

集群架构在精品课程传播模式中的应用

秦文生¹, 苗放¹, 徐松浦², 程小恩¹, 陆宏霞¹

(1. 成都理工大学 信息工程学院, 四川 成都 610059;

2. 成都理工大学 信息管理学院, 四川 成都 610059)

摘要:“创新型人才”培养模式的理论和实践是高等教育改革研究中最重要基础研究领域, 培养创新人才, 优质教育资源是基础。精品课程网站是优质教育资源和信息技术结合的集中体现。为解决精品课程传播中的性能问题, 设计了基于集群架构的传播模式, 经测试可较好提升服务性能, 因此集群架构是一种值得在精品课程网站中采用的传播模式。

关键词:创新型人才培养; 集群结构; 精品课程; 负载调度; 高可用性

中图分类号: TP393.02

文献标识码: A

文章编号: 1673-629X(2008)10-0230-03

Application of Cluster Architecture on Excellent Course Spread Model

QIN Wen-sheng¹, MIAO Fang¹, XU Song-pu², CHENG Xiao-en¹, LU Hong-xia¹

(1. Information Engineering Institute of Chengdu University of Technology, Chengdu 610059, China;

2. Information Management Institute of Chengdu University of Technology, Chengdu 610059, China)

Abstract: The theory and practice of cultivation of innovative talents is the most important research domain in higher education reform. The excellent courses website based on high quality education resources and IT technology. As the development of excellent courses services requires more highly scalable and available server systems. The cluster architecture became a popular solution. Designed an architecture which served excellent course website and benefit for innovative talent cultivation. Through a simulation test, the cluster architecture is beneficial to excellent course spread.

Key words: innovative talent cultivation; cluster architecture; excellent course; load balance; high availability

0 引言

当今时代,人类社会步入了一个科技创新不断涌现的重要时期,创新型人才成为提高综合国力、推动经济社会发展的战略性资源。人才培养的基础重在教育,培养高素质创新人才,优秀的教师、优质的教育资源是关键。如何构建信息时代下国家层次上的精品教育资源? 精品教育资源的传播模式是什么? 这些都是我国高等教育观念与教学改革必须着重研究和思考的重大课题。为此,教育部启动了“国家精品课程建设”工作,相继建立了国家级、省级、校级精品课程网站,旨

在实现优质教育资源共享,促进创新人才培养。近年来,精品课程在人才培养方面起到了有力的支撑作用,为提高学生创新能力奠定了坚实的基础。

然而,随着越来越多的高校师生对精品课程网站进行大量的并发访问,单一服务器的结构模式无法承受超荷的负载,会出现网站响应速度极慢、甚至网站服务器宕机瘫痪等问题;同时单一的结构模式也无法提供自动的故障监测和快速的故障恢复能力。而集群结构模式既可以提供承载高负荷并发访问的能力,还可以提供故障自动恢复能力,能够较好地解决上述问题。因此,设计一个基于集群结构的精品课程传播模式,就成为培养创新人才的一项先行基础工作。

收稿日期:2008-03-01

基金项目:教育部“十一五”教育科学研究规划课题(06A1J0110002);成都理工大学2007年教改项目(JG200718);成都理工大学2007年教改项目(JG200739)

作者简介:秦文生(1980-),男,山西应县人,硕士研究生,主要研究方向为计算机网络、系统管理;苗放,教授,博士生导师,主要研究方向为计算机网络及应用、地理信息系统与遥感技术。

1 集群传播模式架构介绍

当访问数量快速增加时,服务器的CPU、I/O及其它的硬件处理能力很快会成为瓶颈,由于单台服务器的性能总是有限的,简单的提高硬件性能并不能真正解决这个问题。为此,必须采用集群技术才能满足大

量并发访问的需要。

大规模集群一般包括以下几种组件:

- * 负载均衡器(Load Balancer, LB):它是整个集群逻辑上对外提供服务的前置机,用户所有的请求都是面向 LB 的, LB 负责将客户的请求调度到后端的真实服务器上执行。

- * 真实服务器(Real Server, RS):逻辑上是位于负载均衡器后面,并真正给客户返回其所请求内容的服务器。

- * 集群文件系统(Cluster File System):为真实服务器提供一个共享的存储区,存储区可以由 Lustre、ZFS、GFS 等分布式文件系统构成。这样既可以解决各真实服务器上的文件同步问题,又可解决文件 I/O 性能问题。

在集群结构模式中,有几个关键性的技术:①负载均衡调度时的 IP 报文处理过程;②高可用技术的实现;③服务器之间的 Session 同步、TCP 状态同步转移问题。本课题拟就基于 LVS(Linux Virtual Server)的负载均衡 IP 报文处理过程和基于 Linux Heartbeat 的高可用性技术做出研究,从中选择适用于构建精品课程集群架构的部分,实现具备负载均衡(Load Balance)能力和具备高可用性(High Availability)的精品课程网站集群架构设计^[1,2]。

1.1 负载均衡时的 IP 报文处理过程分析

集群模式中,当 LB 收到用户的 IP 请求数据包后,首先,调用 `ip_vs_conn_in_get()` 去 `ip_vs_conn_tab` 表中查找是否存在这样的连接,即 CIP:Port 和 VIP:Port 的对应关系是否存在。如果不存在,那么调用 `ip_vs_schedule()`, 根据调度算法为这个请求分配一个 RS, 然后,调用 `ip_vs_conn_new()` 建立一个新连接, `ip_vs_conn_new()` 将进行一系列初始化操作:设置连接的协议、IP 地址、端口号、协议超时信息,绑定 application helper、RS 和数据报传输函数。最后,调用 `ip_vs_conn_hash()` 将这个连接插入哈希表 `ip_vs_conn_tab` 中,进入 IP 包转发阶段^[3]。

Linux Virtual Server 负载均衡过程涉及到了三种 IP 报文处理模式:①Direct Routing 模式;②IP Tunneling 模式;③NAT 模式,依次对应 `ip_vs_dr_xmit()`、`ip_vs_tunnel_xmit()`、`ip_vs_nat_xmit()` 这三个包处理过程^[4]。

(1)Direct Routing 模式的包处理过程是^[4,5]:

- * 将原 IP 包物理层 MAC 地址改写为 RS 的 MAC 地址;

- * 将新包 CIP(Cliet's IP, 用户端 IP) -> RIP

(Real Server's actual IP, 真实服务器的实际 IP)置于输出队列,经内部网络(必须是同一子网)转发到 RS;

- * RS 收到此包后,将应答内容封装在 VIP(Virtual IP, on director and realservers, 虚拟 IP, 位于调度器和真实服务器上) -> CIP 的包内,直接返回给客户。

(2)IP Tunneling 模式的包处理过程是:

- * 通过 IP 协议,将原 IP 包(三层及以上)的数据封装上 CIP -> RIP 的 IP 包头;

- * 将新包置于输出队列,经内部网络(可以跨子网)或 Internet 转发到 RS;

- * RS 收到此包,利用 Tunneling 功能解包;

- * RS 将应答内容封装在 VIP -> CIP 的包内,直接返回给用户。

(3)NAT 模式的包处理过程是:

- * 将 IP 报头 CIP -> VIP 改写为 CIP -> RIP;

- * 将新包置于输出队列,经内部网络(可以跨子网)转发给 RS;

- * RS 收到此包,将应答内容封装在 RIP -> CIP 包内,返回给 LB;

- * LB 收到 RS 的返回包,再次改写包头为 VIP -> CIP,发送给用户。

从以上分析可以看出,Direct Routing 模式要求 LB 和 RS 必须在同一子网内,这样不利于实现物理上的分布访问,NAT 模式虽然可以实现跨越多子网,但返回给用户的包仍要经过 LB 中转,服务器上的应答数据包会远大于用户的请求数据包,所以这会对 LB 的性能造成较大的影响。而 IP Tunneling 模式既可以在物理上分布的网络环境下工作,回包时又不必经过 LB 中转,不会影响性能,因而是一种值得采用的方式。

1.2 高可用性技术的分析

Heartbeat 使集群的整体服务尽可能可用,如果集群中的主节点发生了故障,那么这段时间内将由次节点代替它。次节点通常是主节点的镜像,所以当它代替主节点时,它可以完全接管其身份,保证用户访问不被中断。

为了保障负载均衡器的高可用性,需要为其建立一个备份机,在负载均衡器和备份机上运行 HeartBeat 程序来实现。Heartbeat 启动后会调用 `master_control_process()` 创建 HeartBeat 的主进程,并调 `read_child()`、`write_child()` 创建读、写消息的子进程,实现对各种信号量和心跳信息的输入、输出处理,检测其它节点的运行状况。`fifo_child()` 子进程将按 FIFO 的方式读入字节流,并将这些字节流转化为 IPC 消息,进行进程间通信。各节点之间的“心跳”信息可以通过串口通信方式、基于 IP 的广播、多播、单播方式进行传播。当备份

机在一定的时间内收不到来自负载调度器上的“心跳”信息时,它就“抢占”掉主调度器的 VIP,接管主调度器进行继续提供服务,整个切换过程可在几秒内完成^[6]。

2 集群传播模式在精品课程网站中的应用

在精品课程网站架构设计中,选择 IP Tunneling 包转发模式,这样只将 LB 集中,RS 可以在任何地理上分布的 LAN/WAN 上运行;LB 只需要改写一次 IP 包头,应答报文由地理上分布的 RS 直接发给用户,这样就减小了 LB 的负载。

精品课程网站的高可用性组件采用 Heartbeat,将 Heartbeat 配置在主 LB 和备 LB 之间,主 LB 出现故障时,备 LB 可以在 1~3 秒内接管主 LB 进行工作。Heartbeat 可以维护节点间心跳并接管失败节点的 IP 地址。如果一个节点失败,它使用“伪造冗余 IP”技术将失败节点的地址添加到工作节点^[7]。

整个精品课程网站的架构如图 1 所示。

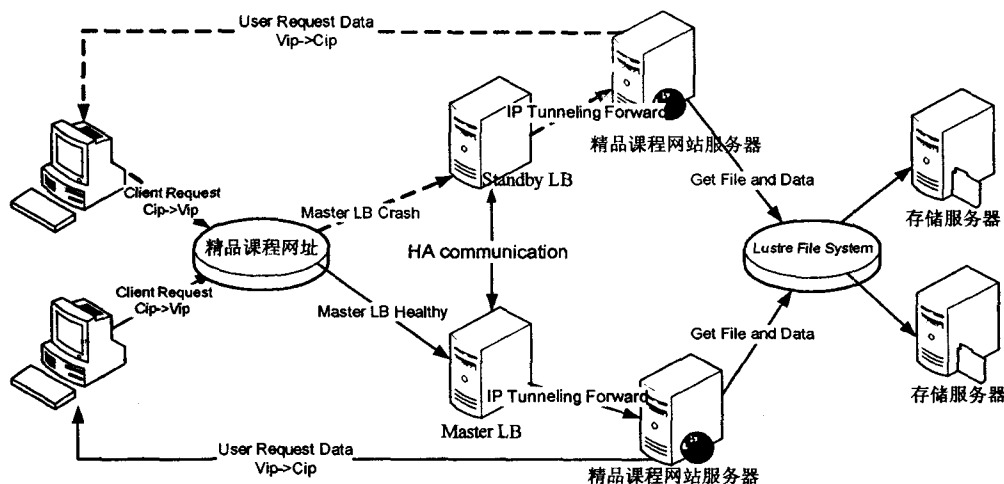


图 1 精品课程集群架构

采用集群结构后,用户访问精品课程网站时,负载调度器根据调度策略把用户的请求经 IP Tunneling 模式分配到后端的某一台真实服务器上,真实服务器应答用户请求,并将应答内容直接返回给用户。采用集群架构后的精品课程网站,相对于之前的单一服务器结构,既可以承受高负载,又具备高可用性,还可以方便地对某个节点进行停机维护,而不引起网站服务中断。对以下三种情况的网站平均响应时间 (Mean Response Time) 分别做了测试对比:①单一服务器 A (采用价值 1.5 万元的 IBM X3250,内存升级到 2G);②单一服务器 B (采用价值 3.2 万元的 IBM X3650,标配);③集群结构中采用 2 台服务器 A 做为真实服务器。

测试数据曲线如图 2 所示。

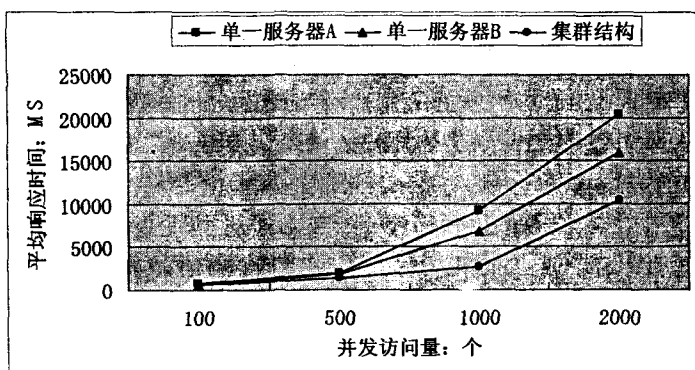


图 2 不同结构下平均响应时间测试对比图

可以看出,集群结构相对于单一服务器 A 和单一服务器 B 的性能都有较为明显的提升,集群结构的响应时间一直优于前 2 种结构,在 1000~2000 并发访问量这个阶段表现尤为突出。因而集群架构可以优化精品课程网站的传播模式,是一种值得采用的架构^[8,9]。

3 结束语

通过研究 LVS 负载调度的 IP 报文处理过程和

Heartbeat 技术,进行精品课程集群架构设计,充分利用了信息技术优势,优化了现有精品课程网站传播模式,为信息时代下的“创新型”人才培养提供了技术保障。随着创新人才培养工作的进一步深入,国家精品课程网站作为其中一个重要因素,势必会得到更多的关注,迎来更多的访问量。

利用开源的集群模式进行网站构架,可以随时在

集群里加入几台服务器来提高访问性能,也可以随时移除某几台服务器进行停机维护,而不会引起网站的临时中断,也不必担心因单一服务器损坏导致的数据丢失。相对于单纯的提升服务器硬件配置和采用 Unix 小型机的解决方案,这种基于开源的集群架构具有高负载性、高可用性、经济性和可伸缩性。故此结构可以较好地用于国家精品课程网站、人才培养模式创新实验区等网站中,为培养“创新型人才”提供信息技术支撑保障。

(下转第 236 页)

-2", "孔 9-3" 及钻井 "孔 9-4" 控制, 其中 "孔 9-4" 不在剖面线上, 需要把孔口坐标投影到 P 点, 把所有钻井数据投影到剖面上。

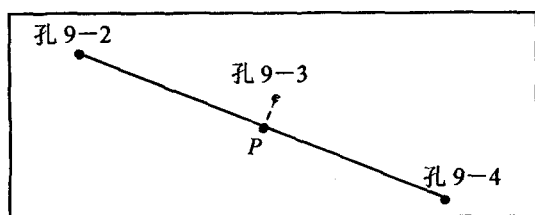


图 5 剖面线平面示意图

系统推理、识别及绘制的基本过程如下:

1) 推理接口将原始数据、用户要求传递给地层类对象。

2) 地层类对象调用类中的相应规则, 进行推理, 得出待识别对象具有“层状”属性, 将消息传递给具有该属性的子类。

3) 子类继续调用内部规则进行推理, 层层递推, 确定识别对象未受构造影响, 并由 ZK_2 中的 C_2t^1 地层在 ZK_1 中没有对应地层, 识别出 C_2t^1 为尖灭地层, 其他地层均为正常层状地层, 把消息分别传给尖灭地层类和正常地层类。

4) 尖灭地层类对象, 调用与尖灭相关的规则, 确定尖灭点的位置, 调用外部函数修改原始数据, 添加尖灭点相关信息。

5) 正常地层类对象, 调用外部函数, 修改原始数据, 添加正常层状地层相关信息。

6) 由输出模块调用已修改的数据, 实现对应地层的连接, 形成如图 6 所示的剖面图。

4 结束语

提出了基于知识的地质体智能识别框架。基于面向对象技术, 解决了复杂知识的结构化表达及嵌入式

的推理机制的问题。以某煤矿实际钻井数据为例, 运用该框架在没有人工参与的情况下, 实现地质体智能识别, 剖面图的自动绘制, 为弥补三维建模中过多依赖专家干预的缺陷, 提高三维建模的自动化程度, 提供了一条有效的解决途径。

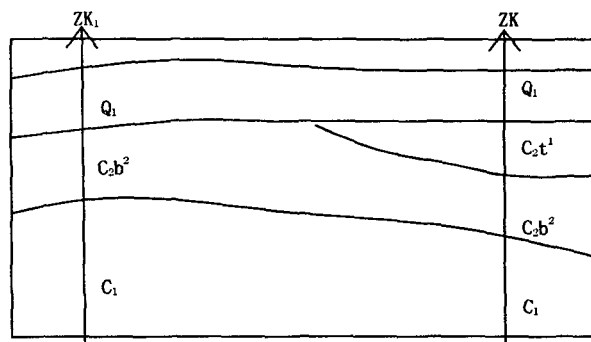


图 6 识别剖面示意图

参考文献:

- [1] 郑贵洲, 申永利. 地质特征三维分析及三维地质模拟现状研究[J]. 地球科学进展, 2004, 19(2): 218-223.
- [2] 徐立明, 牛新生. 地质体三维可视化模拟的现状与展望[J]. 西南民族大学学报: 自然科学版, 2006, 32(1): 151-154.
- [3] 武强, 徐华. 三维地质建模与可视化方法研究[J]. 中国科学(D辑), 2004, 34(1): 54-60.
- [4] 王纯祥, 白世伟, 贺怀建. 三维地层可视化中地质建模研究[J]. 岩石力学与工程学报, 2003, 22(10): 1722-1726.
- [5] 吴键, 曹代勇, 邓爱居, 等. 三维地质建模技术在油田基础地质研究中的应用[J]. 地球科学与环境学报, 2005, 27(2): 52-55.
- [6] 包世泰, 夏斌, 崔学军, 等. 地质三维信息模型研究及其应用[J]. 大地构造与成矿学, 2004, 28(4): 470-476.
- [7] 沙宗尧, 边馥苓. 基于面向对象知识表达的空间推理决策及应用[J]. 遥感学报, 2004, 8(2): 165-171.

(上接第 232 页)

参考文献:

- [1] IBM DeveloperWorks 中国. Linux 集群[EB/OL]. 2001-05-12. <http://www.ibm.com/developerworks/cn/linux/theme/cluster.html>.
- [2] Kopper K. Linux Enterprise Cluster[M]. Sebastopol: O'Reilly, 2005.
- [3] lamafan. LVS 集群系统网络核心原理分析[EB/OL]. 2005-12-19. <http://blog.chinaunix.net/u/553/showart-430386.html>.
- [4] Zhang Wensong. Linux Virtual Servers for Scalable Network Services[EB/OL]. 2007-03-04. <http://www.linuxvirtu->

alservr.org/.

- [5] Sloan J D. High Performance Linux Clusters with OSCAR, Rocks, OpenMosix and MPI[M]. Sebastopol: O'Reilly, 2004.
- [6] Robertson A. The High Availability Linux Project[EB/OL]. 2007-12-21. <http://www.linux-ha.org/>.
- [7] Nntp. Chinaunix 集群版置顶索引[EB/OL]. 2007-11-19. <http://linux.chinaunix.net/bbs/forum-9-1.html>.
- [8] 张媛, 卢泽新, 刘亚萍. NFS over Lustre 性能评测与分析[J]. 计算机工程, 2007, 33(10): 274-276.
- [9] 王莉, 白欣. 实时集群系统设计与性能分析[J]. 计算机工程与应用, 2007, 43(18): 120-122.