

灾备系统

——图书馆数字资源与系统的终极安全对策

□ 郑晓军 / 深圳图书馆 深圳 518036

摘要：任何信息数据和信息系统的灾难发生，都会给图书馆带来不同程度的风险。灾难备份关系到图书馆在突发事件中能否快速扭转困局，关系到图书馆的服务能否持续地进行。因此，灾难备份对于图书馆的信息化和现代化建设，既是技术应用的进步，也是知识积累与传承的保证。

关键词：灾难备份，数据安全，数据保护，备份技术

DOI: 10.3772/j.issn.1673-2286.2011.02.013

随着图书馆全面进入数字图书馆时代，信息数据的安全毫无疑问地成为图书馆业务应用系统持续正常运行的重要前提和保障。虽然图书馆可以从端点准入管理、网络安全、应用系统安全、用户认证安全、存储系统冗余技术保护等多方面、多层次地对信息数据进行不同等级的安全保护，但是发生信息数据的丢失或损坏常常是身不由己。包括水灾、火灾和地震等自然灾害、数字图书馆应用系统软硬件故障，以及人为的误操作等现象并非每天出现，然而一旦发生了，将给图书馆的信息数据带来巨大的损失，甚至造成严重的影响和冲击，导致图书馆对外服务和对内业务管理的瘫痪。战胜灾难的最好方法，就是防患于未然。因此，有必要建立一套灾难备份系统，作为支撑数字图书馆系统的重要组成部分，纳入图书馆网络中心的基础建设中，为信息数据安全提供终极防护服务。

1 灾难的影响和应对方法

灾难的定义很广，概括地说，凡是导致信息数据遭到破坏的意外事件都属于灾难的范畴，包括自然灾害、网络通信中断、业务服务系统死机、服务器故障、工作人员错

误操作、病毒和黑客的恶意攻击，甚至恐怖袭击等等。关于各种灾难发生的比率，有IT业界权威机构作了统计，见表1。

从表1可以看出，仅需本地网

表1 各种灾难发生比率及应对方式

序号	灾难的类型	发生的比率	应对的方式
1	硬件系统故障	44%	本地网络中心保护
2	软件系统故障	32%	本地网络中心保护
3	人为错误操作	14%	本地网络中心保护
4	自然灾害和社会灾难	7%	远程灾备系统保护
5	病毒与黑客入侵	3%	本地网络中心保护

络中心对信息数据进行保护的占了93%，而约占7%的自然灾害和社会灾难才需要依靠远程灾难备份系统。

由于我国的图书馆基本上是以单体建筑居多，因此，图书馆在灾难备份系统建设中，首先考虑的是在本馆网络中心建立信息数据保护措施，以杜绝占93%的灾难发生，

绝大部分图书馆相信已考虑或正考虑采取必要的措施，如冗余、备份、快照等技术和设备，尽可能地完善本馆的信息数据保护系统。

相比远程灾难备份系统的建设来说，图书馆本地网络中心建设信息数据保护系统的成本是较低的，灾难恢复的时间也较短。但是，如果图书馆本地出现自然灾害或社会

灾难,有可能完全摧毁网络中心,而造成图书馆本地所有的保护措施失效。从目前我国图书馆发展情况来看,我国各地正在推进和形成图书馆的总分馆制建设,或建立地区内馆际合作等服务方式,以达到区域性的统一服务。因此,以独立的同城异地为基地,建设地区图书馆统一的或共享的远程灾难备份系统中心正逐步具备条件。

例如,在深圳市,“图书馆之城”项目已全面展开,深圳图书馆、各区馆以及数以百计的社区图书馆,在“深圳市图书馆统一技术平台”服务系统支撑下,共享计算

机软硬件资源、网络资源、文献数据资源以及服务资源,为市民创造了良好的读书环境和提供了方便快捷获取知识的途径。为了保证整个深圳市图书馆业务应用系统和信息数据库的高等级安全,深圳市正计划在坪山新区新规划的“深圳市调剂书库”中建立中型规模的远程灾难备份系统中心,为全市图书馆的业务应用系统和数据库提供灾难备份和恢复的共享服务。虽然需要远程灾难备份系统来应对的灾难只占7%,但作为整个深圳地区图书馆的信息数据和文献资源保护则具有重大的意义和价值。

份,也可以设定每天在凌晨进行一次增量备份,然后每周末凌晨进行全备份。也就是说在尽量短的备份周期内,只可能每天做一次信息数据的备份。因此,如果出现灾难,信息数据的丢失量就是以“天”计算,无法做到“半天级”、“小时级”、“分钟级”等时间段更细致的信息数据保护。这是因为数据备份的时间窗口很大,不可能将繁忙的业务应用系统停机以取得足够长的时间来进行数据备份。

从信息数据恢复角度来看,备份的信息数据并非原有格式,有的甚至是专用的压缩格式,因此数据不能在导入后直接使用,必须进行数据的接收、格式转换和索引重组等,而这一过程往往需要数个小时,甚至更长的时间。

由此看来,磁带备份的技术和方式,对信息数据的保护和恢复能力,不能满足定时开馆的图书馆对数据安全和读者服务的要求,更不能满足365天×24小时服务的现代数字图书馆对实时性越来越高的要求。

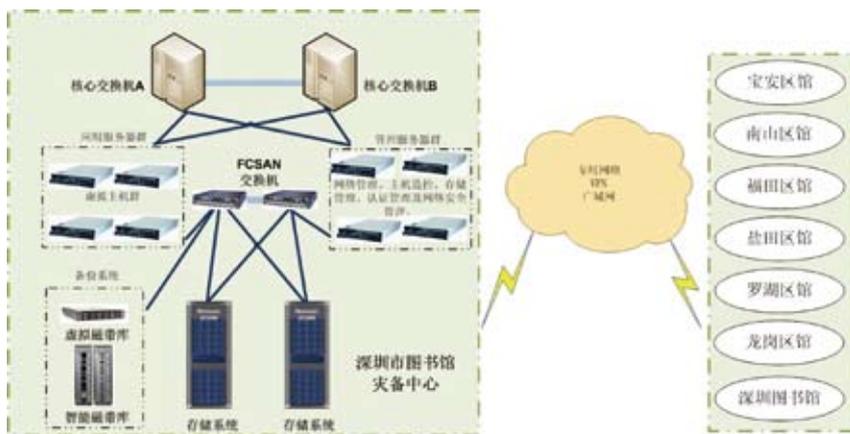


图1 深圳市图书馆灾备系统结构示意图

2 本地信息数据保护措施

适合于图书馆网络中心进行信息数据保护的技术手段和方法方式有很多,诸如磁带备份、磁盘镜像、数据快照等等,但最优秀的无疑是采用连续数据保护技术CDP(Continuous Data Protection)的保护方式。

2.1 磁带备份技术

相信所有的图书馆网络中心都使用过或正在使用最为传统的备份技术,就是磁带备份。虽然磁带备份的历史相当悠久,且具有技术简单、应用广泛、成本低廉等特点,但磁带备份也有备份和恢复时间长、备份数据的实时性和连续性差等不足。

为了不影响图书馆计算机网络应用系统对外服务的性能,同时也考虑到当天数据的完整性,通常网络中心技术人员会在每天闭馆时启动磁带机设备对数据进行全备

2.2 磁盘镜像技术

磁盘镜像技术实际上是一种利用冗余技术来实现当前信息数据完整保存的备份手段,采用磁盘镜像的主要目的是防止存储系统的硬件故障。生产数据和镜像数据被分别存放在不同的两个存储设备中,信息数据的变化被同时写入这两个存储设备,镜像数据就是生产数据的克隆,数据的实时性很高。当其中一个存储设备发生故障时,另一存储设备就会自动实时接管,信息数据的使用者对故障的发生毫无感觉,应用系统完全处于正常运行状

态,可以保证图书馆的服务照常进行。

但是,磁盘镜像也有“软肋”,其主要问题在于发生在生产数据中的病毒、数据结构破坏、数据误删或误改等软错误、软故障无法纠错和防护,而且这些“灾难”也会实时传染到镜像数据中,导致两份数据均不可用。

目前,图书馆网络中心主要在系统软件和应用软件存放的磁盘上采用磁盘镜像技术,一方面是保护系统软件和应用软件不会因为磁盘故障而导致系统崩溃,另一方面是系统软件和应用软件容量较小,一般只需一或两个硬盘就可容纳,镜像的成本较低。而各种各样的数据库容量较大,有的甚至达到几十个TB,镜像的成本太高,图书馆一般不会将存放数据库的硬盘采用磁盘镜像处理。

2.3 连续数据保护(CDP)技术

所谓的连续数据保护技术,就是采用连续数据复制和多点快照保护的方法,对数据的变化进行连续捕获和跟踪,并将变化的数据在生产数据之外独立存放,以确保数据可以恢复到过去的任意时间点。该技术的出现,克服了传统的磁带备份间隔较长、不够精细的缺点,使得传统意义的备份手段逐渐淡出,信息数据的快速恢复和任意时间点恢复成为现实。

从图书馆来说,各种数据库的数据量与日俱增,更关键的是数据的重要性和隐形价值也在逐渐地提高。而对于有些图书馆的核心应用来说,任何数量的信息数据丢失都可能会导致巨大的损失和重大的影

响。因此,为了确保信息数据的高安全性,图书馆网络中心必须对核心的在线业务应用系统,实行在线信息数据实时复制,尽可能多地采用快照等磁盘管理技术维持数据的高可用性,但这势必需要增加较大的投资。

连续数据保护系统可以基于块、文件或应用来实现,它的解决方案具备以下几个基本的特性:数据的改变受到连续的捕获和跟踪;所有的数据改变都保存在一个与主存储地点不同的独立存储地点中;数据恢复点的目标是任意的,而且不需要在实际恢复之前事先定义。

图书馆通过连续数据保护系统中的连续数据快照和复制技术,让信息数据卷自动创建基于增量的多时间点标记,也可以按需制定策略创建时间点标记,为数字图书馆系统等业务信息数据提供多份的异地备份影像。如此,不管是渐变式的软故障,还是站点式的硬故障,都能够快速有效地进行信息数据恢复,将灾难带来的损失降到最低。

采用连续数据保护系统可以为图书馆的业务应用系统提供365天×24小时不停顿运行的保护,使图书馆业务应用系统具有更快的数据库检索、更强的信息数据保护和更高的业务连续服务能力。相比之下,采用连续数据保护系统的总体产出与投入比要高于传统的备份解决方案。

目前,一些国际知名IT厂商推出了连续数据保护的产品或者类似产品,应用也越来越广泛,在我国的省级大型图书馆和高校图书馆已有不少的案例,这些连续数据保护系统主要采用了快照和网络复制相结合的技术,基本上具备了针对MSSQL、DB2、Oracle、MySQL等

流行数据库,提供实时备份、快速恢复、数据库接管、异地备份和集中备份等功能。

3 远程灾难备份

根据上述灾难发生比率的统计,需要远程灾难备份系统应对的自然灾害和社会灾难的发生约占灾难整体的7%,比率较小。因此,图书馆需要特别注意远程灾难备份系统的投入产出比,需要较低的成本实现具有实际意义的灾难备份和灾难恢复。事实上,也少有图书馆有人力、有精力、有经济实力去建立和维护一套远程灾难备份系统,毕竟对图书馆来说,远程灾难备份系统是一项“IT奢侈品”。

在以往的灾难备份系统建设中,图书馆采用的核心系统由于不同厂商在技术上普遍存在兼容性的问题,而导致系统之间很难实现互联互通和资源共享。因此,远程灾难备份中心基本上采用一对一的模式,也就是说图书馆本地网络中心与远程灾难备份中心的硬件资源和应用软件系统是一一对应的关系,本地网络中心有几套生产系统,那么远程灾难备份中心也需要几套相同品牌、相同类型的备份系统与之对应,这就造成了远程灾难备份系统的投资成本与本地网络中心相差不大。远程灾难备份中心与本地网络中心之间的通信线路由于需要高带宽、高安全性和高可靠性,基本上采用独立建设光纤链路模式,或租用电信部门的光纤链路模式,费用相当高昂。同时,远程灾难备份系统在技术和管理上复杂性很高,这就对图书馆建设和管理远程灾难备份中心在人员和技能上提出了很高的要求。由此看来,建设、管理

和维护一个远程灾难备份中心需要耗费大量的人力、物力和财力，我们必须以战略性的眼光，考虑以某个城市或某个区域的多个甚至几十个图书馆来共享远程灾难备份中心的资源，对图书馆群的统一服务、整合文献信息资源、保障应用系统不间断运行等方面都有重要的意义。

从建设一个高性价比的远程灾难备份中心角度考虑，要降低远程灾难备份中心的建设成本，首先必须要解决异构系统平台的兼容性和互通性问题，然后还要解决远程通信链路部署中的传输技术和投入成本问题。

3.1 关于兼容性和互通性问题的解决

在远程灾难备份中心建设中，图书馆毫无疑问地要面对本地网络中心“多国部队”的各种业务应用系统，必须要打通本地网络中心和远程灾难备份中心两端的技术壁垒，实现从封闭系统到开放平台的跨越，从独立系统到集成平台的升级，形成统一的共享资源池。通过标准化的开放平台，实现图书馆本地网络中心与远程灾难备份中心的无缝对接。通过存储虚拟化技术，对主服务器、存储系统和网络系统三个层面进行虚拟化处理，可以有效地解决资源共享问题，消除应用层和存储层资源分配的限制，将多种存储系统整合到统一的资源池中。在本地网络中心通过灾难备份系统统一备份到标准化的远程灾难备份中心的模式，大大地降低了远程灾难备份中心的建设成本。

3.2 远程通信链路部署的问题

考虑到图书馆有效的服务区域范围和图书馆之间文献资源共享的有效达到区域，远程灾难备份中心一般建立在同城范围之内，从建设成本和操作成本上看是最可行的。也就是说，在远程灾难备份中心建设的选址考虑中，图书馆与金融等其他行业是有所不同的。

图书馆本地网络中心与远程灾难备份中心之间的距离，近则几公里，远则几十公里。因此，灾难备份系统对链路的带宽和可靠性的要求往往比较高，传统的FC技术对高带宽能够提供可靠的保证，只要增加FC协议与IP协议的转换设备，直线距离可超十公里以上。另外，基于IP的微扫描增量复制技术，只对变化的数据进行复制，大大地降低了灾难备份系统对带宽的要求，减少了远程灾难备份系统部署的投入成本和维护成本。

3.3 维护管理和技术支持的问题

图书馆在建立远程灾难备份系统时，考虑完善的技术解决方案经过周密的部署不会有很大的漏洞。但是在实施过程中会有不少具体的问题需要考虑和解决，其中两个重要指标RPO（Recovery Point Object，恢复点目标）和RTO（Recovery Time Object，恢复时间目标）是我们必须特别关注的，RPO和RTO两个技术指标越小，也就意味着出现灾难的时候可以让业务继续运作的时间越短，丢失信息数据的时间段也越短。建立灾难备

份系统的最终目标，就是要提高恢复的准确性、缩短恢复时间和缩短业务中断时间。因此图书馆应与主机系统、存储系统、网络系统等厂商甚至包括集成商合作进行基础设施的建设，同时图书馆也必须组建自己的高水平技术支持和维护管理团队，这也是是否有能力建设与维持远程灾难备份系统的关键因素之一。

考虑到灾难备份系统的建设成本、维护成本和技术队伍资源的诸多因素，因此在图书馆领域要有共享远程灾难备份系统的全新建设思路，就是在一个城市或地区中所有或多个图书馆，以资金雄厚、技术力量强大的核心图书馆为依托，进行远程灾难备份系统的共建共享，其意义不仅限于基础设施和设备资源共享，更重要的是必将进一步发展为同城或地区图书馆之间共享数据信息和文献资源的平台，以及统一规则、集中服务和整合异构平台的图书馆综合业务服务应用系统。

4 结束语

灾难备份系统作为图书馆网络中心的重要组成部分，其建设必须要切合实际，在掌握图书馆业务应用需求和现状下，与各共享图书馆的远程通信接入、网络系统架构、服务器系统平台、业务应用和管理系统、存储系统区域网络等各方面多层次的密切结合。充分做好灾难备份系统的需求和技术调研工作，不走错路，不走弯路，为图书馆的文献信息数据提供全面和完善的安全保障，这是图书馆服务大众的基础保证，也是图书馆必须要承担的重要社会责任。

参考文献

- [1] 许志峰. 解读《信息系统灾难恢复规范》[J]. IP领航特刊, 2010.
- [2] 许云松. 统一数据管理——让共享灾备更智慧[J]. IP领航特刊, 2010.
- [3] 带你走进灾难备份的世界[EB/OL]. [2010-09-15]. <http://www.dvpeixun.com/Detail/news-51325.html>
- [4] 主流灾难备份及恢复技术分析[EB/OL]. [2010-09-15]. <http://hf.yesky.com/282/8625282.shtml>.
- [5] 数据大集中和灾难备份技术[EB/OL]. [2010-09-15]. <http://www.docin.com/p-20485746.html>.
- [6] 灾难备份系统在我国市场十分明朗[EB/OL]. [2010-09-15]. <http://tech.watchstor.com/backup-and-archiving-119765.htm>.

作者简介

郑晓军(1962-), 男, 副研究馆员, 大学本科, 深圳图书馆技术中心副主任, 研究方向为计算机应用。E-mail:zxj@szlib.gov.cn

Library Information Data Protection System to Establish the Ultimate Security Strategy

Zheng Xiaojun / Shenzhen Library, Shenzhen, 518036

Abstract: Any information of data and information systems disaster, will bring to the library with different levels of risk. Disaster recovery related to the library in emergencies can quickly turn around, whether related to library services sustainable manner. Therefore, disaster backup information for library construction and modernization of both the progress of technology, knowledge accumulation and inheritance is guaranteed.

Keywords: Disaster recovery, Data security, Data protection, Backup technology

(收稿日期: 2010-09-24)