

文章编号: 2096-1472(2017)-11-30-03

大型网站应用架构对比研究

李绪成, 刘 鑫

(大连东软信息学院软件工程系, 辽宁 大连 116023)

摘 要: 随着使用互联网的人越来越多, 网站本身业务的扩展和用户的增多, 网站面临着高并发访问、大数据的压力, 这时需要对网站进行优化和改进。并发访问量和数据量的瓶颈也成为评价一个大型网站性能的标准。然而怎样改进和优化网站成为一个问题。本文通过研究现今具有代表性的大型网站应用架构, 得出大型网站的通用优化方案, 以及一套可供参考的大型网站基本架构, 为网站设计者提供大型网站优化通用方案、网站改进思路和改进方法。

关键词: 网站架构; 网站优化; 缓存; 大数据; 搜索

中图分类号: TP311 **文献标识码:** A

Comparative Research on the Application Architecture of Large Websites

LI Xucheng, LIU Xin

(Department of Software Engineering, Dalian Neusoft University of Information, Dalian 116023, China)

Abstract: With more and more people using the Internet, the business expansion of websites and the user increase of websites, websites are facing great pressure of high concurrent access and big data, so it is necessary to optimize and improve websites. The bottleneck of concurrent access and data volume is also a standard to evaluate the performance of a large website. However, how to improve and optimize the website has become a problem. By studying the typical application architecture of large websites, the paper proposes the general optimization scheme of large websites and the basic framework of large-scale websites for reference, providing website designers with general solutions, improvement ideas and improvement methods to large websites.

Keywords: website architecture; website optimization; cache; big data; search

1 引言(Introduction)

大型网站分为三大类: 电商网站、视频网站和大数据搜索网站, 通过这三类网站来研究大型网站应用架构。其中最能代表这三类网站的大型网站是: 淘宝、京东、优酷和百度。淘宝和京东是经历过双十一考验的, 其性能是毋庸置疑的; 优酷是中国最大的视频网站之一, 日均播放量上亿, 可见其网站性能很稳定; 百度网站是全球最大的中文搜索引擎、最大的中文网站, 存储了万亿级的网页数量, 其对大数据的处理是值得学习的。这些网站都是研究大型网站的最佳对象。

通过对上述网站进行对比研究得出大型网站典型优化方案, 以及一套可供参考的大型网站基本架构, 为网站设计者提供网站改进思路和改进方法, 以及网站优化方案。

2 网站基本架构(Basic architecture of website)

大型网站架构分为网站前端、应用端和数据库端。每一个部分都有自己的使命, 它们协同合作形成一个完整的网站^[1,2]。大型网站的基本架构如图1所示。

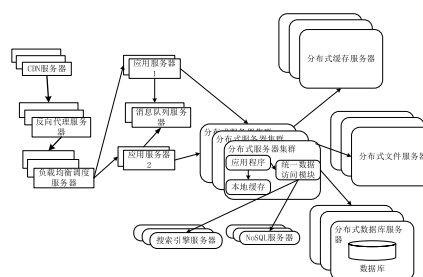


图1 大型网站基本架构图

Fig.1 Basic architecture of large websites

下面从分别从网站前端、应用服务器和数据库三个方面描述改进方法。

3 网站前端(Front end)

网站前端的作用是加快用户访问速度, 过滤恶意请求, 实现负载均衡。

3.1 CDN服务器

CDN是通过在网络各处放置节点服务器所构成的在现有的互联网基础之上的一层智能虚拟网络, 能够实时地根据网络流量和各节点的连接、负载状况, 以及到用户的距离和响

应时间等综合信息将用户的请求重新导向离用户最近的服务节点上,能够加快响应效率,解决网络拥堵问题。

优酷网站建立了比较完善的CDN,可以保证分布在全国各地的用户能访问到离自己最近的服务器,并把服务状况最好的视频服务器地址反馈给用户,让用户可以以最快的速度获取到视频资源,这也是优酷网站比其他视频网站更流畅的原因。

淘宝网站也在全国各地建立了数百个CDN节点。当用户发起请求时,浏览器访问DNS服务器进行DNS域名解析。然后根据用户的IP将访问分配到不同的入口。如果客户的IP属于电信运营商,那么就会被分配到同样是电信的CDN节点,保证用户访问的CDN节点是离用户最近的。

一些前段框架能够为CDN提供便利,例如Angular JS技术^[3]。

3.2 反向代理服务器

通过反向代理服务器处理HTTP请求,在用户和Web服务器之间建立一个屏障,可有效地防止网络攻击,同时还可以实现负载均衡和缓存功能。现在功能比较完善的反向代理服务器有Squid和Nginx^[4,5]。不过Squid不支持多核,但是磁盘缓存容量比较大,支持外部文件读取、热加载和热启动,性能中等;而Nginx支持多核、集群,但不支持外部文件读取,性能较强。反向代理网站架构图如图2所示。

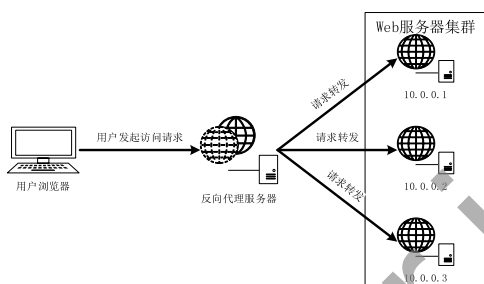


图2 反向代理网站架构图

Fig.2 Reverse proxy website architecture diagram

3.3 负载均衡调度服务器

部署负载均衡调度服务器,专门负责负载均衡,可以将请求分发到应用服务器集群下的任何一个服务器上。随着访问人数的增加,增加服务器就可以,突破了负载压力瓶颈^[6]。

现在比较常用的技术是LVS(Linux Virtual Server, Linux虚拟服务器),LVS技术是为了解决负载均衡而存在的一种技术,它主要的功能是在现有网络结构之上建立一个虚拟的服务器集群系统,用廉价、有效、透明的方法去扩展网络设备和服务器的带宽,增加网站吞吐量,增强对网络数据的处理能力,提高网络的灵活性和可用性。淘宝网站的负载均衡系统采用的是LVS技术。只是这种负载均衡系统的构建是在Linux操作系统上,其他操作系统不行,并且需要重新编译Linux操作系统内核,对系统内核的了解要求很高。LVS集群采用IP负载均衡技术和基于内容请求分发技术,调度器具有很好的吞吐率,将请求均衡地转移到不同的服务器上执行,并且调度器能自动屏蔽掉服务器的故障,从而将一组服务器构成一个高性能的、高可用的虚拟服务器。

4 应用端(Application end)

应用端进行业务处理,业务与业务之间需要通过消息队列实现数据同步,而对于一些热点数据和常用到的数据,需

要缓存到分布式缓存服务器上。

4.1 应用服务器

Web应用程序驻留在应用服务器(Application Server)上,所以应用服务端需要做好业务的拆分,减少业务间的联系,做到高内聚低耦合,即当某个业务无法正常使用时,不影响其他业务的使用。在应用服务器上还应该做好本地缓存,对于热门或常用的数据做好本地缓存。

Session在Web应用中占有很重要的位置。它的主要功能是用来保存用户的状态信息。Session在集群中存在共享的问题。现在解决Session共享问题有三种方法。第一种方法是通过负载均衡设备实现会话保持,第二种方法是采用Session复制,第三种方法则是采用集中式缓存。第二种方式严重影响了集群环境的可伸缩性,不利于集群的横向扩展,即使是采取两两复制的方式也会造成集群内部网络负载严重,还容易造成网络垃圾。

大型网站的业务拆分是很细的,也意味着应用服务器会很多,而所有的应用都会连接数据库,这会导致数据库连接资源不足,从而拒绝服务,所以应该把它们一些共用的业务提取出来,独立部署,由这些可复用的业务连接数据库,提供共用业务服务。而应用系统只需管理用户界面,通过分布式服务调用共用业务服务完成具体业务操作。

4.2 消息队列服务器

对于Web网站,高并发访问是个大问题,异步能很好地解决这个问题。而异步的核心就是消息队列。当用户请求的数据发送到消息队列后就可以立即返回,不需要等待写入数据库完成,这时进程从消息队列中获取数据,再将数据异步写入数据库,并且消息队列服务器处理速度远快于服务器。Notify是淘宝根据自身业务需要量身定制的一款消息中间件,是支撑“双十一”最核心的系统之一。消息系统的核心作用是:解耦、异步和并行。

4.3 分布式缓存服务器

网站的访问特点也符合二八定律,即80%的业务访问在20%的数据上。缓存服务器是专门用来缓存这20%的数据。一般的本地缓存受限于应用服务器内存的大小,而专门的缓存服务器则不受限制,可以将大量的热点数据或者常用到的数据缓存在缓存服务器上,使用分布式进行部署,从而缓解数据库服务器的压力。

Memcached是一个高性能的分布式内存对象缓存系统,它可以有效减轻动态Web应用数据库的压力。它在内存中缓存对象和数据,当请求这些数据时,直接返回,有效地减少了读取数据库的次数,从而提高动态数据库驱动网站的速度。它的客户端可以用任何语言来编写,并且通过Memcached协议与守护进程通信。淘宝的TAIR就是基于Memcached开发的key/value数据库。

5 数据库端(Database end)

数据库端主要是完成数据和文件的存储和提取。大型网站的数据和文件数量是巨大的。存储需要分布式存储,提取需要搜索引擎协助完成。

5.1 分布式文件服务器

分布式文件服务器主要是用来存储图片、视频一类的文件。分布式文件系统(Distributed File System)是指文件系统管理的物理存储资源并不是完全直接连接在本地节点上,

而是利用网络和节点相连接。分布式文件系统支持多台主机通过网络同时访问共享文件和存储目录,实现了多台计算机上的多个用户能够共享文件和存储资源。在分布式文件系统环境下,客户端节点并不直接访问底层的块存储介质,而是使用网络协议进行文件操作交互。因此,分布式文件系统是对存储数据的共享,而不是对存储物理资源的共享。

一个搜索引擎对于访问热点的处理也极其重要。一般的网站对于访问热点传统的解决方案是缓存,包括文件块缓存、数据记录缓存、查询结果缓存;而百度使用分布式文件系统将请求天然地打散到数千个节点上,实现了快速分裂和迁移,缓解了高并发访问热点对数据库服务器的冲击,使网站在高并发访问热点时能够保证用户的体验。

5.2 分布式数据库服务器

大型网站的数据是海量的,一个数据库服务器不可能存储所有的数据,所以需要分布式部署数据库服务器。分布式部署需要对数据库进行合理的拆分。优酷网站在数据库方面做了很多的优化,包括MySQL主从复制、分区和分片,将用户按id以哈希算法进行分组,并且把该用户或者该组用户的数据存储到一个Sharding中。这样当用户数量增加时,只需要增加一台数据库服务器就可以。

使用MySQL做数据库,它有主从复制功能,所以还可以分为主数据库服务器和从数据库服务器。主数据库服务器负责数据的写,而从数据库服务器负责数据的读。其实主从复制原理简单地讲就是主数据库将操作行为写在日志文件里,从数据库通过日志文件和主数据库保持一致。这样做的目的就是实现读写分离,缓解数据库服务器的压力。当数据的存储和检索越来越复杂时,使用NoSQL和搜索引擎也是大型网站解决大数据的方案。

百度网站开发了自己的分布式数据库系统Tera。Tera是一个高性能、可伸缩的结构化数据存储系统,被设计用来管理搜索引擎万亿量级的超链与网页信息。数据需要实时的分析和高效的访问,所以百度使用三维数据模型组织数据(包括行键、列名和时间戳全局排序);并且百度网站还使用了多级缓存系统,将现代服务器硬件内存大、SSD盘和万兆网卡的性能优势充分利用,实现了数据库的高吞吐数据量和数据库的水平扩展。

5.3 搜索引擎服务器

搜索引擎的作用是在数据库查询数据时提供一种有效的查询机制,确保数据能准确地查找。

电商网站的搜索的特点:召回率要求高,电商数据结构化,信息更新时效性高,用户个性化需求高。京东网站的搜索服务有三大系统:爬虫系统、离线信息处理系统、索引系统。商品搜索引擎的核心是建立商品索引,而建立索引需要详细的商品信息数据。京东利用大数据平台的数据库抽取接口和中间件系统,实现了站内商品爬虫系统,用来抽取数据库中的商品信息和及时发现变化的商品信息。然后利用离线信息处理系统来建立商品搜索引擎的待索引数据,包括全量待索引数据和增量待索引数据。索引系统的主要功能是把以商品为维度进行存储的待索引数据,转换为以关键字为维度进行存储的数据,用于搜索引擎上层服务的调用。对于一些比较热门如“衣服”“裤子”之类的查询词,其查询结果会比较多,原始的查询结果就有几千万之多,如果对这些结

果进行一一地处理,那么性能会变得非常差。这时,需要从用户的角度进行分析,一个查询只有排在最前面的结果对用户来说才是有意义的。通过分析用户翻页的次数等数据,可以得到截断保留top N的结果。

百度搜索引擎由四部分组成:蜘蛛系统、监控系统、索引数据库和检索系统。百度网站对搜索请求的处理过程如下:系统在得到搜索关键字后,使用蜘蛛自动搜索超链接库,根据链接从网页库中得到相应的网页,在通过PageRank(一种根据网页之间相互的超链接计算的技术)得到网页排名,然后通过正排计算(将网页分成一个个的关键词)得出正排库,再通过倒排计算(用网页中的关键词来给网页建立另外一个索引库)得出倒排索引,这时候检索系统将结果反馈给用户。百度搜索流程图如图3所示。

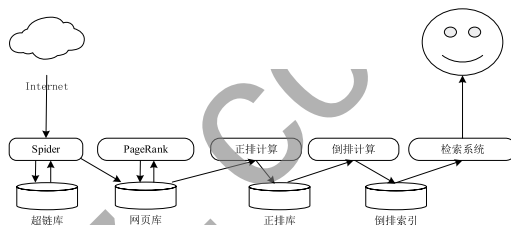


图3 百度搜索流程图

Fig.3 Search flow of Baidu

6 结论(Conclusion)

通过对各大型网站的研究,得出大型网站通用优化方案。网站前端优化主要是减少HTTP请求、使用CDN、使用反向代理、使用负载均衡等;网站应用端优化主要采用集群、异步、多线程、资源复用、数据结构、分布式、缓存等;网站数据端优化主要包括数据冗余、数据库拆分、分布式存储、NoSQL等。

该优化方案能够为网站设计者提供参考。该方案的一些思想在已经在某些系统中使用^[7,8]。

参考文献(References)

- [1] 周强,谢靖,赵华茗.大型网站的架构研究及解决方案[J].计算机科学,2017,44(S1):587-590.
- [2] 张玛丽.大型网站分布式架构的研究和应用[J].山西电子技术,2017(02):73-75;93.
- [3] 董英茹.简谈AngularJS在下一代Web开发中的应用[J].软件工程师,2015,18(05):30-31.
- [4] 孙晓林,张新刚.基于Nginx的网站安全优化方案[J].网络安全技术与应用,2017,(11):43;45.
- [5] 张康.基于Nginx的在线教育平台架构优化研究[D].北京工业大学,2016.
- [6] 马原.基于RPC的高并发网络通信中负载均衡的研究[D].浙江理工大学,2017.
- [7] 付丽梅,邓继禹,贾跃.基于腾讯微校平台的易学习APP设计与实现[J].考试周刊,2017(07):112.
- [8] 付丽梅,刘英鹏,贾跃.基于腾讯微校平台的校园移动办公APP设计与实现[J].信息系统工程,2017(01):156-157.

作者简介:

李绪成(1977—),男,硕士,副教授.研究领域:Java EE,计算机视觉.

刘鑫(1995—),男,本科,工程师.研究领域:软件开发.