Jing. Research on system images storage in embedded virtualization system[D]. Wuhan: Huazhong University of Science and Technology, 2012
- [9] Jonathan R C. A Distributed and Collaborative Dynamic Load Balancer for Virtual Machine[C]//Euro-Par 2010 Parallel Processing Workshops. Berlin Heidelberg: Springer, 2010: 641-648
- [10] Mankiw N G. Principles of Economics 6th Edition [M]. USA: Cengage Learning, 2011
- [11] Sarah G, Roland F, Christophe C. Robust Forecasting with Exponential and Holt-Winters Smoothing[R]. Belgium: Katholieke University Leuven, 2007
- [12] Hung L H. Hype Cycle for the Internet of Things[R]. USA: Gartner, Inc., 2013
- [13] Rodrigo N C, Rajiv R, Anton B, et al. CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms [R]. Cloud Computing and Distributed Systems Laboratory, Australia, 2010
- [14] Baidu Inc [OL]. 2013-10-05. <http://tongji.baidu.com/web/5473605/overview/~mult>
- [15] Tai J Z, Waleed M, Zhang J M, et al. ARA: Adaptive Resource Allocation for Cloud Computing Environments under Bursty Workloads [R]. Boston: Northeastern University, 2011
- [16] Ning F M, Giuliano C, Ludmila C, et al. Burstiness in multi-tier applications: symptoms, causes, and new models[R]. Williamsburg: College of William and Mary, 2008
- [17] 王健宗, 谌炎俊, 谢长生. 面向云存储的 I/O 资源效用优化调度算法研究[J]. *计算机研究与发展*, 2013, 50(8): 1657-1666
Wang J Z, Chen Y J, Xie C S. Research on I/O resource scheduling algorithms for utility optimization towards cloud storage [J]. *Journal of Computer Research and Development*, 2013, 50(8): 1657-1666
- [31] Jamroga W, Bulling N. A logic for reasoning about rational agents[M]//Computational Logic in Multi-Agent Systems-CLIMA. Berlin: Springer, 2007: 42-61
- [32] Bulling N, Jamroga W. Rational play and rational beliefs under uncertainty[C]//8th International Joint Conference on Autonomous Agents & Multiagent Systems/Agent Theories, Architectures, and Languages. 2009: 257-264
- [33] Andreas H, Emiliano L, Dirk W. Logic, Reasoning about Actions Meets Strategic Logics [M]//Logic, Rationality, and Interaction, 4th International Workshop (LORI 2013). Hangzhou, China: Springer Verlag, Tiergartenstrasse 17, Heidelberg, D-69121, Germany, 2013: 162-175
- [34] Nils B, Wojciech J. Comparing variants of strategic ability: how uncertainty and memory influence general properties of games [J]. *Autonomous Agents and Multi-Agent Systems*, 2014, 28(3): 474-518

(上接第 126 页)