

中国科学院网站群平台发布性能优化实践^①

马羽飞, 杜义华

(中国科学院计算机网络信息中心, 北京 100864)

摘要: 针对中国科学院网站群平台承载站点数量激增后产生的页面分发性能瓶颈, 深入分析平台网络架构, 采用 NFS 共享替代 FTP 同步模式、设立中间数据服务器、替换虚拟化平台等方案进行优化及改造, 成功提升平台发布模块吞吐量, 有效解决了平台发布性能问题。

关键词: 发布性能优化; 中科院网站群; NFS 共享; FTP 传输; 架构优化; 数据服务器; 安全

Practice on the Optimization of CAS Website Group Platform Distribution Performance

MA Yu-Fei, DU Yi-Hua

(Computer Network Information Center, Chinese Academy of Sciences, Beijing 100864, China)

Abstract: Focused on the bottleneck of the web page distribution performance caused by the sharp increasing of website quantity carried by CAS website group platform, we have deeply analyzed the network structure of the platform. We optimize and remould the platform by adopting several methods such as using NFS Share mode to replace FTP Synchronous mode, setting up intermedia data server, replacing with virtualization platform, etc. We successfully improve the handling capacity of the release module of the platform, and effectively resolve the problem of the platform distribution performance.

Key words: optimization of distribution performance; CAS website group; NFS share mode; FTP transfer mode; structure optimization; data server; securit

1 引言

中国科学院 2009 年建设应用网络化信息发布平台(网站群平台), 按“集中建站、分级管理、制度约束、服务保障”方式, 建成了以中科院门户网站为主站, 院机关各部门、院属各单位网站为子站的网站群体系^[1]。平台包括内容管理、互动交流、视频播放、全文检索等应用, 具备安全的网络结构和系统配置, 支持百万级页面的高效访问, 为院属单位网站建设提供良好环境和技术支撑。年来, 中科院网站群规模不断扩大, 承载运行站点数从上线之初的 200 余个增加到 800 多个。整体日均访问量从之前不足 10 万次跃升到 300 多万次。平台数据资源量总 400 多万, 站群文件资源总量达 2TB, 平台日均生成分发网页文件 10 万余次。

随着站点数量和信息发布量的递增, 高峰时段平台出现信息发布管理瓶颈。特别是当某站点用户因调整页面模板, 重新发布所有文章或论文专利等资源库

时, 常会出现信息发布延迟、后台管理响应慢, 甚至于出现内容管理系统的假死及宕机等现象, 直接影响各级网站信息员的新闻编发。因此需要研究解决平台发布性能瓶颈, 提升整体稳定性和用户体验。

2 现状分析及优化思路

2.1 网络架构现状

中科院网站群平台采用 B/S 方式进行管理, 管理员或者信息员通过浏览器访问网站群内容管理系统, 进行系统和数据管理, 并将数据存储在网站群数据库中。新闻发布时网站群系统会根据发布任务从数据库中取出数据, 结合模板规则在本地服务器指定目录中生成标准的 HTML 静态页面文件, 并根据网站站点分发设置, 采用 FTP 协议将该目录的发布数据同步到 Web Server 双节点各自目录中, 供网站最终用户浏览访问^[2]。

① 收稿时间:2015-05-14;收到修改稿时间:2015-06-03

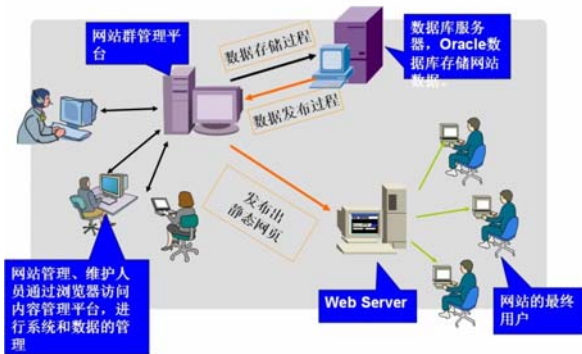


图1 网站群平台发布模式

平台的网络结构^[3]主要分为四个 VLAN 区: Web 发布区、内容管理区、动态交换区、核心数据区。其中信息管理和页面发布涉及 Web 发布区、内容管理区。Web 发布区域由硬件负载均衡设备上加上两台 Apache 构建的 Web Server 物理服务器组成;内容管理区域按需部署于使用开源 KVM 系统虚拟化技术的云计算管理平台^[4,5]系统上。由一个 Apache 负载应用以及两套内容管理系统,共三台虚拟服务器组成集群模式,从而使平台核心区域整体满足了负载均衡、高可用、可扩展的三个基本需求。平台现有架构示意图如下:

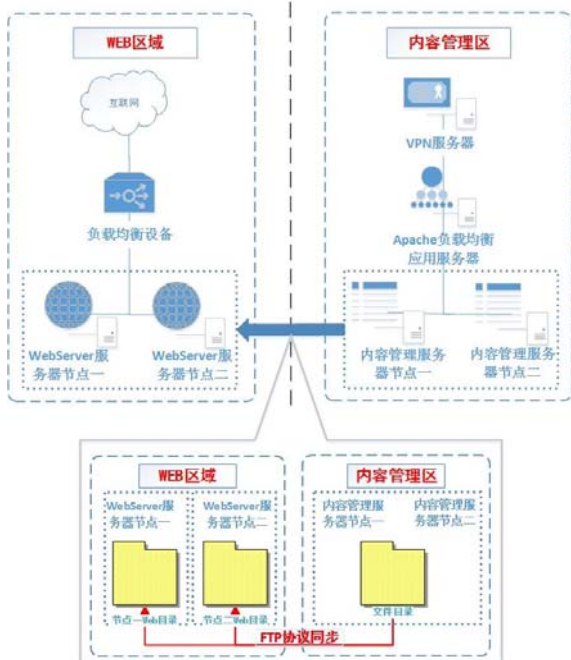


图2 网站群平台发布前端架构示意图

2.2 问题分析

网站群平台发布性能主要涉及服务器负载、网络

流量、应用状态三个方面指标。因此,一方面通过使用 Linux 内核模块和系统命令,对页面发布涉及各服务器、应用性能状态进行检测分析;另一方面,还采用 Cacti 分析工具对网络流量情况进行监控。

通过对监控结果进行分析,发现在内容管理系统同步网页文件到 Web 服务器这个过程中,文件传输队列存在大量网页文件处于等待状态。进一步从网络、操作系统、应用三个层面对此问题进行分析研究,发现此性能瓶颈主要和文件同步模式、文件同步架构、文件 I/O 效率三个方面有关。

2.2.1 FTP 传输不适于数据众多的网页文件

建站初始,根据当时网站群规模,选取文件传输协议^[6](FTP)用来实现文件数据的跨域传输。但随着站群规模的不断扩大,站点网页文件传输次数的增加,FTP 在处理此种数量众多的小网页文件时,传输的性能瓶颈也逐渐暴露出来。其原因在于:FTP 传输数据文件过程,包含繁复的交换握手步骤,并且还需要一个文件副本才可进行传输。

信息员通过网站群平台采编发布新闻时,由于内容管理系统采用实时传输机制,每发布一篇新闻,都会在本地产生成原始的网页数据文件及目录,然后调用 FTP 开始传输,且需要等待服务端接收和反馈完成信息后,才会开始下个传输任务。在此种机制下,当系统负载增高时,单个任务传输时间的增加,累计造成了整体文件传输等待队列的阶梯级递增。同时,对有限的网络、硬件资源,也造成了不必要的消耗。对于中科院网站群来说,由于网站数量多、内容更新快,造成了网页文件传输数量庞大,因此,此种低效的 FTP 传输机制成为网站群平台发布性能瓶颈的主要原因。

2.2.2 发布链路过长影响网络传输效率

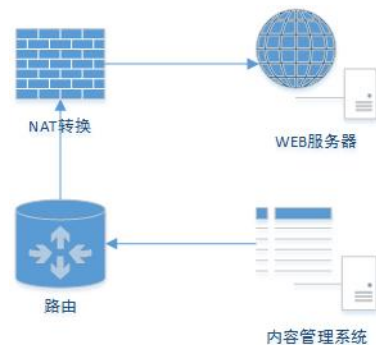


图3 发布数据包流转路径示意图

网站群平台发布的网络架构通过 FTP 实现文件跨域发布. 内容管理系统发出的文件数据包需要经过多层网络设备, 并由防火墙进行 NAT 地址转换, 最终到达 Web 服务器节点上, 整体发布过程涉及链路过于冗长.

在随着网页文件传输基数的增加, 网络流量逐步升高, 网络设备的处理性能也随之逐步降低. 尤其是发布链路中涉及了 NAT 转换技术, 其地址来回转换过程需要耗费更多的网络资源.

这对有繁复握手步骤的 FTP 传输模式来说更是雪上加霜, 直接导致数据包在网络传输的时间成倍增加. 如此繁长的发布链路, 也是造成平台发布性能瓶颈的原因之一.

2.2.3 多虚拟机造成 I/O 性能降低

2012 年, 云计算技术开始普及. 网站群平台对比 XEN、KVM、VMWare 等虚拟化平台, 选取了已成为 Linux 内核的一个模块, 且便于安装、升级、维护的 KVM(Kernel Virtual Machine)平台技术^[7], 来搭建网站群云平台. 由于 KVM 完全利用 Linux 内核来实现 CPU 的调度、内存管理的功能, 可以把 KVM 虚拟机简单看为运行在物理机的抽象层上的一个“特殊”应用程序. 其访问调用 CPU、内存、网络等硬件资源, 都与宿主物理机的管理和分配密切相关.



图 4 虚拟化示意图

但由于虚拟化技术的资源损耗问题, KVM 随着工作负载的增加性能也会逐渐下降. 当多台虚拟机在同一宿主物理机上运行时, 受到多方面因素的综合影响, 则会产生对宿主物理机 CPU、内存、网络、I/O 等资源的争用, 造成虚拟机上的应用发生运行缓慢、假死等情况.

随着网站群规模的扩大, 内容管理系统读写硬盘的频繁度逐渐增加, 也加剧了底层 I/O 资源的争用. 文件读写效率的下降, 也直接影响了等待队列的长度, 产生高并发下平台发布性能的瓶颈.

2.3 优化思路

基于以上分析, 平台发布性能优化的解决思路为: 减少发布过程中, 文件传输等待队列为主. 主要通过替换发布模式、优化发布链路、提升 I/O 性能, 三个方面进行优化改造, 提升目前平台的网页发布性能.

2.3.1 NFS 替换 FTP 提升网页传输效率

网络文件系统(NFS)基于 TCP/IP 协议层, 允许网络中的计算机之间通过 TCP/IP 网络共享资源. 可将远程的计算机磁盘挂载到本地, 像本地磁盘一样操作, 即不必像 FTP 同步文件过程一样的繁杂步骤, 也无需预先生成传输的文件副本. 同时, NFS 在处理文件的增量修改时, 只修改文件中的一个很小的片段, 通过网络传送少量的修改数据, 而非 FTP 协议传输的整个文件.

NFS 模式相比 FTP 传输模式而言, 可以减少文件传输过程中的交互环节, 能最大程度提升小文件的传输效率. NFS 同时支持 TCP 和 UDP 协议进行文件传输, 在可靠性得以保障的网络环境中, 采用 UDP 协议传输文件效率会高于 TCP 协议.

但 NFS 没有用户认证机制, 而且数据在网络上传送的时候是明文传送, 安全性对比 FTP 会有一定程度降低. 同时从数据角度来说, 由于传输过程不需要文件副本, 网页文件数据的冗余度也随之降低. 因此, 采用 NFS 模式后, 需要对发布的网络架构进行优化调整, 并增加网站数据的冗余度, 提升文件传输及数据文件的安全.

2.3.2 优化发布网络架构提升传输效率及安全性

为确保更换 NFS 模式后, 网站群平台整体安全性不降低, 需要增强发布链路的可靠性. 同时, 为提升网络文件传输效率, 在发布流程中, 还应避免涉及防火墙的 NAT 地址转换, 并尽量缩短发布链路长度.

因此, 基于上述优化思路, 需对现有 WEB 区域进行调整, 迁移 WEB 服务器上的网站数据文件, 搭建专享数据服务器, 并用其组建发布流程的 DMZ 区域, 对发布链路 with 访问链路进行间接分割, 同时避免 WEB 区域与内容管理区域有直接交互.

增加数据服务器后, 内容管理系统将网页文件传输至数据服务器, 其发布流程都处于内网之中, 整体发布链路的可靠性得以保障; 且从数据服务器到 WEB 服务器节点的链路, 沿袭了之前防火墙安全策略, 网站群平台整体安全性没有降低. 数据服务器采用双节点模

式,确保了网页文件数据的冗余度.由于NFS缺少用户认证机制,因此在平台改造为NFS模式后,为进一步提升安全,也对NFS共享目录权限进行了相应限制.

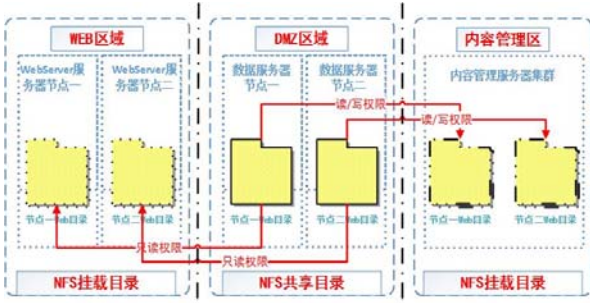


图5 数据服务器共享结构示意图

数据服务器将网页文件目录分别共享给内容管理区域以及WEB区域.内容管理系统到数据服务器需要传输网页文件,对该目录享有读写权限;而WEB服务器只需访问数据,对其设置为只读权限.

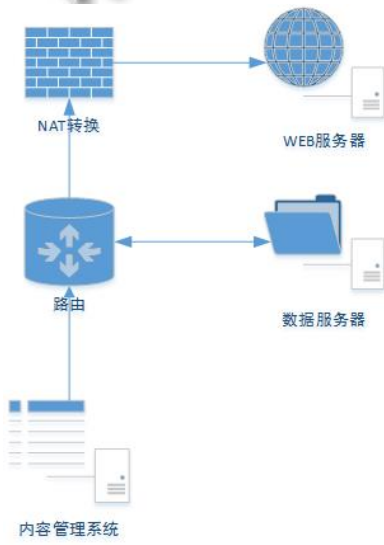


图6 发布架构改造示意图

通过两台数据服务器作为中转后,发布流程只和内容管理系统到数据服务器的链路有关,这样不仅避免了涉及NAT地址转换,也有效缩短了发布链路长度.在网络流量增高时,缓解了链路上数据拥堵的情况.

2.3.3 物理服务器替换虚拟服务器解决关键设备I/O瓶颈问题

对比分析现有虚拟化技术,有些虚拟化平台的性能已接近物理机,但实施改造成本过高.通过多方面考虑分析,以解决平台核心内容管理系统的性能、稳定性

等为首要目标,决定采用物理服务器直接替换虚拟服务器.一方面,有效避免资源竞争对内容管理系统造成的干扰及影响;另一方面,不仅降低了改造实施风险,也在一定程度上缩减了对平台核心业务的影响.

3 优化实践过程

中科院网站群平台为生产系统,面向各网站信息管理员及社会公众提供7×24小时不间断服务,优化改造不能影响或中断服务,操作过程需要严格参照操作规程.

实践中制定了详细实施方案及进度控制措施,将优化过程划分为:平行迁移内容管理系统阶段、修改分发模式阶段、增加数据服务器以及调整网络结构阶段.且针对各阶段环节,制定详细应急预案并搭建测试环境进行模拟测试.其中,从管控进度及质量方面,还需要注意如下技术要点:

①迁移大数据量文件时,应采用先压缩复制,再遍历更新的文件同步方式.

由于网站群平台涉及的数据文件量大、复制周期过长,应先对旧文件目录采用压缩模式复制数据,再使用Rsync遍历更新复制过程中产生的新数据文件.这样,可以有效提升迁移效率,缩短工作周期.

②调整站点分发模式时,可定制脚本批量对其进行更新.

平台中站点数已达800多个,如在内容管理系统中手工逐一修改,不仅其效率低下,而且修改时间过长会影响期间数据的更新.更由于系统内部逻辑复杂,不能简单、直接的对配置文件或数据表进行更改,因此,需要定制更新脚本,调用系统本身提供的webservice接口,进行批量更新,降低实施风险.

③服务器安全设置时采用“白名单”策略.在对新增服务器安全加固阶段时,为保障新增设备不影响整体平台安全性,采用白名单安全设计思路,有效控制某台机器上运行的可执行文件,禁止未知程序.对比“黑名单”策略,“白名单”策略对服务器限制更加严格,可以主动抵御部分未知攻击,进一步提升了平台整体安全性.

4 优化效果

4.1 平台架构安全性对比

调整后平台的网络架构示意图如下.

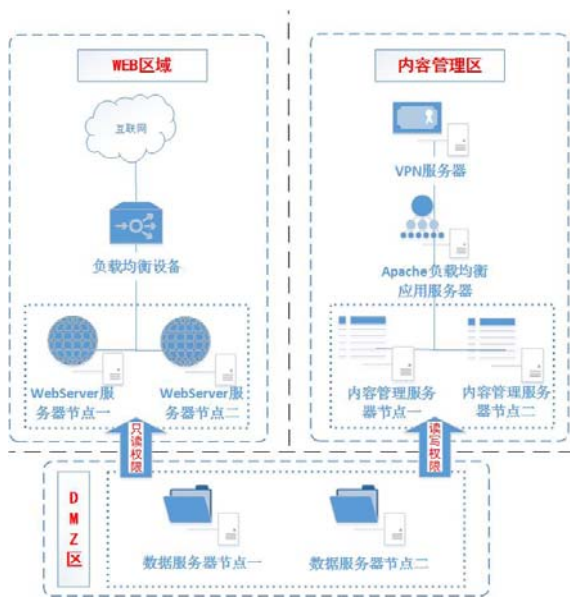


图7 平台发布优化架构示意图

其中网站信息管理员操作时,先通过VPN登录内容管理区域进行信息发布操作时,内容管理系统会自动调用发布模块将生成的页面文件写至DMZ区域中的两台数据服务器上,并通过防火墙NAT策略将文件共享给Web Server读取。互联网公众用户访问网站前台页面时,访问请求经过硬件负载设备轮询至不同Web Server节点服务器上,Apache应用通过本机读取DMZ区域中数据服务器的共享目录,将请求的页面最终返回给访客。

调整后的网络结构中,新增DMZ区域和两台数据服务器,一方面作为WEB区域和内容管理区域的一个安全缓冲,可降低通过Web Server服务器直接对平台内容管理区域进行攻击的可操作性,另一方面,当前端Web访问并发加剧后,可动态在WEB区域增加Web Server节点服务器,直接挂载数据服务器的共享目录,从而便捷快速扩充集群规模,提升集群增提服务能力,能有一定抵御减缓DDos攻击作用。

4.2 平台发布性能对比

为测证调整后平台的发布性能,采用黑盒测试方法,通过测试基准生成工具,模拟系统实际运行情况下的常规操作,记录文档发布过程时间。同时,使用iostat、vmstat等工具对服务器性能进行全程监控。

具体测试方案为:在内容管理系统中指定测试站点,发布同一栏目下的前500篇(前20页)文档,选取最直观的文档发布响应时间作为评判对比点,分别进

行记录对比,其测试结果如下:

操作事务	并发用户	平均响应(S)优化之前	平均响应(S)优化之后
栏目增量发布	50	0.66	0.61
	100	2.79	1.64
	150	3.04	2.67

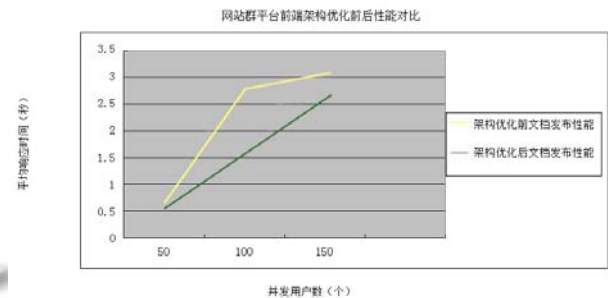


图8 平台架构优化发布性能对比图

从发布响应时间对比中可以直观看出,优化后平台在高并发访问情况下,系统各并发数下的响应时间均有所降低。监测服务器在发布过程中的状态,其等待输入输出的CPU时间百分比值比优化前有所下降,其I/O效率有所提升。检查系统发布后的页面文件,均没有缺失、重复或错乱情况。

在整体测试过程中,应用发布功能运转正常,系统、服务器状态稳定,平台发布性能优化效果显著。

5 结语

中科院网站群平台已于2015年3月优化发布性能完毕,调整后有效支持日常日均10万以上文件的生成分发,站群整体运行稳定。

参考文献

- 杜义华,张亚.中国科学院网站群建设的系统工程方法.计算机系统应用,2012,21(1):17-20.
- 杜义华,张亚.网站信息管理发布系统设计与应用.计算机系统应用,2005,14(1):9-11.
- 杜义华.中国科学院网站系统的整合探讨.计算机系统应用,2005,14(11):71-74.
- 丛培民.云计算架构下的网站群应用实践.计算机系统应用,2012,21(2):5-8.
- 丛培民,龚立武.基于云计算的网站群架构及安全性涉及与实践.第26次全国计算机安全学术交流会,2011,(09):60-63.
- 文件传输协议.Wikipedia,http://zh.wikipedia.org/wiki/文件传输协议,2015.
- 任永杰,单海涛.kvm虚拟化技术:实战与原理解析.机械工业出版社,2013.