

文章编号:1001-9081(2006)07-1732-03

## 一种高性能的附网刻录服务器的研究与实现

黄振中, 张 青

(河南工业大学 化学工业职业学院, 河南 郑州 450042)

(hnhxhuang@163.com)

**摘 要:**为了解决网络共享刻录机的需要,设计实现了一个新的基于 iSCSI 协议的附网刻录服务器(Network-attached Recorder Server, NRS)。NRS 服务器采用虚拟刻录机技术同 iSCSI 协议相结合,既实现了刻录机的网络透明映射,也提高了网络共享的性能。采用多线程和磁盘文件缓存技术,使多用户可同时共享刻录机,也缩短了用户刻录的响应时间。详细介绍了这种 NRS 服务器的设计方案,并进行了相应的试验测试和性能分析。

**关键词:**附网刻录服务器; iSCSI; SAMBA

**中图分类号:** TP303 **文献标识码:** A

## Research and implementation of a high performance network-attached recorder server

HUANG Zhen-zhong, ZHANG Qing

(Chemical College, Henan University of Technology, Zhengzhou Henan 450042, China)

**Abstract:** In order to meet the need of sharing network recorder, a new Network-attached Recorder Server(NRS) based on iSCSI protocol was designed. NRS combined virtual recorder technology with iSCSI protocol, realized the network transparent map of recorder, and improved the performance of the network share greatly. It adopted multi-thread and disc file buffer technology to share recorder of multi-users, and to short the response time of recorder. The design of NRS was described in detail, and the corresponding testing and performance analysis were carried out.

**Key words:** network-attached recorder server; iSCSI; SAMBA

针对光盘刻录共享服务的特点,设计开发了一种专用的高性能附网刻录服务器(Network-attached Recorder Server, NRS),在很大程度上解决了这一问题。NRS 服务器是一个裁剪后的瘦服务器,它本身没有通用服务器那样复杂的操作系统和硬件连接,只完成 NRS 服务器与客户机之间的数据传送,使客户机能以硬盘的访问速度来共享服务器上的刻录机,方便了用户。在这种新的刻录体系结构中,刻录机不再通过 I/O 总线附属于某个特定的服务器或客户机,而是通过网络接口与网络直接相连,由用户通过网络访问。与附属于服务器的刻录机相比,NRS 服务器具有非常好的共享性。

### 1 NRS 的总体设计

NRS 系统具有“即插即用”的特性,只需要一根网线挂接到以太网上,就可供网络上的用户方便地共享。连接到网络上的任何客户端计算机,只要安装上 NRS 的客户端程序,即可在客户端虚拟一个刻录机,并把此刻录机映射到网络上的 NRS 服务器的刻录机。这种虚拟技术对用户完全透明,用户使用这个虚拟的刻录机就如同使用自己计算机上物理安装的刻录机一样,而且对应用程序也是透明的,任何第三方的刻录程序可以透明地使用此虚拟刻录机,大大方便了用户的使用。

针对光盘刻录机网络共享的特定应用,NRS 系统采用 iSCSI 协议进行网络数据传输,实现了光盘刻录机的网络共享。主要是通过采用虚拟技术,在客户端虚拟一个 SCSI 光盘刻录机,并将接收到的 SCSI 命令通过 iSCSI 协议传输到 NRS 服务器。NRS 服务器采用光盘虚拟技术,用户访问远程服务

器上的光盘完全同使用本地光盘一样。

从图 1 可以看出,客户端用户的刻录请求通过操作系统发给虚拟刻录机,虚拟刻录机驱动层再将请求通过 iSCSI 协议转发给 NRS 服务器,NRS 服务器处理完请求后再通过 iSCSI 协议将处理结果发给客户端的虚拟刻录机,最后由虚拟刻录机将结果返回给用户刻录程序。

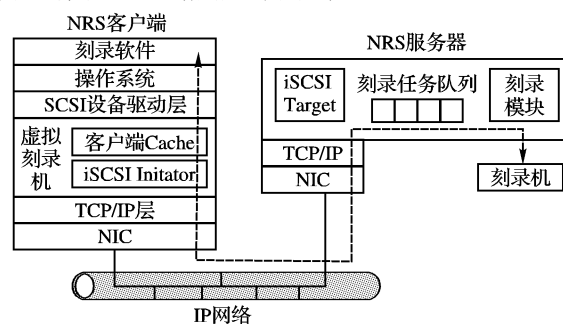


图 1 NRS 系统架构

按照功能 NRS 可以划分为四个模块,客户端包含客户端 Cache 模块和 iSCSI Initiator 模块,服务器包含 iSCSI Target 模块和刻录模块。

NRS 客户端模块的主要功能是在用户的客户端计算机上虚拟一个光盘刻录机驱动器。用户访问此光盘刻录机驱动器上的光盘时,操作系统就调用光盘刻录机客户端提供的服务接口;此接口接收到操作系统的请求后,将它转发给 NRS 服务器,并将 NRS 服务器处理的结果返回给操作系统;操作

收稿日期:2006-01-04

作者简介:黄振中(1973-),男,河南郑州人,讲师,主要研究方向:数据库、网络; 张青(1973-),女,河南郑州人,讲师,硕士,主要研究方向:数据库、网络。

系统将此结果返回给用户刻录程序,从而完成虚拟光盘刻录机的工作。光盘刻录机客户端模块还负责按照光盘组织 Cache,实现应用程序刻录同网络传输的并行工作,在网络传输协议上采用 iSCSI 协议。iSCSI Initiator 模块接受 SCSI 命令,封装成 iSCSI 命令,通过网络的 TCP/IP 层发给服务器。

iSCSI Target 模块能够同时接受多个客户端的刻录请求,采用多线程的方式支持多用户同时访问,对每一个用户连接,都创建一个对应的处理线程来处理此用户的请求,将要刻录的数据以本地文件的形式保存起来,当用户需要刻录的数据完全接受到时,iSCSI Target 模块以刻录任务队列的方式将刻录请求发给刻录模块。刻录模块接收到刻录请求后,不断地检测刻录机,发现可刻录的光盘后就将用户要刻录的数据一次性的刻录到光盘上。为了提高刻录的成功率,只有用户的刻录数据全部传输到服务器后,刻录模块才进行真正的刻录工作,防止刻录过程中因网络故障导致刻录的失败。多个客户端的刻录请求是采用 FCFS 的调度策略进行刻录的。

## 2 iSCSI 协议的分析

iSCSI 是 SCSI 和以太网两种成熟的技术融合的产物,是构建 IP-SAN 的一种新兴技术。iSCSI 协议定义的是 SCSI 到 IP 的映射,即将启动器的 SCSI 命令封装成 IP 数据包,在 IP 网络上传输,到达目标器节点后,再恢复成封装前的 SCSI 命令,从而实现 SCSI 命令在 IP 网络上的直接、透明传输。iSCSI 协议实质是在一个 IP 隧道上运行 SCSI 命令,是一种 SCSI 远程过程调用模型。iSCSI 启动器封装 SCSI 命令和数据,使之能够被 TCP/IP 协议处理,并通过以太网向目标器发出 I/O 请求。iSCSI 目标器直接连接存储设备,解析收到的数据和命令,并处理。

实现 iSCSI 主要有以下三种方式:1) 纯软件方式。采用标准的网卡,iSCSI 和 TCP/IP 协议栈功能层都由主机 CPU 完成。由于采用的是标准的网卡,因此这种方式的硬件成本最低,但由于 iSCSI 和 TCP/IP 层功能都由主机 CPU 完成,主机的运行开销就会增加,从而造成主机系统性能的下降。2) 智能 iSCSI 网卡实现方式。采用特定的智能网卡,iSCSI 层的功能由主机来完成,而 TCP/IP 协议栈功能由网卡来完成。和方式 1) 相比,部分降低了主机的运行开销,但是需要购买专用的智能网卡,增加了用户的投资。3) iSCSI HBA 卡实现方式。采用主机总线适配器的方式,iSCSI 层和 TCP/IP 协议栈功能均由该主机总线适配器来完成。对主机的 CPU 的需求最少,但是需要购买专用的 iSCSI HBA 卡,增加了用户的投资。

NRS 系统采用 iSCSI 的第一种实现方式,这样可以充分利用用户的现有投资,不需要购买其他的硬件设备。而且我们的 NRS 系统的客户端是用户的终端计算机,随着 CPU 性能的不提高,其性能已经大大提高了,完全能够采用软件方式实现。

## 3 客户端 Cache 的结构设计

为了提高 NRS 系统客户端的性能,在客户端设置一个专用的后台写线程,通过设置一个写 cache,刻录程序采用写回的方式将数据写到 cache 中,后台写线程再将要刻录的数据异步传输到 NRS 服务器,从而能够使刻录程序写 cache 同后台写线程传输数据并行工作,减少系统的响应时间。

图 2 为设置后台写线程的刻录程序刻录流程,首先刻录程序向操作系统发送刻录数据,如步骤(1);操作系统接收到刻录程序的数据后,将它转发给虚拟刻录机程序,虚拟刻录机程序接收到操作系统的请求后,分配写 cache,并将此数据写入 cache,如步骤(2);然后虚拟刻录机程序向操作系统返回写的

结果,如步骤(3);操作系统将此结果返回刻录程序,如步骤(4);同时,虚拟刻录机程序启动后台写线程,准备传输刻录数据,如步骤(5);后台写线程接收到请求后,就调用 TCP 的网络接口,将写数据发给网络 TCP 层,如步骤(6);网络 TCP 层再通过网络将数据发送给 NRS 服务器,如步骤(7);NRS 服务器接收到客户端的数据后,将请求的数据写入本地磁盘,并将结果返回给客户端的网络 TCP 层,如步骤(8);网络 TCP 层再将接收到的结果返回后台写线程,如步骤(9)。

在这个过程中,刻录程序写 cache 同后台写线程传输数据是并行工作的,当刻录程序在写 cache 时,后台写线程能够将 cache 中的数据传输到 NRS 服务器,二者是并行工作。也就是图中的步骤(1)~(4)同步骤(6)~(9)可以并行工作。

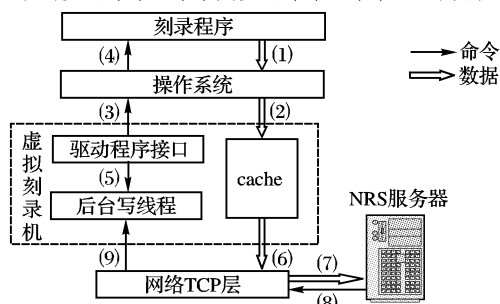


图2 设置预取线程的用户访问处理流程

## 4 刻录模块的结构设计

为了提高 NRS 服务器的效率,NRS 服务器支持多个客户端同时刻录,服务器采用多线程的方式支持多个客户端,如图 3,服务器有一个主线程负责监听客户端连接请求,每当一个客户端连接到服务器,主线程就为此客户端启动一个专用的客户端刻录处理线程,此线程负责处理客户端的刻录请求。因为刻录的速度相对网络比较慢,为了减少用户的响应时间,NRS 系统采用缓存技术;又因为刻录的数据量比较大,内存无法完全缓存,我们采用磁盘文件作为缓存。客户端刻录处理线程接受客户端的刻录数据,按照顺序保存到一个磁盘文件中,只有到此客户端的所有刻录数据全部传输完成,客户端刻录处理线程才关闭文件,并将此刻录任务提交刻录处理线程进行真正的刻录工作。

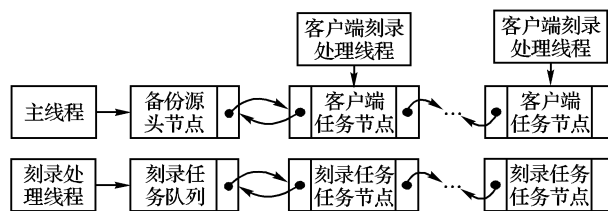


图3 客户端请求在服务器上的处理过程

为了实现多用户的共享刻录,所有已经传输完成的刻录任务组织成一个刻录任务队列,刻录处理线程采用 FCFS 的策略依次刻录各个任务。刻录处理线程是专门进行刻录工作的线程,它循环检测服务器的刻录机,如果发现刻录机中有空白的可刻录光盘,就从刻录任务队列中取最早的任务节点,将此节点指示的磁盘缓存文件刻录到光盘中。当刻录机无可刻录的光盘或者刻录任务队列为空,刻录处理线程才睡眠等待刻录任务的唤醒。

## 5 性能测试与结果分析

### 5.1 性能测试平台

为了验证 NRS 服务器的性能,本文在 100M 以太网上对其进行了性能测试,其客户端和服务端采用的都是自己开

发的程序。测试是在两台通过 IP 交换机连接起来的 PC 机上进行的,其中一台是客户机,另一台是服务器。两机的配置大致相同: Intel Celeron 1.7GHz CPU; 256M RAM; 硬盘: Maxtor 60GB ATA Disk, 最大数据传输率 40MB/s; 100M 网卡; 操作系统: RedHat 9.0。为进行对比,在此平台上分别测试了 SAMBA 服务器和 NRS 服务器。

## 5.2 测试数据及分析

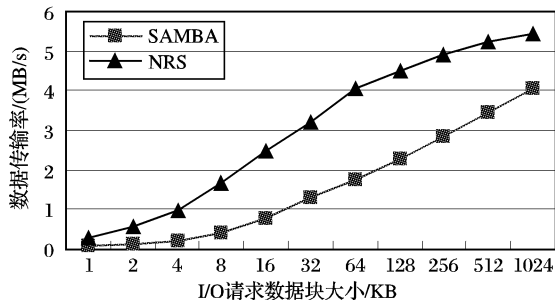


图4 数据传输率曲线

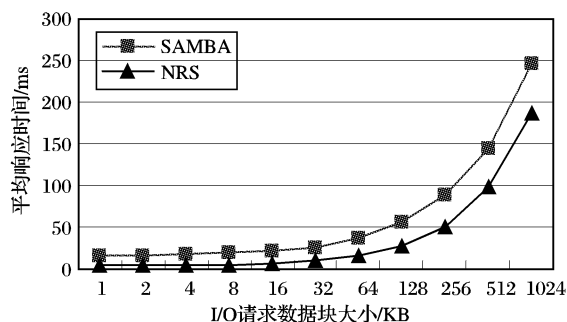


图5 平均响应时间曲线

图4,图5分别为NRS服务器和SAMBA服务器的吞吐率曲线和平均响应时间曲线。测试结果显示:随着I/O请求块大小的增加,两种服务器的数据传输率也都随之增大。总体上NRS服务器的数据传输率比SAMBA高30%左右,而NRS服务器平均响应时间则明显低于SAMBA。NRS服务器能够达到如此好的性能,主要是NRS系统采用iSCSI协议,iSCSI协议采用数据块级的网络共享,要比SAMBA的文件级共享性能高得多,而且NRS系统采用磁盘文件作为缓存,采用后台刻录技术,在用户提交完成后进行后台刻录,这样既提高了共享的方便程度,也缩短了用户的刻录响应时间。这种共享方式同共享网络打印机方式一致,大大方便了用户的使用。

## 参考文献:

- [1] GIBSON GA. Network attached storage architecture[J]. Communications of the ACM, 2000, 43(11): 37-45.
- [2] NAGLE D, GANGER G, BUTLER J, et al. Network Support for Network-Attached Storage[A]. Hot Interconnects'1999[C], 1999.
- [3] LEE L-W, SCHEUERMANN P. File Assignment in Parallel I/O Systems with Minimal Variance of Service Time[J]. IEEE Computer, 2000, 49(2): 127-140.
- [4] BODEN NJ, COHEN D, FELDERMAN RE, et al. Myrinet: A Gigabit-per-Second Local Area Network[J]. IEEE Micro, 1995, 15(1): 29-36.
- [5] GIBSON G. Cost-effective high-bandwidth storage architecture[A]. Proceedings of ACM ASPLOS[C], 1998.
- [6] SACHS M, LEFF A, SEVIGNY D. LAN and I/O Convergence: A Survey of the Issues[J]. IEEE Computer, 1994, 27(12): 24-33.
- [7] ORENSTEIN G, SHURTLEFF G. Integration Scenarios for iSCSI and Fibre Channel[Z]. SNIA IP Storage Forum, 2002.
- [8] Intel iSCSI protect[EB/OL]. <https://sourceforge.net/projects/intel-iscsi>, 2002.

(上接第1731页)

等于零的高斯白噪声。图2和图3分别给出了四种算法的译码性能、迭代次数比较曲线。仿真结果表明:改进算法比其他硬判决算法的性能有较大提高,明显接近BP算法;同时改进算法的迭代次数明显低于BP算法,且没有浮点和乘法运算。

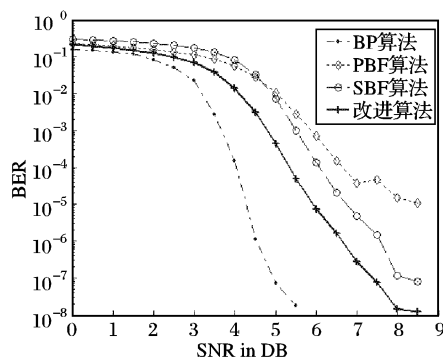


图2 (1008,504,3,6)码不同译码算法下的误码率(Bit Error Rate, BER)比较

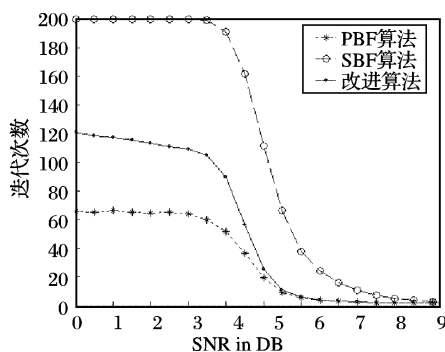


图3 (1008,504,3,6)码不同译码算法下的迭代次数的比较

## 4 结语

本文结合统计译码思想,基于并行BF算法提出了一种新的硬判决译码的实现方案:在明显改进译码性能的同时,平均迭代次数仍保持较低水平,实现了译码性能和算法复杂性的有效均衡。由于统计译码仅仅是对LDPC码校验矩阵的扩充,因此其他各种改进的BF算法都可以与之结合,用以研究更为高效的译码方案。而且统计译码中构造校验子集的方法,既可推广用于非规则LDPC码,也可作为进一步优化LDPC编码方案的参考依据。

## 参考文献:

- [1] GALLAGER RG. Low-Density Parity-Check Codes[D]. Cambridge, MA: MIT Press, 1963.
- [2] MIADINOVIC N, FOSSORIER MPC. Improved Bit-Flipping Decoding of Low-Density Parity-Check Codes[J]. IEEE Transactions on Information Theory, 2005, 51(4): 1594-1606.
- [3] ZHANG J, FOSSORIER MPC. A modified weighted bit-flipping decoding of low-density parity check code[J]. IEEE Communication Letters, 2004, 8(3): 165-167.
- [4] CHAN AM, KSCHISCHANG FR. A Simple Taboo based Soft Decision Decoding Algorithm for Expander Codes[J]. IEEE Communication Letters, 1998, 2(7): 183-185.
- [5] KOU Y, LIN S, FOSSORIER MPC. Low density parity check codes based on finite geometries: A rediscovery and new results[J]. IEEE Transactions on Information Theory, 2001, 47(7): 2711-2736.
- [6] NOUH A, BANIHASHEMI A. Bootstrap decoding of low-density parity check codes[J]. IEEE Communication Letters, 2002, 6(9): 391-393.
- [7] LIU Z, PADOS DA. A decoding algorithm for finite-geometry LDPC codes[J]. IEEE Transactions on Communications, 2005, 53(3): 415-420.
- [8] AL JA. A new Class of Attacks On McEliece Public-Key and Related Cryptosystems[A]. The 2001 Canadian Workshop On Information Theory[C], 2001.