

网站架构演变技术研究

张建超 , 张连堂

(河南大学计算机与信息工程学院, 开封 475004)

摘要: 网站架构过程中采取的方案, 直接决定其可用性、负载能力以及并发性能。从网站演变脉络入手, 通过研究网站开发时所需要的相关技术, 从应用的角度探讨解决网站面临的高负载和高并发问题的方法, 给出一种综合性能较高的网站构架方案。

关键词: 页面缓存; 负载均衡; 高并发

0 引言

可用性是决定网站或企业网络成功的重要因素之一, 信息架构是网站设计中可用性的重要因素之一。当然还包括其他因素, 但是, 忽视信息架构就会让你身陷危险之中^[1]。本文从网站架构演变过程和技术来讲起, 分析网站架构的演变和相关技术, 并给出高性能架构网站。

1 网站架构演变

1.1 数据与 Web 页分离缓存页面

当一小型网站随访问量增加逐渐系统的压力发现越来越大, 响应速度越来越慢, 而压力比较明显的问题是数据库和应用互相影响, 数据库易出现问题, 原因是访问数据库的操作太多, 导致数据连接竞争激烈, 响应变慢。但数据库连接又不能过多, 连接过多数据库机器压力会很高, 因此采用缓存机制来减少数据库连接资源的竞争和对数据库读取的压力。首先会选择采用 Squid 等类似的机制来将系统中相对静态的页面(例如时间较长更新的页面)进行缓存也可以采用将页面静态化的方案, 这样程序可以不作修改, 就能够很好地减少对 Web Server 的压力以及减少库连接的竞争, 数据连接资源于是开始采用 Squid 来做相对静态的页面的缓存, 将应用和数据库从物理上分离。



图 1 分离数据

1.2 在 Web Server 上实现页面片段和数据缓存

增加缓存后, 系统整体速度提升, Web Server 的压力也下降, 随着访问量的增加, 系统又开始变慢, 现在把那些动态页面里相对静态的部分也缓存起来, 因此采用类似 ESI 之类的页面片段缓存策略, 于是开始采用 ESI 来做动态页面中相对静态的片段部分的缓存。系统中存在一些重复获取数据信息的地方, 像获取用户信息等, 将这些数据信息也缓存起来, 于是将这些数据缓存到本地内存, 系统响应速度恢复正常, 数据库的压力也降低了, 如图 2 数据缓存所示。

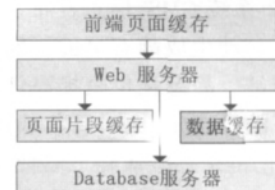


图 2 数据缓存

收稿日期: 2012-01-12 修稿日期: 2012-02-02

作者简介: 张建超(1983-), 男, 河南沈丘人, 硕士, 研究方向为计算机网络教育

1.3 增加 Web Server 和数据分库

Web Server 机器的压力在高峰期会上升到比较高,这时增加一台 Web Server,同时数据库写入、更新的这些操作部分、数据库连接的资源竞争非常激烈,表间关联是一种是一种强制措施,建立后,对父表(Parent Table)和子表(Child Table)的插入更新删除操作均要占用系统的开销^[2],导致了系统变慢。解决的方案有数据库集群和分库策略,集群方面有些是 SQL 数据库支持的并不是很好,因此分库会成为比较普遍的策略,分库也就意味着要对原有程序进行修改,修改实现分库后,系统速度比以前快,如图 3 增加服务器所示。

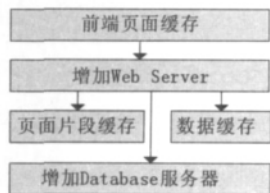


图 3 增加服务器

1.4 数据分表、DAL 和分布式缓存

数据访问过程中,随着系统的不断运行,数据量开始大幅度增长,分库后查询仍然慢,于是按照分库的思想做分表,增加一个通用的框架来实现分库分表的数据访问,因为数据量太大,导致现不可能将缓存存在本地,同步的方式,需要采用分布式缓存方案,大量的数据缓存转移到分布式缓存上。

2 网站架构技术

2.1 实现 HTML 静态

前台实现完全的静态化,HTML 静态页面效率会最高,消耗最小。频繁更新的网站,手动去实现比较困难,这样用信息发布系统 CMS,信息发布系统可以实现最简单的信息录入自动生成静态页面。而交互性要求很高的社区类型网站,静态化也是提高性能的必要手段。如果不能很好的静态化,面对高负载访问,就不能承受。HTML 静态化也是某些缓存策略使用的手段,对于系统中频繁使用数据库查询,可以使用。

2.2 大容量数据与页面分离

对于 Web 服务器,最消耗资源是大容量数据,为了降低提供页面访问请求的服务器系统压力,可以将大容量数据与页面进行分离,这是一般会采用的策略,

都有独立的大容量数据服务器。在大容量数据服务器上,进行不同的优化配置,例如 Apache 在配置内容类型的时候可以尽量少支持,尽可能少的下载模式,保证更高的系统消耗和执行效率。

2.3 数据库集群和库表散列

当网站需要使用数据库,有大量的数据库访问请求,数据库的瓶颈很快就能显现出来,这时一台数据库将很快无法满足应用,于是需要使用数据库集群。数据库集群由于在架构、成本、扩张性方面都会受到所采用 DB 类型的限制,需要从应用程序上来改善系统架构,库表散列是常用有效的解决方案。河南省招生办公室高招网报志愿系统就是采用了这样的架构,将用户信息进行数据库分离,然后 ID 进行散列数据库和表,最终可以在配置文件中简单的配置便能让系统随时增加一台低成本的数据库进来补充系统性能。

2.4 动态数据缓存

网站开发中缓存是非常重要的。就是将动态数据存储到缓存文件中,动态网页直接调用这些文件,不必再访问数据库,许多网站都大量使用这种缓存技术。架构方面的缓存,对 Apache 比较熟悉的人都能知道 Apache 提供了自己的缓存模块,也可以使用外加的 Squid 进行缓存,这两种方式均可以有效地提高 Apache 的访问响应能力。网站程序开发方面的缓存,常用的缓存方案是 Memcache,可以在 Web 开发中使用,可以实时把数据、对象等内容进行缓存,策略灵活。

2.5 负载均衡

负载均衡是建立在网络结构之上的一种调度策略,能有效地扩展服务器带宽和增加吞吐量,增强网络数据处理能力^[3]。硬件四层交换使用第三层和第四层信息包的报头信息,根据应用区间识别业务流,将整个区间段的业务流分配到合适的服务器进行处理。第四层交换功能就象是虚 IP,指向物理服务器。软件四层交换负载均衡的使用策略就是在软件或者硬件四层交换的基础上搭建 Squid 集群,在很多网站包括搜索引擎上被采用,这样的架构低成本、高性能还有很强的扩张性,随时往架构里面增减节点都非常容易。

2.6 分页数据存取

数据分页存取 ASP.NET DataGrid 提供了非常好的能力:数据分页支持。启用 DataGrid 中的分页功能,则每次只显示固定数量的记录。此外,分页用户界面也会

显示在 DataGrid 底部用于导航记录。分页用户界面允许向前向后导航所显示的记录,一次显示固定数量的记录。DataGrid 分页需要将所有数据绑定到此栅格控件(Grid)。当记录数不断增加,此应用程序的性能便会变得慢,因为每次请求所要发送的数据会越来越多。

3 高性能网站架构方案

基于网站架构演变技术分析,一个高性能的网站具有前台静态页面缓存,为增加网站生动灵活的同时,大容量数据是一个网站不可少的一部分,例如大量图片,这时大容量数据应与页面分离。如果网站用用到数据库数据量大比较时(例如十万条数据以上),应采用数据库集群和库表散列。当然反复用到的数据要缓存起来,以便减少数据库访问的次数,防止数据库的瓶颈问题出现。对于用户访问量比较大时,同时在线人数十万以上的网站,要从硬件和软件两方面保证负载均衡。数据访问做到 DAL 的读写分离存储问题。同时做好网站内部分页数据存取,注意多页数据的分页存取。下面给出了高性能网站架构,采用廉价的存储方案,部署分布式应用程序和采用廉价的服务器群。如图 4 高性能服务器架构所示。

4 结 语

Web 站点不是只靠新技术就可以方便快捷地搭建起来的。丰富的开发经验和设计经验、良好的团队合作,甚至需要对硬件一定程度的了解都是构建起一个

成功的大规模可扩展的 Web 站点所必需的^[4]。网站架构在实践应用中还要考虑大量数据存取、数据并发、文件存储、数据关系处理、数据索引、分布式处理及安全性、数据同步和集群处理。

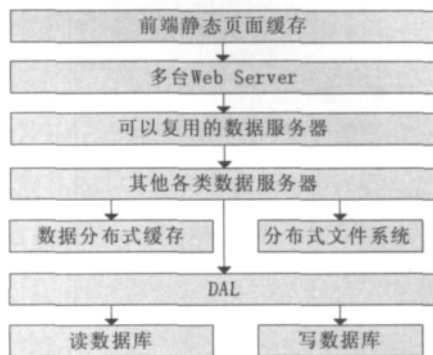


图 4 高性能服务器架构

参考文献

- [1] Peter Morville Louis Rosenfeld. Web 信息架构:设计大型网站(第 3 版). 陈建勋译. 电子工业出版社,2010(8)
- [2] 孙继红. 基于 .NET 三层架构的网上商城设计与实现. 电脑编程技巧与维护,2008(16)
- [3] 周莹莲,刘甫. 服务器负载均衡技术研究[J]. 计算机与数字工程,2010(9)
- [4] Cal Henderson. 构建可扩展的 Web 站点. 徐宁译. 电子工业出版社,2008(3)

Research on the Website Architecture Evolution Technology

ZHANG Jian-chao , ZHANG Lian-tang

(College of Computer and Information Engineering, Henan University, Kaifeng 475001)

Abstract: The solution is applied in Website architecture, it directly determines the Website of the availability, load capacity and concurrency. Starting from the Website evolution, by studying the related technology required in developing Website, gives a high performance Website framework from the application point of view on solving the high load and high concurrency problems, the site development faces.

Keywords: Page Cache; Load Balance; High Concurrency



知网查重限时 7折 最高可优惠 120元

本科定稿，硕博定稿，查重结果与学校一致

立即检测

免费论文查重: <http://www.paperyy.com>

3亿免费文献下载: <http://www.ixueshu.com>

超值论文自动降重: http://www.paperyy.com/reduce_repetition

PPT免费模版下载: <http://ppt.ixueshu.com>

阅读此文的还阅读了:

1. [浅析大型网站的基础架构](#)
2. [云桌面技术架构与发展研究](#)
3. [网站备案技术研究](#)
4. [钢铁行业SAP/MES架构的演变及技术路线](#)
5. [网站架构中集团型多分公司网站的策划、架构和实施](#)
6. [网站架构演变技术研究](#)
7. [跨境电商团队架构的演变](#)
8. [网站动态应用系统技术平台架构分析](#)
9. [各国监管架构及演变](#)
10. [浅谈ASP.NET网站建设中MVC与三层架构技术](#)
11. [网站管理分层架构技术](#)
12. [银行组织架构或向平台型架构演变](#)
13. [新媒体网站安全技术架构与安全体系的建立](#)
14. [政府服务型网站架构的研究](#)
15. [电力营销技术系统架构研究](#)
16. [大型网站的架构研究及解决方案](#)
17. [云计算技术在网站群安全架构中的应用研究](#)
18. [设计新闻网站技术架构的要点](#)
19. [网站后端应用系统架构优化](#)
20. [大型网站架构演变和知识体系](#)
21. [浅谈网站架构建设](#)
22. [大型网站应用技术架构演变的研究](#)
23. [唯品会峰值系统架构演变](#)
24. [地方新闻门户网站架构研究](#)
25. [政府网站群“无锡模式”的技术架构解析](#)

- [26. 高校Web架构的安全技术研究](#)
- [27. 大型网站架构演变和知识体系](#)
- [28. 大型网站的架构模式研究](#)
- [29. 新媒体网站安全技术架构与安全体系建立研究](#)
- [30. 网站架构的优化](#)
- [31. 基于网站群架构的校园门户网站建设研究](#)
- [32. 网站前端技术及其对网站性能的影响研究](#)
- [33. 云计算架构及其技术研究](#)
- [34. 基于Web标准的精品课程教学网站技术架构研究](#)
- [35. 高校档案网站技术架构方案之比较研究](#)
- [36. BBC公共服务数字电视的技术架构演变](#)
- [37. 网站图片的优化技术研究](#)
- [38. 大型网站高性能架构研究](#)
- [39. 企业组织架构演变路径分析](#)
- [40. LAMP网站架构方案与实施探讨](#)
- [41. 云计算及其架构技术研究](#)
- [42. 基于LAMP技术架构的SNS网站系统的开发与研究](#)
- [43. 灾备架构研究与技术实现](#)
- [44. 对外汉语教育网站的架构与内容分析研究](#)
- [45. 大型网站架构中主动防御的代理WEB技术研究](#)
- [46. 基于LAMP技术架构的网站方案研究](#)
- [47. 基于Hybrid技术架构的APP技术方案研究](#)
- [48. 基于Asp.net技术的网站开发架构设计](#)
- [49. 开放式教学资源库网站架构研究](#)
- [50. 基于FNMP架构的高校网站建设方案研究](#)