

检验检疫融合存储数据容灾应用研究

张宗平¹, 秦磊华²

(1. 广东出入境检验检疫局信息中心, 广东 广州 510623;

2. 华中科技大学 计算机科学与技术学院, 湖北 武汉 430074)

摘要:文中首先介绍了笔者所在单位目前存在IP SAN和FC SAN的应用带来的管理需求。为建立快速高效的灾难恢复解决方案,保障业务系统的连续运行使保障检验检疫工作正常,针对应用不同存储系统的特点,探索一种可以融合多种不同存储的容灾技术,提出并浅析了在可兼容FC及IP存储(包括iSCSI和NAS)的网络存储基础架构上,研究融合存储数据容灾的可行性,并利用系统原型完成融合存储数据容灾应用的研究。对今后基于融合存储的数据容灾研究有较好的理论借鉴意义和实际参考价值。

关键词:检验检疫;融合存储;数据容灾

中图分类号:T333.4

文献标识码:A

文章编号:1673-629X(2014)08-0172-03

doi:10.3969/j.issn.1673-629X.2014.08.040

Research on Storage Data Disaster Recovery Application Integrated Inspection and Quarantine

ZHANG Zong-ping¹, QIN Lei-hua²

(1. Information Center of Guangdong Entry/Exit Inspection and Quarantine Bureau,

Guangzhou 510623, China;

2. School of Computer Science and Technology, Huazhong University of Science and Technology,

Wuhan 430074, China)

Abstract: In this paper, firstly introduce the application of the existing IP SAN and FC SAN to bring the management demand. In order to build a rapid and efficient disaster recovery solution, ensuring continuous operation of business system is the guarantee of inspection and quarantine work, according to the characteristics of different storage system, explore a disaster recovery technology which can merge a variety of different storage, in compatible with FC and IP storage (including iSCSI and NAS) architecture based on network storage, studying the feasibility of fusion storage data disaster recovery, and making use of the system prototype to complet fusion of storage data disaster recovery applications. For the future research based on the data fusion disaster recovery storage has good theoretical significance and practical reference value.

Key words: inspection and quarantine; integrated storage; data disaster tolerance

0 引言

随着检验检疫工作的信息化程度不断提高,业务流程对数据的依赖性也在日益增加。目前,笔者所在单位在综合业务管理系统、日常办公、财务管理、人事管理、标准网、档案管理、固定资产管理以及各科研生产管理等诸多方面全面使用信息系统。特别是近几年来,业务系统的信息化发展迅速,导致信息系统的数

量迅速增长,业务数据的重要程度日益增加,任何重要信息的丢失都将会给科研生产带来难以估计的损失和影响。因此,如何保证核心科研生产系统的业务连续性和数据的安全性,如何建立快速高效的灾难恢复解决方案,保障业务系统的连续运行是保障检验检疫工作正常的必要条件。

目前在用的存储系统中往往包含FC SAN和IP SAN,备份要兼顾FC SAN和IP SAN的特点和应用要

收稿日期:2013-10-26

修回日期:2014-01-27

网络出版时间:2014-05-29

基金项目:国家质检总局科技计划项目(2012IK256)

作者简介:张宗平(1977-),男,广东雄人,高级工程师,系统分析师,研究方向为DSS系统相关及OLTP业务系统建设、维护,Web应用系统开发、数据交换及嵌入式系统应用技术等;秦磊华,副教授,博士,研究方向为计算机网络和网络存储系统。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20140529.1623.001.html>

求。笔者单位的核心业务系统等重要系统使用的是 FC SAN 存储网络,而辖区内下属各分支机构的相关应用使用的是 IP SAN 存储网络,但 FC SAN 和 IP SAN 这两种存储网络因物理介质和传输协议的不同导致信息孤岛现象。为了实现对现有信息化系统数据容灾,研究设计基于 FC SAN 和 IP SAN 两种存储网络融合存储系统的数据容灾,研究融合存储环境下的数据备份机制,提高融合存储系统中数据的可用性。

1 基于融合存储的数据容灾设计目标

现行主流的 SAN 主要采用 FC 或 IP 方式连接网络存储,其连接方式较为单一,同样缺点也显而易见。虽然采用 FC 连接方式的 SAN 在数据传输速度上有优势,但其缺点是价格昂贵而且传输距离较短;而采用 IP 方式连接的 SAN 虽然在可扩展性和传输距离上都有较大的优势,但是数据传输速度仍是其瓶颈。分析两种连接方式的优缺点后设计融合存储平台,综合两种连接方式的优势较好地解决了这一问题^[1]。

借助融合存储平台就很好地 IP SAN 和 FC SAN 的互联互通的融合性问题解决了。存储服务器通过 IP、FC 等不同连接方式将不同的远程存储资源映射到平台中进行统一虚拟化,并通过 IP、FC 等不同连接方式对外提供块级存储服务,并且可以在不中断服务的情况下动态地对统一虚拟化存储资源进行配置与管理^[2]。

基于融合存储的灾难备份系统的体系结构和实现方法,在可兼容 FC 及 IP 存储(包括 iSCSI 和 NAS)的大规模可扩展网络存储基础架构上^[3],研究通过建立网络架构级的智能灾难备份存储网络,提高整个网络存储系统的可用性和可靠性,主要的研究内容有以下三方面。

1.1 基于融合存储的存储系统架构

通过对融合 IP 和 FC 的大规模可扩展存储系统的组织与架构的研究,提出建立融合 FC、iSCSI 和 NAS 的大规模网络存储系统的设计构想以及实施核心存储交换的思路,满足 IPv4 及 IPv6 环境下构建大规模存储系统的需要^[4],以便解决目前 IP SAN 核心互连能力不足、管理复杂以及设备匮乏等问题。本项目拟构建的网络存储系统是一个支持各类存储应用的和谐统一的网络存储大系统,构建能在 IPv6 环境下的物理存储基础设施之上,由既分立工作又互相联系的动态可调的多个虚拟存储子网络构成。系统通过若干安全融合控制器将普通 IP 交换机、FC 交换机、本地及远程分支机构的存储设施以及各种存储服务连接起来,形成的每个虚拟存储子网络可进行服务器端口、存储端口、交换端口的动态配置,利用安全融合控制器中的互连功能,

可在必须进行通信的虚拟存储子网络间配置数据迁移通道,提供数据服务能力^[5]。

1.2 多层存储容灾策略与故障检测技术

通过对故障检测技术的研究,对融合存储系统中不同层次的故障(存储子系统层、融合存储控制器层、存储网络链路层)进行分析,并进行针对性处理,最大限度地保障存储服务连续性,实现存储与容灾间的融合。通过对融合存储控制器层、存储池层和链路层各层次中可能出现的各种故障的假设,以及按故障发生后对系统的影响进行分析和分类,运用相对应的存储迁移策略进行服务迁移。待所发生的故障修复后,实施存储服务回迁^[6]。

1.3 基于融合存储与存储容灾策略的容灾技术

通过研究融合存储系统环境下的网络架构级智能灾难备份的体系结构,利用主站点与灾难备份站点之间建立的多层存储容灾策略与故障检测机制,借助服务 IP 地址的接管、存储服务接管以及自动存储服务重定向等措施,将存储应用和存储设备有机结合在一起,并将基于逻辑卷的数据远程复制、数据快照与恢复、服务故障检测与接管以及存储应用自动转接服务等灾难备份服务智能部署其中,最大限度地降低宕机时间,提高应用系统的可用性,满足系统服务连续性的要求。

2 基于融合存储的容灾设计实现思路

融合存储系统环境下的容灾设计思路主要从代理端的设计、基于 Web 的 Client 端设计、融合式存储容灾原型系统环境设计和融合式存储容灾原型系统工作流程设计四个方面进行研究。

2.1 代理端的设计—系统资源的 SMI-S 服务化

所谓系统资源的 SMI-S 服务化,就是将系统中需要管理的资源,使用符合 SMI-S 的中间件进行描述,以使其能够被符合 SMI-S 规范的管理软件统一管理。这个符合 SMI-S 的中间件就是 CIM。简单的说,在文中系统资源的 SMI-S 服务化,就是具体的各个子系统 Provider 的实现^[7]。

针对本项目的存储系统,主要目的是对要管理的子系统实现底层存储设备、系统状态、逻辑卷管理和网络链路管理。对应起来,在相应的系统各节点上部署并实现 DiskProvider、OperationSystemProvider、Storage VolumeProvider、FCProvider 以及 iSCISProvider。基于融合存储控制器的存储容灾系统设计架构如图 1 所示。

2.2 基于 Web 的 Client 端设计

基于 Web 的 Client 采用 PHP 和 Apache 作为开发应用平台,采用 Mysql 搭建数据库,整个 Web 服务器架设在 ATCA 上。管理员通过浏览器对存储控制器 ATCA 上的应用服务器进行操作,实现远程管理。

Web 服务器架设在 ATCA 上,服务器中 PHP 代码最终会调用 CIMClient 与本地和远程的 CIMServer 端交互获取信息,并将获取的信息通过 Web 的方式呈现给管理者^[8]。

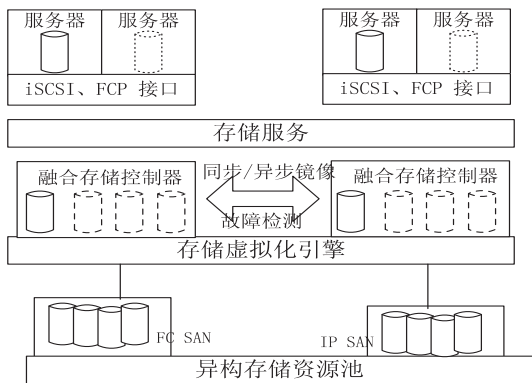


图 1 基于融合存储控制器的存储容灾系统设计架构

2.3 融合式存储容灾原型系统环境设计

所设计的基于融合存储控制器的融合式存储容灾原型共分为四层,从下至上分别为存储池层、融合存储控制器层、应用服务器层和客户端层。基于本地和异地的使用,设计又分为本地存储容灾站点和异地存储容灾站点两部分,两者之间通过广域网模拟器相连。

其中,在本地站点中,FC、iSCSI 和 NAS 服务器分别支持 FC、iSCSI 和 FTP 协议,通过本地存储控制器,使用本地或异地存储池中的存储资源。存储控制器通过数据复制功能将关键应用数据复制到异地存储容灾站点中的存储池中^[9]。

当本地站点发生故障时,通过故障检测机制及存储服务切换策略,应用服务器通过异地站点中的存储控制器可继续使用相应的存储资源。其中,应用数据在故障发生前已通过数据复制保存到了异地站点。通过这样的方式,存储容灾系统就能确保对应用服务器提供连续的存储服务的能力^[10]。

2.4 融合式存储容灾原型系统工作流程设计

整个融合存储容灾系统在本地、异地各设计部署一个融合存储子网站点,两个存储子网的存储环境是对称的^[11]。应用通过融合存储控制器使用存储服务,同时本地与异地环境之间的融合存储控制器协同工作,实现容灾功能。融合存储控制器通过数据复制功能保护应用数据不丢失。本地融合存储控制器和异地融合存储控制器之间通过故障监测,互相监视对方的服务状态,以便在一方发生故障时,接管对方的存储服务,从而避免应用环境由于本地存储环境故障而导致应用服务长时间中断。

存储容灾系统典型工作流程如下^[12-13]:

1) 当本地存储站点发生灾难时:

(1) 通过故障监测手段检测到灾难的发生时间节

点;

(2) 异地融合存储控制器对相应存储资源进行初始化后进入存储服务接管状态;

(3) 应用服务切换到异地融合存储控制器,通过异地融合存储控制器继续使用存储资源。关键应用数据将通过数据复制功能复制到异地存储站点,从而使整个应用服务可以延续。

2) 当本地存储站点从灾难恢复过来时:

(1) 本地融合存储控制器重新发出心跳恢复信息,异地融合存储控制器可检测到本地融合存储控制器已从灾难中恢复;

(2) 异地融合存储控制器为存储服务回切做好准备,如将灾难期间产生的应用数据同步到本地存储站点、传输应用服务状态等;

(3) 当本地融合存储控制器获得相应的数据、状态信息后,便做好了再次提供存储服务的准备;

(4) 本地融合存储控制器向应用服务发出存储服务回滚指令,应用服务重新切换到本地融合存储控制器,系统恢复到正常工作状态。

通过上述设计思路,在融合存储控制器的基础上实现存储容灾系统。

3 结束语

通过基于融合存储系统的数据容灾的应用研究,可实现支持异构存储架构的核心存储交换平台,构成模块化、可扩展、可动态存储分区的网络存储体系,辅以全方位的具备远程数据镜像、数据快照与恢复、广域高可用服务诊断与接管等功能的存储容灾手段,为检验检疫数据中心及各类信息系统提供统一的、满足服务质量要求的容灾网络存储服务。但容灾存储系统在商业领域有着广泛的应用历史和前景,目前,所研发的系统原型,虽然理论上能够满足用户的指标要求,但是,距离真实上线应用还有一段距离,系统的稳定性及技术风险还存在,但对于指导新型的备份系统研究有一定的应用价值和指导意义。

参考文献:

- [1] 汪琦晔. 融合平台下 SAN 设备的访问控制设计与实现 [D]. 武汉:华中科技大学,2008.
- [2] 陈焱雄. 融合式存储系统的数据复制研究与设计 [D]. 武汉:华中科技大学,2009.
- [3] 秦磊华,苏彦君,张宗平. 光纤通道存储区域网扩展研究 [J]. 计算机工程,2007,33(24):109-111.
- [4] 李 瑞. IP 与 FC 融合式存储系统的安全体系设计与实现 [D]. 武汉:华中科技大学,2011.
- [5] 傅湘林,谢长生,曹 强,等. 一种融合 NAS 和 SAN 技术的

(下转第 178 页)

示,横轴表示 GPX 芯片同步测试的平均值,纵轴表示 TDC 每个通道测试结果的平均值。实验表明各通道间的测量误差最大为 50 ps,在 $\pm 1\text{LSB}$ 以内。

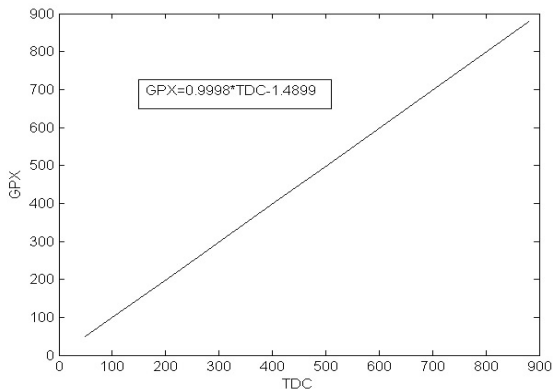


图 5 TDC 精度曲线拟合分析

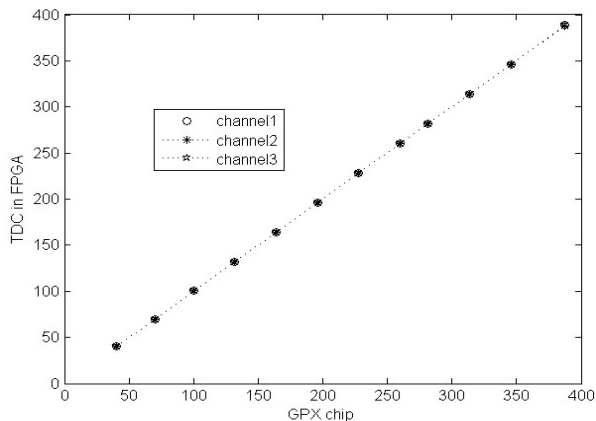


图 6 TDC 多通道一致性分析

4 结束语

文中设计了一款应用于激光测距系统的多通道时间数字转换电路,通过与 GPX 高精度测试芯片的对比测试,准确地检验了 TDC 的性能。实验数据表明,校准后测量精度可达 58 ps,通道间测量误差在 1LSB 以内,支持多回波模式的流水线测量,满足大数据量的实

时测量及精度应用需求,适用于阵列式激光测距系统。

参考文献:

- [1] 张 婷,张珂殊.基于 Matlab 的相位式激光测距研究[J].激光与红外,2010,40(1):22-27.
- [2] Song Jian, An Qi, Liu Shubin. A high-resolution time-to-digital converter implemented in field-programmable-gate-arrays [J]. IEEE Transactions on Nuclear Science, 2006, 53(1):236-241.
- [3] Quiros-Olozbal A, Barrientos-Villar J M, de los Angeles Cifredo-Chacon M. Reconfiguration-based time-to-digital converter for Virtex FPGAs [C]//Proc of international conference on field programmable logic and application. Prague: IEEE, 2009:439-443.
- [4] 朱小兰,李惠军,赵守磊,等.新型高精度大量程 TDC 核的设计与实现[J].仪表技术与传感器,2010(3):75-77.
- [5] 徐 振.基于 TDC 的时间间隔测量技术的研究[D].哈尔滨:哈尔滨工业大学,2008.
- [6] 王福源,杨玉叶,时 伟,等.高分辨率时间数字转换电路的 PLD 实现[J].半导体技术,2006,31(6):452-455.
- [7] 许春香,时 伟,黄传金,等.基于 FPGA 的高精度时间数字转换电路设计[J].微计算机信息,2009,25(2):208-210.
- [8] 郑 熠,李惠军,赵守磊,等.时间数字转换(TDC)IP 核设计与低功耗优化[J].电力系统及其自动化学报,2010,22(3):138-141.
- [9] 叶 超,冯 莉,欧阳艳晶.基于 FPGA 的精密时间间隔测量仪设计[J].信息与电子工程,2009,7(2):159-163.
- [10] 冯志辉,刘恩海.用于激光测距的高精度时间数字转换电路[J].光学精密工程,2010,18(12):2665-2671.
- [11] Virtex-6 FPGA configurable logic block user guide [EB/OL]. 2013. <http://www.xilinx.com>.
- [12] Virtex-6 FPGA data sheet DC and switching characteristics [EB/OL]. 2013. <http://www.xilinx.com>.
- [13] 康晓文,刘亚强,崔均健,等.基于 FPGA 进位链 TDC 延时模型的建立与性能测试[J].核电子学与探测技术,2011,31(3):267-273.

(上接第 174 页)

- 存储网络系统[J].计算机科学,2003,30(2):79-82.
- [6] 蔡昭权,张宗平,罗 伟,等. FCIP 跨层流量控制的建模与仿真[J].计算机与现代化,2012(12):51-54.
 - [7] Cegiela R. Selecting technology for disaster recovery [C]//Proceeding of international conference on dependability of computer system. [s. l.]:[s. n.], 2006:160-167.
 - [8] 蒋丽娟.基于 SAN/NAS 网络架构的存储整合技术[J].计算机系统应用,2011,20(11):154-157.
 - [9] 郑 轲.存储系统中 iSCSI 和 FC 链路融合技术研究与实践[D].武汉:华中科技大学,2008.

- [10] Devianov B, Toueg S. Failure detector service for dependable computing [C]//Proceedings of the first international conference on dependable system and networks. [s. l.]:[s. n.], 2000:14-15.
- [11] 厉 剑,廉国斌,黄 栋.数据容灾系统与 CDP 技术[J].计算机技术与发展,2009,19(1):168-171.
- [12] Schmidt K. High availability and disaster recovery: concepts, design and implementation [M]. [s. l.]:Springer, 2006.
- [13] Pasin M, Fontaine S, Bouchenak S. Failure detection in large scale system: a survey [C]//Proc of NOMS workshops 2008. [s. l.]:IEEE, 2008:165-168.

检验检疫融合存储数据容灾应用研究


作者:

张宗平, 秦磊华, ZHANG Zong-ping, QIN Lei-hua

作者单位:

张宗平, ZHANG Zong-ping(广东出入境检验检疫局信息中心, 广东 广州, 510623), 秦磊华, QIN Lei-hua(华中科技大学 计算机科学与技术学院, 湖北 武汉, 430074)

刊名:

计算机技术与发展 

英文刊名:

Computer Technology and Development

年, 卷(期):

2014(8)

本文链接: http://d.wanfangdata.com.cn/Periodical_wjtz201408040.aspx