

校园网云盘系统构建研究

李军, 劳凤丹, 邹仁明

(中国农业大学 网络中心, 北京 100083)

摘 要: 分析了校园网云盘系统的意义及需求, 提出了一种结合开源的 miniyun、LVS 负载均衡、MooseFS 分布式存储和 MySQL 主从复制读写分离等技术构建高可用云盘系统的技术方案。在此基础上构建了中国农业大学云盘系统, 并投入实际应用。应用结果表明, 本系统具有低成本、高可用、动态可扩展的特点, 能满足在校师生对云盘的使用需求。

关键词: 云盘; MooseFS 分布式存储; MySQL 主从复制; 高可用

中图分类号: TP393

文献标识码: A

文章编号: 1000-436X(2013)Z2-0133-05

Research of cloud disk system construction in campus network

LI Jun, LAO Feng-dan, ZOU Ren-ming

(Network Center, China Agricultural University, Beijing 100083, China)

Abstract: According to the significance and demand of cloud disk system in campus network, a solution of combining the miniyun open source project, LVS(Linux virtual server) load balancing, Moosefs distribute file system and MySQL master-slaver replication read/write splitting was proposed to construct high availability cloud disk system. A cloud disk system was built and put into practice based on this solution in China agricultural vniversity. The application results shows that it meets the demand for cloud disk in campus network with characteristics of low cost, dynamic scalable and high availability.

Key words: cloud disk; MooseFS distribute file system; MySQL master-slaver replication; high avaiability

1 引言

随着计算机及智能移动终端设备的应用普及, 高校师生在教学、科研、办公及生活中产生了大量数据, 如教学课件、多媒体素材、办公文档和实验数据等。传统的数据存储、备份和共享方式包括 U 盘、移动硬盘、邮箱、FTP 等, U 盘和移动硬盘存在硬件故障及病毒感染的隐患; 而邮箱和 FTP 等共享手段在使用的便捷程度以及跨平台方面具有局限性。因此, 校园网内数据可靠存储和快速无缝共享成为高校信息化建设亟需解决的问题。

云服务以其廉价、高可用、动态易扩展等特性得到广泛关注和应用研究^[1]。其中云盘服务为用户提供一个容量动态可扩展的网络存储空间, 通过多终端(PC、移动设备)接入 Internet 方便、快捷地上传/下载个人文件数据(如文档、图片、音乐、视

频等)。云盘中的用户数据存储于云存储中心, 云存储中心采用高可用和动态可扩展的云技术架构, 可保障系统的可靠性和数据的安全性。近年来, 互联网上出现了一些商业的或免费的公网云盘系统, 例如 Google 云端网盘、微软 SkyDriver、Rapidshare 网盘、百度云网盘、金山快盘、华为网盘 Dbank、360 云盘等。校园网用户使用公网云盘存在如下不便: 国外公网云盘的使用受限于国外网站的访问速度, 用户实际体验并不好。国内公网云盘上传/下载速度相对较快, 但受限于高校校园网出口带宽普遍紧张, 速度无法保证, 用户需支付上网流量且免费的存储容量有限。

若在校园网内部署云盘系统, 可充分利用高速校园网(万兆主干、百兆到桌面)的优势, 在满足师生对免费大容量在线存储需求的同时带来良好的上传/下载体验。因此, 校园网云盘建设具体意义如下。

1) 替代传统数据存储备份介质。校园网云盘系统为单校园网用户提供数十 GB 甚至更大的存储容量,可在一定程度上替代传统的 U 盘和移动硬盘,不但为师生节约这类硬件的购买成本,还可避免个人电脑因使用 U 盘及移动硬盘感染病毒的危险。

2) 数据安全可靠存储。用户数据若存储于本地存储介质(电脑硬盘、移动硬盘、U 盘等),在出现硬件故障时会造成数据毁坏或丢失;数据若存储于云盘中,通过云存储技术中的多副本机制可进行数据冗余存储,从而保障数据的可靠性,同时通过加密技术进行加密存储,保障数据的安全性。

3) 多终端数据同步和共享。能通过主流上网平台的客户端同步云盘中的数据,确保各终端数据的一致性。文件可根据算法生成下载链接分享给其他用户,实现数据共享。

4) 降低校园网出口带宽压力。校内云盘上传下载速度较公网云盘更具优势,在校内使用不会产生校园网出口流量,可减轻校园出口带宽的压力。

基于此,本文通过整合开源的 miniyun、LVS 负载均衡、MooseFS 分布式存储和 MySQL 主从复制读写分离技术,构建了一个适用于校园网的低成本、高可用、动态可扩展的云盘系统并已在校园网中部署使用。

2 校园网云盘系统需求

目前,国内高校校园网用户数普遍在 2 万及以上的规模,云盘系统需能为每个校园网用户提供数十 GB 的存储容量才具使用价值和吸引力。因此无论存储容量还是系统并发负载均需满足实际需求。此外,用户账号能与学校网关账号或数字化校园账号进行对接实现统一认证。在用户体验方面,需实现跨平台数据同步。校园网云盘系统建设具体需求如下。

1) 弹性云存储平台。以用户总数为 2 万,每个用户 30 GB 的存储容量计算,存储总需求将达 600 TB。若使用传统的 SAN 或 NAS 存储,显然成本过于昂贵。在建设资金投入有限的情况下,云存储平台需满足低成本的需求,还需支持动态弹性可扩展和平滑扩容。而且在数据存储时需对数据进行切块、加密、多副本存储,以保障数据的可靠性和安全性。

2) 高并发,高可用。云盘服务普及后,尤其是与服务器端定时通信的云盘 Windows 客户端广泛使用后,系统稳定并发数将达千甚至万的级别,系统的前端 Web 和后端数据库都需要按实际需求设计成能应对高并发的高可用的集群架构。

3) 账号统一认证。通过简单数据库访问接口连接高校已有的网关系统或数字化校园系统,以网关系统或数字化校园系统用户库为依据进行用户登录认证,在学生或教职工离校注销账号后清除用户数据释放存储空间。

4) 多终端访问。除了 Web 访问,还需支持客户端访问,客户端覆盖目前主流上网终端平台,如 Windows、Mac OS、Android、iPad、iPhone 等。客户端将云盘虚拟成本地磁盘进行管理,支持复制、粘贴、移动及数据同步。其中 Windows 和 Mac OS 2 类客户端可充分利用本地存储空间在本地存储一份数据副本,当有数据更改时才进行数据同步,以减轻服务器端的 I/O 压力。而 Android、iPad、iPhone 等移动终端设备本地存储空间有限,客户端中只显示文件树列表,而不在于本地存储数据副本,在有需要时可选择下载所需数据。

3 校园网云盘系统部署的关键技术

在综合考虑云盘功能性需求、成本以及软件成熟度等多方面因素后,本文选择对国内开源的云盘系统 miniyun 进行二次开发和应用部署。miniyun 是一个 Web 端程序基于 PHP 开源而客户端程序闭源的项目。它提供 Windows、Android、MacOS、iPad/iPhone 4 类终端设备的客户端,基本涵盖了目前主流的上网终端平台。对其进行账号统一认证二次开发后,便可在小范围内测试应用,若大范围部署应用尚需解决高并发和海量存储等问题。

3.1 系统架构

为解决云盘系统高并发和海量存储等问题,云盘系统架构设计如图 1 所示,共分为 4 部分:LVS 负载均衡层、Nginx Web 集群层、MooseFS 云存储层和 MySQL 主从集群层。LVS 负载均衡层采用 LVS+Keepalived 对 Nginx Web 服务器群进行请求分发、负载均衡及故障检测。Nginx Web 集群层用以部署 miniyun 应用,提供 Web 服务及与客户端通信。云存储层采用 MooseFS 分布式文件系统进行构建用以存储云盘系统中的文件数据。数据库采用 MySQL 主从架构并使用 MySQL-Proxy 进行读写分

离。其中 Nginx Web 集群层、MooseFS 云存储层和 MySQL 主从集群层均能根据实际负载需求进行横向平滑扩容。

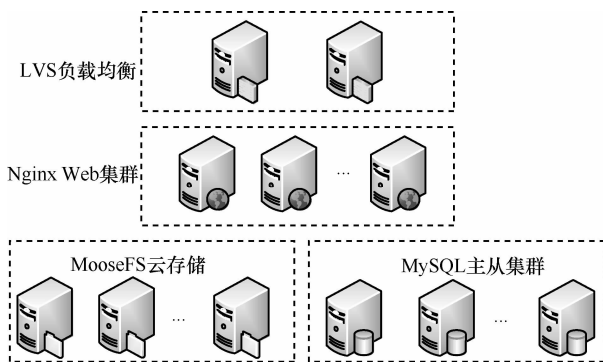


图 1 校园网盘系统架构

3.2 LVS+Keepalived+Nginx 高可用集群

本文在 Web 端采用 LVS+Keepalived+Nginx 架构部署 miniyun 校园网盘系统，保证 Web 端的高可用和能应对云盘系统高负载的上传/下载请求。LVS (Linux virtual server) 是一个开源的高可用负载均衡技术，它能有效地应对高负载的网络应用^[2]。LVS 由一个前端的负载调度器 (LB, load balancer) 和多个后端的 Web 真实服务器 (RS, real server) 所组成。LVS 集群对用户是透明的，对外入口是一个统一的虚拟 IP。当用户请求发生后，LB 按照一定的转发算法将请求分发到后端的 RS^[3]，RS 的应答数据不需要汇集到 LB 进行转发，而是直接发送给用户。由于 LB 只负责请求分发，所有的操作都会在 Linux 的内核完成，因此 LB 负载小、吞吐率高。此外，由于应答数据不经过 LB 转发，LVS 将具有网络流量分流的特性。在 LVS 中，当系统总负载超过后端 RS 群所能处理的能力时，可以增加 RS 数量以满足不断增长的并发负载。

为了解决 LB 的单点故障和 RS 的故障检测，结合开源的高可用软件 Keepalived 为 LB 配置一个备份节点 LB-Backup，使其与 LB 互为主备。Keepalived 工作在 LVS 之上，它能有效地对 LB 和 RS 进行故障检测，并在 LB 和 LB-Backup 间进行失效切换。当 RS 故障失效时，它能及时将失效 RS 剔除出集群，避免其影响系统正常响应，并在 RS 故障节点恢复正常后自动将其加入到 LVS 集群中。上述操作均由 Keepalived 自动完成，管理员只需人工修复故障节点。

Nginx 是一个高性能的 Web 服务器，采用内核 Poll 模型，在硬件资源一定的情况下比传统的

Apache 服务器能处理更多并发连接^[4]。在 LVS 集群中将 Nginx 和 PHP-FPM 部署在 RS 节点上解析 PHP，为云盘提供 Web 服务，并处理客户端的通信请求。

3.3 MooseFS 云存储平台

为解决校园网盘服务所需的海量存储 (上百 TB)，本文采用 MooseFS (MFS) 分布式存储系统搭建校园云存储平台^[5]。MFS 是一个 PB 级别的分布式存储系统，它的工作原理类似于 Google 的 GFS，能将多台存储服务器的在逻辑上虚拟成一个大容量存储池为业务系统提供存储服务，并支持切块存储 (默认 64 MB) 及多副本技术。MFS 中共有 4 类节点，存储和管理元数据的 MFS-master、存储实际数据块的 MFS-chunk、备份元数据的 MFS-logger 以及挂载存储的 MFS-client，系统逻辑结构如图 2 所示。当系统存储容量不能满足实际需求时，可通过新增 MFS-Chunk 实现平滑扩容。MFS 支持 Posix 协议，MFS-client 能通过 Linux 内核 fuse 模块将 MFS 所构建出来的大容量存储池挂载到本地文件系统中使用。

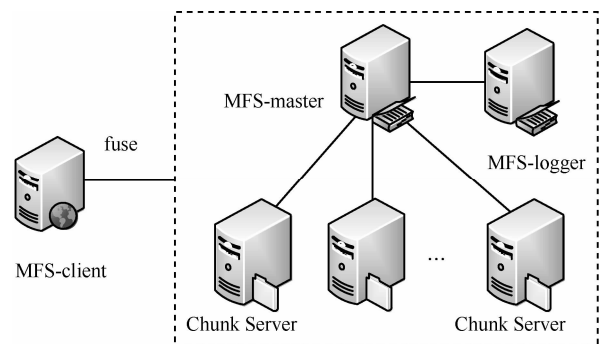


图 2 MFS 系统结构

3.4 MySQL 数据库优化及集群

在校园网盘系统中，用户个人文件数普遍上千甚至过万，即系统总存储的文件数或数据库中文件表数据记录可能达到千万甚至亿级别。同时，云盘客户端每隔 10 s 与服务器端通信一次进行数据库查询。在大数据记录表中频繁查询，会使 MySQL 的查询响应成为系统瓶颈。针对此问题，本文在数据库层进行优化如下。

- 1) MySQL 服务器硬盘采用高转速的 SAS 盘，并做 RAID1+0 以提升磁盘寻道能力。
- 2) MySQL 数据库软件使用最高稳定版 (目前是 MySQL 5.6.12 版) 编译安装，存储引擎使用查询性能最好的 MyISAM。
- 3) 对大记录数据表 (如文件表) 按用户 ID 进行分区存储。

4) 对所有耗时及大记录的数据表查询建立索引以避免全表扫描。

5) 采用 MySQL 主从复制结合 MySQL-Proxy 读写分离构建 MySQL 集群架构。

其中, MySQL 主从复制、读写分离的集群架构如图 3 所示, Web 应用中数据库的读写请求都集中汇集到 MySQL-Proxy 连接池, 由 MySQL-Proxy 按请求类型进行分发^[6]。若请求类型是写数据库操作, MySQL-Proxy 会将请求转发到 MySQL-Master; 若是读数据库操作, 则均衡地转发到每个 MySQL-Slaver。云盘应用的数据库操作具有读多写少的特点, 配置一个 MySQL-Master 和多个 MySQL-Slaver, 它们之间进行主从复制(replication), 确保 MySQL-Slaver 和 MySQL-Master 中的数据保持一致。当 MySQL 读数据的负载超过 MySQL 所能处理的能力时, 可以增加

MySQL-Slaver 数以满足负载需求。

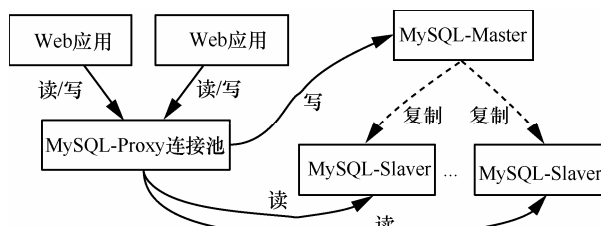


图 3 MySQL 集群架构

4 云盘系统实际运行效果

基于上述方案, 本文构建了中国农业大学云盘系统, 并已于 2013 年 6 月初上线运行。Web 界面如图 4 所示, Windows 客户端界面如图 5 所示。中国农业大学校园网用户 2 万余人, 云盘系统为每个用户提

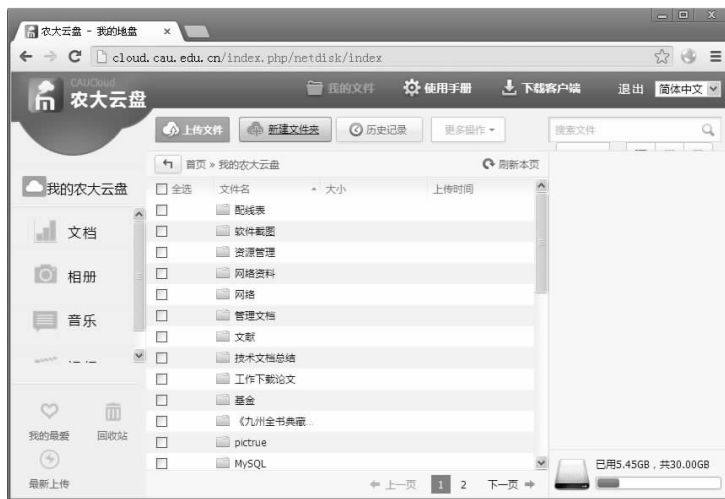


图 4 中国农业大学云盘系统 Web 端

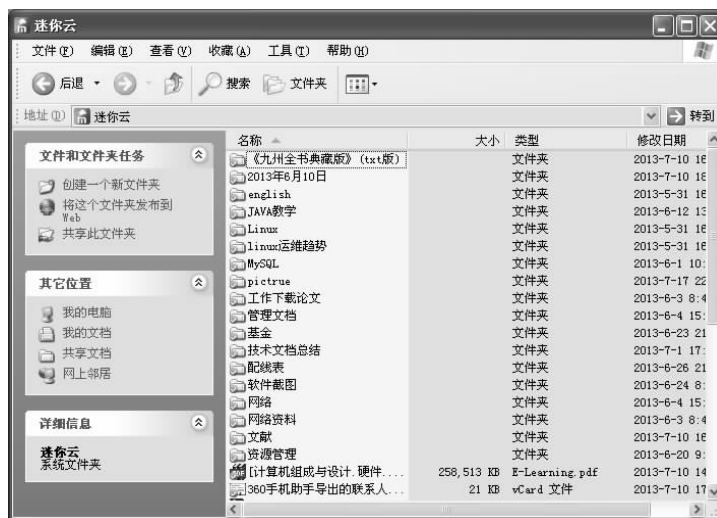


图 5 中国农业大学云盘 Windows 客户端

供 30 GB 的存储容量。目前建设的 MFS 云存储平台总容量为 93 TB, 实际消耗存储容量 6 TB, 后期随存储实际消耗再动态增加存储服务器进行扩容。系统使用用户数为 3 600 余人, 稳定并发为 200 余人, 存储总文件数 57 万个。Web 集群和 MySQL 集群中的各服务器 load 值均小于 1, 整体运行情况良好。

5 结束语

本文基于校园网盘应用的需求, 使用 miniyun 构建云盘应用, 采用 LVS+Nginx 解决 Web 端负载问题, MooseFS 构建云存储平台解决海量文件存储问题, 采用 MySQL 主从复制、读写分离的架构解决 MySQL 读写瓶颈问题。本文所设计的方案具有如下特点: 1) 低成本, 使用 MFS 组织普通服务器挂载廉价的 SATA 盘实现海量云存储平台的搭建; 2) 高可用, 系统架构中各层均存在冗余机制; 3) 动态可扩展, Web 集群层、MFS 和 MySQL 集群都能依据需求进行动态扩展。

参考文献:

- [1] 张进宝, 黄荣怀, 张连刚. 智慧教育云服务: 教育信息化服务新模式[J]. 开放教育研究, 2012, 18(3):20-26.
ZHANG J B, HUANG R H, ZHANG L G. Smart education cloud: new service model of ICT in education[J]. Open Education Research, 2012, 18(3):20-26.
- [2] 章文嵩. 可伸缩网络服务的研究与实现[D]. 长沙: 国防科学技术大学, 2000.
ZHANG W S. The Study and Implementation of Scalable Network Services[D]. Changsha: National University of Defense Technology, 2000.
- [3] 赵攀, 周刚, 魏正曦等. LVS 集群技术在校园网中的研究与设计[J]. 四川理工学院学报, 2009, 22(2):31-35.
ZHAO P, ZHOU G, WEI Z X, et al. Study and design of LVS cluster in campus net[J]. Journal of Sichuan University of Science & Engineering, 2009, 22(2): 31-35.
- [4] 凌质亿, 刘哲星, 曹蕾. 高并发环境下 Apache 与 Nginx 的 I/O 性能比较[J]. 计算机系统应用, 2013, 22(6):204-208.
LING Z Y, LIU Z X, CAO L. I/O performance comparison of apache and nginx in high concurrence environment[J]. Computer Systems & Applications, 2013, 22(6): 204-208.
- [5] 李军, 劳凤丹, 邹仁明. 基于 MooseFS 的云存储系统构建及应用研究[J]. 武汉大学学报, 2012, 58(S1):206-210.
LI J, LAO F, ZOU R M. Research on construction and application of cloud storage based on MooseFS[J]. Journal of Wuhan University, 2012, 58(S1):206-210.
- [6] 祝雄锋. 数据库集群中间件 MySQL Proxy 研究与分析[D]. 武汉: 武汉理工大学, 2011.
ZHU X F. Research and Analysis on Database Cluster Middleware MySQL Proxy[D]. Wuhan: Wuhan University of Technology, 2011.
- [7] Frequently asked questions: MooseFS fault tolerant distributed file system[EB/OL]. <http://www.moosefs.org/moosefs-faq.html#cpu>, 2013.

作者简介:



李军 (1987-), 男, 湖南郴州人, 硕士, 中国农业大学网络中心助理工程师, 主要研究方向为计算机网络技术及应用、云计算。



劳凤丹 (1975-), 女, 壮族, 广西凌云人, 博士, 中国农业大学网络中心高级工程师, 主要研究方向为计算机网络技术及应用、农业信息化等。



邹仁明 (1965-), 男, 江苏扬州人, 学士, 中国农业大学网络中心高级工程师, 主要研究方向为计算机网络技术及应用、网络安全、云计算。