

复杂数据环境下备份技术的比较研究

Comparatively Study on Data Backup Technology in a Complicated Data Center

王印红 王 昆 (中国矿业大学(北京)管理学院 北京 100083)

摘 要:在一个复杂数据存储环境中,为确保服务器的运行安全,需针对于不同服务器的数据量和备份要求,选择恰当的备份方式,本文就给定的一个复杂数据存储环境中,探讨服务器的三种备份方式,即 LAN-FREE、SERVER-FREE 和 LAN 的适用性,并比较各个备份方式的优缺点。

关键词:数据存储 备份技术 LAN-FREE SERVER-FREE

1 引言

数据备份技术是企业数据中心安全性的关键所在,各种意外事件,如硬故障、软件 BUG、人为操作或环境因素都有可能对数据的损坏与丢失。数据备份就是当上述意外事件发生时,将丢失的数据恢复,维持企业业务的连续性。但在一个复杂的数据中心环境中,服务器的连接方式、数据量、性能都和数据备份方式存在不同的联系,针对复杂环境中的各种服务器,采用恰当的备份技术是规划和建设一个高效的、适用的数据备份系统前提和有效保证。

2 一个典型的复杂数据环境

企业数据中心环境变得越来越复杂,不同的系统平台、混合的存储网络、海量的数据存储等都给数据备份提出了一个更高的要求。本文给定一个复杂数据存储环境(如图 2-1): 1. 异构的平台,即环境中存在多个操作系统。2. 多个厂商的数据库及应用软件。3. 混合的数据存储架构即服务器与盘阵的连接即有传统的 DAS 架构,又有 NAS 和 SAN 架构。4. 不同的厂商的存储设备。在这种复杂的环境中,虽然各个厂商的数据库有自己的备份系统,但不支持磁带的备份,即必要通过第三方备份软件才能备份到磁带库中,各个操作系统也有自己的备份系统,但无法实现跨平台的集中式管理,无法共享备份资源。因此本文选择 Symantec 的 Netbackup 作为备份平台的建构软件和管理软件,选择 NBU 主要是基于如下的原因:一是它的易管理性,

具有图形化管理,设备及作业的监控,日志和故障的诊断等。二是开放性,具有多种操作系统、数据库、磁带库、磁盘阵列的支持等。三是性能优势,在技术上运用 Flashbackup、Staging、Synthetics 等提高备份速度。当然还有它备份市场上占有最高的市场占有率。选择第三方备份软件的目的是在一个复杂的数据中心环境,构建一个统一的数据备份管理界面,实现自动的备份管理,统一共享备份资源,实现数据快速恢复。在硬件的规划上可以考虑加装一台备份服务器 B_S,一台支持光纤的磁带库,磁带库内含 6-8 台 LTO2 的磁带机以满足同时 6 个以上作业的备份需要,连接方式同见图 1。

表 1 数据中心服务器备份要求、数据量及应用说明

服务器	系统平台	数据量级	对备份窗口的要求	应用说明
HA_S1	AIX	T 级	晚上,10 小时内	可高用服务器和 ORACLE
HA_S2	AIX	T 级	晚上,10 小时内	可高用服务器和 ORACLE
S1	SOLARIS	T 级	晚上,10 小时内	数据库服务器 ORACLE
S2	SOLARIS	几百 G	晚上	应用服务器
NAS_S	LINUX	T 级	无	文件服务器
C1	WINDOWS	几十 G	无	其他小型服务器如 MAIL 服务器 EXCHANG
C2	WINDOWS	几十 G	无	同上
B_S	AIX			备份服务器

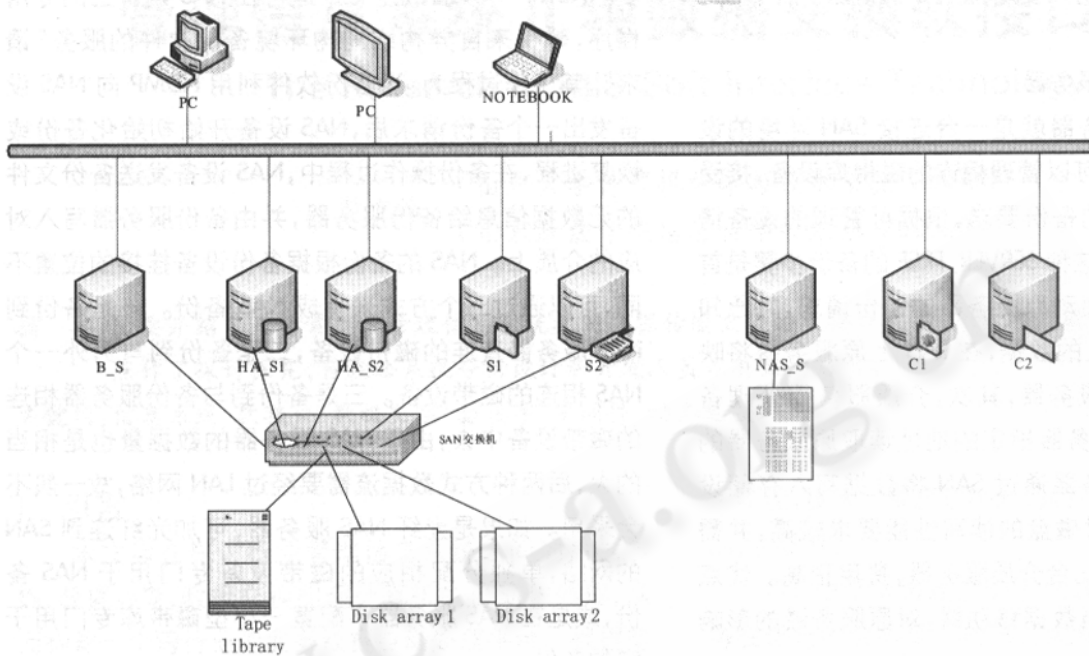


图 1 信息中心复杂数据存储网络图

3 服务器数据备份方式的比较与选择

数据的备份方式目前比较常用的有三种:LAN 备份、LAN_FREE 备份、SERVER_FREE 备份,三种备份方式适用于不同备份需求的服务器。

3.1 数据量较小的 C1 和 C2 的备份

在数据环境中,有一些数据量不大的服务器,像企业里的 MAIL 服务器,WEB 服务器,数据在几十 GB 左右,数据重要性不如数据库、应用服务器,但数据同样需要备份,如上述环境中的 C1、C2,针对此类的数据备份需要,可以采用传统的 LAN 备份方式。LAN 备份是一种传统的备份方式,它是明显特点就是备份数据流经过 LAN 网络,它的优点是实施方便,软、硬件费用较低。缺点是备份速度慢,如果是百兆网络,备份速度只有 6-7M/S,只能备份小数据量服务器,由于数据传输依靠 TIP/IP 封装,数据在网络上传输,占用较大带宽,影响终端客户对其他服务器的访问性能。

3.2 S2 服务器的备份

S2 服务器的数据量达到几百个 G,如果用网络的备份方式备份的话,网络会是备份速度的瓶颈,严重影响

备份的性能,使服务器的备份很长,并会占用磁带机的读写时间,与其他的备份形成资源争用。因此考虑到以上的因素,对于 S2 服务器的备份采用 LAN_FREE 的备份方式,LAN_FREE 备份的方式数据流从 S2 到交换机,交换机到备份服务器,最后再到磁带库。LAN_FREE

备份的好处是备份数据流不会占用网络带宽,提升备份速度,减少备份窗口,提高磁带机的利用率。缺点是要基于 SAN 的网络,硬件建构成本较高,实现的软件费用也稍高。

3.3 HA_S1、HA_S2 和 S1 服务器的备份

对于 HA_S1、HA_S2 和 S1 服务器,由于数据量十分庞大,且对性能的要求较高,如果采用 S2 服务器的备份方式,也可勉强完成,但考虑到减少备份作业对服务器性能的影响,进一步提高备份速度,应采用 SERVER_FREE 的备份方式,这种备份方式实现的途径有三种:备用服务器方式、专用介质服务器、第三方复制设备。

3.3.1 备用服务器方式

此方法是将要备份的服务器上的数据转移到另一台备用的服务器,以达到节省原服务器上的各种资源目的,所要备份的数据的快照是在备用的服务器上创建的,备份对原服务器没有影响。对于备用服务器方式,原服务器和备用服务器必须是共享存储设备,对磁盘的要求会比较高,要求支持分割镜像和数据复制两种类型的数据共享机制。此种备份方式的数据流是从镜像的磁盘到备用服务器,然后再到磁带库,对原服务

器的几乎没有影响。缺点是对盘阵的要求比较高,支持数据共享和镜像,另外要增加一个备份服务器,费用较高。

3.3.2 NBU 的介质服务器

NBU 的介质服务器就是一台连接 SAN 环境的设备管理器,此服务器可以管理相连的磁带库设备,接受主备份服务器发出的备份要求,根据可管理的设备情况,完成备份作业。这种 SERVER_FREE 的备份步骤是首先由主备份服务器启动原服务器的备份请求,并通知原服务器映射磁盘上的数据,在 LAN 上原服务器将映射信息发送给介质服务器,其次,介质服务器处理备份,通过 SAN 从原服务器指定的地址读取原服务器的数据,最后,介质服务器通过 SAN 将数据写入存储设备中。这备份方式对磁盘的读写性能要求较高,并需支持快照技术,增加一台介质服务器,费用稍高。优点是由介质服务器充当数据移动器,对原服务器的影响较小,且备份速度快。

3.3.3 第三方复制设备

第三方复制设备就是 SAN 上的备份代理,它代表备份应用程序进行操作。第三方复制设备使用 SCSI 扩展复制命令,从挂接到光纤通道的磁盘中接收备份数据,并将这些数据发送到存储设备,第三方复制设备有时称为复制管理器、第三方复制引擎或数据移动器。在 SAN 硬件配置中,第三方复制设备可以作为网桥、路由器或存储设备的一部分实现。第三方复制设备发起备份的步骤是:首先主备份服务器在 LAN 上启动原服务器的备份作业,并通知原服务器映射快照数据,其次原服务器将映射信息发送给备份服务器,备份服务器通过 SAN 将第三方复制命令发送到第三方复制设备,最后第三方复制设备从 SAN 挂接的磁盘中读取原服务器的信息,并将数据写入 SAN 相连的磁带设备中去。第三方复制设备实现 SERVER_FREE 的好处就是利用第三方设备处理 I/O 的能力,解放原服务器和介质服务器,减少对原服务器性能的影响。缺点是对 WINDOWS 系统的支持不好,并要求在 SAN 的网络中有第三方复制功能的设备。

3.4 NAS_S 服务器的备份

NAS 服务器是使用 HTTP、FTP、CIFS 或是 NFS 协议为客户机提供文件的 NAS 主机系统,NAS 设备的备份一般是通过 NDMP(网络数据管理协议)来进行,NBU

的备份程序通过该协议控制 NDMP 主机的备份和恢复。NDMP 可以看成是一个运行在 NAS 设备上的专用程序,等待来自异构或同构环境备份软件的服务“请求”,其工作过程为:当备份软件利用 NDMP 向 NAS 设备发出一个备份请求后,NAS 设备开始初始化备份或恢复进程,在备份操作过程中,NAS 设备发送备份文件的元数据信息给备份服务器,并由备份服务器写入对应的介质上。NAS 的备份根据备份设备挂接的位置不同,可以通过三个方式来完成 NAS 备份。一是备份到 NAS 服务器直连的磁带设备,二是备份到与另外一个 NAS 相连的磁带设备。三是备份到与备份服务器相连的磁带设备中去,由于 NAS 服务器的数据量也是相当的大,后两种方式数据流需要经过 LAN 网络,故一般不会采用。如果是光纤 NAS 服务器,可加光纤连到 SAN 的网络,单独分配相应的磁带设备专门用于 NAS 备份,或是在 NAS 服务器下配置一小型磁带库专门用于它的备份。

4 结论

对一个复杂数据中心来说,备份系统的建设对保障业务系统的安全运行,保障由于各种异外原因导致的数据损坏和丢失来说具有重要的意义,用户在确定备份系统的方案时,应该就数据保护的要求、数据运行环境作需求分析,在分析的基础上确定数据备份方案,本文在给定一个复杂的数据环境下,有针对性的研究数据备份技术和优劣势分析,对于信息中心的数据备份及容灾系统建设来说有较强的借鉴与指导作用。

参考文献

- 1 Ann L Chervenak, VIVEKANNAD VELLANKI, ZACHARY KURMAS. Protecting File System: A Survey of Backup Technique[R]. 2002.
- 2 韩德志、谢长生、李怀阳,存储备份技术探析[J],计算机应用,2004,6.
- 3 劳眷,灾难备份和恢复技术及解决方案分析[J],计算机应用与软件,2005,3.
- 4 曾庆辉,基于 Agent 的存储区域网络(SAN)智能管理技术[J],计算机系统应用,2005,6.
- 5 童强国、胡越明,异构平台下集中式网络备份框架研究[J],计算机应用与软件,2004,11.