

集群服务器响应延时预测及其负载调度控制

Response Delay Predictive and Load Scheduling Control for the Cluster Server

于国防 王耀才 庄立运 贾栋清 (中国矿业大学 信电学院 江苏徐州 221008)

摘要:单纯以负载均衡为目标的 Web 服务器集群调度不仅增加了算法的复杂度,在集群超载的情况下也无法保证服务的质量,而以响应延时为基础的均衡算法则可确保更好的服务质量。综合运用多级模糊系统与人工神经网络,既基于服务器系统的负载状况,又兼顾 Web 请求文件的类型与大小,对响应延时进行模糊预测,并在此基础上提出了一种基于响应延时阈值的集群服务器负载控制算法。构建仿真环境对所提出的参数预测及算法的有效性进行了验证。

关键词:集群服务器 响应延时预测 多级模糊系统 人工神经网络 负载调度

1 引言

目前,Web 服务器集群的负载调度目标主要基于各 Web 服务器的负载均衡^[1,2]和最小响应延时^[3]两类。针对后者,响应延时的估计或预测成为调度算法实现的前提,但是,由于集群服务器配置的异构性及其负载指标和请求文件类型的多样性,所以不易设计对应的系统模型或采用具体的数学表达式来获取该延时参数的预测值。本文采用模糊控制与神经网络技术研究多类型、变文件大小请求的响应延时预测,并将其用于基于响应延时阈值的集群服务器控制算法中。

2 响应延时的影响因素及其模糊集划分

2.1 负载指标及其模糊集划分

有多种信息可以用来作为系统的负载指标,如,CPU 利用率、内存利用率、硬盘传输量以及网络流量^[4],而近年来 CPU 的速度增长迅速,使得 CPU 和内存通信之间的瓶颈较为突出,内存空间的不足可能导致频繁的页面交换,这时访问延迟将会成千倍地增加,所以本文以 CPU 利用率、内存利用率以及网络流量作为系统的综合负载指标。

分别以 c 、 m 和 n 表示 CPU 利用率、内存利用率以及网络流量,创建三个负载指标模糊集,每个均匀划分为 5 个模糊区间,记为 $A_i^x, i=1, \dots, 5; x=c/m/n$ 。

2.2 请求文件及其模糊集划分

基于因特网的请求文件类型有很多,不失一般性,在此主要考虑 ASP、PHP 及 ASP.NET 三类动态网页,而将静态网页(html/html)的服务响应单独考虑。由于动态网页所实现功能的复杂性,而不易根据其文件大小或源代码分析对响应延时进行比较准确的量化判断,但考虑到它们也要经过磁盘文件读取和 Web 服务器的文件解析,所以,可将它们先以静态网页(大小以其文件本身及要随其下载的音频、视频、图片和 Flash 流等文件的总和计算。其下载资源可通过包含它们的网页源文件主体或特别设置的文件头解析获得)按文件大小进行响应延时的定量预测,然后再附加上按动态网页分类的定性响应延时预测。

在此以 f 表示上述三类文件类型,创建一个包含 3 个模糊区间的单值输入模糊集(图略),记为 $A_i^f, i=1, 2, 3$ 。而以 s 表示网页文件大小,为了能在一定程度上减少规则数,并能随着文件的由小到大变化而采用由细到粗的控制灵敏度,在此对文件大小采用单向伸展式三角形模糊集^[5]划分,如图 1 所示。

3 响应时间的模糊预测与在线修正

由上述的模糊集划分可知,如果利用传统的模糊系统? 则规则数为 $5 \times 5 \times 5 \times 8 \times 3 = 3000$,而多级模糊系统^[6,7]具有良好的性能,其规则数仅与变量数线性相关。本文在注重减少规则数的同时,也兼顾系统

设计和规则规划的难易度。首先设计一个对应输入 c 、 m 、 n 和 s 的两级模糊, 而将对应 f 的输入单独设计为一个位居第 2 级的单输入模糊系统, 最后将两个 2 级系统的输出取代数和, 如图 2 所示。

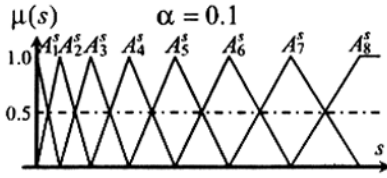


图 1 文件大小模糊集

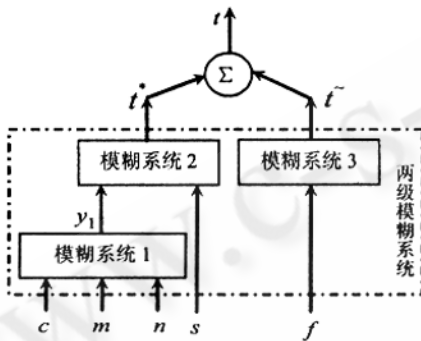


图 2 两级模糊系统框图

以下设计两级模糊系统。

首先将 y_1 分为 5 个模糊集 (B_1, B_2, B_3, B_4, B_5), 类似于图 1 (此略), 它既是模糊系统 1 的输出, 又是模糊系统 2 的输入。则此系统的规则数为 $5 \times 5 \times 5 + 5 \times 8 + 3 = 168$, 可见, 规则数减少了 94%, 且适当增加文件大小模糊集, 也不至于规则数量骤增; 再将 y_2 分为 40 个 (5×8) 单值模糊集 (T_1, T_2, \dots, T_{40}), 如图 3 所示, 表示根据系统当前负载状况所预测的响应时间 (t_j)。

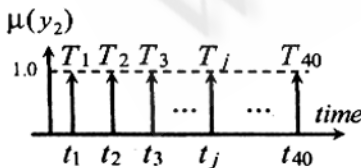


图 3 输出 y_2 的单值模糊集

同理, 将 y_3 分为 3 个单值模糊集 (T_1, T_2, T_3) (类似图 3, 此略)。

3.1 响应时间的模糊预测

根据以上构建的多级模糊系统, 设计以下模糊规则:

(1) 模糊系统 1

$$R_1^1: \text{IF } (c = A_1^c) \text{ AND } (m = A_1^m) \text{ AND } (n = A_1^n) \text{ THEN } (y_1 = B_1^1)$$

...

$$R_1^2: \text{IF } (c = A_2^c) \text{ AND } (m = A_2^m) \text{ AND } (n = A_2^n) \text{ THEN } (y_1 = B_2^1)$$

...

$$R_{125}^1: \text{IF } (c = A_5^c) \text{ AND } (m = A_5^m) \text{ AND } (n = A_5^n) \text{ THEN}$$

($y_1 = B_{125}^1$) (注: 规则中的 B_{125}^1 为对应 y_1 模糊集中 B_1, B_2, B_3, B_4 或 B_5 的一个输入)

(2) 模糊系统 2

$$R_2^1: \text{IF } (s = A_1^s) \text{ AND } (y_1 = B_1) \text{ THEN } (y_2 = T_1)$$

...

$$R_2^2: \text{IF } (s = A_2^s) \text{ AND } (y_1 = B_2) \text{ THEN } (y_2 = T_2)$$

...

$$R_{40}^2: \text{IF } (s = A_8^s) \text{ AND } (y_1 = B_5) \text{ THEN } (y_2 = T_{40})$$

采用中心平均解模糊, 并考虑到对所有输入的单值模糊隶属度之合为 1, 则, 系统的解模糊输出为

$$t^* = \frac{\sum_{i=1}^{40} t_i^* \mu_{R_i^2}(s, y_1)}{\sum_{i=1}^{40} \mu_{R_i^2}(s, y_1)} = \sum_{i=1}^{40} t_i^* \mu_{R_i^2}(s, y_1) \quad (1)$$

根据乘积推理, 式 (1) 中

$$\mu_{R_i^2} = \mu_1(s) * \mu_1(y_1) \quad (2)$$

(3) 模糊系统 3

$$R_3^1: \text{IF } (s = A_1^s) \text{ THEN } (y_3 = T_1)$$

$$R_3^2: \text{IF } (s = A_2^s) \text{ THEN } (y_3 = T_2)$$

$$R_3^3: \text{IF } (s = A_3^s) \text{ THEN } (y_3 = T_3)$$

同理可得模糊系统 3 的解模糊输出

$$t^- = \sum_{i=1}^3 t_i^- \mu_{R_i^3}(3)$$

3.2 预测响应时间的在线修正

由上文设计可知, 响应时间还是一个尚未确定的参数, 很难用一个具体的数学模型来表示之, 文献^[3]利用模糊控制和人工神经网络^[8,9], 不但较好地解决了该参数的预测问题, 而且也增强了系统的自适应能力。在此分别运用两个神经网络对多级模糊系统预测的响应时间进行学习修正。

(1) 基于文件大小的响应时间预测

首先,根据文件的大小,运用一个神经网络对该级模糊系统的响应时间预测值进行学习修正,如图 4 所示。

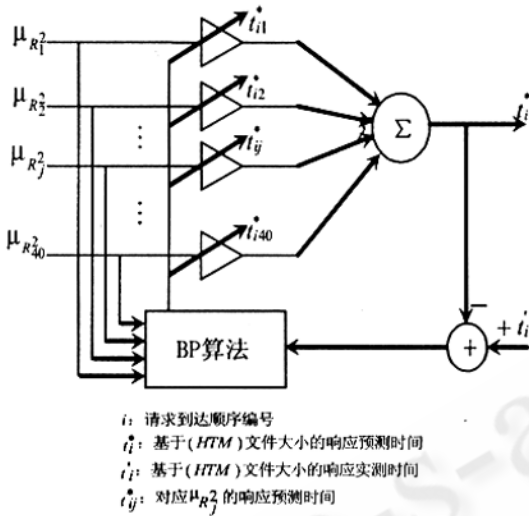


图 4 基于文件大小的响应时间预测

该部分的预测响应总时间为

$$t_i^* = \sum_{j=1}^{40} t_{ij}^* \mu_{R_j}^2 \quad (4)$$

采用误差反向传播 (BP) 算法修正预测时间,如式 (5),其中, $\eta(1+p)$ 为对应误差率 p 的学习步长,目的是根据误差状况动态调节学习速率,在此设 $\eta = p = 0.1$ 。

$$t_{(i+1)}^* = \begin{cases} t_{ij}^* + \eta(1+p)\mu_{R_j}^2(t_i^* - t_i'), & | \frac{t_i^*}{t_i'} - 1 | > p \\ t_{ij}^*, & \text{其它} \end{cases}$$

(5)

(2) 基于文件类型的响应时间预测

其次,根据文件的类型,运用另一个神经网络对该级模糊系统的响应时间预测值进行学习修正(类似图 4,此略)。该部分的预测响应总时间为

$$t_i^- = \sum_{j=1}^3 t_{ij}^- \mu_{R_j}^- \quad (6)$$

同样,采用误差反向传播 (BP) 算法,如式 (7),其中, η 和 p 的同上文。

$$t_{(i+1)}^- = \begin{cases} t_{ij}^- + \eta(1+p)\mu_{R_j}^-(t_i^- + t_i^+ - t_i), & | \frac{t_i^-}{t_i - t_i^+} - 1 | > p \\ t_{ij}^-, & \text{其它} \end{cases} \quad (7)$$

4 实验及分析

为了仿真检验非平衡负载下的负载调度,简化实验环境,在此通过 3 台客户机,直接向由 3 台服务器构成的服务器集群模拟并发随机请求。其中,三台 Web 服务器的软件配置分别为: S1、S2: Windows2000 Server + IIS 5.1, S3: Windows2000 Server + Apache 2.0。网络连接使用 10/100M 自适应交换机。

实验前先设计如表 1 所示的测试文件, DB 为包含 1 万条记录的 SQL Server 数据库的一个 3 字段数据表。为了便于请求发送控制,在每个网页中编写相应的请求控制脚本,但是,各客户机的请求发送不同:

- (1) #1 Browser 随机选取 htmN 文件 ($N=1,2,\dots,5$,下同) 轮询请求 #1 Server、#2 Server、#3 Server;
- (2) #2 Browser 随机选取 aspN 和 aspxN 文件按 40%、60% 轮询请求 #1 Server、#2 Server;
- (3) #3 Browser 随机选取 phpN 文件请求 #3 Server。

表 1 测试文件

类型	文件名/文件设置				
HTM	htm1	htm2	htm3	htm4	htm5
	50k	100k	200k	500k	2M
ASP	asp1	asp2	asp3	asp4	asp5
	50k	100k+DB	200k	500k+DB	2M
.NET	aspx1	aspx2	aspx3	aspx4	aspx5
	50k	100k+DB	200k	500k+DB	2M
PHP	php1	php2	php3	php4	php5
	50k	100k+DB	200k	500k+DB	2M

各服务器均配置负载调度器,调度器间相互配合、协同工作。采用基于响应延时阈值的控制算法,即,只有当本服务器预测该请求的响应延时阈值大于响应延时阈值时,才将其转发给负载相对最轻的服务器。在此设定响应延时阈值为 500ms(实际应用可按照设定规则动态调整)。仿真过程中分时段逐级增加客户请求量,最后统计整理服务器日志,得到客户的平均响应延时和服务器负载分布统计(限于篇幅,在此仅给出 CPU 使用率统计),分别如图 5、图 6 所示。

尽管仿真时发送给各服务器的请求量很不平衡,从图中可以看出,经调度转发后的负载也不尽均衡,但集群的最终响应却能使客户都能基本得到比较满意的

服务质量,即实际响应延时被基本控制在响应延时阈值内(从图 5 可知,集群的响应延时控制误差约为 $50/500=10\%$),达到了以服务质量为目标的集群调度控制。

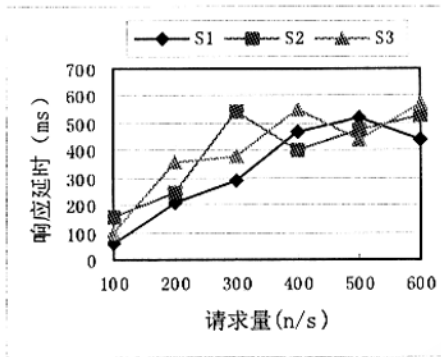


图 5 客户平均响应延时

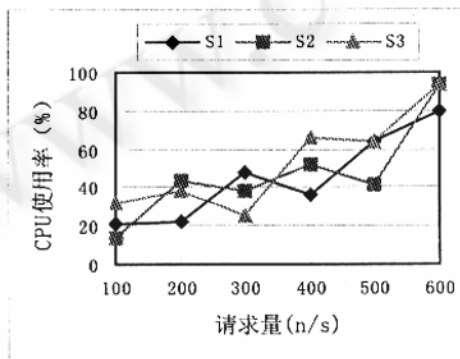


图 6 服务器负载分布

5 结语

与基于服务器负载轻重的调度策略相比,基于响应延时预测的服务器集群负载调度具有更明确的服务质量控制目标,更能满足实际应用的需要。但是,毋庸置疑,其预测的影响因素相对更多,复杂性也更高。所以本文考虑,如果网站建设中统一规划,并将文件预测的响应时间^[10]记录于网页头信息中,那么,实际应用中,在服务器负载信息的基础上,就可以将该网页头信息作为响应延时权重参数,对响应延时加以预测,这是我们下一步拟将进行的研究。

参考文献

- 1 Chin Wen Cheong, Ramachandran V. Genetic based Web cluster dynamic load balancing in fuzzy environment. The Fourth International Conference/Exhibition on Volume 2, 14 - 17 May 2000: 714 - 719.
- 2 Cardellini V, Colajanni M, Yu P. S. Dynamic load balancing on Web - server systems. [J]. IEEE Internet Computing, May - June 1999, 3(3): 28 - 39.
- 3 Borzemski L, Zlatwornicki K: A fuzzy adaptive request distribution algorithm for cluster - based Web systems [J]. Parallel, Distributed and Network - Based Processing, 2003. Proceedings. Eleventh Euro micro Conference on 5 - 7, 2003. 02: 119 - 126.
- 4 郑洪源、周良、吴家祺, WEB 服务器集群系统中负载均衡的设计与实现 [J], 南京航空航天大学学报, 2006. 38(3): 349 - 350.
- 5 于国防、王耀才、张宁, 一种伸缩式三角形模糊集的研究 [J], 计算机工程 2007. 04.
- 6 王立新、王迎军, 模糊系统与模糊控制 [M], 北京: 清华大学出版, 2003: 227 - 229.
- 7 Waratt Rattasiri, Saman K Halgamuge. Computationally Advantageous and Stable Hierarchical Fuzzy Systems for Active Suspension [J]. IEEE Trans Industrial Electronics, 2003, 50(1): 48 - 61.
- 8 Rong - Jing Wai, Faa - Jeng Lin, Rou - Yong Duan. Robust Fuzzy Neural Network Control for Linear Ceramic Motor Drive via Backstopping Design Technology [J]. IEEE Fuzzy System, 2002, 10(1): 102 - 112.
- 9 K M Bossley, M Brown, C J Harris. Neuro - fuzzy model construction for the modeling of non - linear processes [C]. Proc. 3rd European Control Conference, Roma, Italy, 1995: 2438 - 2443.
- 10 于国防、王耀才, 基于内容自适应的 Web QoS 提升技术研究 [J], 计算机工程与应用, 2006. 42(12): 156 - 158.