

一种两阶段自适应的 IMS-Based IPTV 应用 服务器负载均衡策略^①

李凌南^{1,2} 雷为民¹ 林镜华¹ 白松¹ (1.中国科学院沈阳计算技术研究所 辽宁 沈阳 110004;
2. 中国科学院研究生院 计算与通信工程学院 北京 100049)

摘要: 随着 IPTV 用户数目的增长, IPTV 应用服务器的负担愈加繁重。在这种情况下, 集群系统逐渐成为构建高性能和高服务质量的一种有效的应用服务器体系结构。保证应用服务器各节点的负载均衡, 可以有效提高系统的吞吐量, 提高 IPTV 服务质量。针对负载均衡问题, 提出了一种基于 IPTV 用户行为特点的, 两阶段自适应负载均衡策略, 负载均衡器通过实时监控应用服务器集群内各节点的负载情况, 动态修改各节点负载分配权值。算法在保持应用服务器各节点负载均衡的同时, 维护用户会话的完整性, 保证了服务质量 QoS(Quality of Services)。仿真结果显示此算法具有较好的性能, 能在保证服务质量的前提下有效地提高系统吞吐量。

关键词: IMS; IPTV; 会话; 负载均衡; 服务质量

A Two-Stage Adaptive Load Balancing Strategy for IMS-Based IPTV Application Server

LI Ling-Nan^{1,2}, LEI Wei-Min¹, LIN Jing-Hua¹, BAI Song¹

(1.Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110004, China;
2.College of Computing & Communication Engineering, Graduate University of the Chinese Academy of Sciences, Beijing 100049, China)

Abstract: As the IPTV user number continues to grow, the IPTV Application Servers increases and eventually stats to overload. In this case, the cluster system has become an effective application server that provides high performance and a high quality of services. It keeps the load in balance, and it can enhance the system throughput effectively. It can also improve the IPTV quality of service. For the load balancing problem, a two-phase adaptive load-balancing strategy, which is based on the characteristics of IPTV users' behaviors, is proposed. This load balancer will dynamically modify the load dispatch weight of each node in the application server cluster system by a real time watch of each node's loading degree. This algorithm keeps a load-balanced situation among all the application servers. It maintains the integrality of users' session, and it also ensures the QoS (Quality of Services). The experimental results of this paper show that this algorithm has a good performance, and can improve the system's throughput on the basis of guaranteed QoS.

Keywords: IMS; IPTV; session; load balancing; quality of services

1 引言

IPTV 业务融合了音频、视频和数据等媒体传输, 是一项重要的电信增值业务, 也是因特网、电信网和

有线电视网三网融合的典型业务。在三网融合达到互联互通、资源共享的趋势下, IPTV 业务平台将会广泛地得到应用。IMS(IP Multimedia Subsystem)即 IP

^① 收稿时间:2010-01-28;收到修改稿时间:2010-03-16

多媒体子系统,是下一代网络的一种主要架构方式,基于IMS的IPTV业务体系结构可以重用IMS中用户集中管理、统一鉴权计费、QoS保证及基于IP的策略控制等共同基础架构,还可在IPTV业务中集成Presence/XDM等功能模块以拓展IPTV好友电视、在线电视讨论等新的增值业务模式。近些年Open IPTV Forum、ETIS TISPAN和ITU-T等国际标准组织、Fokus、IBM等研究机构都在加紧制定基于IMS的IPTV业务相关技术规范,华为、中兴、NEC等各大厂商也在积极参与相关规范讨论和产品研发。

在IMS-Based IPTV业务体系结构中,IPTV业务的实现和执行逻辑功能主要集中于IPTV应用服务器(Application Server, AS)上。为了支持大规模的用户量,IPTV应用服务器必须能够在保证服务质量的前提下,提高自身性能。文章分析了IMS-Based IPTV体系结构中可能出现的性能瓶颈问题,在介绍负载均衡相关工作的基础上,结合IPTV应用服务器的特点,提出了一种负载均衡策略以提高IPTV业务的可用性。

2 相关工作

服务器集群系统是被广泛采用的一种有效的体系结构,通过负载均衡策略在各个节点之间进行合理的任务分配和任务调度,集群中的每一台服务器共同承载应用负担,避免造成单个服务器过载,并且在单个服务器出现故障时进行自动故障切换,从而提高了系统的可靠性和可用性。

有关负载均衡的问题,WEB服务器方面的研究已经有很多成果。例如基于DNS技术的负载均衡,肖军弼等^[1]应用DNS的Round Robin功能,实现WEB集群服务的负载均衡并验证了其方法的正确性。郭成城等^[2]针对Web服务器集群系统中负载动态变化特性,提出了一种临界加速递减动态请求负载分配算法,通过负载权值的等效变换更准确地反映集群中单台服务器的当前负载状态,通过临界递减因子来有效抑制服务器可能出现的“拒绝访问”现象,使访问负载的分布更均匀。此外还有基于NAT的负载均衡,轮转法、最小连接数法^[3]等等。

这些有关负载均衡的研究是基于HTTP协议的。在IMS-Based IPTV业务体系结构中,Core-IMS对用户和信令进行集中管理,以及应用服务器使用会话初始协议SIP(Session Initiation Protocol)控制信令

流程,这使得IPTV应用服务器的负载均衡问题同WEB服务器相比具有新的特性。对比两种协议,SIP协议是同会话(session)紧密结合的,当对到来的请求消息进行负载均衡分配的时候,需要充分考虑到会话的结构,保持会话的完整性。SIP协议在设计时加入的超时重传机制也带来了额外的负载问题,J. Sun等^[4]在研究SIP应用服务器过载情况时,发现基于UDP方式传输SIP消息时,由于SIP协议自身设计的超时重传机制,当造成应用服务器过载的请求爆发现象消失时,依然还存在持久重传的现象。对比两种服务应用,WEB应用主要进行文字、图片等内容的下载,每个HTTP链接持续时间相对较短,而IPTV应用是一种典型的电信级业务,对于响应时间要求较为严格,每个SIP会话持续时间相对较长,故相对WEB应用而言,IPTV应用具有更严格的QoS需求。IPTV应用服务器负载均衡算法需要解决如何保持用户会话完整性和较为均匀地分配服务器集群各节点的负载的问题。

3 体系架构及业务流程分析

ITU-T和ETSI TISPAN电信标准组织都制定了IMS-based IPTV体系架构及相关规范,后者可操作性较强,文章主要参考后者制定的一系列规范。

3.1 IPTV应用服务器的位置与功能

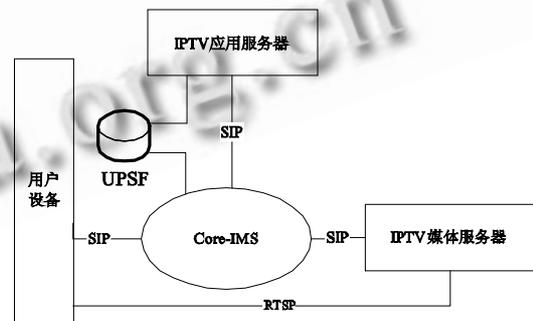


图1 IMS-Based IPTV体系架构

如图1所示,IMS-based IPTV体系架构^[5]下,各个功能模块通过Core-IMS互联,用户终端设备(User Equipment, UE)、IPTV应用服务器、IPTV媒体服务器(Media Server, MS)使用SIP协议通过Core-IMS进行信令交互。用户配置功能实体(User Profile Server Function, UPSF)储存了用户配置信息、用户终端设备参数特性等等。用户终端设备UE通过Core-IMS向IPTV应用服务器发起请求,IPTV

应用服务器对该用户进行认证,同 UPSF 交互查询用户的权限,确定是否允许向用户提供该应用,并且同 IPTV 媒体服务器交互,选择相应的 IPTV 媒体功能,才能为用户提供直播节目(BC)、点播节目(CoD)、个人视频录制(nPVR)的主要业务。IPTV 应用服务器还应为 UE 提供服务发现功能、服务选择功能。即生成服务接入信息,提供个性化服务发现,提供电子节目菜单,提供服务选择信息,为用户设定网络侧参数以使用户激活某种业务。在会话初始建立过程和会话改变的过程中, IPTV 应用服务器必须进行业务授权,也就是检查确认 IPTV 用户的配置信息来决定允许或者拒绝用户接入服务。此外还要使用在线计费系统^[6]进行信用限制,和信用控制,以及在相应的业务中选择合适的 IPTV 媒体功能实体。可见 IPTV 应用服务器主要承担了信令控制的任务。

从 IMS-Based IPTV 系统的架构方面分析系统瓶颈问题,由于系统中 Core-IMS 对信令的集中管理,使得 Core-IMS 本身成为可能的系统性能瓶颈,此问题由 Core-IMS 内相关功能实体产生,此处可以把 Core-IMS 看作黑匣子,直接分析 UE、IPTV 应用服务器、IPTV 媒体服务器的信令交互。IPTV 媒体服务器也是可能存在的性能瓶颈,这涉及到媒体分发策略等。相对于以信令控制为主的 IPTV 应用服务器的瓶颈问题,以上两个瓶颈问题文章暂不讨论。

3.2 IPTV 应用服务器信令流程

IPTV 应用服务器使用 SIP 协议来控制各种业务,以 IPTV 的基本应用之一——直播节目为例^[7],图 2 显示了用户申请观看直播节目的过程。UE 发起观看直播节目的 INVITE 请求,该请求中包含了节目地址等信息,这些信息需要在会话过程中得到授权认证。Core-IMS 收到这个请求后,识别出这是一个发起观看直播节目的请求,并且在 UE 和 Core-IMS 之间的网络中进行资源预留工作。接下来,Core-IMS 将 INVITE 请求转发给应用服务器,应用服务器对该用户的请求进行鉴权和授权,有可能改变用户请求消息中的直播节目地址。若用户有权观看,则把成功的消息 200OK 返回给 Core-IMS,Core-IMS 转发该消息至 UE。UE 通过 Core-IMS 返回 ACK 给应用服务器,并且加入应用服务器返回 200OK 消息中给出的组播地址。最后,直播节目内容从媒体服务器下发到用户,用户可以观看节目。

以上分析表明用户申请观看直播节目的过程是一个完整的会话过程,在集群系统中,多个应用服务器节点的存在可能会引起以下两个问题:1)一个完整的会话过程中的消息被分配到不同的应用服务器节点中去,无法保持会话的完整性^[8];2)若负载分配不均衡,可能导致某一个服务器节点过载运行,某些消息处理时间过长,甚至超时。这可能会引起严重的后果——使得正在进行的会话被破坏,从而引起极坏的用户体验。文献^[4]指出在 SIP 应用服务器过载时,消息超时可能会使 UE 不断重发请求,引发持续性的消息重传,加重了应用服务器和网络的负担,甚至在突发流量停止后,重传的消息依然充斥网络。从 IPTV 用户的行为分析,IPTV 用户会话时间持续较长,一个用户接入系统后可能长时间观看同一频道的电视节目。在某一时段,例如假日时段,系统可能会达到用户数量高峰或者频道切换信令数量高峰。这要求 IPTV 应用服务器必须有良好的负载分配策略以保证服务可用性以提供高质量的用户体验。

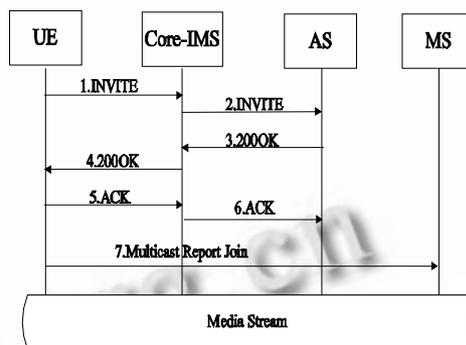


图 2 用户观看直播节目的会话建立过程

4 IPTV应用服务器负载均衡策略

4.1 IPTV 应用服务器负载均衡框架

如图 3 所示,负载均衡器位于 IPTV 应用服务器集群前端,通过 Core-IMS 转发的 SIP 消息首先进入负载均衡器的前端队列,负载均衡器依照两阶段自适应负载均衡策略,将消息队列中的 SIP 消息分发到最适宜的应用服务器节点,每个应用服务器节点都各自维护他们的消息队列。

4.2 相关参数

IPTV 应用服务器集群中的节点个数为 M ,各个节点对应的权重为 $W_{k,m}$ ($k=1,2,3,\dots;m=1,2,\dots,M$),负载均衡器按照各节点对应的权重按照一定的概率分发

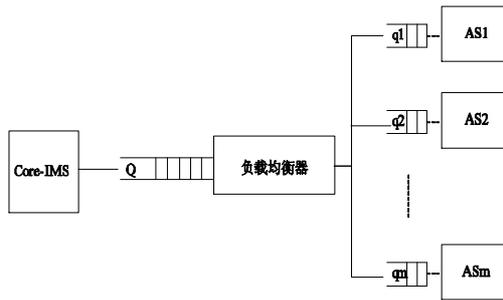


图 3 IPTV 应用服务器负载均衡框架

消息。在计算节点权重时，应考虑到多方面的评价因素，此处采用第 m 个节点服务器的 CPU 占用率 C_m 、内存占用率 M_m 、带宽占用率 B_m 和维护会话的队列长度占用率 L_m 四个因素来共同评价，对于每个节点来说，这四方面评价因素的影响因子 $w_{j,m}(j=1,2,\dots,4; m=1,2,\dots,M)$ 可能是不同的，由用户自己定义，由上可知，每个节点的权重 $W_{k,m}$ 定义为：

$$W_{k,m} = F(C_m, M_m, B_m, L_m) \tag{1}$$

$$= 1 - (w_{1,m} \times C_m + w_{2,m} \times M_m + w_{3,m} \times B_m + w_{4,m} \times L_m)$$

其中 $\sum_{i=1}^4 w_{i,m} = 1, \sum_{m=1}^M (W_{k,m} / M) \leq 1$ 。

将 m 个节点的权重按照升序排列后，计算权重累加向量

$$SUM_{k,m} = \sum_{j=1}^m W_{k,j} \quad (k=1,2,3,\dots; m=1,2,\dots,M) \tag{2}$$

各个节点的权重有可能是动态变化的，但是在 T_k 时间间隔内保持不变， k 值记录了节点权重整体调整的次数。

每个节点需要维护会话的平均队长 $\overline{L_{k,m}}$ ，

$$\overline{L_m} = L_m \times (1 - a) + L_m \times a \tag{3}$$

其中， L_m 为新消息到达之前队列长度， L_m 为当前实际队列长度，为队列权值，是当前队列实际长度对平均队列长度的贡献。选取合适的对平均队列长度有一定的影响，如果过大，突发流量会使得平均队长变化较大，造成抖动，如果过小，那么流量变化对平均队长贡献过小，无法反映当前的负载状况。

4.3 两阶段自适应负载均衡策略

两阶段自适应负载均衡策略由静态负载分配策略、动态权值调整策略、权值调整触发策略三部分组成。

4.3.1 静态负载分配策略

在 T_k 时间间隔内，各个节点对应的权重为 $W_{k,m}$

是固定不变的，采用 Hash 映射方法按照权重把 SIP 消息分配到各节点。此处采用的 Hash 函数的映射关系是从 SIP 消息中的 Call-ID 头域值集合 C 到具有 N 个表项的节点表 N ，如下式

$$H(\cdot): C \rightarrow N \tag{4}$$

由于不同会话请求消息其包含的 Call-ID 头域值是不相同的，所以同一会话内的消息，可以根据 SIP 消息中的 Call-ID 头域值判断，只要把含有相同 Call-ID 值的 SIP 请求发往同一个后台服务器就可以保证会话的完整性。

静态负载分配步骤如下：

1) 负载均衡器从消息队列中取出一个新的 SIP 消息，根据 Call-ID 进行 Hash，如果该用户已经创建会话，并且记录了已分配的服务器为 r ，则转(5)；如果负载均衡器未记录曾为该用户创建会话，转(2)；

2) 把 M 个服务器的分配权值按照 $W_{k,m}$ 从小到大排列，并记录每个对应的服务器编号 S_m ，形成权重服务器对 $(W_{k,m}, S_m)$ 。

3) 按照式(2)计算权重累加向量 $SUM_{k,m}$ ，形成分配待选对 $(SUM_{k,m}, S_m)$ ；

4) 产生一个 0 至 $SUM_{k,m}$ 之间的随机数 $RandNum(r)$ ，根据 $RandNum(r)$ 隶属的值区间，当 $RandNum(r)$ 满足 $SUM_r \leq RandNum(r) < SUM_{r+1}$ 时，选择该值所在的分配对 (SUM_r, S_r) ，选择该请求的目标服务器为第 r 个服务器，转(5)；

5) 把新的 SIP 请求发送到第 r 个服务器，完成消息的分发。

4.3.2 动态权值调整触发策略

很多自适应负载均衡方法中采用了周期性轮询、报告的触发策略方法，虽然这种方法比较简单易行，无需收集太多的节点信息，然而相对传统的通话性质 SIP 应用而言，IPTV 应用会话持续性时间较长，这也使得周期本身难以设定。本策略监测节点维护会话的平均队长，以此作为触发策略，一旦某个节点的维护会话的平均队长超过预设的阈值，则触发动态权值调整。每个节点按照式(3)计算自己维护会话的平均队长。由负载分配均衡器动态监测各个节点的输入队列平均队长，当某一个队列长度超过预设的阈值时，进行动态权值调整。

4.3.3 动态权值调整策略

动态权值调整步骤如下：

1) 获取第 m 个节点的 CPU 占用率 C_m 、内存占用率 M_m 、带宽占用率 B_m 和维护会话的队列长度占用率 L_m ;

2) 依据预先设置对应各个资源的权重分配 $w_{j,m}(j=1,2,\dots,4; m=1,2,\dots,M)$, 按照式(1)计算出自身的 SIP 请求分配权重 $W_{k,m}$;

4.3.4 两阶段自适应负载均衡策略

主要思想是, 在没有节点报告负载过重的情况下, 按照各节点的权值进行静态负载均衡分发消息, 当有节点负载过重后, 触发系统进行动态权值调整, 再进入新一轮的静态负载均衡分发消息过程。系统通过动态权值调整, 和静态负载分发两个阶段, 达到各节点的负载均衡。步骤描述如下

1) 系统初始化, 令 $k=1$, 计算出各个节点权重 $W_{k,m}$;

2) 按照各个节点权重 $W_{k,m}$, 调用静态负载分配策略;

3) 负载均衡器监测各个节点的输入队列平均队长, 若某一个节点平均队列长度超过预设的阈值时, 转 4), 否则转 2);

4) 调用动态权值调整策略, 令 $k=k+1$, 计算出各个节点权重 $W_{k,m}$, 转 2)。

5 仿真分析

5.1 仿真环境

实验环境包括 1 台充当负载均衡器的服务器和 4 台 AS 服务器以及 5 台普通的 PC 机作为客户端, 所有的设备均连接到交换机上并组成内网。设备的硬件配置如下:

表 1 仿真环境设备硬件配置表

设备	CPU	内存	网络
负载均衡器	Intel(R) Xeon(R) CPU 3065 @ 2.33GHz, 2000 MHz; CPU 3065 @ 2.33GHz, 2000 MHz	2GB	1G bps
AS 服务器	Pentium (R) D CPU 2.8GHz 2.97GHz	1GB	100M bps
客户端	Pentium (R) D CPU 2.8GHz 2.97GHz	1GB	100M bps

在这里, 负载均衡器使用较好的服务器运行, 是为了让其不成为系统的瓶颈, 从而能够从运行结果更

准确的分析模型的效率。为了更好的分析此模型的效率, 使用加权轮转算法 (WRR) 和临界加速递减法 (MDC) 来进行对比, 同时对他们进行一些修改, 使得同一个 Session 的所有消息都被分配到同一个 AS 服务器上, 以维持会话的完整性。

客户端运行基于 IMS 的 IPTV 服务请求模拟软件, 会话发起时间为符合指数分布的 Poisson 事件。

5.2 结果分析

基于 IMS 的 IPTV 主要给用户直播节目 (BC)、点播节目 (CoD)、个人视频录制 (nPVR), 其中 CoD 和 N-PVR 无论在任何时候都需要等待媒体服务器完成处理之后才能进行返回给用户, 但在这里不考虑媒体服务器的处理能力, 因此使用 BC 业务来测试模型性能。IPTV 提供的业务对于用户来说, 较小的延迟能够提供较好的 QoS, 在此用响应延迟来衡量算法的效率。此响应延迟是指用户发起服务请求直至服务器处理完成并返回结果之间的时间。测试包括三项内容: 1) 测试改进的加权轮转算法 (WRR); 2) 测试改进的临界加速递减法 (MDC); 3) 测试两阶段自适应负载均衡算法 (TFA)。结果如图 4 所示。

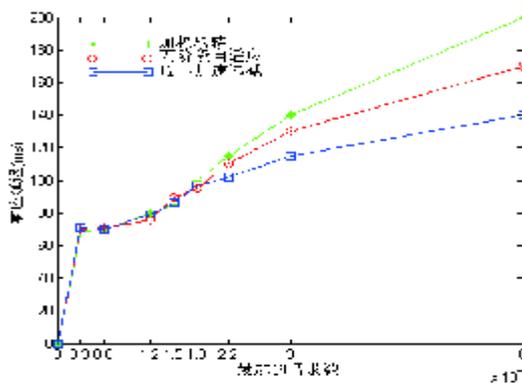


图 4 测试结果

从测试结果可以看出, 在发起的请求数量不大的情况下 (<18000), 各种算法之间的差异不明显。当发起的请求数量逐渐增加的时候, 加权轮转法的响应时间逐渐增加; 两阶段自适应算法在此阶段优于加权轮转算法, 但是次于临界加速递减算法。但是临界加速递减算法在达到临界区的时候将拒绝客户端的一些请求, 因此造成了拒绝服务, 影响了服务质量。因此两阶段自适应算法在保证服务质量的情况下减少了在大量请求的情况下的响应时间。

6 结论

随着 IPTV 用户数目的增长, IPTV 应用服务器的负担愈加繁重, 因此引入集群系统来构建高性能和高服务质量应用服务器。为了保证应用服务器各节点的负载均衡, 提出了一种基于 IPTV 用户行为特点的, 两阶段自适应负载均衡策略, 负载均衡器通过实时监控应用服务器集群内各节点的负载情况, 动态修改各节点负载分配权值。由于基于 IMS 的 IPTV 的会话具有时间长、同一会话的请求需要被同一个服务器处理, 因此改进了加权轮转算法和临界加速递减算法, 使得其能够保持会话的完整性, 在此基础上进行对比仿真试验。仿真结果表明, 两阶段自适应算法在保证服务质量的情况下表现出较好的性能。

参考文献

- 1 肖军弼, 王宇. 应用 DNS 实现 WEB 集群服务的负载均衡. 计算机系统应用, 2003, 12(10): 59 - 62.
- 2 郭成城, 晏蒲柳. 一种异构 Web 服务器集群动态负载均衡算法. 计算机学报, 2005, 28(2): 179 - 183.
- 3 Emiliano C, Salvatore T. Static and Dynamic scheduling algorithm for scalable Web server farm. Proc. of the IEEE 9th Euro micro Work
- shop on Parallel and Distributed Processing. USA: 2001: 369 - 376.
- 4 Sun J, Hu JF, Tian RX, Yang B. Flow Management for SIP Application Servers. ICC 2007 proceedings. 2007: 646 - 652.
- 5 Telecommunications and Internet converged Services and Protocols for Advanced Networking(TISPAN). IPTV Architecture; IPTV functions supported by the IMS subsystem. ETSI TS 182 027 Ver.2.0.0. 2008 - 02.
- 6 Telecommunications and Internet converged Services and Protocols for Advanced Networking(TISPAN). Charging management. ES 282 010 Ver.2.0.6. 2008 - 04.
- 7 Telecommunications and Internet converged Services and Protocols for Advanced Networking(TISPAN). IMS-based IPTV stage 3 specification. ETSI TS 183 063 Ver.2.1.0. 2008 - 06.
- 8 Chen YJ, Lu XC, Sun ZG. MSF: A Session-Oriented Adaptive Load Balancing Algorithm. 2007 IFIP International Conference on Network and Parallel Computing Workshops. Dalian, China: 2007: 657 - 663.