

电子档案备份应用系统研究

○顾彦彦

电子档案备份研究是随着现代信息技术的快速发展,以及屡屡发生的档案消失、受损事件警示下,而引起档案界人士对档案数据安全保管及备份问题的关注。尽管早在2005年,我国国家互联网信息办公室就明确提出:在重要信息系统领域要加强灾难备份工作,并明确指出灾难备份为我国信息安全保障体系的重要组成部分。^[1]然而对电子档案安全保管的实际研究则始于近两年,随着我国档案信息化的不断发展,电子档案数量的不断增加,且日渐占据主要地位,电子档案备份作为最直接最有效的手段以保证电子档案安全而倍受瞩目。

1 电子档案备份涵义及必要性

1.1 电子档案备份涵义

电子档案备份是指为了能够恢复电子档案数据而将其进行定期备份的过程,是电子档案安全防御的最后一道防线。其目的是根据某种数据冗余机制,实现现行电子档案与电子档案备份的隔离,当正在使用的电子档案被破坏时,能够通过电子档案备份进行及时完整的恢复。^[2]电子档案备份的实现需要有存储介质、备份软件工具和备份策略方式三个部分构成。

1.2 电子档案备份的必要性

电子档案备份可以在电子档案因自然灾害、病毒入侵、硬件故障、人员误操作等造成档案数据破坏丢失时,用来完整地恢复原始档案数据,从而能够提高电子档案的容灾能力,进一步保障电子档案的安全。

首先,电子档案备份有助于保证档案信息资源“永久长存。”档案信息资源不仅对科学研究有历史查考作用,而且为现代化建设提供凭证参考价值,同时还发挥着保存民族记忆的“印记”作用,然而其本身的原始性、唯一性特点决定了一旦遭到损毁,将永远消失。电子档案备份从源头上避免了档案“永久”消失问题的发生。^[3]

其次,电子档案备份适应了档案信息化建设的步伐。电子档案备份研究工作的开展是以档案信息化建设为基础平台,涉及到电子档案归档管理,档案数字化等档案信息化建设内容。同时,电子档案备份研究也促进了档案信息化建设。

最后,电子档案备份是档案安全体系建设的需求体现。电子档案备份因其可以有效保障电子档案安全,对预防意外自然灾害,网络病毒攻击,人为误操作等造成的电子档案消失,受破坏等发挥着重要保障作用,从

而成为提高电子档案安全保障能力的新模式,迎合了现代档案安全体系建设工作的需求。

1.3 电子档案备份系统的设计要求

电子档案是档案的网络化表现形式,与纸质档案有着同样重要的保管利用价值,然而由于电子档案自身的易修改性、不安全性等因素的存在,要求电子档案备份系统不仅要全面考虑电子档案的数据海量、容灾能力要求高的特点,而且应满足一套完善的备份系统标准,即稳定性、全面性、容错性、实时性及自动备份等。此外,电子档案数据备份的高性能和操作简单也应考虑在内。^[4]下面将对电子档案备份方式、存储介质及软件工具进行一一选择,以设计出符合电子档案备份系统要求的电子档案备份系统。

2 电子档案备份的类别选择

电子档案备份有多种形式,可以分别按照不同方式进行分类。

2.1 按照电子档案备份的策略分类

按照电子档案备份的策略可分为完全备份、增量备份和差分备份。^[5]

(1)完全备份就是每次都整个电子档案系统进行全部备份,这种备份方式对电子档案进行完整的拷贝,当恢复时可以一次性快速恢复。但是,因为每次备份都是对全部电子档案备份,会占用很大硬件存储空间,同时备份的时间也比较长,因为随着备份次数的增多,备份之间会产生许多重复的部分,特别是当电子档案数据是海量的时候,这就明显增加了存储成本,并且备份效率低。

(2)增量备份每次只备份与上一次备份的电子档案中不同的数据,这些数据可能是新增加的或者是修改过的。这种方式可以避免完全备份的那种因为每次都要备份重复数据而耗费大量的存储空间和时间,但是其恢复比较复杂,要将之前的所有备份通过时间逐个进行处理,才能恢复出所需的数据,且恢复的效率很低。

(3)差分备份同时吸收完全备份和增量备份的优点,主要针对现在相对于上一次全备份的基础上增加了新的数据或者修改了数据,才进行备份。这种备份方式是基于全备份方式的,应用在全备份后的后续备份工作中。它避免了以上两种备份方式的确定,不必备份大量重复数据,从而节省了备份的存储空间和时间,而且在对电子档案进行恢复的时候,因为基于完全备份,所以速度很快,从而使得备份和恢复同时具有较高的

效率。

2.2 按照电子档案备份的地理位置分类

按照电子档案备份的地理位置不同可分为本地备份和异地备份。

(1)本地备份是将电子档案的备份保存在本地的存储介质上,如直接备份在原来计算机硬盘上的专门备份区域,或者在其他本地的存储介质上。这种备份方式比较方便快捷,但是电子档案无法抵抗本地的自然灾害和人为破坏,这就需要与异地备份进行结合,才能实现电子档案备份的有效性。

(2)异地备份就是将电子档案的备份保存在异地的存储介质上,两地间的距离最好保持在500公里以上,使得其可以避免受到本地灾害破坏,提高电子档案的容灾能力。但同时也增加了备份复杂性,因为要备份在异地,需要将备份数据通过网络传输或者人工直接传送存储备份的硬件。^[6]

2.3 按照电子档案备份的响应方式分类

按照电子档案备份是否实时对电子档案的增加和修改可将电子档案备份分为冷备份和热备份2种,有时候也被人们称为离线备份和在线备份。^[7]

(1)离线备份即冷备份,在进行备份时,电子档案服务器将停止用户使用,即不进行电子档案的增加和修改,专门进行备份工作,这就使得用户在电子档案备份这段时间,不能对档案进行修改和增加等操作,所以电子档案备份一般在晚上或凌晨进行,尽量避免影响电子档案的及时更新与修改。不过这种备份方式可以避免由于用户操作进行更新修改电子档案而造成的数据不一致问题。

(2)在线备份即热备份,与冷备份相对应,热备份在进行电子档案备份的时候不停止用户对电子档案的操作,可以实时地对电子档案进行备份。但同时也增加了电子档案备份的技术复杂性,因为要实现电子档案备份与用户对档案的更新操作之间的实时同步,就必须解决数据不一致的问题。此外因为是实时备份,还会占用相当一部分的系统内存和网络资源,这就需要先进的硬件和网络支持。

3 电子档案备份存储介质与软件工具的选择

3.1 电子档案备份存储介质的选择

目前,市面上常用的电子档案备份存储介质有磁盘阵列、光盘库和磁带库3种,因各有其优缺点,根据不同电子档案管理单位的条件可以选择不同的存储介质,也可以进行结合,以最大化满足电子档案备份需求。

(1)磁盘阵列

磁盘阵列是将多块硬盘通过特殊的计算机技术组合成一个阵列,通常为保证兼容性和一致性,这些硬盘类型和大小都是相同的。^[8]作为一种常用的数据存储设备,磁盘阵列能够实现快速便捷的存取电子档案数据,且具有高可靠性和安全性,能够实现实时的电子档案备份,但是磁盘的价格相对较高,用其保存海量的电

子档案会使得备份成本很高。

(2)光盘库

光盘库是由多个光盘通过光盘驱动器并联组成。这些光盘是可擦写光盘,可以利用光盘驱动软件通过光盘刻录机将电子档案备份数据刻录到光盘上。光盘的容量一般很小,相对硬盘来说性价比不高,同时光盘不易经常用来读写,因为刻录机稳定性不高。所以可以在光盘上可用来存放不经常访问和修改增加的电子档案备份。^[9]

(3)磁带库

磁带存取数据速度比较慢,且具有不可修改性。但是磁带价格低廉,特别适合不经常查看的海量电子档案的定期或永久备份。

3.2 电子档案备份软件工具的选择

当前数据备份软件多种多样,所用技术也各有不同。根据电子档案备份应用系统的需求以及本地与异地备份结合的设计,本文主要选择了2款主流的数据备份管理工具 Veritas Backup Exec 和 VERITAS Storage Replicator(VSR)。

Veritas Backup Exec 数据备份软件是由 VERITAS 软件公司推出的目前市场上唯一中文版的备份程序,^[10]提供直观友好的备份管理界面,用户可以自定义备份方式和存储介质空间分配,并可以实时监控备份过程。在本地使用 Veritas Backup Exe 支持智能化备份策略和磁带多路复用功能,能够进行高效快速稳定的备份和自动搜寻恢复。同时其还具有病毒木马查杀防御功能,能够在备份的同时,进一步进行电子档案的安全检测,防止不安全的档案进入档案备份系统。

VSR 是一款基于文件系统的数据复制软件,能够通过远程方式进行异地数据备份。^[11]在远程异地进行电子档案备份会面临到与本地数据不一致和网络传输故障破坏电子档案完整性的问题,而 VSR 数据复制软件高可靠性的实时备份文件功能和多地分布同步数据拷贝功能能够很好的解决这些问题,同时 VSR 提供调节备份过程中网络带宽和选择复制数据的功能,使得用户能够有选择的备份电子档案数据,并根据网络利用率灵活调节所使用的网络带宽大小,这就能够在一定程度上保证数据备份的效率,此外其用户图形界面非常友好,操作简单,使得远程备份监控端的工作人员易于学习掌握,简化了工作量,所以选择 VERITAS Storage Replicator 作为远程异地电子档案备份的软件。

4 电子档案备份的应用系统设计

根据以上所分析的各类型电子档案备份方式及所选择的存储介质、软件工具,笔者对电子档案备份系统进行了设计应用,以弥补当前我国对电子档案备份认识不足及技术缺陷的问题。

4.1 电子档案备份的总体系统结构

考虑到电子档案的极其重要性,电子档案备份必须要对病毒攻击、人为破坏特别是自然灾害具有较强

的抵御能力,所以电子档案备份的总体系统结构采用本地备份和异地备份相结合的结构,即在本地和异地同时对原始电子档案进行备份。如果电子档案只备份在本地,当本地发生地震、洪水等不可避免的自然灾害时,备份也将被毁坏,起不到备份的作用。而增加异地备份,即使一地的备份数据被破坏,仍然可以利用另外一地的备份进行恢复电子档案。这样在一定程度上实现了电子档案备份的双重防护,增强电子档案的安全性。电子档案备份的总体系统结构如图1所示。

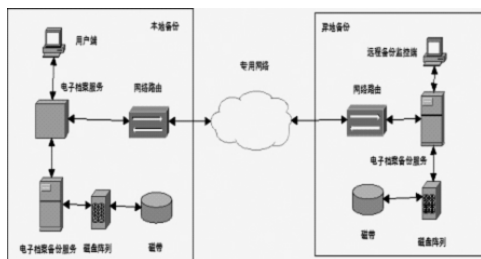


图1 电子档案备份的总体系统结构

图1中本地备份是在电子档案服务器所在地进行的备份,当进行本地备份时,电子档案备份服务器,首先根据制定的备份策略将电子档案从电子档案服务器中备份到磁盘阵列,因为磁盘的存取速度要比磁带快很多,所以首先备份到磁盘阵列中来提高备份效率,减少对电子档案服务器的影响,这之后再从磁盘阵列中的备份拷贝到磁带中进行永久保存。这样的设计有两个优点,一是磁盘存取速度快,提高了备份效率,同时如果只是恢复近期的数据,可以直接从磁盘上恢复,加快了恢复的速度;二是磁带虽然存取数据效率低,但价格便宜、存储容量大,节省了成本,且磁盘向磁带拷贝数据时已经不需要电子档案服务器的参与,并不会影响到用户的使用。

异地备份是在与电子档案服务器不同的地方进行备份,两地应当尽量有较远的距离,才能避免自然灾害同时在两地发生,当进行异地备份时,电子档案服务器利用网络路由通过专用网络与在异地的电子档案备份服务器进行连接,发送电子档案数据。异地的远程备份监控端可以控制异地备份并查看备份情况,同样,为了实现效率和成本的平衡,异地的电子档案备份服务器也是先将电子档案备份到磁盘阵列中,而后再永久存储到磁带或光盘中。

4.2 电子档案备份系统的内部策略

本文中所设计的电子档案备份系统的内部策略采用完全备份与差分备份相结合的方式,在本地通过 Veritas Backup Exe 数据备份管理软件备份,在远程异地通过 VERITAS Storage Replicator 数据复制软件来实现电子档案双地备份。电子档案备份系统的内部策略实现流程如图2所示。

系统首先通过数据备份管理软件对电子档案进行完全备份,而后初始化触发器。触发器开始计时,当时

间没到达设定的一个周期时将循环等待,至到触发器触发下一步动作,表明已经过了一个周期时间,触发器触发系统将现有的电子档案文件与先前的完全备份进行比较。如果没有不同表明电子档案没有做新的修改或者增加,无需进行备份,则初始化周期触发器进入下一个周期等待;如果有不同的部分,则需要将这部分进行差分备份,在本地通过 Veritas Backup Exe 数据备份管理软件备份,在远程异地通过 VERITAS Storage Replicator 数据复制软件进行备份。最后初始化周期触发器循环进入下一轮的备份周期等待。

5 小结

本系统采用完全备份与差分备份相结合的方式能够综合存储容量少、备份和恢复速度快的优点,使得备份和恢复同时具有较高的效率。同时稳定的数据备份软件的应用也保证了系统高效安全地运行,极大满足了电子档案备份系统的设计要求,具有重要的开发价值。

参考文献:

- [1]袁晓瑛.浅论电子档案信息安全保障[J].科技创新导报.2009(16):33.
- [2]郝晨辉.浅谈电子政务环境下档案信息灾难备份系统建设[J].档案学研究.2005(4):12.
- [3]徐震.数字档案馆数据备份系统方案的规划[J].兰台世界.2006(9):7-11.
- [4]肖红.浅析图书馆数据备份管理[J].科技情报开发与经济.2006(3):54.
- [5]黄晶.数据备份系统的研究与实现[D].武汉:华中科技大学.2008:3-8.
- [6]汪中夏,刘伟.数据恢复高级技术[M].北京:电子工业出版社.2006(7):34.
- [7]牛云,徐庆.数据备份与灾难恢复[M].北京:机械工业出版社.2004(2):28-30.
- [8]常丽.电子档案存储载体的选择及保管探析[J].办公室业务.2006(3):2-3.
- [9]陈勇.电子档案载体与信息保护新探[J].科技档案.2006(2):11-15.
- [10]孟涛,徐力.数据备份的技术探讨[J].网络安全与应用.2004(8):15-17.
- [11]景晓丽.档案管理系统之数据备份与恢复[J].兰台内外.2008(4):45-56.

(作者单位:广西民族大学管理学院)

责任编辑/张杰

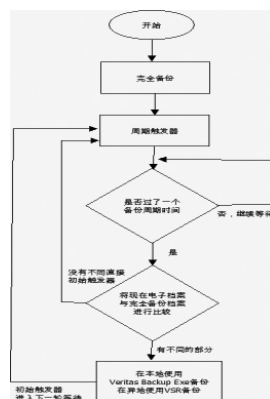


图2 电子档案备份系统的内部策略实现流程。