

基于大规模网络信息发送的任务调度策略研究

董海量^{1,2}, 胡华平^{1,2}, 刘建锋^{1,2}, 虞欣平²

(1 国防科技大学 计算机学院, 湖南 长沙 410073; 2 61070 部队, 福建 福州 350003)

摘 要: 针对电子邮件、即时通讯等网络通讯工具的流行, 在研究大规模网络信息发送系统组成及工作流程的基础上, 提出了一种基于响应时间的任务调度策略(STSRT), 并对其进行了仿真验证. 仿真结果表明: 文中提出的任务调度策略具有较好的性能.

关键词: 信息群发; 响应时间; 任务调度

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2010)08-0165-03

The Study of Strategy for Task Scheduling Based on Large-Scale Message Group Sending

DONG Hai-liang^{1,2}, HU Hua-ping^{1,2}, LIU Jian-feng^{1,2}, YU Xin-ping²

(1 School of Computer Science, National University of Defense Technology, Changsha 410073, China;
2 61070 Army, Fuzhou 350003, China)

Abstract: Many network communication tools such as E-mail, Instant Messenger (IM) and so on, are very popular today. In this paper, we put forward a strategy for task scheduling based on responding time (STSRT) after the constitution and workflow of a large-scale message group sending system (LSMGSS) are discussed. The experiment shows that this strategy can raise the quality of the cluster using by LSMGSS.

Key words: message group sending; responding time; task scheduling

1 引言

大规模网络信息群发系统定义为可在较短时间内(比如数分钟或者数小时)内向十万数量级以上的用户群体完成信息发送任务^[1-2]. 为此, 系统的发送服务器一般采用集群服务器. 这是由于集群服务器具有性价比高、工作稳定高效等特点^[3]. 集群服务器技术典型的见于 Web 服务, 在这类应用中, 用户请求非常频繁, 因此对于用户任务的调度一般采用负载均衡算法将用户请求合理地调度到集群中当前负载较轻的服务器上. 负载均衡算法^[4-6]的好坏对集群服务器性能有重大影响. 文中研究的大规模网络信息发送系统中, 用户任务少, 单个任务执行时间长, 与一般的 Web 服务器集群具有很大的不同, 因

此目前比较成熟的负载均衡算法均不适合处理这类需求. 文中针对任务的特点研究合适的任务调度技术, 目标是使发送任务在最短时间内完成.

2 大规模网络信息发送系统组成及工作流程

单个用户日常使用网络通讯工具的过程可概括为注册(或者登录)网络通讯工具、选择要发送的对象、编辑要发送的信息、确定并发送的过程. 这个过程可由计算机自动生成相关网络报文的方式进行模拟. 大规模网络信息发送系统的原理就是使用计算机根据发送源帐号和目标帐号以及发送信息批量生成发送报文, 模拟很多人同时向多个目标用户发送信息的情景.

根据此原理,文中设计了大规模网络信息发送系统结构,如图 1 所示.

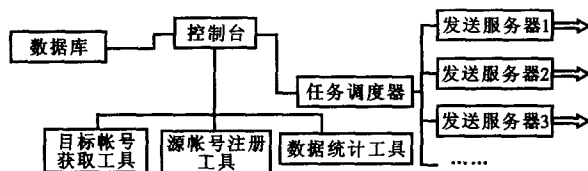


图 1 大规模网络信息发送系统组成

系统由四个部分组成:控制台,发送服务器集群,数据库和辅助工具.控制台提供用户控制界面,让用户能够操作整个系统;发送服务器集群包括任务调度器和一组发送服务器,任务调度器接收到控制台发出的任务后根据任务调度策略向发送服务器分配任务;数据库存储目标帐号和待发送信息等;辅助工具包括数据分析工具、源帐号注册工具和目标帐号获取工具等.对于手机短信这种典型的基于移动通信网的应用,发送服务器集群还将连接一个短信 modem 池.

该系统典型的工作流程是:(1)使用源帐号注册器申请大量合法网络通讯工具帐号,使用目标帐号获取工具提取大量发送目标,所得帐号均存储于数据库中;(2)从数据库中选择发送目标;(3)从数据库中选择发送内容,或者临时编辑发送内容;(4)核对发送目标、发送消息后并确定发送;(5)发送服务器集群完成发送任务并将结果信息回送控制台.为了提高工作效率,第一步可以利用空闲时间预先完成.

系统中各个功能模块互相配合,形成了从帐号获取到信息最终发送的完整解决方案.

从系统的构成还可以看出,实现大规模网络信息发送系统涉及多项关键技术,如控制台与发送服务器集群之间的通信与控制技术、发送服务器集群的任务调度技术、大容量数据库技术、发送目标批量获取技术等.

3 任务调度策略

3.1 任务调度过程分析

控制台传给任务调度器的发送任务文件一般由以下信息组成:必要的通信控制信息、发送源帐号、目标帐号以及发送信息等.任务调度器接收到任务文件后,可以把目标帐号分成若干块,其余信息不变,从而形成若干个子任务,并分配给集群中的发送服务器.发送服务器在接收到任务文件后会生成大批量发送报文,根据目标帐号的多少需要一定的时间,少则数分钟,多则以小时计.任务调度过程就是

任务调度器接收到任务文件后根据需要拆分任务文件并把子任务文件分发给相应发送服务器的过程.

为了讨论方便,提出如下形式化描述:

某次发送任务总负载为 α , 服务器集群为 $S = \{S_1, S_2, \dots, S_n\}$, n 为服务器数量.

定义 1 服务器执行速度:某服务器接收到的任务量与执行过程所花费的时间之比称为该服务器的执行速度,记为 $E_i (i = 1, 2, \dots, n)$.

定义 2 发送任务完成时间:是从任务调度器接收到任务开始直到集群中每台服务器均完成所分配任务的时间,记为 t ,显然,发送任务完成时间取决于集群中完成任务最慢的服务器.

集群中服务器数量一定时,关于任务执行时间有如下论断:

推论 1 当任务调度器按照集群中各服务器的执行速度比分割发送任务,并把分割后的任务传给相定的服务器时,任务完成时间最短.

证明:对于某次发送任务 α ,按照上述任务分割

方法,服务器 S_i 分到的任务大小为 $l_i = \alpha \frac{E_i}{\sum_{i=1}^n E_i}$,该

服务器完成任务所需时间为 $t_i = \frac{l_i}{E_i} = \frac{\alpha}{\sum_{i=1}^n E_i}$,结果

与 i 无关,即各服务器将在相同的时间 t 内完成.下面证明 t 就是完成任务的最短时间:

考察集群中任意一台服务器 S_i ,其分得的任务为 l'_i ,完成任务的时间记为 t'_i ,如果 $l'_i < l_i$,则 $t'_i = \frac{l'_i}{E_i} > \frac{l_i}{E_i} = t$;如果 $l'_i < l_i$,则必存在服务器 $S_j (j \neq i)$ 使得 $l'_j > l_j$, $t'_j = \frac{l'_j}{E_j} > \frac{l_j}{E_j} = t$;如果 $l'_i = l_i$,则 $t'_i = t$,发送任务完成的时间必定不小于 t .

得证.

当发送服务器分布于异地时,任务调度器与发送服务器之间通信的网络延迟会比较大,传输任务文件需要一定的时间.假设任务调度器向一台服务器分发任务的平均时间为 ρ ,对于任务 α ,假设使用单台服务器完成所需时间为 T ,使用 n 台服务器完成的时间为 t ,由于分割任务共需要传输任务 m 次 (通常 $m \geq n$),则 $t \approx \frac{T}{n} + m\rho \geq \frac{T}{n} + n\rho \geq 2\sqrt{\rho T}$,当 $\frac{T}{n} = n\rho$ 即 $n = \sqrt{\frac{T}{\rho}}$ 且 $m = n$ 时, t 取最小值 $2\sqrt{\rho T}$,因此有如下结论:

推论2 如果任务调度器与发送服务器之间的通信延迟比较大,那么集群中服务器数量为 $\sqrt{\frac{T}{\rho}}$ 并且分割任务次数尽量少时,发送任务完成时间最短。

随着网络基础建设的进步,现阶段网络带宽多数情况下比较稳定,即 ρ 一般较小,所以 $\sqrt{\frac{T}{\rho}}$ 可以取较大值。

3.2 集群服务器网络环境分类

根据大规模网络信息发送系统的应用环境,可将集群服务器所处的网络环境分为三类:

(1) 本地专用服务器集群. 该集群处于本地局域网中,特点是网络带宽大,服务器性能稳定,表现在任务调度器与发送服务器之间通信稳定、执行发送任务速度基本恒定并且容易预测任务执行时间。

(2) 远程专用发送服务器集群. 发送服务器位于异地,任务调度器与发送服务器之间通信延迟比较大,服务器本身性能比较稳定。

(3) 分布式服务器集群. 该集群中服务器地点比较分散,甚至跨网分布,服务器本身可能正在运行其他任务,仅利用空闲资源执行发送任务,性能不稳定. 这种环境与现在流行的利用低耦合度大量普通个人计算机执行云计算的情形比较相近。

3.3 基于响应时间的任务调度策略(STSRT)

根据以上讨论结果,设计基于响应时间的任务调度策略(Strategy for Task Scheduling Based on Responding Time, STSRT)如下:

发送任务到达时,任务调度器判断各个服务器的执行速度,按照各服务器执行速度的比值分割任务文件,确保各服务器收到的任务量与其执行速度相匹配。

此时,各服务器将在几乎相同的时间内完成各自任务,使得整个发送任务在最短时间内完成。

针对三种集群服务器网络环境,将任务调度策略进一步细化,提出三种调度策略:

(1) 任务调度器根据集群中各服务器的执行速度比值分割任务文件,并发送给相应的服务器. 各服务器的执行速度可通过发送测试任务文件或首个任务文件的执行时间来确定,也可以查看历史记录确定。

(2) 任务调度器将每次收到的发送任务分配给集群中剩余执行时间最少的服务器. 一旦某台服务器进入空闲状态并且没有新任务到达时,任务调度器查询所有仍在执行任务服务器的剩余执行时间,

若超过预先设置的阈值,则把剩余执行时间最长的服务器分出一半剩余任务让该空闲服务器来执行,如此反复直至所有服务器完成任务。

(3) 任务调度器按照给定的参数把发送任务分成等粒度的小任务,从而形成一个待发送任务集合. 每当集群中的服务器有空闲时,就从待发送任务集合中选择一个任务分配给该服务器. 当待发送任务集合为空后,查询仍在执行任务的服务器的任务执行情况,对于执行速度异常缓慢的服务器剥夺其剩余任务分配给空闲服务器执行。

策略(1)适用于网络环境以及服务器性能比较稳定的情形,如第(1)类集群服务器环境,调度策略实现非常简单. 策略(2)和策略(3)均适用于网络环境以及服务器性能不够稳定的情况,如第(2)类和第(3)类集群服务器环境。

4 实验仿真

模拟一个手机短信群发系统,集群中服务器数量为6,每台服务器连接一个多端口短信 modem. 控制台先后发出8个任务,目标号码的数量分别是2000、5000、6000、4000、8000、7000、6000、5000,分别使用加权轮转法(WRR)、基于响应时间的任务调度策略(1)~(3)进行仿真。

仿真结果表明,使用常见的负载均衡算法 WRR 对集群环境并不适合. 在文中提出的三种任务调度策略中,策略(2)具有最好的性能. 策略(2)的优势在于任务调度器的任务分发次数最少. 由于每一次任务分发都牵涉到握手通信、文件传输以及一定量的数据迁移,因此应尽量减少任务分发次数。

对比三种任务调度策略,策略(1)和策略(3)的任务分发次数比较固定,不会因服务器性能的波动而发生改变;策略(1)仅适用于网络环境和服务器性能非常稳定的情形;策略(2)任务分发次数在一般情况下相对较少,在服务器性能波动比较剧烈的情况下,任务分发次数将显著增加。

综合比较,策略(2)的性能相对较高. 为了进一步减少任务分发时间,提出任务文件预调度机制对策略(2)进一步改进,即在执行最快的服务器进入空闲状态之前启动任务调度,从而隐藏了任务分发时间。

5 结束语

在研究大规模网络信息发送系统组成及 workflow
(下转第172页)

5 结束语

文中提出了一种基于 Web 服务的信任评估模型,针对这种模型设计了 Web 服务的体系结构并说明了服务流程.该信任评估模型能对 Web 服务提供商做出真实的评价,能有效防止服务提供商恶意提高自己的信任度,降低他人信任度的恶意行为,保护了服务使用者的利益.

参考文献:

- [1] 李研,周明辉,李瑞超,等.一种考虑 QoS 数据可信性的服务选择方法[J].软件学报,2008,19(10):2620-2627.
- [2] 肖道举,杨剑,陈晓苏.Web 服务安全保障机制研究[J].华中科技大学学报:自然科学版,2004,32(4):50-52.
- [3] George Spanoudakis, Stephane LoPresti. Web service trust: towards a dynamic assessment framework[C]// the Proceedings of the 4th International Conference on Availability, Reliability and Security. New York: ACM, California: IEEE Computer Society Press, 2009.
- [4] William Conner, Arun Lyengar. A trust management framework for service-oriented environments[C]// 18th International World Wide Web Conference. New York: ACM, 2009.
- [5] 苏锦钿,郭荷清,刘森.Web 服务中的信任和声誉评估模型[J].计算机工程与应用,2006,42(21):127-130.
- [6] 胡旭东,曾国荪,陈波.一种基于非功能属性决策的可信 Web 服务发现模型[J].计算机科学,2009,36(2):95-98.
- [7] 李海华,杜小勇,田萱.一种能力属性增强的 Web 服务信任评估模型[J].计算机学报,2008,31(8):1471-1477.
- [8] 罗俊海,范明钰.基于博弈的 MANETs 信任模型研究[J].计算机研究与发展,2008,45(10):1704-1710.
- [9] 吴鹏,吴国新,方群.一种基于概率统计方法的 P2P 系统信任评估模型[J].计算机研究与发展,2008,45(3):408-416.
- [10] 张琳,王汝传,张永平.一种基于模糊集合的可用于网格环境的信任评估模型[J].电子学报,2008,36(5):862-868.
- [11] 林剑柠,吴慧中.基于主观逻辑理念的网络信任模型分析[J].计算机研究与发展,2007,44(8):1365-1370.
- [12] 刘莉平,葛志辉.Grid 环境下基于实体行为的信任评估模型[J].计算机科学,2008,25(7):2020-2022.

作者简介:

刘彬男,(1973-),副教授.研究方向为网络计算、计算机图形学.
张仁津男,(1963-),教授.研究方向为软件方法学、组件技术、对象技术.

(上接第 167 页)

程的基础上,提出了一种基于响应时间的任务调度策略(STSRT),并对其进行了仿真验证.下一步的工作主要包括:在真实的应用环境中对任务调度策略进行进一步优化、大规模网络信息发送原型系统的实现.

参考文献:

- [1] 刘伯红,蔡丽.短信群发系统的设计与实现[J].微型电脑应用,2004,20(9):22-23.
- [2] 陈慧玲,帅立国,姜昌金.基于 http 协议的论坛群发技术的研究[J].计算机技术与发展,2007,17(3):38-43.
- [3] An Youngbae, Park Daeyeon. Scalable web server cluster system using content-aware packet filtering[C]// International Conference on Parallel and Distributed Processing Techniques and Applications. San Jose, 2003:721-727.
- [4] Hantian Wei, Xiaomei Xiong. The research and implement of an efficient load balancing algorithm[C]// Internet Conference on Services Science, Management and Engineering. Hong Kong, 2009:416-419.
- [5] Shailendra S Aote, Kharat M U. A game-theoretic model for dynamic load balancing in distributed systems[C]// International Conference on Advances in Computing, Communication and Control. Thailand, 2009:235-238.
- [6] Umit V Catalyurek, Erik G Boman, Karen D Devine, et al. A repartitioning hypergraph model for dynamic load balancing[J]. Parallel and Distributed Computing, 2009, 69(8):711-724.

作者简介:

董海量男,(1984-),硕士研究生.研究方向为网络安全、密码学.
胡华平男,(1967-),博士后,教授,博士生导师.研究方向为信息安全与密码学.
刘建锋男,(1984-),硕士研究生.研究方向为网络安全、密码学.
虞欣平男,(1963-),高级工程师.研究方向为网络工程与保障.