数据容灾系统与CDP技术

厉 剑,廉国斌,黄 栋
(上海交通大学 信息安全工程学院,上海 200240)

摘 要:数据容灾系统向来被认为是信息安全领域中的一项重要内容。为帮助大家更全面、更深入地了解数据容灾系统的相关知识与技术,文中对传统的数据容灾系统及其所采用的相关技术进行了系统的整理与介绍,对数据备份与数据复制两项技术做了重点论述,对目前最热门的技术CDP作了较深入的探讨。结果表明,当前及今后一段时间内,传统的数据容灾技术在某些应用领域中依然比较适合,而CDP作为一种新型的技术理念,在容灾领域中扮演着越来越重要的角色。随着其技术的日臻成熟完善,最终将会完全取代传统的容灾技术。

关键词:容灾;备份;数据恢复;网络存储;持续数据保护
中图分类号:TP309.2 **文献标识码**:A **文章编号**:1673-629X(2009)01-0168-04

Data Disaster-Tolerant System and CDP Technology

LI Jian, LIAN Guo-bin, HUANG Dong
(School of Information Security, Shanghai Jiaotong University, Shanghai 200240, China)

Abstract: Building data disaster-tolerant system has been regarded as the most important thing in information security field. The purpose of this article is that can realize all techniques and knowledge related to DDTS. Traditional DDTS and primary techniques are introduced, and emphasis is placed on data backup and data copy, then the most advanced technology-CDP is discussed. A conclusion is reached that traditional disaster-tolerant technologies are still suitable for most applications in the next period of time, and as a new kind of technology, CDP is playing an increasingly important role. With its continuous improvement, CDP will eventually replace the traditional disaster-tolerant technology.

Key words: disaster-tolerant; backup; data recovery; net storage; CDP

0 引 言

在当今信息化时代,数据和信息已经成为各行各业的业务基础和命脉。但当存储数据和信息的计算机系统在遭受诸如火灾、水灾、地震、战争等不可抗拒的自然灾难以及电脑病毒、黑客入侵、系统掉电、网络通信失败和各种软、硬件错误的时候,人们不得不重新审视以信息系统为核心业务载体所面临的高度风险,因此数据容灾系统越来越受到人们的重视。

1 数据容灾系统概述

数据容灾系统[1](Data Disaster-tolerant System, DDTS),即灾难发生时,确保计算机信息系统能够恢复,并降低数据和信息损失的系统。广义上,所有与业务连续性相关的内容都纳入容灾。对于IT而言,则是提供一个能防止用户业务系统遭受各种灾难影响和破坏的计算机系统。狭义上是指,除了生产站点以外,用户另外建立的冗余站点,当灾难发生,生产站点受到破坏时,冗余站点可以接管用户正常的业务,达到业务连续性的目的。数据容灾系统是数据存储备份的最高层次。

1.1 衡量数据容灾系统的性能指标

衡量DDTS的应急能力和系统保护能力有两个指标RPO[1](Recovery Point Object,恢复点目标)和RTO(Recovery Time Object,恢复时间目标)。RPO侧重于数据丢失,RTO侧重于服务丢失,二者没有必然联系。

RPO以时间为单位,即系统和数据必须恢复到的时间点要求,它表示系统能够容忍的最大数据丢失量。系统容忍丢失的数据量越小,RPO的值越小。RTO以时间为单位,即信息系统或业务功能从停止到完全恢复的时间要求。RTO表示系统能够容忍的服务停止的最长时间。系统服务的紧迫性要求越高,RTO的值越小。

收稿日期:2008-04-26
作者简介:厉 剑(1980-),男,山东日照人,硕士,主要研究方向为网络安全、加密技术;廉国斌,工程师,主要研究方向为RFID防碰撞算法研究;黄 栋,研究员,主要研究方向为密码算法。

1.2 容灾系统的分类

容灾系统可分为在线式与离线式[2]两种。在线式容灾要求生产中心和灾备中心同时工作,数据从生产中心实时复制传送到灾备中心。当生产中心出现故障时,灾备中心自动接管并继续提供服务,其关键是数据复制技术。离线式容灾是将数据通过备份系统备份到磁带,将磁带在异地保存,由备份软件来实现备份和磁带的管理,其主要依靠备份技术实现。缺点是:从磁带上恢复数据慢;备份窗口内的数据都会丢失;实时性差。

2 传统容灾技术

传统数据容灾技术包括数据备份、数据复制和数据管理技术等,而应用容灾技术包括灾难检测、进程迁移和系统恢复技术等。

2.1 数据备份技术

数据备份就是把数据从生产系统备份到备份系统介质中的过程。数据备份技术最初是备份到本地磁带,随着网络发展,备份技术有了新的发展。

2.1.1 主机备份

每台主机都配备专用的存储磁盘或磁带系统,主机中的数据备份到位于本地的专用磁带驱动器或资源库中[3]。缺点:即使一台磁带驱动器处于空闲状态,另一台主机也不能使用它进行备份,磁带资源利用率低。不同的操作系统平台使用的备份恢复程序一般不相同,使得备份工作和对资源的总体管理变得更加复杂。

2.1.2 网络备份

磁带资源由一个主备份/恢复服务器控制,而备份和恢复进程则由一些复杂程序来控制[3]。主备份服务器接收其它服务器通过局域网或广域网发来的数据,并将其存入公用磁盘或磁带系统中。这种集中存储方式提高了磁带资源的利用率,同时也使企业在高性能磁带驱动器和资源库的投资取得更好的回报。缺点是网络带宽将成为备份和恢复进程中的瓶颈,制约快速备份和恢复的能力。另外,通过局域网或广域网主链路传输备份和恢复数据将占据大量带宽,增加网络负荷,影响其性能。

2.1.3 专有存储网络备份

专有存储网络备份可以分为 LAN－Free 备份和 Server－Free 备份。

LAN－Free[4],如图 1,是把局域网排除在备份和恢复进程之外的备份方案。用户只需将磁带驱动器和资源库连接到 SAN 网络中,各服务器就可绕过局域网把数据直接备份到共享的磁带库中。局域网只承担各服务器之间的通信(而不是数据传输)任务。先进的备份和恢复程序仍然被用来完成进程控制和数据跟踪的工作。这种分工使得存储设备、服务器和局域网资源得到更有效的利用。

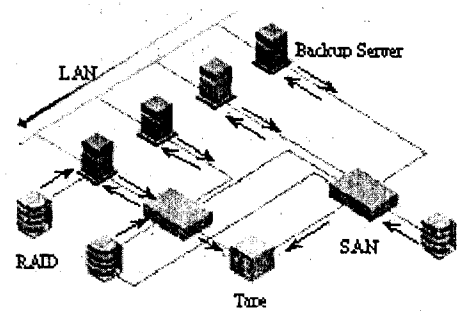

图 1　LAN－Free 原理

Server－Free[4],如图 2,数据无需通过主服务器直接在存储设备之间进行传输(如从磁盘到磁带机),这种依赖于一项仍在发展的新技术第三方拷贝(3rd－Party Copy),这种技术被用于 SAN 网络设备、主机系统和存储设备(有待实现)当中。独立于服务器的备份和恢复方案可降低主机的 CPU 占用率,提高操作系统效率。这项技术目前仍在发展完善阶段,进一步同操作系统、数据库及应用程序相结合,以支持应用级的备份和恢复。

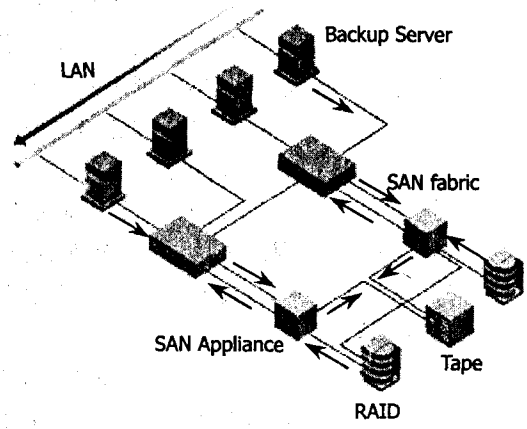

图 2　Server－Free 原理

2.2 数据复制技术

如图 3 所示。数据复制技术是通过不断将生产系统的数据复制到另外一个不同的备份系统中,其实现可以分为三种:服务器层、交换机层和存储层[5]。

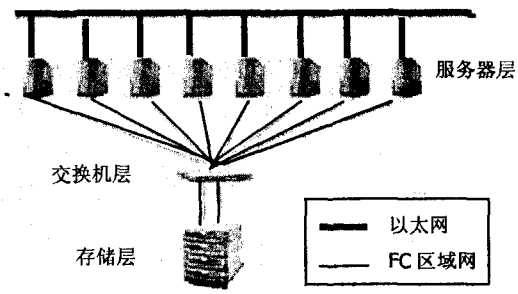

图 3　数据复制层次

2.2.1 服务器层

在生产中心和灾备中心的服务器上安装专用的数据复制软件,以实现远程复制,二端必须有网络连接作为数据通道。这种方式投入较少,可兼容不同品牌的服务器和存储设备,较适合硬件组成复杂的用户。

2.2.2 交换机层

在生产中心和灾备中心都要配备这种交换机,交换机之间用专用链路连接。由于交换机可以管理和复制的数据是存放在存储层内的,因此,用户需要将生产中心的数据都存储在与交换机所连接的存储设备中。

2.2.3 存储层

在生产中心和灾备中心各配备一套存储系统。若距离在几十公里之内,可在两中心的存储交换机通过光纤连接;若距离在 200 公里内,可增加 DWDM 等设备进行光纤连接;超过 200 公里,则可增加存储路由器进行协议转换,通过 WAN 或 Internet 连接。存储层的数据复制技术已经很成熟,且对应用服务器的性能基本无影响,是容灾方案的主流选择。

2.3 灾难检测技术

灾难检测的目的是尽早地发现生产系统端的灾难,目前采用 Heart Beat[6]技术。

该技术有两种实现方法,一是生产系统端在空闲时每隔一段时间向外广播自身的状态,检测系统若在给定的时间段内没有收到状态信号,则表明发生了灾难。二是检测系统每隔一段时间对生产系统进行一次检测,若在给定的时间内生产系统无响应,则表明发生了灾难。心跳技术的关键是心跳检测的间隔周期,太短系统开销会很大;过长则无法及时发现灾难,应根据不同的容灾等级和容灾要求,设定适当的检测周期。

2.4 进程迁移技术

进程迁移是为保持生产系统业务的连续性,通过提取一台处理机的进程状态,在另一台处理机上根据进程状态再生该进程,从其"断点"继续运行下去的技术[7]。目前主要采用基于本地机群的进程迁移,可分为如下三类:

* 应用级迁移:实现简单,可移植性好,但需了解应用程序语义并可能需对应用程序进行修改或重编译,透明性较差。

* 用户级迁移:实现较简单,软件开发和维护也较为容易,但由于在用户级无法获得 Kernel 的所有状态,对于某些内核态进程,无法进行迁移。另外由于打破 Kernel 空间和 User 空间之间边界从而获得 Kernel 服务,需要巨大的开销,因此效率远远低于内核级实现。

* 内核级迁移:可充分利用 OS 的功能,全面获取进程和 OS 的状态,效率较高。但由于需要对 OS 进行修改,实现较为复杂。

3 数据容灾系统新技术——CDP

3.1 CDP 技术概述

全球网络存储工业协会(SNIA)对 CDP(Continuous Data Protection,持续数据保护)的定义是:一种能独立对主要数据进行持续捕捉或跟踪数据修改,并保存变化,从而实现从过去的任何非预设点恢复的方法[8]。它的最大优势在于对数据的保护是连续性的,而且可以快速恢复,从根本上解决传统备份中低恢复能力和非精细时间策略的先天弱点。

CDP 技术包括 Near CDP 和 True CDP。前者只能恢复部分指定时间点的数据,类似存储系统的逻辑快照,后者可以恢复指定时间段内的任何一个时间点。

3.2 CDP 技术原理

首先利用连续或间隔型的复制策略,实现用户系统内的包括系统数据在内的数据连续复制,以确保灾害发生时,数据恢复到最新的时间点,同时,采用复制的时间点快照技术,连续产生多个时间点固定影像,这些影像不仅时间精细化,而且是直接可用的 Image,直接 Mount 即可使用。对于大量的渐进性故障,如病毒的侵袭、人工的误操作、软件的 BUG 等发生时,只要找到影像,进行简单的空间重定位,就会重现被破坏的数据。

从操作方式来看,CDP 解决方案的可分为基于块的、基于文档的和基于应用[8]。基于块的解决方案位于物理储存或逻辑卷管理层之上。当数据块被写入主存储器时,写入的数据副本就被 CDP 系统捕获并存储到一个单独地点中。基于文档的解决方案位于文档系统之上,它能够捕获文档系统数据和元数据事件。基于应用的 CDP 解决方案设计则直接位于受保护的特定应用之中。基于块和文档的 CDP 解决方案能够利用一种相同的通用方法来支持多种不同的应用。基于应用的 CDP 则只为某种应用提供 CDP 能力。

3.3 CDP 解决方案特点

3.3.1 主机代理程式

为了在数据发生变化时进行访问,一些 CDP 解决方案需要在受保护的主机上安装一种特别的"代理程式"软件。另一些 CDP 解决方案则使用已内置到受保护主机或网络中的数据传输协议(如 NFS、CIFS、FC 或 iSCSI)来实现[8]。

3.3.2 精细程度

不同的 CDP 解决方案提供不同精细程度的恢复能力,精细程度可分为:卷组、单个卷或文档系统、单个

文档夹或文档组、单个文档或应用对象(如电子邮件或日历项目)。

3.3.3 恢复时间应用集成

在进行恢复时,CDP解决方案能够识别出该应用的先前历史中最优化或最重要的恢复点。这类应用集成是完全自动的,也是可扩展的。

3.3.4 数据库连续保护

CDP解决方案一般都支持常见数据库环境(如Oracle或SQL)的连续保护。支持的意思是该解决方案经过了厂商的全面测试和认证,而且还会向用户提供相关文档内容。

3.3.5 库架构

大多数CDP解决方案是将任何数据的变化存储在单独的地点,形成存储库架构,而且这种存储库是局域网、广域网或存储区域网上明确的专用节点。

3.3.6 库复制

一些CDP解决方案还提供将CDP库复制到另外一个远程库的能力。这样就能够提供更高的灵活性,防止主CDP库可能出现损坏或丢失对恢复能力产生影响。

3.4 CDP发展方向

随着CDP应用范围的扩大和人们认知的深入,CDP技术将会作为在线数据的重要保护手段而独立开辟一条新的通道。其发展方向概括为三点:

第一,CDP的连续和系统保护范畴将继续延伸,从目前基于微软的各类操作系统平台延伸到更多企业级所采用的UNIX平台。

第二,在精细点恢复的技术上,拉杆式日志恢复技术将使精细点的恢复超越最近的快照点。该技术将为一些高端的、以秒级错误恢复为目标的用户带来真正的数据保护方案。

第三,继续完善历史数据的存档机制、在线数据和离线数据的分级保护体系。可利用在线多时间点数据在后台自动提取的serverless备份技术,将近期各时间点数据在不影响应用的情况下存档到VTL虚拟带库中或者存档到物理磁带库中,从而实现数据的离线保存能力和长时间历史数据的保管。

4 结束语

今后,人们的信息系统将面临着越来越多的人为的或自然的不确定因素的威胁,因此,融合现有技术,不断发展新技术,构建更加安全、可靠的数据容灾系统,已经成为当今世界共同关注的课题。

参考文献:

[1] 崔可升,王玉春.建设容灾系统的几点思考[J].计算机系统应用,2003(7):26-29.

[2] 王树鹏,云晓春,余翔湛,等.容灾的理论与关键技术分析[J].计算机工程与应用,2004(28):54-58.

[3] 李兆玉,韦世红,李　鹏.容灾系统的建设方案研究[J].重庆邮电学院学报:自然科学版,2005(4):35-37.

[4] 杜　宁,姚玉坤,黄　伟.浅析基于SAN的容灾实现[J].山东通信技术,2006(2):9-11.

[5] 胡　勇,罗维荣.容灾备份技术架构浅析[J].电子政务,2006(9):15-16.

[6] 妙全兴,武海鹰.一种高性价比的网络容灾与高可用集群的设计[J].微机发展(现更名:计算机技术与发展),2003,13(9):41-42.

[7] Benjamin B, Shao M. Optimal Redundancy Allocation for Information Technology Disaster Recovery in the Network Economy[J]. IEEE Transactions on Dependable and Secure Computing, 2005, 3:262-267.

[8] 颜　军.CDP带来存储新气象[N].计算机世界报,2006-04-03(12).

(上接第164页)

[D].青岛:中国海洋大学,2005.

[4] 云辽飞.基于角色的访问控制技术在统一设备管理中心系统中的应用[D].西安:西安电子科技大学,2007.

[5] 范玉顺.工作流管理技术基础[M].北京:清华大学出版社,2001.

[6] 杨冬梅,张亦军.基于角色的工作流模型研究与应用[J].电脑开发与应用,2006(11):22-24.

(上接第167页)

on Trusted Computing[C]//Proceedings of the Fifth International Conference on Machine Learning and Cybernetics. Dalian:[s.n.],2006:2776-2781.

[4] 魏　乐,叶剑新,黄　健.可信计算的初步研究[J].科技资讯,2007,13:166-167.

[5] 肖　曦,韩　军,汪伦伟.可信计算平台关键机制研究[J].信息工程大学学报,2007,8(2):217-220.

[6] 靳蓓蓓,张仕斌.可信计算平台及其研究现状[J].长春大学学报,2007,17(2):45-49.

[7] Smith S W. Trusted Computing Platforms: Design and Applications[M].冯登国,徐　震,张立武,译.北京:清华大学出版社,2006:148-151.

[8] 邢启江,肖　政,侯紫峰,等.一种基于TPM芯片的计算机安全体系结构[J].计算机工程,2007,33(15):152-154.

[9] Price A. More Secure Computing[EB/OL]. 2006. http:\\www.trustedcomputinggroup.org\TCGBackgrounder_revised_amp_oct_17_06.pdf.