

网络负载均衡策略的研究与应用^{*}

The Study and Application of Balance Tactics for Network Load

孙玉强¹ 徐立新² 吕海莲¹ 冯乃勤¹

(河南师范大学计算机科学系 新乡 453002)¹ (河南机电高专电子信息工程系 新乡 453002)²

Abstract This paper puts forward several kinds of basic thought and the applications of balance tactics of network load. These balance tactics raise the answer speed of network server, and increase the utilization efficiency of server and other network resources, and avoid the questions such as one some inefficacy of critical position of network emergence etc.

Keywords Network, Balance of network load, Tactics, Application

1 引言

目前,无论是对企业内部网、校园网还是对互联网服务运营商如 ISP、ICP、IDC 等,新的应用和服务层出不穷,使得 IP 网络的数据流量剧增,尤其是对网络的各个核心部分,都提出了越来越高的性能要求。如何提供一种廉价、有效的方法来扩展服务器带宽及增加其吞吐量,加强网络数据处理的能力,提高 IP 网络的灵活性和可用性的问题就摆在了我们面前。

在现在的 IP 网络中,网络的数据流量分布并不均匀,例如,电子商务网站属于计算集中型的应用,服务器计算负荷会很大;网络中的数据库服务属于读写频繁的应用,存储系统的压力比较大;网络中的视频服务属于传输量大的应用,数据在网络接口上的流量比较大;网络中的路由器和防火墙等设备属于网络的门户,访问量大的应用易对其形成瓶颈。想要消除这些瓶颈,除了升级设备、改变网络拓扑结构外,采用相应的负载均衡策略,不失为一种好的办法。

网络负载均衡技术即是根据实际的响应时间制定优先级交付决策,从而实现高性能、智能化流量管理,达到最佳的服务器集群性能。它主要完成以下任务:解决网络拥塞问题,实现服务就近提供与其地理位置无关;提高服务器响应速度;提高服务器及其他资源的利用效率;避免网络关键部位出现单点失效等。

2 几种网络负载均衡策略与技术的应用

网络负载均衡策略有多种多样的形式,广义上的负载均衡既可以设置专门的网关、负载均衡器,也可以通过一些专用软件与协议来实现。在 ISO/OSI 七层协议模型中的第二、三、四、七层都可有相应的流量控制和负载均衡策略,在数据链路层上,可以用基于数据包的目的 MAC 地址选择不同链路;在网络层,可以用基于 IP 地址的分配方式,将同一请求的数据流分配到相同功能的多个节点;而第四层、第七层交换技术,本身便是一种基于访问流量的控制方式,起到了很好的负载均衡作用。

对一个网络的负载均衡问题,具体情况要看对网络瓶颈所在之处的具体分析,大体包括采用传输链路聚合、采用更高层网络交换技术、设置服务器群集、利用 DNS 服务、利用协议本身的均衡功能及采用负载均衡设备六种策略和技术。

2.1 传输链路聚合技术

当企业的网络内部对高带宽应用需求不断增大时,例如 Web 访问、FTP 传输及内部网连接,核心部位的数据接口将产生瓶颈问题,瓶颈延长了客户应用请求的响应时间。传统的解决瓶颈问题采用的对策是提高服务器链路的容量,使其超出目前的需求,例如可以由快速以太网升级到千兆以太网。采用升级技术是一种长远的、有前景的解决方案,但需花费大量的金钱和时间进行升级,在这种情况下,传输链路聚合技术为消除传输链路上的瓶颈与不安全因素提供了成本低廉的解决方案。

传输链路聚合技术的基本思想是将多个线路的传输容量融合成一个单一的逻辑连接。当原有的线路满足不了需求,而单一线路的升级又太昂贵或难以实现时,就要采用多线路的解决方案了。目前有 5 种传输链路聚合技术可以将多条线路“捆绑”起来。

1. 同步 IMUX 系统工作在 T1/E1 的比特层,利用多个同步的 DS1 信道传输数据,来实现负载均衡。

2. IMA 是另外一种多线路的反向多路复用技术,工作在信元级,能够运行在使用 ATM 路由器的平台上。

3. 用路由器来实现多线路是一种流行的链路聚合技术,路由器可以根据已知的目的地址的缓冲(cache)大小,将分组分配给各个平行的链路,也可以采用循环分配的方法来向线路分发分组。

4. 多重链路 PPP 又称 MLP,是应用于使用 PPP 封装数据链路的路由器负载均衡技术。MP 可以将大的 PPP 数据包分解成小的数据段,再将其分发给平行的多个线路,还可以根据当前的链路利用率来动态地分配拨号线路。这样做尽管速度很慢,因为数据包分段和附加的缓冲都增加时延,但可以在低速的线路上运行得很好。

5. 对于 IP 系统,还可以考虑采用 VRRP(虚拟路由冗余协议),VRRP 可以生成一个虚拟缺省的网关地址,当主路由器无法接通时,备用路由器就会采用这个地址,使 LAN 通信得以继续。

总之,当主要线路的性能必须提高而单条线路的升级又不可行时,可以采用传输链路聚合技术。

2.2 更高层网络交换技术

1)第四层交换技术 端到端性能和服务质量要求对所有

^{*} 本项目受河南省自然科学基金资助。孙玉强 教授,西安电子科技大学在读博士,主要研究方向为网络与软件工程与环境等。

连网设备进行负载均衡,以保证客户机与服务器之间数据平滑地流动。二层与三层交换技术在解决局域网和互联网络的带宽及容量问题上发挥了很好的作用,但是,这可能还不够,还需要更多的性能,而这正是第四层交换的用武之地。

第四层交换技术利用第三层和第四层包头中的信息来识别应用数据流会话,这些信息包括 TCP/UDP 端口号、标记应用会话开始与结束的“SYN/FIN”位以及 IP 源/目的地址。利用这些信息,四层交换应可以做出向何处转发会话传输流的智能决定。对于使用多种不同系统来支持一种应用的 IDC、ISP、ICP,四层交换的作用就尤为重要。

当前局域网交换机所谓的第四层交换技术,就是按照 IP 地址和 TCP 端口进行虚拟连接的交换,直接将数据包发送到目的计算机的相应端口。通过交换机就能将来自外部的初始连接请求,分别与内部的多个地址相联系,此后就能对这些已经建立的虚拟连接进行交换。因此,一些具备第四层交换能力的局域网交换机,就能作为一个硬件负载均衡器,完成服务器的负载均衡,其网络传输速度和交换速度远远超过普通的数据包转发。然而,正因为它是使用硬件实现的,因此也不够灵活,仅仅能够处理几种最标准的应用层协议的负载均衡问题,如 HTTP 协议等。

第四层交换除了负载均衡功能外还支持其它功能,如基于应用类型和用户 ID 的传输流控制功能。采用多级排队技术,第四层交换可以根据应用来标记传输流以及为传输流分配优先级。此外,四层交换直接安放在服务器前端,它了解应用会话内容和用户权限,因而使得它成为防止非授权访问服务器的理想平台。

2) 第七层交换技术 在高可用性和负载均衡方面,第七层交换可以实现有效的数据流优化和智能化负载均衡。通过利用由应用返回给最终用户的第七层信息,用户不仅能验证是否在发送正确的内容,而且还能打开网络上传送的数据包而不用考虑 IP 地址或端口,并根据包中的信息做出负载均衡决定。

从本质上讲,这种智能性迁移超越了第四层的功能,具有第四层功能的设备是无法识别流过此端口的不同类型的传输流,因此,它们对所有传输流同等对待。可是传输流并不都是相同的,对于负载均衡设备来说,能够知道流过此端口的数据是流媒体还是简单请求非常有用,不少具有第四层功能的设备以同样的方式对待这两种类型的数据,因而可能会将流媒体数据发送到无法做出响应的服务器,导致错误的信息和时延。

而第七层的智能性能够实现进一步的控制,即对所有传输流和内容的控制,这种具有第七层认知功能的技术,保证了不同类型的传输流可以被赋予不同的优先级。具有第七层认知功能的设备不是依赖路由设备或应用来识别差别服务、通用开放策略服务或其它服务的传输流的,它可以对传输流进行过滤并分配优先级,这就使你不必依赖应用或网络设备来达到这些目的了。Web 内容交换技术即是第七层交换技术,也提供了一种对访问流量的高层控制方式。

2.3 设置服务器群集

网络服务器必须具备提供大量并发访问服务的能力,其处理能力和 I/O 能力已经成为提供服务的瓶颈。如果客户的增多导致通信量超出了服务器能承受的范围,那么其结果必然会形成瓶颈。显然,单台服务器有限的性能不可能解决这个问题,但若能将多台这样的服务器组成一个系统,并通过软件

技术将所有请求平均分配给所有服务器,这就是利用服务器群集实现负载均衡的最初基本设计思想。

传统的服务器群集通常以光纤镜像卡进行主从方式备份,但这却使两台服务器的投资只得到一台服务器的性能。新的负载均衡策略是通过 LSANT (Load Sharing Network Address Transfer) 将多台服务器网卡的不同 IP 地址翻译成一个 VIP (Virtual IP) 地址,使得每台服务器均时时处于工作状态。单一计算机可以提供有限级别的服务器可靠性和可伸缩性,但是,通过将两台或多台高性能服务器的主机连成群集,网络负载均衡就能够提供关键任务服务器所需的可靠性和性能。

建立一个高负载的 Web 站点,必须使用多服务器的分布式结构,从多个服务器的组成结构上来分,又可分为非对称的分布式结构和对称的分布式结构。非对称的分布式结构中每个服务器起到的作用是不同的,例如一台服务器用于提供静态网页,而另一台用于提供动态网页等等。这样就使得网页设计时需要考虑不同服务器之间的关系,一旦要改变服务器之间的关系,就会使得某些网页出现连接错误,不利于维护,可扩展性也较差。能进行负载均衡的网络设计结构应为对称式结构,在对称的分布式结构中每台服务器都具备等价的地位,都可以单独对外提供服务而无需其他服务器的辅助,将外部发送来的请求均匀分配到对称结构中的每台服务器上,接收到连接请求的服务器都独立回应客户的请求。在这种结构中,由于建立内容完全一致的 Web 服务器并不困难,因此,负载均衡技术就成为建立一个高负载 Web 站点的关键性技术。

2.4 利用 DNS 服务的 IP 地址定向功能

由于网络的数据流量多集中在诸如 Web、FTP、E-mail 等访问服务器一端,因此对访问服务器的负载均衡技术尤为重要。

DNS 服务是一种简单而有效的解决访问服务器负载均衡的方法,但应注意其所存在的二个问题,一是它不能区分服务器的差异,也不能反映服务器的当前运行状态,因此,当使用 DNS 负载均衡的时候,必须尽量保证不同的客户计算机能均匀获得不同的地址。由于 DNS 数据具备刷新时间标志,一旦超过这个时间限制,其他 DNS 服务器就需要和这个服务器交互以重新获得地址数据,这就为获得不同的 IP 地址提供了可能性,因此为了使地址能随机分配,就应使刷新时间尽量短,不同地方的 DNS 服务器能更新对应的地址,达到随机获得地址,然而将过期时间设置得过短,也会使 DNS 流量大增,从而造成额外的网络问题。DNS 负载均衡的另一个问题是,一旦某个服务器出现故障,即使及时修改了 DNS 设置,还是要等待足够的时间(刷新时间)才能发挥作用,在此期间保存了故障服务器地址的客户计算机将不能正常访问服务器。

2.5 利用协议本身的负载均衡支持功能

高层协议中有的协议本身支持与负载均衡相关的功能,例如 HTTP 协议中的重定向能力。HTTP 协议运行于 TCP 连接的最高层,客户端通过端口号 80 的 TCP 服务直接连接到服务器,然后通过 TCP 连接向服务器端发送一个 HTTP 请求。在服务器分清客户端所需的网页和资源之前,至少要进行四次 TCP 的数据包交换请求。由于外部负载均衡设备要把进入的请求分配给多个服务器,因此它只能在 TCP 连接时建立且 HTTP 请求通过后才能确定如何进行负载的平衡。当一个网站的点击率达到每秒上百甚至上千次时,TCP 连接、HTTP 报头信息以及进程的时延已经变得很重要了。在

HTTP 请求和报头中有很多对负载均衡有用的信息,首先,我们可以从这些信息中获知客户端所请求的 URL 和网页,利用这个信息负载均衡设备就可以将所有的图像请求引导到一个图像服务器,或者根据 URL 的数据库查询内容调用 CGI 程序,将请求引导到一个专用的高性能数据库服务器。注意唯一能局限这些信息获取的因素是负载均衡设备本身的灵活程度。如果网络管理员熟悉 Web 内容交换技术,他可以仅仅根据 HTTP 报头的 cookie 字段来使用 Web 内容交换技术,改善对特定客户的服务,如果能从 HTTP 请求中找到一些规律,还可以充分利用它作出各种决策。除了 TCP 连接表的问题外,如何查找合适的 HTTP 报头信息以及作出负载均衡决策的过程,是影响 Web 内容交换技术性能的重要问题。

这种负载均衡技术依赖于特定的协议,因此其使用范围有限。

2.6 利用负载均衡设备的功能

目前,许多厂商推出了专用于平衡服务器负载的负载均衡器。目前负载均衡器生产商有: Intel、Alteon Web、Arrow Point、Coyote Point、F5 Networks、Foundry Networks、HydraWeb 以及 RADWare 等。所提供的解决方案主要包括如下几大方面:

1. 服务器群集的负载均衡,主要提供 Web Server Director,它可以监视所有的用户请求,并在可用的服务器群之间进行智能化的负荷分配,从而可以提供容错、冗余、优化和可扩展性能。

2. 高速缓存服务器的负载均衡,它可以提供优化的 Internet 访问和存储资源使用率,同时,也使整个服务器群的性能得以最大程度的发挥。

3. 防火墙的负载均衡,例如, RadWare 公司提供的 FireProof 是一种动态负载均衡系统,可有效地管理多个防火墙

和其他安全设备上的流量。它使用的算法能够监视客户的数量和每个防火墙上的负载,并在各单元之间动态地平均分配流量,同时还可兼顾呼入和呼出的流量。

4. 链路的负载均衡,将多个线路的传输容量融合成一个单一的逻辑连接,适用于具有多个链接的网络。

在最新的负载均衡产品中,智能化越来越明显。一些智能化的负载均衡器能够检测到像数据库错误、服务器不可用等信息,从而采取措施使会话恢复和重定向服务器,使电子商务能够得以顺利进行。另外多址负载均衡器可以对客户发来的访问请求进行解析,计算出最佳地址,然后将该地址返回客户,使客户自动连接到对其请求来说最佳的数据中心。

结束语 对于一个网络的负载均衡策略和技术的应用,从网络的不同层次可以采用不同的策略和技术,具体要依据网络的瓶颈所在。网络负载均衡技术的采用,使得自动故障恢复得以实现,服务的时间延长,24 小时×7 天可靠性和持续运行成为可能,同时负载均衡技术也为那些焦急等待大量数据和文件请求响应的客户提供了更快的响应时间。

参考文献

- 1 沈中林,等. Linux 下集群虚拟服务器的研究与设计[J]. 计算机工程与应用,2000. 11
- 2 赵水宁,等. 多 Web 服务器负载均衡技术的研究[J]. 电信科学,2001. 7
- 3 姚耀文,等. 基于 Linux 的服务器群集方案[J]. 计算机工程,2001. 4
- 4 胡季敏,等. 使用动态负载均衡技术的 LINUX 高性能集群服务器研究[J]. 微型电脑应用,2001. 4
- 5 吕健,等. Linux 环境下的 Web 服务器负载均衡技术初探[J]. 科技情报开发与经济,2001. 2
- 2 Sahami M, et al. A Bayesian Approach to Filtering E-Mail. <http://robotics.stanford.edu/users/sahami/papers-dir/.spam.ps>,1998
- 3 Sahami M. Using Machine Learning to Improve Information Access:[PhD thesis]. Stanford University,Dec. 1998. 11~29,170~180
- 4 Rennie J D M. ifile: An Application of Machine Learning to E-Mail Filtering. <http://www.cs.cmu.edu/~jr6b/papers/ifile98.ps>,1998
- 5 Cohen W W. Learning rules that classify E-Mail. In: Proc. of the AAAI Spring Symposium on Machine Learning in Information Access,1996
- 6 Payne T. Learning Email Filtering Rules with Magi A Mail Agent Interface, MSc Thesis, University of Aberdeen, Scotland, 1994
- 7 边肇祺,张学工等编著. 模式识别. 北京:清华大学出版社,第二版,1999. 9~16
- 8 Lewis D D. Feature Selection and Feature Extraction for Text Categorization. <http://www.research.att.com/~lewis/chronobib.html/lewis92e.ps>,1992
- 9 Nigam K, et al. Using EM to Classify Text from Labeled and Unlabeled Document. <http://www-2.cs.emu.edu/pepole/mccallum/emcat-mlj2000.ps>,1998
- 10 McCallum A, Nigam K. A Comparison of Event Model for Naive Bayes Text Classification. <http://www-2.cs.cmu.edu/people/mccallum/multinomial-aaai98w.ps>,1998

参考文献

- 1 林亚平. 概率分析进化算法及其研究进展. 计算机研究与发展,2001,37(1):43~49