

一种基于负载流量矩阵的均衡网关技术研究^{*}

秦 勇

(茂名学院, 广东 茂名 525000)

摘 要: 为了达到最佳的负载均衡效果, 将负载分给多个服务器分担。负载控制器需要根据网络流量状态和服务器资源来分配负载, 这就需要应用优化的负载分配策略, 达到平均分配负载的目的。讨论了特定流量条件下网关数目的计算模型, 分布式路由策略实现负载分配的策略的方法, 并研究了基于网络地址转换在 FreeBSD unix 环境下实现负载均衡网关的方法。

关键词: 负载均衡; NAT; FreeBSD; 网关

中图分类号: TP393.03 **文献标识码:** A **文章编号:** 0529-6579 (2002) S1-0028-04

服务器需要具备提供大量并发访问服务的能力, 服务器的处理和 I/O 能力成为了提供服务的瓶颈。由于单台服务器的性能总是有限的, 必须采用多服务器和负载均衡技术才能满足大量并发访问的需要。

DNS 负载均衡技术是在 DNS 中为多个地址配置同一个名字, 因而查询这个名字的客户机将得到其中一个地址, 从而使得不同的客户访问不同的服务器, 达到负载均衡的目的^[1]。DNS 负载均衡是一种简单而有效的方法, 但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。

反向代理服务器可以将请求转发给内部 Web 服务器, 如果代理服务器能够将请求均匀转发给多台内部服务器, 就能达到负载均衡的目的。反向代理方式下能应用优化的负载均衡策略, 每次访问最空闲的内部服务器来提供服务。但是随着并发连接数量的增加, 代理服务器本身的负载也变得非常大, 最后反向代理服务器本身会成为服务的瓶颈。

有的协议内部支持与负载均衡相关的功能, 例如 HTTP 协议中的重定向能力等, 但它依赖于特定协议, 因此使用范围有限。

根据现有的这些负载均衡技术, 使用软件方式实现网络地址转换的负载均衡的方式即可以弥补硬件负载均衡器的不灵活, 又可应用优化的均衡策略来实现后端服务器负载分担的最优状态。

支持负载均衡的地址转换网关中可以将一个外部 IP 地址映射为多个内部 IP 地址, 对每次 TCP 连接请求动态使用其中一个内部地址^[2]。

1 负载均衡策略

为了将负载均匀地分配给内部的多个服务器上, 就需要应用一定的负载均衡策略。本文采用启发式搜索算法与遗传算法相结合的方法, 将邻域的解进行比较, 使之始终沿着优化解的方向搜索, 从而得出网关出口线路带宽、路由数目和服务器数目的优化配比方案。

1.1 网关划分原则

(1) 子网间的通信流量尽可能小, 以减少通过网络互连设备的开销;

(2) 子网内的流量尽可能大, 这表明子网内的资源利用率高, 内聚力强;

(3) 流经各网关的流量应尽量趋于平衡以防止因新增网络设备而导致子网性能急剧下降。

1.2 网络负载流量矩阵

网关划分的一个重要依据是网络中任意一组路由和出口网关间的通信量, 因此采用网络流量矩阵描述网络中各流量的分布是一种有效的方法。假设网络中有 n 个网关服务器, 则网络流量矩阵可表征如下:

$$\delta = (a_{ij}) \quad i, j = 1, 2, \dots, n \quad (1)$$

即:

^{*} 收稿日期: 2002-03-21

作者简介: 秦勇 (1970-), 男, 讲师, E-mail: qy@mmcg.edu.cn

$$\delta = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ & & \cdots & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \quad \begin{aligned} a_{ij} &= 0, i = j \\ a_{ij} &= a_{ji}, i \neq j \end{aligned}$$

a_{ij} 代表第 i 站点与第 j 站点间链路的总流量, 因此 $a_{ij} = a_{ji}$ 成立。通过流量矩阵, 可计算出网络的总流量 γ 和每个服务器的输入 / 输出的总流量 ϵ :

$$\gamma = \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_{ij} \tag{2}$$

$$\epsilon = \sum_{j=1}^n a_{ij} \tag{3}$$

1.3 请求的复杂性和分布式路由对策

在一个现实网络中, 客户发起的服务请求类型是多种多样的。在某一特定时间片内, 流经某网关的流量是时间的函数。如果链路堵塞概率加大, 网络的可用性就降低。因此必须找出一个网络流量分布的统计学模型作为路由策略的理论依据。

在对给定拓扑结构的校园网络链路堵塞概率的加权平均值进行了进一步地研究后, 提出了一种考虑了网络链路生存概率的链路加权堵塞概率, 并得到网络链路容量分配的分布式路由对策。

Sanso^[3] 提出了一种评价网络堵塞概率的综合指标, 即网络链路堵塞概率的加权平均值。设每个网关有一条链路相连。考虑一个有 N 个节点和 L 条链路的校园网络, 令每一条链路的流量分别为 $a_1, a_2, \cdots, a_L$, 每一条链路的堵塞概率分别为 $b_1, b_2, \cdots, b_L$, 则可得到网络链路堵塞概率的加权平均值 B 为:

$$B = \sum_{i=1}^L \left(\frac{a_i b_i}{A} \right) \tag{4}$$

其中, A 为所有链路流量之和, 即

$$A = \sum_{i=1}^L a_i \tag{5}$$

通常情况下, 这个指标能够很好地反映整个网络的堵塞概率, B 越大则网络的堵塞概率越大, 网络的通过量越小; B 越小则网络的堵塞概率越小, 网络的通过量越大。我们在具体网络的研究中得到的大量结果也证实了这一点。

对网络链路堵塞概率的加权平均值作进一步的研究, 可得到更有意义的结果。对于 L 个正数为 $b_1, b_2, \cdots, b_L$ 的加权平均值

$$\frac{\rho_1 b_1 + \rho_2 b_2 + \cdots + \rho_L b_L}{\rho_1 + \rho_2 + \cdots + \rho_L} \text{有} \quad b_{21}^{\rho_1} b_{22}^{\rho_2} \cdots b_{LL}^{\rho_L} \leq \left(\frac{\rho_1 b_1 + \rho_2 b_2 + \cdots + \rho_L b_L}{\rho_1 + \rho_2 + \cdots + \rho_L} \right)^{\rho_1 + \rho_2 + \cdots + \rho_L} \tag{6}$$

等号当且仅当 $b_1 = b_2 = \cdots = b_L$ 时成立。

令 $b_i (i = 1, \cdots, L)$ 为网络链路的堵塞概率, $\rho_i (i = 1, \cdots, L)$ 为网络链路的归一化流量, 即 a_i / A , 由不等式 (6) 可得, 网络只有在每条链路的堵塞概率相等时, 网络链路堵塞概率的加权平均值最小, 一般情况下此时网络的通过量最大。

2 实现方法及实验结果

我们在多出口的固定网络拓扑的中型校园网上对网关负载均衡进行了大量的研究。服务器系统由使用快速以太网连接起来的多台 FreeBSD unix 系统组成。每台后端服务器上运行一个守护进程来动态进行网络地址翻译, 而中心控制网关通过静态路由策略控制流经各网关服务器的网络流量, 以进行正确的负载分配。

2.1 支持负载均衡的网关

在 FreeBSD 系统下, 提供了 divert 接口以支持网络地址转换能力。IP 数据包通过系统内核的 ipfw 过滤功能被发送到 divert 接口中, 以便外部守护进程 natd 能接收原始数据包, 处理之后再发回系统内核进行正常的 IP 分发^[4]。

因此根据 FreeBSD 的地址转换结构, 由于它是软件实现的方式, 很容易支持非标准的协议及更优化的反向负载均衡策略, 具备很大的灵活性。

2.2 实验及分析

(1) NAT 的实现

在 FreeBSD 4.4 unix 环境下重新编译的内核, 使它有 2 块网卡, 并加入下面两项 options:

```
options IPFIREWALL
options IPDIVERT
natd 8668 /divert      #启动 NAT 服务
定义分组过滤规则:
ipfw f flush
ipfw add 1000 divert natd all from any to any via fxp1
ipfw add 1100 pass all from any to any
natd interface fxp1
```

(2) 策略路由

ip route 162. 105. 0. 0 255. 255. 0. 0 default gateway ; 指向默认出口路由器
...
ip route 137. 189. 0. 0 255. 255. 0. 0 a1 ; 指向第 1 个网关服务器
...
ip route 143. 89. 0. 0 255. 255. 0. 0 ai ; 指向第 i 个网关服务器
...
ip route 211. 96. 0. 0 255. 255. 0. 0 an ; 指向最后一个网关服务器
...

(3) 流量监控和分析

图 1 分别是校园网出口 24 h 流量监控分析图, 由 mrtg 软件测得。我们由式 (6) 提供的期望值算法对链路 1—3 应用分布式路由策略。进行统一路由分割, 在无失效部件状态时通过量较大的网络其期望值较大, 相应地, 可行性也较大。所以只需将网络链路容量调整到其无失效部件状态时通过量最大, 便可使得网络的可行性最大。因此重点对链路 2 进行了二次路由分割, 增加了多台网关服务器。每台服务器从中心策略路由控制器“获得”流量概率期望值近似相等的路由表, 以突破网关服务器本身端口数目的限制和 CPU、磁盘 I/O 的瓶颈, 实现了二次流量负载均衡。按平均加权值考虑, 每台网关服务器可带 2 000 个用户。

3 结 论

1. 2 讨论了校园网内部各子网流量分配模型并推导了网关数目的计算公式, 1. 3 讨论了在网络链

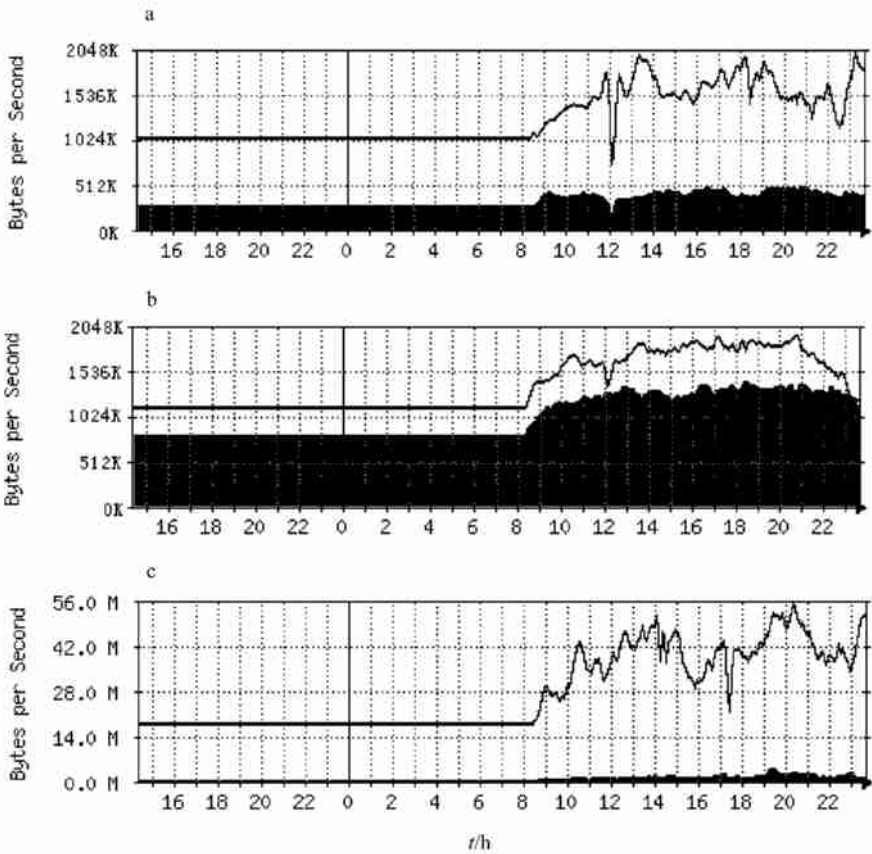


图 1 校园网出口流量监控分析图

(a) 出口为 Cernet; (b) 出口为 Cnuninet; (c) 出口为 GDhednet

路为一个不失效部件前提下流向校园网外部流量的统计学模型，并根据 1. 2 的计算结果正确应用分布式路由策略。使用基于 NAT 的 FreeBSD 网关技术和分布式静态路由策略实现了整个网络负载的均衡。

在固定网络拓扑模型下，通过对网络链路的容量进行分配，以使网络的可靠性最大。

参考文献：

[1] KATA E, BUTLER M, McGRATH R A scalable HTTP

server; the ncsa prototype[J] . Computer Networks and ISDN Systems, 1994, 27: 155— 164.
[2] McCLOGHRIE K. Management information base for network management of TCP/IP-based internets; MIB-II - Network Working Group[M] . Request for Comments; 121.
[3] <http://www.faqs.org/rfcs/rfc1631.html>, 1991, 3[EB/OL] .
[3] SANSO B, et al On the evaluation of telecommunication network reliability using routing models[J] . IEEE Trans Communs, 1991, 39(10) : 1494— 1501.
[4] 王波. FreeBSD 使用大全[M] . 北京: 机械工业出版社, 1999: 284— 293.

The Balanced Gateway Research of FreeBSD Load Based on NAT

QING Yong

(Maoming College, Guangdong, Maoming 525000, China)

Abstract: In order to achieve the optimum load balancing effect, the load should be shared by some servers. The load control machines have to distribute the load according to the state of network outflow and the sources of the servers, so it needs to apply the optimizing load distributed strategy to achieve the goal of average distributed load. This paper has discussed the computing model on the condition of specific outflow, the method of the load distributed strategy that is achieved by distributed route strategy, but also has researched the method of achieving load balancing gateway, which is based on the network address transfer in the FreeBSD Unix.

Key words: load balancing; NAT; FreeBSD; gateway