

陕西电信存储虚拟化整合技术^①

杨云峰, 李 喆

(陕西电信公司, 西安 710075)

摘 要: 本文介绍了存储虚拟化的相关概念和技术, 重点阐述了虚拟化的设计思想和各种实现方式的对比以及在陕西电信项目建设中的技术选型与实现方式, 在将理论变成实际应用上所需的各种技术和物理环节储备也进行了描述; 该技术的使用, 盘活了企业的固定资产降低生产成本, 更重要的是实现了数据的动态管理, 异构存储的统一化管理。

关键字: 虚拟池; SAN 网络; 存储控制设备; LUN

Storage Virtualization Integration with Technology on Shaanxi Telecom

YANG Yun-Feng, LI Zhe

(Shannxi Telecom Branch, Xi'an 710075, China)

Abstract: The paper introduces storage virtualization related concepts and techniques, and emphatically expounds the design thought and virtualization of realization of comparison and the construction project in shanxi telecommunication technology selection and realization ways, in practical application of theoretical into all kinds of technical and physical link reserves are also described, The use of technology, revitalize the fixed assets of enterprises to reduce the production cost, more important is to realize the dynamic management of heterogeneous data storage and management of unification

Keywords: virtual storage pool; SAN network; storage control device; logic; unit number

1 引言

陕西电信 IT 基础硬件平台作为陕西电信整个 IT 系统的核心和基础, 其重要性不言而喻。为了对基础 IT 资源的最优配置, 提高资源利用率, 同时使整个硬件平台资源进行统一规划、统一部署, 实现企业平台在可靠性、可用性和安全性等方面得到全面增强, 陕西电信借建设全省 CRM 应用的契机, 提出了建设“支撑应用滚动式发展”基础平台的理念。在这一理念的指导下, 陕西电信希望建设一个将主机、网络、存储全面整合的可扩展的架构, 并进一步将存储平台建成一个虚拟的存储池, 从而构造出一个跨越主机、存储的基础平台模型, 最终实现按需部署应用系统、按需供应存储空间, 并根据数据的生命周期, 实现数据分级存储, 降低系统整体拥有成本。

2 陕西电信系统现状及存储虚拟化技术原理

2.1 系统现状

由于历史原因, 陕西电信以前的 IT 系统基础平台都是采用分散建设的方式, 各自为政, 因此形成了一个信息孤岛, 导致数据不能共享, 存储设备利用率较低, 基础平台的可扩展性和弹性较差, 无法及时响应业务的快速变化。陕西电信 IT 系统的主机采用了 HP Superdome 和 IBM AIX p 系列, 应用主要包括 BSS、OSS、MSS、EDA 的综合受理平台、综合统计决策分析平台、统一客户资料平台、EAI 平台、传输网管集中监控系统、大客户网络服务保障系统、电子运维系统和资金管控系统等, 按照业务的重要程度, 对数据进行分块化, 数据存储分级化部署; 将 A 类业务系统数据放在日立的高端 USP 存储, 并做了最高的 RAID 0+1 数据保护, B 类业务放在日立的中端存储

^① 收稿时间:2010-04-28;收到修改稿时间:2010-05-30

AMS1000,其他类业务按照当年的建设情况分别存放于 HP 存储,IBM 存储和 EMC 多个不同厂商的异构存储设备中。

连接方式:采用 SAN 网络通过不同性能等级的 SAN 交换机搭建起主机与存储组成的局域网

存在问题:随着近些年电信业务的不断发展和演变,有的系统数据量快速增长,对所在存储空间要求较多,有的系统数据量需求较少比较稳定,反应到存储级即不同类型的存储资源利用率不同,利用率低的存储空间由于硬件技术的异构性限制无法给业务数据增长快的系统分配,造成存储空间资源的使用率不均,无法保护企业的成本投资,盘活现有资源,再者,对于异构存储,管理方式不同,无法实现统一的存储管理机制,增加了管理人员的技术难度存在一定的系统风险。

2.2 存储虚拟化技术介绍

新的存储发展趋势已经打破了传统的存储观念,将控制器部分与磁盘部分分离,通过存储虚拟化实现一个虚拟存储池,消除了设备之间的物理差异,池中的所有资源被统一管理,统一使用,按需分配,提高了使用率,合理利旧设备,按照数据的属性和访问特点进行部署,提高了业务支撑、系统扩展、数据访问的能力。

虚拟存储池:所有存储被视为一个存储池,按业务性质灵活供应逻辑卷,统一的数据管理,提高效率,提供灵活的数据复制。

虚拟存储的三种实现方式:基于主机层,基于网络层的虚拟化存储技术,基于控制器的虚拟化存储技术三种实现方式比较:

存储控制器的虚拟化--提供通用管理和功能,降低了复杂性;可动态调整缓存,虚拟化端口可扩展性较好,多结点组成高可用性集群,共享全局高速缓存,传输 I/O 不需要中转,可以标准数据安全,外部存储格式不变,与服务器连接方式可以随意改变,采用存储控制器现有工具和大容量缓存,实现数据在线移动,可以改善性能较差的外部存储。

网络的虚拟化--需要额外的 SAN 交换机,PC 服务器和软件,网络 PC 服务器成为瓶颈,2 结点 PC 服务器集群,缓存不能共享,不能做到实时切换,在虚拟化引擎中必须对 I/O 进行重定向,在数据移动功能方面需要在虚拟化引擎中开发数据移动工具,并利用有限的

缓存资源实现数据移动功能,增加了 SAN 的复杂性,有限的缓存对性能有一定影响。

主机的虚拟化--卷映射由主机处理器完成,位于主机层,消耗主机资源,执行效率层看由于是软件实现,指令处理效率偏低,需要安装代理,可靠性相对较差安全级别较低,逻辑分区较少.指令处理效率偏低,需要安装代理,可靠性相对较差安全级别较低,逻辑分区较少。

2.3 存储虚拟化技术比较结论

基于主机层的虚拟化存储案例很少,弊端也非常突出:

① 对于主机资源的消耗巨大;

② 安装复杂、对于卷管理需要重新划分;

③ 存在单点故障;如果软件发生问题,那么将导致存储的不可用;

④ 难以回退;存储管理将绑定一种 Volume Manager Software 或者类似虚拟化软件,无法替换,存在较大的隐患。

基于网络层的虚拟化方案不是很成熟,分为带内带外的技术流派,特别是带外产品,由于技术所限,多停留在研究测试阶段。

基于存储控制器的虚拟化技术,在全球已经有广泛的实施案例,而且对于应用无需修改。

3 陕西电信存储虚拟化的整合与实现

鉴于以上虚拟化技术的优缺点,陕西电信经过全面调研和方案对比,最终选择了基于存储控制器的虚拟化技术,采用日立公司(HDS)的 USP 存储作为存储控制设备。其技术实现思想如下:存储控制设备与存储池子系统 Storage Pool 集成在一起,组成 SAN Appliance。可以看到在该工程中存储控制设备在主机与存储池数据交换的过程中起到核心作用,由它内嵌的存储管理系统将存储池中的物理硬盘虚拟为逻辑存储单元(LUN),并进行端口映射(指定某一个 LUN 能被哪些端口所见),主机端将各可见的存储单元映射为操作系统可识别的盘符。当主机向 SAN Appliance 写入数据时,用户只需要将数据写入位置指定为自己映射的盘符(LUN),数据经过存储控制设备的高速并行端口,先写入高速缓存,存储控制设备中的存储管理系统自动完成目标位置由 LUN 到物理硬盘的转换,在此过程中用户见到的只是虚拟逻辑单元,而不关心每个

LUN 的具体物理组织结构。

3.1 存储虚拟化实现目标

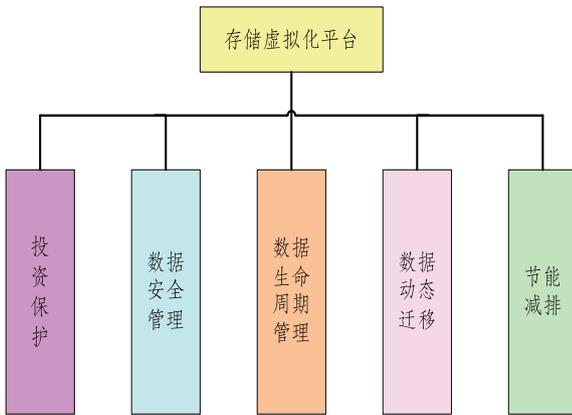


图 1 虚拟化目标

设备投资保护：能够将同构或异构存储通过虚拟化引擎都虚拟化到存储池中，统一分配资源，提高资源使用率，利旧原先的存储设备。

数据安全：可以在虚拟池中实现本地数据镜像，生产数据迁移和远程数据复制，保证重要业务数据的安全。

数据生命周期管理：根据数据管理规则对于不同响应时间，保存期限的数据分级存储在高端，中端和低端存储中，即能达到响应时间要求，又能够节约高端存储资源，充分利用资源。

数据动态管理：跨级存储数据迁入能够改进效能；在异构存储数据迁移实现系统升级；跨级存储数据迁出实现历史归档。

3.2 存储虚拟化整合

主机通过 SAN 网络访问存储设备，使用 HDS 存储虚拟化引擎挂接异构存储设备即将原有的 EMC Symmetrix 8530、EMC Symmetrix 8830 和 HP EVA 等外部存储系统与 HDS 存储虚拟化引擎对接，将这些存储设备虚拟化，形成存储虚拟池，对存储设备进行统一管理。

3.2.1 技术实现方式

外部存储设备通过 external port 映射到 HDS 虚拟化引擎，作为 ELUN，映射关系可以是 1 对 1 的，也可以是多对一的^[1]。主机从虚拟化引擎看到的设备都是同样的 LUN，并不区分各种同构或异构设备，这样就实现了存储设备的虚拟化。当服务器主机从 USP 的

FC 接口获得这些外部存储的存储空间时，感觉就是获得 USP 存储空间一样，并可以使用到 USP 的 Cache、CSW 等性能。这样不仅可以对原有的存储系统进行投资保护，而且可通过分级存储，将一些边缘数据和测试数据放在外部存储上，实现数据的统一管理。

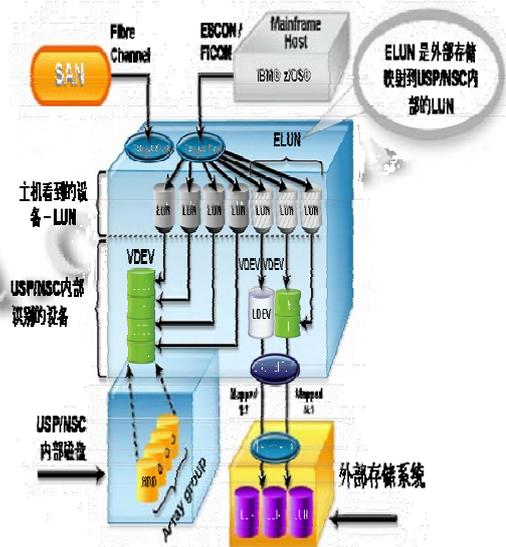


图 2 虚拟化整合原理图

USP 的光纤端口可以与外部存储的光纤端口直连，或通过 SAN 交换机连接，利用原有的光纤交换机。

3.2.2 物理结构

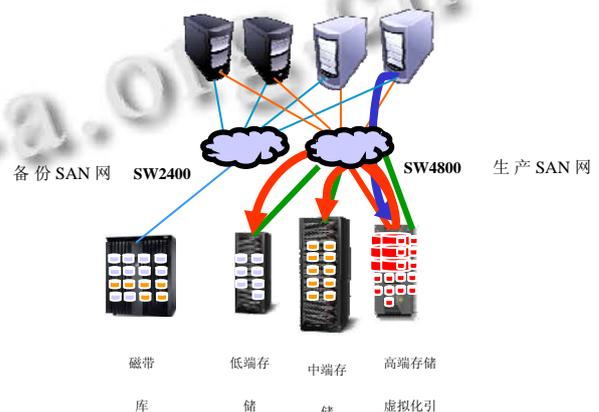


图 3 物理结构图

数据库主机通过 SW24000 接入备份 SAN 网络，通过 SW48000 接入生产 SAN 网络，两个 SAN 网络物理隔离。存储虚拟池中所有设备连接到生产 SAN 网络中，虚拟化引擎也通过生产 SAN 网络连接其他存储设备，主机访问 HDS 高端存储如蓝色线标注路径，访问

中端及低端存储如红色线标注路径。

生产 SAN 网络中划分主机 ZONE, 存储 ZONE 和虚拟化 ZONE。主机连接端口划入主机 ZONE, 高端存储连接端口划入存储 ZONE, 虚拟化引擎和中低端存储连接端口划入虚拟化 ZONE。各 ZONE 之间相互隔离。

这样最终以存储虚拟化技术为基础实现了对原有的存储孤岛和分散的存储局域网进行整合。同时, 根据数据的生命周期, 建立分级存储架构, 将存储平台建成一个虚拟的存储池, 实现数据的分级管理和应用。

4 结论

在新的存储架构中, HDS USP 承担起了一级存储的职责, 主要存放关键业务数据, 而原来的 EMC Symmetrix 磁盘阵列则作为二级存储, 主要负责非关键性数据的备份、归档。这也就形成了分级存储架构, 可以根据数据在生命周期中价值的不同, 将数据存储

在最适当的存储设备上, HDS USP 可以把挂接在其后的不同品牌的磁盘阵列整合到一个统一的存储池中, 从而提高了存储设备的利用率, 方便数据共享和管理, 以及设备利旧。通过实践, 深切地感受到了采用存储虚拟化的好处: 首先, 可以将原有的存储孤岛进行整合, 通过存储虚拟化技术构建一个统一的虚拟存储池, 对存储资源进行统一管理, 并按需分配, 从而大大提高了系统的可扩展性; 通过存储虚拟化的整合功能, 实现了原有存储设备的再利用, 有效地保护了原有设备投资; 可以构建分级存储和管理体系, 将各应用系统的数据进行分级, 分为联机访问数据、近线访问数据、归档数据, 然后将其存储在不同成本的存储模型上, 按照数据的生命周期特性进行管理, 从而降低了整体成本。

参考文献

- 1 张治国. 陕西电信存储整合成功经验. 北京: Hitachi Data System, 2009.

(上接第 69 页)

5 小结

赤平足识别技术中特征提取及测量一般都在人工操作下进行, 虽然专业人员具有较强的先验知识, 但特征点定义和参数测量都会产生不同程度误差, 测量参数精度也不能准确把握。如果把先验知识与自动计算相结合, 精度会明显提高, 从而减少误判率。对上述所提取的形态特征参数与常规人工识别数据进行对比, 准确率已达到了公安刑侦技术鉴定的要求。而且对同一赤平足图像进行数十次识别操作所得到的数据误差率很小, 精度已达到 95% 以上。通过交互系统识别所得到的数据可以为自动识别程序设计提供对比的技术数据, 对自动识别算法研究建立可靠的技术保障。同时用所取得的数据来建立赤平足生物形态特征数据

库也是奠定侦破工作的基础数据。

参考文献

- 1 云宝璠, 童莉, 宗家进. 赤足迹生物特征识别研究初探. 公安部刑事侦查局等. 全国足迹检验鉴定学术研讨会论文集. 北京: 中国人民公安大学出版社, 2007. 31-35.
- 2 黄群. 赤脚足迹的统计分析. 辽宁警专学报, 2005, (1): 5-9.
- 3 葛哲学. 精通 MATLAB. 北京: 电子工业出版社, 2008: 27-175.
- 4 王爱玲, 叶明生, 邓秋香. MATLAB R2007 图像处理技术与应用. 北京: 电子工业出版社, 2008. 30-165.
- 5 葛哲学, 等. 小波分析理论与 MATLAB R2007 实现. 北京: 电子工业出版社, 2007. 51-446.
- 6 杨文强. 统计学—科学与工程应用. 北京: 清华大学出版社, 2007. 306-328.