

基于IP的体全息存储系统的设计与实现

吴明, 吴非, 谢长生

(华中科技大学计算机学院外存储国家重点实验室, 武汉 430074)

摘要: 体全息存储技术不同于传统的光存储、磁存储等二维存储技术, 与其它传统大容量存储技术相比, 具有存储容量大、数据传输率高、存取时间短以及可快速进行图像匹配和内容相关寻址操作等特点, 使其有可能成为下一代海量存储设备。但体全息存储系统要成为性能优良的计算机数字信息存储系统, 充分发挥体全息存储系统独特的优势, 满足实用化的要求, 必须使其传输速率达到 100MB/s。该文结合 IP 网络存储技术, 将 IP 和全息结合起来设计了体全息存储高速数据传输通道。

关键词: 体全息存储; 网络存储; 网络处理器; Linux; 千兆以太网

Design and Implementation on IP Network Technology for Volume Holographic Data Storage

WU Ming, WU Fei, XIE Changsheng

(National Storage System Laboratory, Huazhong University of Science and Technology, Wuhan 430074)

【Abstract】 Volume holographic storage is different from the other 2D storage such as optical or magnetic. It has the potential to offer high density, fast data readout rate, and associative content addressable storage as compared with other conventional mass data storage technology. However, in order to be a digital data storage system for computer, and maintain the requirement of application, it is very important for volume holographic storage to attain the goal of 100MB/s. This paper designs the high-speed data channel for volume holographic storage based on the IP network technology.

【Key words】 Volume holographic data storage; Network attached storage; Network processor; Linux; Gigabit Ethernet

1 概述

人们对于存储扩容、升级速度和共享的需求, 已经远远超出了对 CPU 和服务器的升级速度。Internet 技术的发展、电子商务和自动化管理的应用, 使得网络中的数据迅速膨胀。而网络数据和信息的急剧增加也正在带来存储市场的爆炸性增长。

传统的存储设备主要有以下两方面的问题:

(1)传统的存储技术主要是基于二维面存储技术, 如磁存储、传统光盘存储和半导体存储等。虽然这些技术仍在不断地改进以满足对存储系统更大和更快等要求, 然而这些存储手段正逐步接近其物理极限。目前正在迅速发展的一种有别于传统技术的方法就是三维存储技术, 如体全息存储等。体全息存储具有存储密度大、高的数据传输速率和快的读出时间等优点。在体全息存储技术中, 数据按“页面”方式并行读写, 与传统存储技术中按“位”操作的串行方式显著不同。以页为单位存储和利用内在的并行性恢复数据, 读出数率可高达 1GB/s, 而仅仅受到 I/O 设备 (SLM、探测器阵列) 和传输通道速率的限制。

(2)传统的网络存储设备都是将 RAID 硬盘阵列直接连接到网络系统的服务器上, 这种形式的网络存储结构称为 DAS (Direct Attached Storage)。在这种结构下, 对存储设备的访问实际上是通过服务器来完成, 存储设备本身并非严格意义上的网络设备。

真正的网络存储设备应该是基于 IP 的 NAS (Network Attached Storage)。NAS 将存储设备通过标准的 IP 网络拓扑

结构连接到一群计算机上, 所以 NAS 在适用性方面具有不少优势。首先, NAS 可以无需服务器直接上网, 不依赖通用的操作系统, 而是采用一个面向用户设计的、专门用于数据存储的简化操作系统, 内置了与 IP 网络连接所需的协议, 因此使整个系统的管理和设置较为简单, 其次 NAS 是真正即插即用的产品, 并且物理位置灵活, 可放置在工作组内, 也可放在其他地点与 IP 网络连接。

所以, 体全息存储系统作为一种新型的网络存储系统, 为了达到实用性要求, 在其接口方面需要解决两方面问题:

(1)存储系统的传输速度要求达到 100MB/s; (2)为了网络访问而需要的网络接口标准。

2 体全息存储原理

图 1 所示是一个典型的体全息存储系统。图中的两个透镜作傅立叶变换操作。在传统的磁盘存储和光盘存储中, 一个信息位是由介质表面物理性质的改变 (如消融的凹点或磁畴的翻转等) 来表示的。而在体全息存储中, 一位的信息分布在整个记录体中, 在记录介质上没有同信息位一一对应的微小元素。体全息存储是以一整页的形式进行的, 记录介质是厚的感光光学材料, 存储的是两束相干激光束在存储材料中相遇而产生的光学干涉图样。

基金项目: 国家“973”计划基金资助项目(G1999033006)

作者简介: 吴明(1968—), 男, 博士生, 主研方向: 计算机系统结构大容量数据存储; 吴非, 博士生、讲师; 谢长生, 教授
收稿日期: 2005-08-18 **E-mail:** limitw@dragonhk.com

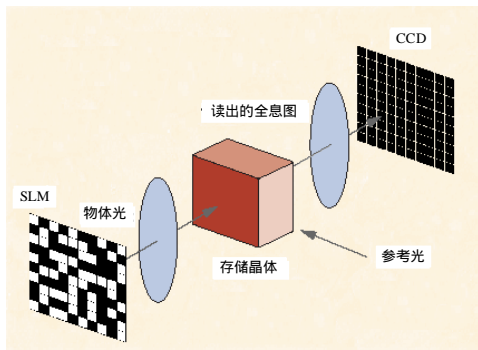


图1 典型的体全息存储系统示意图

通常情况下，体全息存储使用的两束光是由一束激光分离而成：一束经过空间调制而携带信息，称为物光；另一束直接到达存储体，称为参考光。存储时，一个二维的振幅型空间光调制器SLM (spatial light modulator)，以二维的‘0’和‘1’所组成的二值数据页代表欲存储的信息。物光通过SLM携带存储信息经过傅里叶变换后与参考光在记录介质中发生干涉，光学干涉图样引起感光材料发生化学或物理变化。感光材料在吸收率、折射率或厚度上相应的变化就作为干涉图样的复制品存储起来^[1-3]。

数据读出时，读出光照射体全息记录介质，衍射光经傅里叶变换后成像于探测器阵列(CCD)表面，探测器阵列根据各像素接收到的光强大小将光信号转变为电信号。体全息存储技术固有的面向页面的特性，使数据按“页面”方式并行读写，与传统存储技术中按“位”操作的串行方式显著不同。

3 体全息存储通信及控制接口系统的设计

3.1 作为标准的网络设备的体全息存储系统应用

图2给出了作为标准的网络设备的体全息存储系统的应用框图，计算机通过千兆网卡和光纤，体全息存储系统通过千兆通信接口和光纤，经标准的千兆交换机或路由器连接。这样，计算机可以通过高速千兆以太网对体全息存储系统进行读写操作。

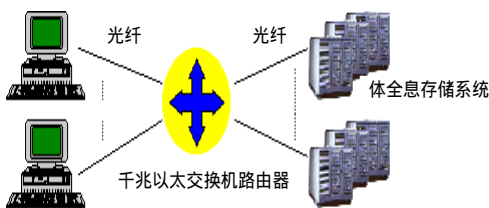


图2 体全息存储系统在标准以太网络中的应用

3.2 体全息存储系统通信及控制接口系统的总体设计

本系统设计共分以下5个模块：

- (1)千兆以太网接口模块。完成标准千兆以太网上数据流的收发处理。
- (2)通信协议及文件系统处理模块。基于Linux嵌入式操作系统，完成标准的TCP/IP协议栈和网络文件服务系统(NFS)的处理^[4-6]。
- (3)编解码处理模块。针对体全息存储的读写特点，完成数据的RS、交错、差分的编解码^[7-9]。
- (4)CCD读控制模块。按照系统读要求，控制CCD读取体存储器的一页数据，并将数据编解码模块进行处理。
- (5)SLM写控制模块。按照系统写要求，将数据编解码模

块输出的数据转换成标准的RGB信号，控制SLM写入体存储器一页。

图3给出了本系统组成。

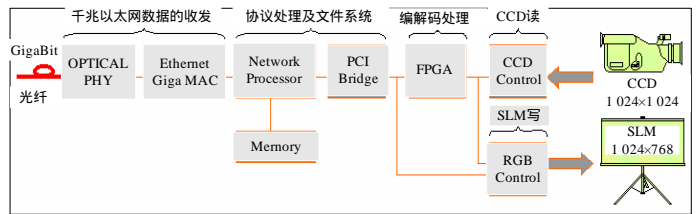


图3 体全息存储系统的通信及控制模块

3.3 体全息存储系统应用软件及协议栈

本系统软件及协议包括两方面：

(1)网络中的计算机，采用标准Windows操作系统和远程过程调用(RPC)软件包来完成客户程序的编写。常用的RPC软件包是Sun RPC，它的一个版本是建立在API基础上和TCP/UDP交换数据。

(2)网络中的体全息存储系统侧，采用开放源码基于网络处理器的Linux操作系统和TCP/IP协议栈来完成NFS服务器的功能。

图4描述了应用软件和协议栈组成。

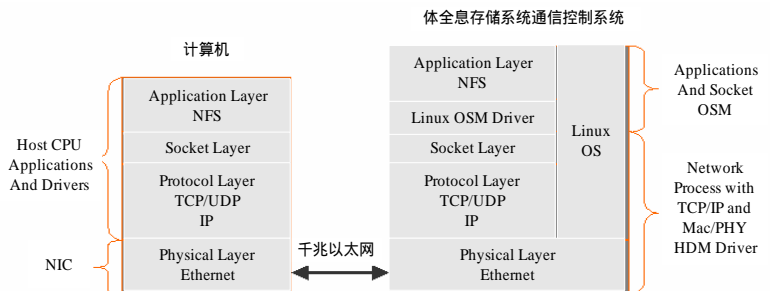


图4 体全息存储系统的协议栈组成

图5描述了NFS客户和NFS服务器的典型配置

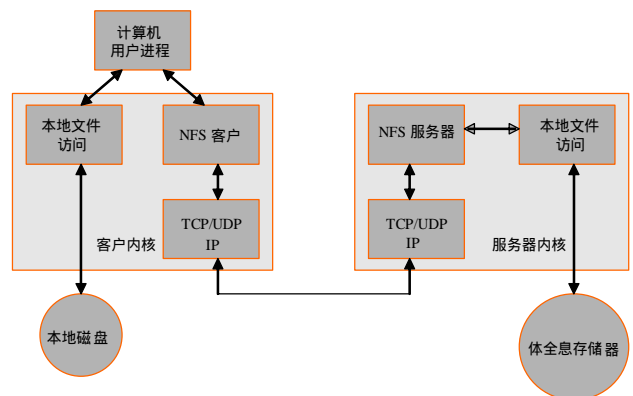


图5 NFS客户和NFS服务器的典型配置

4 总结

针对体全息存储系统接口的设计与研究，本文有两方面的创新：

- (1)将IP网络存储技术引入体全息存储中，充分发挥了体全息存储的存储密度大、高的数据传输速率和快的读出时间等优点，扩展了体全息存储的应用领域。
- (2)通信与控制接口的硬件设计采用当今世界上最先进的Intel IXP1200网络处理器为核心^[10,11]，软件采用Linux操作系统，TCP/IP协议栈及NFS文件管理。整个系统方案的设计具有世界领先水平。

(下转第270页)