

一种快速负载均衡设备 KLBD 的设计

刘绍南¹ 王 鹏¹ 刘雪梅²

(装备指挥技术学院试验工程系 北京怀柔101416)¹ (中南大学艺术学院 长沙410083)²

摘 要 随着网络技术的发展,服务器的负载越来越重,因而负载均衡技术得到了空前发展。本文提出了一种基于数据链路层的快速转接网络包的负载均衡设备 KLBD 的设计方法与过程,并就其关键模块进行了分析。

关键词 负载均衡,数据链路层,Linux

Design of a Fast Load Balancing Equipment KLBD

LIU Shao-Nan¹ WANG Peng¹ LIU Xue-Mei²

(Department of Test Engineering, Institute of Command and Technology of Equipment, Beijing Huairou 101416)¹

(Art College of Central South University, Changsha 410083)²

Abstract With the development of network technology, the burden of servers is becoming more and more heavy, so the technology of load balancing is widely developed. A designing method and process of load balancing equipment KLBD based on fast link network packages of data link layer is brought out in this paper, and the key modules are also analyzed.

Keywords Load balancing, Data Link layer, Linux

当前,无论在企业网、园区网还是在广域网如 Internet 上,业务量的发展都超出了过去最乐观的估计,上网热潮风起云涌,新的应用层出不穷,即使按照当时最优配置建设的网络,也很快会感到吃不消。尤其是各个网站的核心部分,其数据流量和计算强度之大,使得单一设备根本无法承担。然而如何在完成同样功能的多个网络设备之间实现合理的业务量分配,不致于出现一台设备过忙、而别的设备却未充分发挥处理能力的情况,这就成了一个比较大的科研问题,负载均衡机制也因此应运而生。

负载均衡建立在现有网站结构之上,它提供了一种廉价有效的方法扩展服务器带宽和增加吞吐量,加强网络数据处理能力,提高网络的灵活性和可用性。它主要完成以下任务:解决网络拥塞问题,服务就近提供,实现地理位置无关性;为用户提供更好的访问质量;提高服务器响应速度;提高服务器及其他资源的利用效率;避免了网络关键部位出现单点失效。

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。对一个网络的负载均衡应用,从网络的不同层次入手,根据网络瓶颈所在进行具体分析。从客户端应用为起点纵向分析,参考 OSI 的分层模型,我们把负载均衡技术的实现分为客户端负载均衡技术、应用服务器技术、高层协议交换、网络接入协议交换等几

种方式。

本设计为基于数据链路层的快速负载均衡器产品。

1 硬件设计

系统硬件采用 SBS 的 PC104, 主板为 SysCentreModuleTM /6231, 网卡采用 SBS 的 SysExpanModuleTm/Ethernet, KLBD 采用多网卡结构,按照流量的大小可以扩充网卡来平衡每一块网卡的负载。其硬件结构如图1所示。

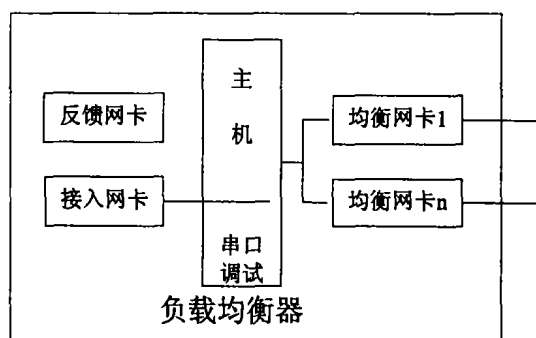


图1 负载均衡器硬件结构

从结构图中可以看出,系统采用多网卡结构, SBS 的 PC104 模块很方便扩充,具体网卡配置多少可以根据系统规模的大小确定,确定以后,通过在系统的串口终端运行配置程序进行配置。

2 软件设计

系统的软件环境为 linux 操作系统。系统的软件模块结构如图2。

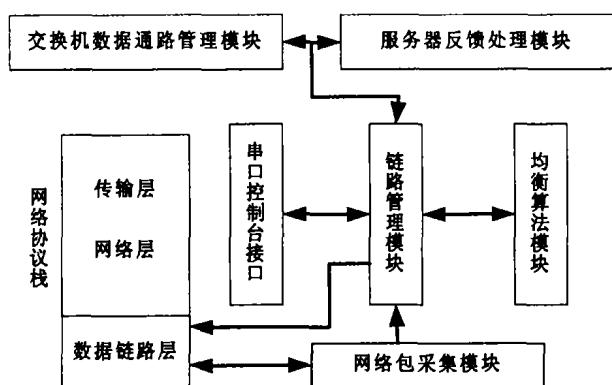


图2 KLBD 的软件结构

从图2可以看出,整个负载均衡软件的结构分成6个模块:网络包采集模块、链路管理模块、均衡算法模块、交换机数据通路管理模块、服务器反馈处理模块、串口控制台接口。这6个模块的功能分述如下。

图1负载均衡器硬件结构网络包采集模块:负责完成对网络请求包的采集。它需要分析包的特性、类型,并按 TCP、UDP 等进行分辑处理。在 linux 内核的网卡驱动模块中,当网络包被网卡中断驱动程序接收到时,都规范为 sk_buff 结构,并通过 netif_rx 函数送到系统的接收队列。我们就在这时对接收队列中的包进行处理,是 TCP 和 SYN 连接包,送到链路管理模块进行链路登记和管理,将经过处理后的包直接送到由均衡算法模块确定的相应网卡的发送队列,这样所有的包就不需要经过上层的网络进行转接,而达到快速转接的目的;当接收到的是 FIN 包时结束链路。如果是 UDP 包就按照原地址进行管理,使之与同一台服务器通信。

链路管理模块:负责对各类型包进行链路管理,当然主要是对 TCP 包的链路管理。当网络包采集模块有新的链路请求时,根据均衡算法模块确定。那台服务器对此请求进行服务后,链路管理模块需要对新的链路请求包中的 MAC 地址进行替换,然后管理新的链路。

均衡算法模块:目前的均衡算法比较多,常用的有基于链路的、基于数据量的、基于服务器忙闲的等。系统可以通过串口终端进行配置,默认为基于链路的。

交换机数据通路管理模块:从结构图中可以看出,如果每一个链路建立后,它的数据通讯还通过原路进行,那么负载均衡器就成为了通讯的瓶颈。因此需要将实际服务时的通讯包分配给其达到他的高速通讯设备来完成,即当将负载均衡器配置在服务量很大的网络上时,需要有交换机作为数据通路。所以

在建立链路时,需要对链路建立之后的数据交换包进行地址伪装,以使其达到从交换机进出的目的。KLBD 系统可以通过串口终端进行配置,默认为基于负载均衡器本身进行通讯。

服务器反馈处理模块:接收所联接的服务器组内的高可用系统的反馈信息,根据反馈信息,确定目前正在正常运行的服务器,从而确定正确的负载均衡策略。

串口控制台接口:可以通过它改变负载均衡器的配置。

3 系统特性分析

从软硬件的设计可以看出,系统有以下三个特性:第一、系统配置了一块反馈网卡和服务器反馈处理模块,使负载均衡系统具备了根据服务器的运行情况进行动态调配链路的功能,当某台服务器运行出错、掉电、甚至某个服务出错等出现时,负载均衡器会相应将其提供的服务分配给其他服务。第二、网络包采集模块设计在数据链路层,提高了效率,数据链路层基本可以说是可以控制网络包的最低层网络协议,在这时对网络包进行控制,使网络包减少了流转的时间,达到快速转接的目的。第三、采用交换机数据通路管理模块,使负载均衡器仅仅负责数据的链路管理,而不负责数据的流转,大量数据的流转通过交换机进行,这样处理以后,负载均衡可以达到线速转接的目的。

4 基于 KLBD 的 Web 网站结构

从上面的分析可以看出,用上面的网络结构来构建的网络还存在一个问题,即负载均衡器将会成为网络的瓶颈,因为访问服务器的数据包和服务器的回答包即双向包都要通过负载均衡器。但我们可以看到,如果用负载均衡器只作为负载均衡使用,而不作为数据交换使用,这就是说,它仅仅管理访问服务器的链路,而不作为数据通路,这样一来,它的吞吐率就将提高几个数量级,当然它取决于使用的网卡和主机的速度。这样的网络结构如图3。

结束语 负载均衡的思路下多台服务器为对称方式,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。然后通过某种负载分担技术,将外部发送来的请求均匀分配到对称结构中的某一台服务器上,而接收到请求的服务器都独立回应客户机的请求。由于建立内容完全一致的 Web 服务器并不复杂,可以使用服务器同步更新或者共享存储空间等方法来完成,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。应该说现在很多网站采用了很多类型的负载均衡产品,加上网络基础设施的快速发展,网络拥挤

得到了一定程度的缓解,但随着网络的快速发展,老的负载均衡产品不能满足新的需要,KLBD 的设计

较好地适应和满足了当前网络快速发展的需要。

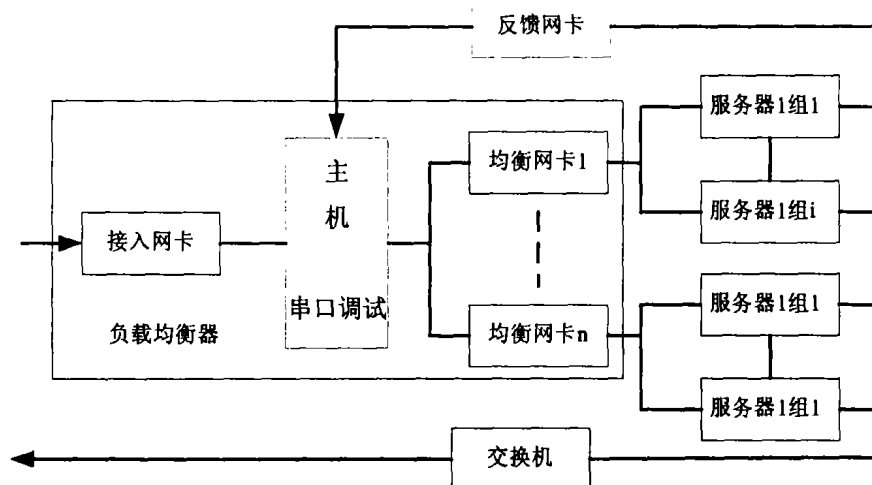


图3 基于 KLBD 的网络结构

参考文献

1 王波. 使用网络地址转换实现多服务器负载均衡[Z]. <http://www.lslnet.com> 2000-9-23

2 李冬. 动态负载均衡 DNS 简介[Z]. <http://www.lslnet.com> 2000-3-31

(上接第180页)

中 F_T 是相似度的阈值。 F_T 的大小影响聚类的结果,根据试验的结果选择 $F_T=0.237$ 能得出较好的聚类。

Step4. 输出聚类结果。

Step5. 根据聚类结果综合生成索引页面。

5 试验结果

本文采用的数据文件,是从网站下载的自由匿名 Web 日志。数据文件从 <http://www.fin.ne.jp/~yokubota/mail-log/log.txt> 处下载^[5],它记录了2003年9月30日00:05:16到16:03:13所有的访问信息。共有331个不同的访问 IP 对243个页面进行了1421次的访问。我们从中抽取出了1103次有效的会话过程。

我们在上述数据文件中使用 CHAMELEON 算法进行了聚类,表2给出了部分聚类的结果。通过对结果的观察,我们发现 CHAMELEON 算法的聚类结果是合理的相似的页面被聚集在一起。根据聚类的结果我们可以生成每一个类的索引页面。

结束语 本文就 CHAMELEON 算法应用于自适应 Web 站点的构建,给出了距离公式等算法实现的具体细节,并在此基础上针对具体的数据进行了程序实现。初步的试验表明,该方法能应用于各种不同情况的 Web 日志,它能发现不同的聚类结果,效果明显优于其它算法。

表2

C_1	/~yokubota/giga/index.html /~yokubota/giga/giga11.html /~yokubota/giga/Win98&METips&Tricks.html ...
C_2	/~yokubota/mandsui2/mandsui2.shtml /~yokubota/mandsui1/mands1e.shtml /~yokubota/mandsui2/syou18.mid ...
C_3	/~yokubota/jv/j-tech1.html /~yokubota/jv/index.html /~yokubota/jv/j-syxed.html ...
...	...

参考文献

- 1 Perkowitz M, Etzioni O. Towards Adaptive Web Sites: Conceptual Framework and Case Study. *Artificial Intelligence*, 2000, 118: 245~247
- 2 Karypis G, Han E-H, Kumar V. CHAMELEON: A hierarchical clustering algorithm using dynamic modeling. *COMPUTER*, 1999, 32: 68~75
- 3 Karypis G, Kumar V. Multilevel k-way partitioning schema for irregular graphs. *Journal of Parallel and Distributed Computing*, 1998, 48(1): 96~129
- 4 Kantardzic M. 数据挖掘: 概念、模型、方法和算法. 冯四清, 等译. 北京: 清华大学出版社, 2003
- 5 王晔, 李德毅. 自适应 Web 站点的访问数据聚类方法. *中国人工智能进展*. 北京: 清华大学出版社, 2001