

李刚, 杨奕, 孙路强, 栗连弟, 刘磊, 周利霞, 王晓磊, 丁晶, 2013. iSCSI 技术在天津地震数据存储中的应用. 震灾防御技术, 8 (3): 319—325.

iSCSI技术在天津地震数据存储中的应用¹

李 刚 杨 奕 孙路强 栗连弟 刘 磊 周利霞 王晓磊 丁 晶

(天津市地震局, 天津 300201)

摘要 天津市地震局在进行专业数据存储时, 采用了基于 iSCSI 技术的 IP SAN 网络, 与原有 FC SAN 网络结合, 形成了多种应用模式、不同服务等级的数据存储网络。本文详细介绍了基于 EMC CX4-120 系统的 iSCSI 存储技术, 以及其在地震专业数据存储中的应用情况, 分析对比了几种存储技术的数据交换效率, 并为今后 iSCSI 存储技术在行业中的应用提出了建议。

关键词: iSCSI IP SAN 地震 数据

引言

天津市地震局信息网络系统在“十一五”天津市地震安全基础工程项目建设中, 购置了基于 FC SAN 和 iSCSI 架构的 EMC CX4-120 存储系统, 它是 EMC CX4 系列中为部门级应用或中型企业提供的网络存储产品, 具备“五个 9”的可用性并有多种技术创新, 如全自动存储分层、FAST 缓存、闪存驱动器、64 位操作系统和多核处理器等, CX4-120 的容量可扩展到 120TB。CX4-120 通过 iSCSI 实现的 IP SAN 存储网络环境(曾薇等, 2011), 可以大幅降低统一存储系统的建设与应用成本, 实现基于普通 IP 网络的数据存储能力。本文将详细阐述 iSCSI 技术在天津市地震局数据存储环境中的应用情况。

1 iSCSI 介绍

1.1 iSCSI 介绍

iSCSI (互联网小型计算机系统接口) 技术是一种基于 IP Storage 理论的新型存储技术, 它将存储行业广泛应用的 SCSI (small computer system interface) 接口技术与 IP 网络技术结合, 从而在 IP 网络上构建出 SAN 存储区域网, 它是一种在 Internet 协议网络上, 特别是以太网上进行数据块传输的标准。iSCSI 由 Cisco 和 IBM 两家发起, 并且得到了 IP 存储技术拥护者的大力支持, 是一个供硬件设备使用的可以在 IP 协议上层运行的 SCSI 指令集。

推动 iSCSI 协议发展的最主要的动力是人们希望能够在 IP 网络上使用业已部署的大量 DAS 存储设备。通过 iSCSI 协议(蒋晓山, 2007), 这些存储设备可以为更多的用户和应用使用, 同时可以让这些简单 DAS 存储设备支持只有高级存储设备才能够支持的备份、镜像、

¹ 基金项目 地震科技星火项目“天津地震观测数据可视化展示与发布系统初步研究与应用(XH12003Y)”

[收稿日期] 2013-02-17

[作者简介] 李刚, 男, 生于 1978 年。高级工程师。主要从事地震信息网络规划与建设工作。E-mail: ligang@tjdzj.com

灾难恢复等高级存储应用。

1.2 iSCSI 存储网络的构成

iSCSI 结构基于客户/服务器模式，iSCSI 在 TCP/IP 网络上的主机系统（启动器 initiator）和存储设备（目标器 target）之间进行大量数据的封装和可靠传输。

iSCSI 存储网络中主要包括：提供 iSCSI 协议支持的存储设备和具备 iSCSI 协议的主机系统。主机系统必须具备：①Initiator 驱动程序+普通网卡；②硬件 TOE HBA 卡；③iSCSI HBA 卡三者之一的 iSCSI 软件或硬件卡，实现与存储设备的数据交换。同时还要有具备数据交换的 IP 网络系统，目前以千兆或万兆以太网为主。图 1 为 iSCSI 网络结构示意图。

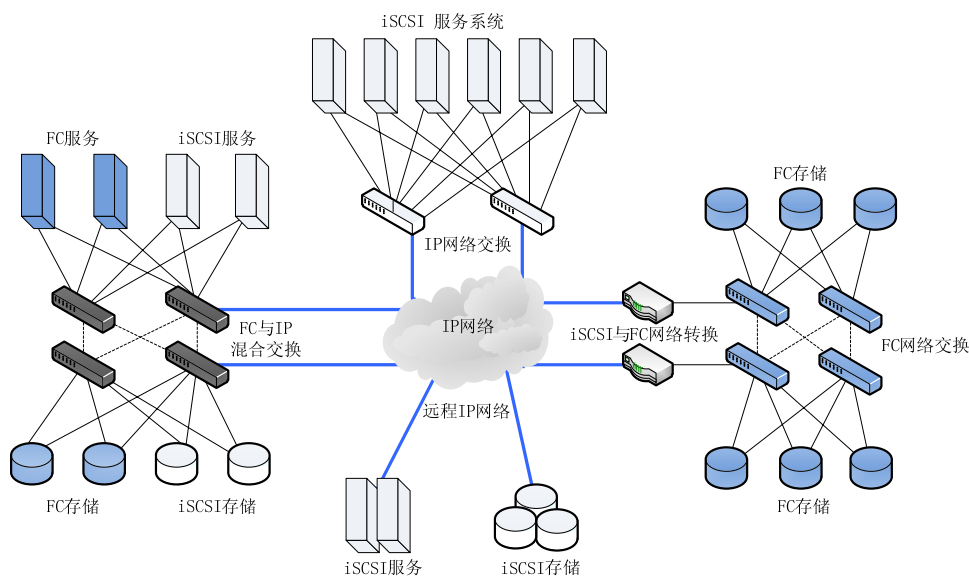


图 1 iSCSI 网络结构示意图

Fig. 1 Frame structure of the iSCSI network

1.3 iSCSI 的优缺点

iSCSI 技术的优点和成本优势主要表现在以下几个方面（张志宏，2011）：

（1）硬件成本低。构建 iSCSI 存储网络，除了存储设备外，交换机、线缆、网络接口卡都是标准的以太网配件，价格相对来说比较低廉。同时，iSCSI 还可以在现有的网络上直接安装，不需要更改网络体系，可以最大程度地节约投入。

（2）操作简单，维护方便。对 iSCSI 存储网络的管理，实际上就是对以太网设备的管理，只需花费少量的资金去培训 iSCSI 存储网络管理员。当 iSCSI 存储网络出现故障时，问题定位及解决也比较容易。

（3）扩充性强。对于已经构建的 iSCSI 存储网络来说，增加 iSCSI 存储设备和服务器都将变得简单且无需改变网络的体系结构。

（4）带宽和性能强。iSCSI 存储网络的访问带宽依赖以太网带宽。随着千兆、万兆以太网的普及和应用，iSCSI 存储网络会达到甚至超过 FC（Fiber Channel，光纤通道）存储网络的带宽和性能。

(5) 突破距离限制(姜楠等, 2010)。iSCSI 存储网络使用的是以太网, 因而在服务器和存储设备空间布局上的限制就会少了很多, 甚至可以跨越地区和国家。

虽然 iSCSI 有上面众多的优势, 但从目前的实际情况来分析, iSCSI 还有如下问题需要改进或优化:

(1) 噪声碰撞问题。iSCSI 利用的是 IP 网络, 其中充斥着各类数据及噪声, 所以碰撞情形在所难免, 在数据传输的过程中, 容易导致延迟发生, 影响传输效能, 甚至数据的正确性。

(2) 效能瓶颈问题。主要包括: 传输频宽问题、流量控制问题、I/O 端口的速度限制问题以及软件 iSCSI Initiator 的效率问题。其中 iSCSI 的效能问题将是其在未来一段时间内要进行突破的重点工作。

(3) 硬件 iSCSI 适配卡成本高。要想让 IP SAN 整体效能有好的表现, 就必须添置成本较高的 iSCSI HBA 卡或 TOE HBA 卡(TCP Offload Engine), 从而使 SAN 网络的整体成本大幅攀升。

(4) 安全性问题。IP 网络环境复杂, 相对封闭或独立的 FC SAN 环境而言, iSCSI 构成的 IP SAN 环境, 其安全性也有更多的工作需要加强。

2 iSCSI 技术在天津地震数据存储中的应用

2.1 网络系统概况

通过“十五”及“十一五”项目的建设, 天津市地震局组建了具备 58 条 10M 或 100M 光纤, 54 个节点间的高速互联 MPLS VPN 区域网络通信系统, 这些线路与节点组成了 4 个 MPLS VPN 分组, 如图 2 所示。

在天津市地震局网络中心与滨海地震台通过 100M MPLS VPN 网络, 利用 iSCSI 技术实现了部分数据在两地间的实时互备, 如图 3 所示。

2.2 存储需求

天津市地震局在“十五”及“十一五”期间, 改扩建了多类地震观测系统(王建国等, 2010), 监测类型不断扩展, 每年产生约 5TB 的观测数据。以测震系统为例, 除了连续波形数据外, 天津地震台网每年需要备份与归档的本地波形数据约 1TB, 以往此类数据主要通过刻录光盘形式进行存储, 在硬盘存储成本下降后, 测震台网采用分时段存储模式, 一般在线归档数据保存时间为 4 个月, 其他历史数据还是采用刻录光盘进行备份(李刚等, 2012a; 2012b)。此种方式虽然可以满足一般需求, 但在人们对数据依赖度不断增长的要求下, 还是不太适应信息化系统发展与服务的需要。

在不降低或不增加运营成本的前提下, 为了实现天津市地震局内数据的统一存储管理, 满足数据长期在线服务, 作者采用 iSCSI 技术为测震系统建立了一套地震事件波形数据存储系统, 实现了天津地震台站归档波形数据在线 5—10 年服务的能力, 系统具备 RAID5 级别存储安全, 并可快速挂接到不同服务器实现对外服务。

2.3 建设过程与主要命令

在建设地震事件波形数据存储系统时, 由于地震归档波形数据只需要提供文件上传下载服务, 因此, 作者通过 1 台普通的机架服务器作为服务设备, 配置为单 CPU, 1G 内存, 2 个 1000M 网卡, 操作系统为 Suse10.2, 从 EMC 阵列上为此服务器划分 5TB 存储空间用于数

据的在线存储。

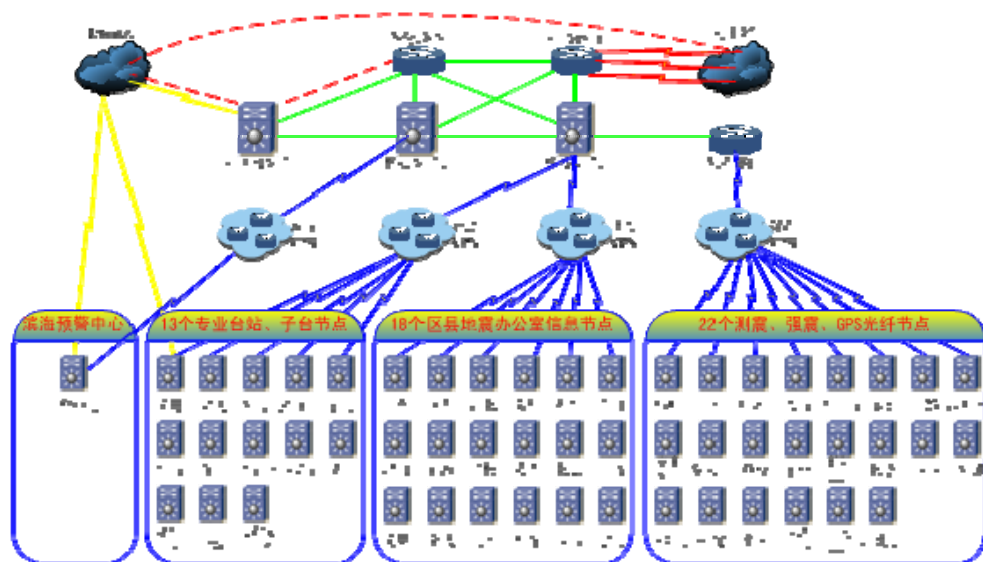


图2 天津市地震局 MPLS VPN 网络拓扑

Fig. 2 Topology of the MPLS VPN

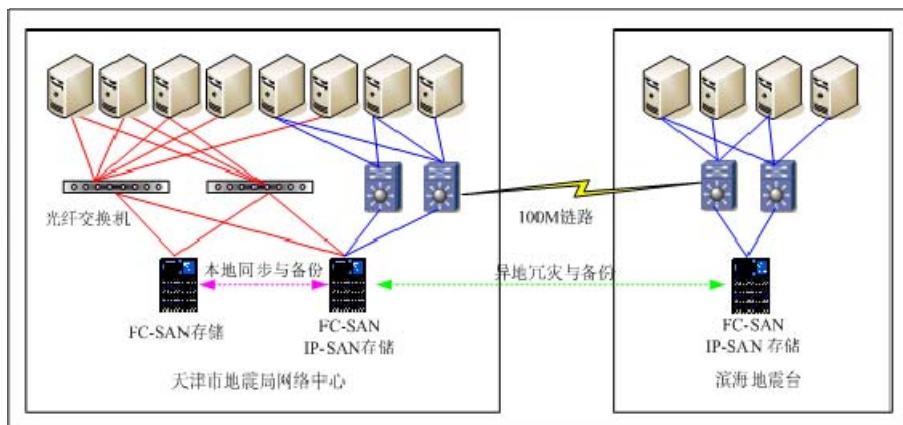


图3 异地数据备份

Fig. 3 Data backup offsite

服务器的2个网卡用做提供网络服务与实现 EMC CX4-120 iSCSI 网络的连接, 实现网络服务与数据交换的独立处理。系统连接拓扑如图4所示。

服务器与 EMC 的连接过程, 可以通过相应的 Yast 图形界面来完成, 下面主要介绍通过命令行方式进行的操作。

(1) 进入 SuSe10 SP2 系统, 通过“rpm -qa | grep iscsi”查看当前系统是否已经安装了 iSCSI 启动器, 如果没有, 从 SuSe10 SP2 系统第一张安装盘中的“\suse\i586”目录下找到“open-iscsi-2.0.707-0.44.i586.rpm”安装包, 执行: rpm -ivh open-iscsi-2.0.707-0.44.i586.rpm 进

行安装。

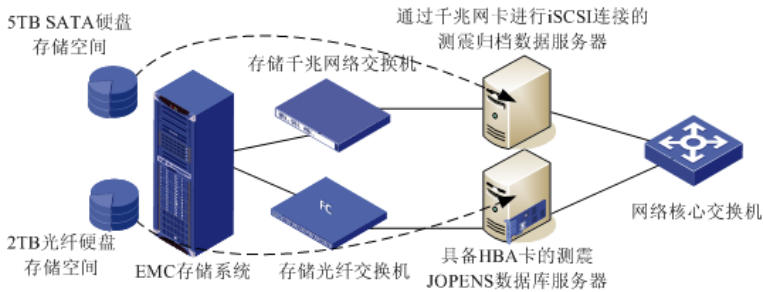


图 4 iSCSI 系统连接示意图

Fig. 4 The iSCSI system connection

(2) 通过/etc/init.d/open-iscsi start 启动 iSCSI 服务。

(3) 执行 iscsiadm -m discovery -t st -p targetip 查找 EMC 存储。如: iscsiadm -m discovery -t st -p 10.12.3.251。在查找到 EMC 存储设备后, 在 EMC 管理系统 Navisphere 中, 为服务器与相应的 LUN 建立 Storage Group, 使服务器与 LUN 具备建立数据交换的条件。

(4) 执行 iscsiadm -m node -p targetip -l 将本机注册到 EMC 存储。如: iscsiadm -m node -p 10.12.3.251 -l。完成在 EMC 上的注册后, 在服务器上通过 fdisk -l 可以查看到对应的 LUN 已经分配给服务器, 建议此时通过 Yast 中的 Partitioner 将新的 LUN 空间添加到系统中, 并确保在/etc/fstab 文件中有对应的记录信息, 以确保在完成所有操作后重新启动系统时, 相应的 LUN 会被自动挂载到服务器上。

(5) 设置 iSCSI 自启动。如: chkconfig -s open-iscsi on -level 235。

(6) 编辑/etc/iscsid.conf 文件, 设置自动登录到 EMC 存储器。如: node.startup=automatic。

(7) 修改/etc/init.d/rc, 在文件末尾“exit 0”前面添加:

```
/sbin/iscsiadm -m node -p 10.12.3.251 -l
```

```
/usr/sbin/hot_add
```

```
/bin/mount -a
```

设置 iSCSI 设备开机自动挂载相应的 LUN。

完成上述 7 个步骤后, 服务器已经具备通过 iSCSI 协议自动与 EMC 进行空间挂载的能力, 每次重新启动服务器时, EMC 上为服务器分配的 LUN, 会被服务器以设置好的方式装载到系统中。

2.4 系统测试

在测试 iSCSI 协议的应用效果时, 主要是通过文件传输效率来做对比分析。测试文件分为三种, 即 50 个 20M 文件、10 个 50M 文件和 1 个 600M 文件。测试存储目标分为四种, 即服务器本地 SATA 硬盘、iSCSI 连接的 SATA 硬盘空间、FC 连接的 SATA 硬盘空间和 FC 连接的光纤硬盘空间。文件传输方式采用 SCP, 服务器网络传输速率分为 100M 和 1000M 两种。

测试过程:

(1) 在测试服务器上通过如下命令建立不同类型和数量的测试文件

```
dd if=/dev/zero of=F2050.tmp bs=20M count=1 建立 20M 测试文件
```

(2) 通过 SCP 命令, 将生成的测试文件传送到不同形式的存储空间, 测试结果如表 1 和

表 2 所示。

表 1 不同存储模式数据传输时间对比 (100M/s)

Table 1 Comparison of the data transmission time in different storage modes at 100M/s

存储模式	20M 文件 50 个 传输时间 (s)	50M 文件 10 个 传输时间 (s)	600M 文件 1 个 传输时间 (s)
服务器 7200 转 SATA 硬盘	98	45	53
EMC iSCSI 7200 转 SATA 硬盘	92	45	53
EMC FC 7200 转 SATA 硬盘	89	45	53
EMC FC 15k 转光纤硬盘	88	45	53

表 2 不同存储模式数据传输时间对比 (1000M/s)

Table 2 Comparison of the data transmission time in different storage modes at 1000M/s

存储模式	20M 文件 50 个 传输时间 (s)	50M 文件 10 个 传输时间 (s)	600M 文件 1 个 传输时间 (s)
服务器 7200 转 SATA 硬盘	42	17	13
EMC iSCSI 7200 转 SATA 硬盘	41	16	12
EMC FC 7200 转 SATA 硬盘	40	18	13
EMC FC 15k 转光纤硬盘	42	16	13

通过测试可以看出,在 100M、1000M 的网络环境内,FC 存储与 iSCSI 存储提供的数据交换能力主要还是限制在通讯网络带宽上,与存储模式、硬盘类型、硬盘转速的关系不大,即在 1000M 带宽网络环境以内,iSCSI 作为文件服务类应用是满足用户需要的。通过测试证明,为测震系统建设的波形数据归档在线服务在当前网络带宽下是满足需要的。

2.5 应用分析

采用 iSCSI 技术后,天津测震台网每天定时将波形文件归档为台站卷、台网卷后,将 2.8GB 数据存储到本地与滨海地震台远程磁盘阵列中,本地存储时间约 1 分钟,通过 100M 光纤存储到滨海地震台约 5 分钟。通过 iSCSI 技术,即解决了本地磁盘空间不足,数据保存时间短的问题,又降低了系统的建设与运行成本。系统投入使用后,对于地震事件文件、连续波形文件的查询读取更加方便快捷,同时实现了数据存储多方式、多途径,降低了单点故障造成的数据丢失。

3 结论

iSCSI 技术的发展与应用是未来存储网络的一个重要方向,随着 10Gb 以太网的逐步普及,iSCSI 的数据交换带宽也将超过 FC SAN 中的 8G 主流带宽,其应用前景十分看好。并且 iSCSI 技术上的许多优势,在数据灾备、远程存储与数据共享(李刚等,2012a; 2012b)、存储与服务虚拟化(李刚等,2011)等方面,也有着值得我们继续深入研究与学习的内容。

目前作者只是对基于文件的地震数据进行了 iSCSI 应用存储,今后将在更多的业务应用中对 iSCSI 技术进行分析研究,使这项技术能为地震系统提供更好的服务,同时也降低行业内存储网络的整体运营成本。

致谢: 本项工作得到了天津市地震局各级领导的支持与帮助, 特别是从事监测预报工作的同事们, 在此表示衷心的感谢!

参考文献

- 蒋晓山, 2007. IP SAN 存储技术面向地震信息网的发展前景分析. 山西地震, **4**: 16—18, 42.
- 姜楠, 李战怀, 余承龙等, 2010. 基于 iSCSI 的多数据保护技术容灾系统框架研究. 科学技术与工程, **10**: 2502—2506.
- 李刚, 周利霞, 王晓磊等, 2011. 基于 VMware Esxi 软件的服务器虚拟化技术在天津地震监测网络中的应用. 震灾防御技术, **6** (3): 311—318.
- 李刚, 栗连第, 孙路强等, 2012a. EMC 存储系统在天津测震台网 JOPENS 数据库服务中的应用. 地震地磁观测与研究, **33** (1): 100—104.
- 李刚, 孙晶岩, 卞真付等, 2012b. MPLS VPN 高速区域网络在天津地震监测系统中的应用. 震灾防御技术, **7** (1): 92—99.
- 王建国, 姚会琴, 高逊等, 2010. 天津市地震前兆台网的运行监控与维护管理. 大地测量与地球动力学, **30** (增刊 1): 111—115.
- 曾薇, 杨乐, 谭颖, 2011. 网络存储技术在地震数据存储中的应用. 震灾防御技术, **6** (3): 335—342.
- 张志宏, 2011. 构建下一代数据中心的基础架构. 海峡科学, **6**: 63—65, 68.

The Application of iSCSI Technology in Tianjin Seismic Data Storage

Li Gang, Yang Yi, Sun Luqiang, Li Liandi, Liu Lei, Zhou Lixia,
Wang Xiaolei and Ding Jing

(Earthquake Administration of Tianjin Municipality, Tianjin 300201, China)

Abstract In seismic data storage Earthquake Administration of Tianjin Municipality, IP SAN network based on the technology of iSCSI is adopted. By combining with the original FC SAN network, a variety of application mode, data classified with different levels of service network are developed. This paper introduces in detail the iSCSI storage technology of the system based on EMC CX4-120, its application in seismic data storage situation, analyzes several kinds of storage technology in data exchange efficiency, and proposes the application of iSCSI storage technology application in industry for the future.

Key words: iSCSI; IP SAN; Seismic; Data