

第 1 章

Nginx 简介

1.1 常用的 Web 服务器简介

Web 服务器也称为 WWW (WORLD WIDE WEB) 服务器、HTTP 服务器, 其主要功能是提供网上信息浏览服务。

Unix 和 Linux 平台下的常用 Web 服务器有 Apache、Nginx、Lighttpd、Tomcat、IBM WebSphere 等, 其中应用最广泛的是 Apache。而 Windows NT/2000/2003 平台下最常用的服务器则是微软公司的 IIS (Internet Information Server)。

1.1.1 Apache 服务器

Apache 仍然是世界上用得最多的 Web 服务器, 市场占有率达 60% 左右。它源于 NCSAhttpd 服务器, 在 NCSA WWW 服务器项目停止后, 那些使用 NCSA WWW 服务器的人们开始交换用于此服务器的补丁, 这也是 Apache 名称的由来 (pache 补丁)。世界上很多著名的网站都是 Apache 的用户, 它的优势主要在于源代码开放、有一支开放的开发队伍、支持跨平台的应用 (可以运行在几乎所有的 Unix、Windows、Linux 系统平台上), 以及其可移植性等。Apache 的模块支持非常丰富, 虽在速度、性能上不及其他轻量级 Web 服务器, 但是属于重量级产品, 所消耗的内存也比其他 Web 服务器要高。

官方网站: <http://httpd.apache.org/>。

1.1.2 Lighttpd 服务器

Lighttpd 是由一个德国人写的开源软件，其目标是提供一个专门针对高性能网站，安全、快速、兼容性好并且灵活的 Web Server 环境。它具有内存开销低、CPU 占用率低、效能好，以及模块丰富等特点。支持 FastCGI、CGI、Auth、输出压缩（output compress）、URL 重写及 Alias 等重要功能。Lighttpd 跟 Nginx 一样，也是一款轻量级 Web 服务器，是 Nginx 的竞争对手之一。

官方网站：<http://www.lighttpd.net/>。

1.1.3 Tomcat 服务器

Tomcat 是一个开放源代码、运行 servlet 和 JSP Web 应用程序的基于 Java 的 Web 应用软件容器。Tomcat Server 是根据 servlet 和 JSP 规范执行的，因此也可以说 Tomcat Server 实行了 Apache-Jakarta 规范，且比绝大多数商业应用软件服务器要好。但是，Tomcat 对静态文件、高并发的处理比较弱。

官方网站：<http://tomcat.apache.org>。

1.1.4 IBM WebSphere 服务器

WebSphere Application Server 是一种功能完善、开放的 Web 应用程序服务器，是 IBM 电子商务计划的核心部分，它基于 Java 的应用环境，建立、部署和管理 Internet 和 Intranet Web 应用程序。这一整套产品目前已进行了扩展，以适应 Web 应用程序服务器的需要，范围从简单到高级，直到企业级。据 IBM 官方网站介绍，有 10 000 多个企业正在使用 IBM WebSphere，相对于其他流行的 Web 服务器而言，应用的数量很少。

官方网站：<http://www.ibm.com/developerworks/cn/websphere/>。

1.1.5 Microsoft IIS

Microsoft 的 Web 服务器产品为 Internet Information Server (IIS)，IIS 是允许在公共 Intranet 或 Internet 上发布信息的 Web 服务器。它是目前最流行的 Web 服务器产品，很多著名的网站都是建立在 IIS 平台上的。IIS 提供了一个图形界面的管理工具，称为 Internet 服务管理器，可用于监视配置和控制 Internet 服务。

IIS 是一种 Web 服务组件，其中包括 Web 服务器、FTP 服务器、NNTP 服务器和 SMTP 服务器，分别用于网页浏览、文件传输、新闻服务和邮件发送等方面，它使得在网络（包括互联网和局

域网)上发布信息成了一件很容易的事。它提供 ISAPI (Intranet Server API) 作为扩展 Web 服务器功能的编程接口;同时,它还提供一个 Internet 数据库连接器,可以实现对数据库的查询和更新。

IIS 只能运行在 Microsoft Windows 平台、Linux/Unix 平台上,因此须要购买商业的 Windows Server 操作系统。

1.2 Nginx 的发展

Nginx (“engine x”)是俄罗斯人 Igor Sysoev (伊戈尔·塞索耶夫)编写的一款高性能的 HTTP 和反向代理服务器。Nginx 能够选择高效的 epoll (Linux 2.6 内核)、kqueue (FreeBSD)、eventport (Solaris 10)作为网络 I/O 模型,在高连接并发的情况下,Nginx 是 Apache 服务器不错的替代品,它能够支持高达 50 000 个并发连接数的响应,而内存、CPU 等系统资源消耗却非常低,运行非常稳定。

Nginx 已经在俄罗斯最大的门户网站—— Rambler Media (www.rambler.ru)上运行了 3 年时间,同时俄罗斯超过 20% 的虚拟主机平台采用 Nginx 作为反向代理服务器。

在国内,已经有新浪博客、新浪播客、网易新闻、六间房、56.com、Discuz!官方论坛、水木社区、豆瓣、YUPOO 相册、海内 SNS、迅雷在线等多家网站使用 Nginx 作为 Web 服务器或反向代理服务器。

图 1-1 是 Netcraft 公司统计的从 1995 年 8 月至 2009 年 1 月各 Web 服务器的市场占有率曲线图。

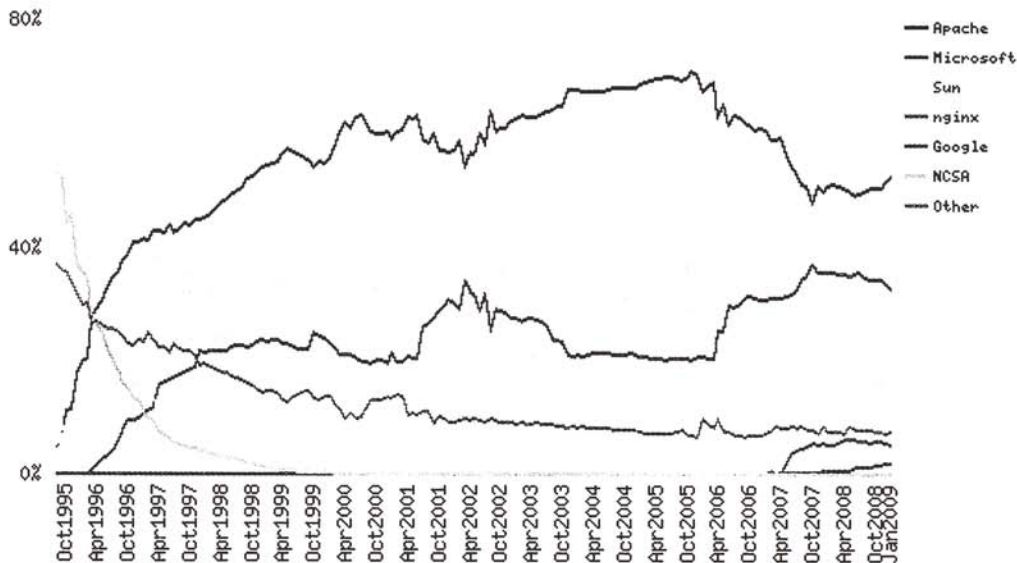


图 1-1 Web 服务器的市场占有率曲线图

2009 年 1 月，对 185 497 213 个网站进行了抽样调查，发现除去 Google 自己开发的仅供自己使用的 GWS、GFE 服务器外，排在前两位的分别是 Apache、Microsoft IIS，而 Nginx 已经超过 Lighttpd，位居第三，详见表 1-1。

表 1-1 Nginx 的市场占有率排名

Web 服务器	2008 年 12 月	占有率	2009 年 1 月	占有率	占有率变化
Apache	95 678 052	51.24%	96 947 298	52.26%	1.02%
Microsoft IIS	63 126 940	33.81%	61 038 371	32.91%	-0.90%
Google	10 455 103	5.60%	9 868 819	5.32%	-0.28%
Nginx	3 354 329	1.80%	3 462 551	1.87%	0.07%
Lighttpd	3 046 333	1.63%	2 989 416	1.61%	-0.02%

1.3 选择 Nginx 的理由

选择 Nginx 有如下一些理由。

1.3.1 它可以高并发连接

官方测试 Nginx 能够支撑 5 万并发连接，在实际生产环境中可支撑 2~4 万并发连接数。这得益于 Nginx 使用了最新的 epoll (Linux 2.6 内核) 和 kqueue (freebsd) 网络 I/O 模型，而 Apache 使用的则是传统的 select 模型，其比较稳定的 Prefork 模式为多进程模式，需要经常派生子进程，所消耗的 CPU 等服务器资源要比 Nginx 高得多。

笔者曾完成 6 台 Web Server 从 Apache 到 Nginx 服务器的迁移 (这 6 台 Web Server 搭建的是一个日均 2 500 万 PV 的分类信息网站，迁移前每台服务器的平均系统负载为 50~60、CPU 使用率为 70%~90%，迁移后平均系统负载为 1~4，CPU 使用率为 20%~40%，见图 1-2。

```
top - 17:59:33 up 10 days,  3:10,  3 users,  load average: 2.69, 2.17, 2.57
Tasks: 300 total,   1 running, 299 sleeping,   0 stopped,   0 zombie
Cpu(s): 30.1% us,  5.0% sy,  0.0% ni, 64.1% id,  0.1% wa,  0.1% hi,  0.7% si
Mem:  8230016k total, 3850368k used, 2379648k free, 588616k buffers
Swap: 2031608k total,   508k used, 2031100k free, 1755208k cached
```

图 1-2 迁移到 Nginx 后的系统负载与 CPU 使用率图

由此可见，处理大量连接的读写，Apache 所采用的 select 网络 I/O 模型非常低效。

Libevent (一个事件触发的网络库，适用于 Windows、Linux、bsd 等多种平台，内部使用 select、epoll、kqueue 等系统调用管理事件机制) 的一张测试结果图，对 select、epoll、kqueue 等作了清

晰的对比, 结果自然是 epoll (Linux 2.6 内核) 和 kqueue (freebsd) 胜出, 见图 1-3。

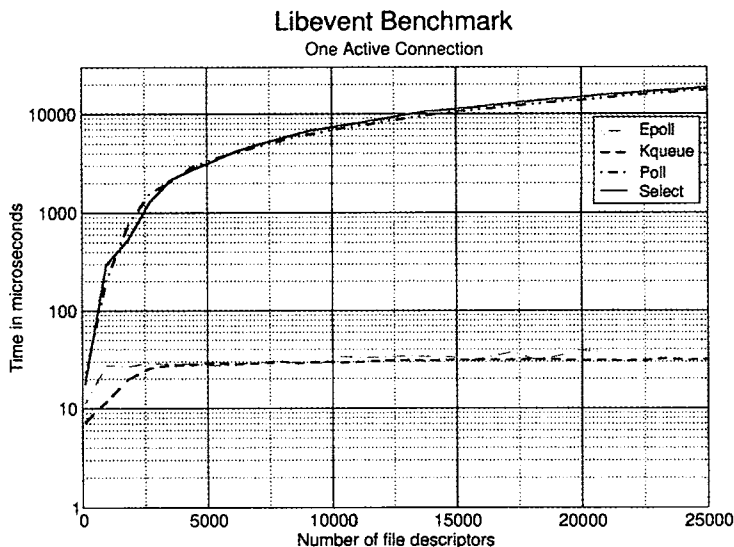


图 1-3 epoll、kqueue、select 等网络 I/O 模型性能测试对比图

图 1-4 显示了在实际的生产环境中, Nginx 支撑高达 28 000 的活动并发连接数。

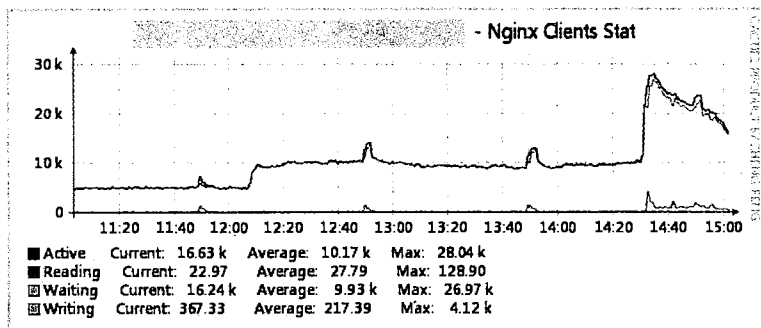


图 1-4 生产环境下的 Nginx 活动并发连接数据统计图

1.3.2 内存消耗少

Nginx + PHP (FastCGI) 服务器在 3 万并发连接下, 开启的 10 个 Nginx 进程消耗 150MB 内存 ($15\text{MB} \times 10 = 150\text{MB}$), 开启的 64 个 php-cgi 进程消耗 1280MB 内存 ($20\text{MB} \times 64 = 1280\text{MB}$), 加上系统自身消耗的内存, 总共消耗不到 2GB 的内存。如果服务器内存较小, 完全可以只开启 25 个 php-cgi 进程, 这样 php-cgi 消耗的总内存数才 500MB。用 Webbench 做压力测试, 在 3 万

并发连接下，访问 Nginx + PHP (FastCGI) 服务器的 PHP 程序，运行速度仍然飞快。

在实际的生产环境下，两台 Nginx + PHP5 (FastCGI) 服务器运行多个复杂性一般的纯 PHP 动态程序，从 Nginx 的日志可以统计出，单台 Nginx + PHP5 (FastCGI) 服务器处理 PHP 动态程序的能力已经超过“700 次请求/秒”（见图 1-5），相当于每天可以承受 6 000 万（ $700 \times 60 \times 60 \times 24 = 60\,480\,000$ ）的访问量，而服务器的系统负载也不算高（见图 1-6）。

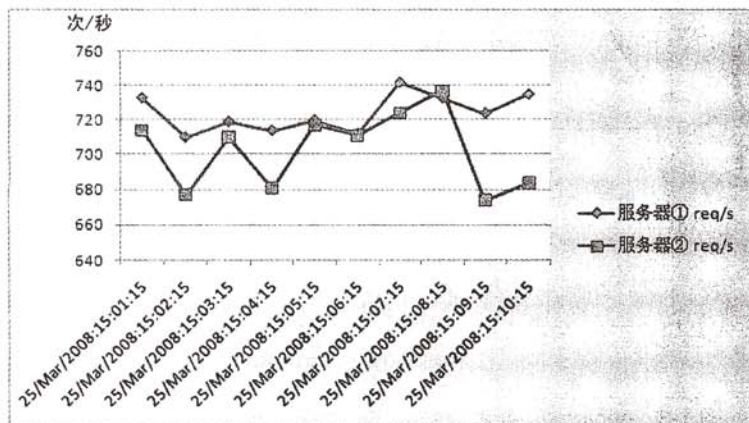


图 1-5 生产环境下的 Nginx+PHP 动态程序处理速度统计图

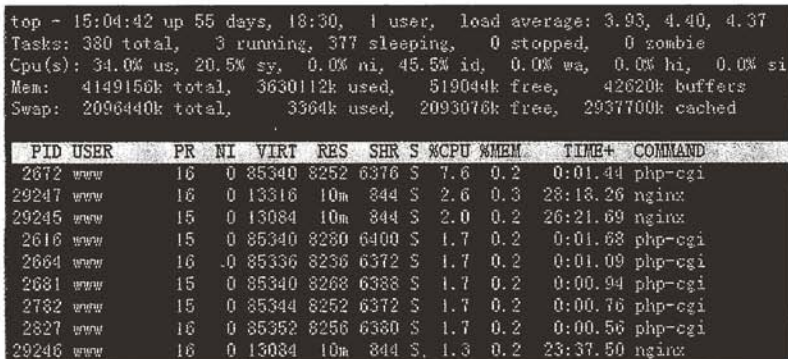


图 1-6 生产环境下的 Nginx+PHP 系统负载与 CPU 使用率图

以下是这两台服务器的配置清单及运行的程序说明：

服务器①：DELL PowerEdge 1950（两个 Intel(R) Xeon(R) 双核 CPU 5120 @ 1.86GHz，4GB 内存）。

服务器②：DELL PowerEdge 1950（一个 Intel(R) Xeon(R) 双核 CPU 5140 @ 2.33GHz，4GB 内存）。

Web 服务器: CentOS Linux 4.4 + Nginx 0.5.35 + PHP 5.2.6RC2 (300 FastCGI Processes, unix-domain socket, with XCache)。

PHP 程序内容: 大量 Memcached 读写操作, 少量 MySQL 读操作, 大量文件队列写操作。

请求数统计方式: 从 Nginx 访问日志中, 统计每分钟的 第 15 秒共有多少条日志记录。

同等硬件环境下, Nginx 的处理能力相当于 Apache 的 5~10 倍。

1.3.3 成本低廉

购买 F5 BIG-IP、NetScaler 等硬件负载均衡交换机需要十多万甚至几十万人民币。而 Nginx 为开源软件, 采用的是 2-clause BSD-like 协议, 可以免费使用, 并且可用于商业用途。

BSD 开源协议是一个给使用者很大自由的协议。协议指出可以自由使用、修改源代码, 也可以将修改后的代码作为开源或专有软件再发布。当你发布使用了 BSD 协议的代码, 或者以 BSD 协议代码为基础做二次开发时, 须满足三个条件:

- (1) 如果再发布的产品中包含源代码, 则源代码中必须带有原来代码中的 BSD 协议。
- (2) 如果再发布的是二进制类库/软件, 则需要在类库/软件的文档和版权声明中包含原来代码中的 BSD 协议。
- (3) 不可以用开源代码的作者/机构名字和原来产品的名字做市场推广。

BSD 代码鼓励代码共享, 但须尊重代码作者的著作权。BSD 由于允许使用者修改和重新发布代码, 也允许使用或在 BSD 代码上开发商业软件, 并进行发布和销售, 因此它是对商业集成很友好的协议。很多的公司、企业在选用开源产品的时候都会首选 BSD 协议, 因为可以完全控制这些第三方的代码, 在必要的时候可以修改或二次开发。

Nginx 所采用的 2-clause BSD-like license 衍生自 BSD 协议, 也就是删掉了 BSD 协议的第 3 个条件——“不可以用开源代码的作者/机构名字和原来产品的名字做市场推广”。

1.3.4 其他理由

配置文件非常简单

网络跟程序一样通俗易懂, 即使非专业系统管理员也能看懂。

支持 Rewrite 重写规则

能够根据域名、URL 的不同, 将 HTTP 请求分到不同的后端服务器群组。

内置的健康检查功能

如果 Nginx Proxy 后端的某台 Web 服务器宕机了，不会影响前端访问。

节省带宽

支持 GZIP 压缩，可以添加浏览器本地缓存的 Header 头。

稳定性高

用于反向代理，宕机的概率微乎其微。

支持热部署

Nginx 支持热部署。它的启动特别容易，并且几乎可以 7 天×24 小时不间断地运行，即使运行数月也不需要重新启动。你还能够在不间断服务的情况下，对软件版本进行升级。

1.4 Nginx 与 Apache、Lighttpd 的综合对比

表 1-2 是 Nginx 与 Apache、Lighttpd 的综合对比情况。

表 1-2 Nginx 与 Apache、Lighttpd 的综合对比

Web 服务器	Nginx	Apache	Lighttpd
反向代理	非常好	好	一般
Rewrite 规则	非常好	好	一般
FastCGI	好	差	非常好
热部署	支持	不支持	不支持
系统压力比较	很小	小	很大
稳定性	非常好	好	一般
安全性	一般	好	一般
技术资料	很少	非常多	一般
静态文件处理	非常好	一般	好
虚拟主机	支持	支持	支持
内存消耗	非常小	很大	非常小

通过表 1-2 可以看出，Nginx 在反向代理、Rewrite 规则、稳定性、静态文件处理、内存消耗等方面，表现出了很强的优势，选用 Nginx 取代传统的 Apache 服务器，将会获得多方面的性能提升。