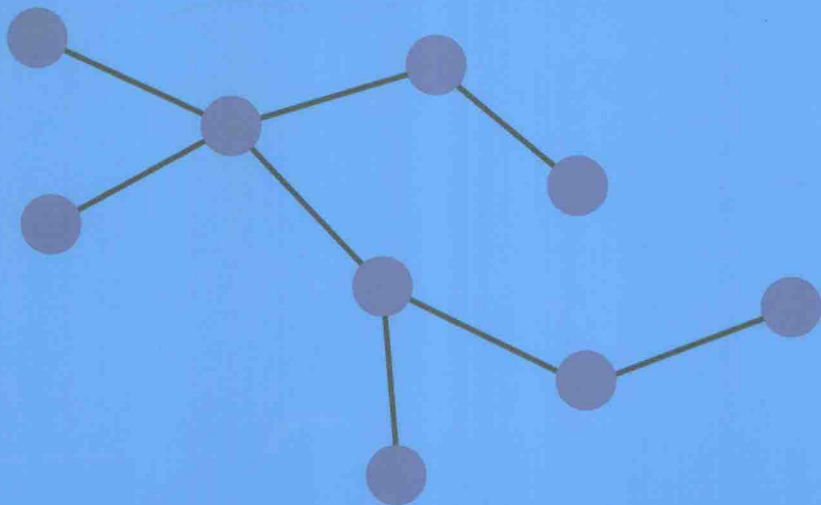


高等院校计算机任务驱动教改教材

数据安全与灾备管理

贾如春 周晓花 主 编
陈新华 王宏旭 吴 粟 副主编
赵克林 康 乐 主 审



清华大学出版社



高等院校计算机任务驱动教改教材



数据安全与灾备管理

贾如春 周晓花 主 编
陈新华 王宏旭 吴 粟 副主编

清华大学出版社
北 京

内 容 简 介

本书基于项目化教学方式编写而成,并引入实际企业案例。主要包括:数据存储基础、存储应用环境、数据存储技术、RAID 技术与应用、虚拟磁带库技术、数据灾备与恢复技术、虚拟化技术、灾备系统设计与典型案例分析、数据中心安全运维、云计算应用与云灾备、大数据存储。全书以社会调查、企业岗位需求为基准,以一家真实企业公司在不同发展时期的 IT 架构建设、数据安全与灾备为背景,从数据灾备系统建设到数据中心、数据安全维护进行了详细讲解,并配合详细的案例分析与实训内容,本着“学生能学、教师能教、企业能用”的原则,注重理论知识与实践技能的掌握,不断培养学习者的自信心和自学能力,力求学习者通过学习对此技术领域形成一个比较全面的知识体系,并且掌握数据安全与灾备相关的主流技术和产品管理,从而提高自身技能与分析解决问题的能力,将对知识的理解和实践应用有机地结合为一体,更好地应用于社会服务中。

本书既适合作为计算机相关专业本科、专科学生的教材,也可以作为相关工程技术人员的学习用书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数据安全与灾备管理/贾如春,周晓花主编. —北京:清华大学出版社,2016

高等院校计算机任务驱动教改教材

ISBN 978-7-302-44507-4

I. ①数… II. ①贾… ②周… III. ①数据处理—安全技术—高等学校—教材 IV. ①TP274

中国版本图书馆 CIP 数据核字(2016)第 171848 号

责任编辑:张龙卿

封面设计:徐日强

责任校对:李 梅

责任印制:李红英

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62770175-4278

印 装 者:北京国马印刷厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:23.25

字 数:562 千字

版 次:2016 年 8 月第 1 版

印 次:2016 年 8 月第 1 次印刷

印 数:1~2000

定 价:49.00 元

产品编号:070241-01

前 言

随着“互联网+”、大数据时代的到来,人类已经从信息技术(Information Technology, IT)时代逐步走向数据技术(Data Technology, DT)时代。业务应用数据的日益更新越来越深刻地影响着企业的经营管理模式,这当中影响最大的就是企业的信息管理模式。一方面极大地促进了企业的发展,另一方面又对海量信息数据的存储和管理提出了新的挑战。对于数据量很大的企业级数据来说,性能好坏非常重要,直接影响到业务的质量。因此如何有效使用数据已成为 IT 系统中极其重要的基础部分,并已成为计算机领域里相对独立的分支学科。

大数据时代的到来,也对数据安全提出了更高的要求,对于信息化应用而言,灾备系统的建设已成为热点。随着 2007 年《信息系统灾难恢复规范》(GB/T 20988—2007)正式成为国家标准,许多用户对数据灾备从观望、徘徊转向实际应用。随着虚拟化的广泛应用、云计算的出现,以及数据量每年的高速增长,数据类型和来源的多样化,使得在如此复杂的环境下如何高效、安全、可靠、完整地使用和保存宝贵的数据资料已成为当务之急,只有使用高性能计算机网络存储系统,才能从根本上解决企业日益增长的需要。而自动化的存储管理手段,不仅可以解决现有企业关键数据的存储和管理的需求,而且可以同时为网络上各种工作站的用户提供数据的备份,减轻系统管理员的负担,有效地保护宝贵的数据及人力资源,这样在不幸遇到灾难后,就可以很迅速地恢复数据,使整个系统在最短的时间内重新投入正常运行。

本书响应国家安全号召,坚持“自主开发、安全可控”的原则,基于国内众多知名企业如精容数安、人大金仓、聚比特科技等具有多年计算机系统设计、存储系统管理、数据安全、数据备份系统建设等真实项目案例累积的经验基础上所提供的资料,致力于为用户提供最完整的解决方案,不仅能够支持企业自身的操作系统平台,而且能够支持 COMPAQ、IBM、Microsoft、SGI、SUN 等各种主流操作系统平台和多种数据库和应用平台,涵盖从高可用性、远程/异地容灾、数据整合到网络数据备份、零停机备份以及长期归档等数据管理和数据保护的解决方案。

本书内容概况如下。

(1) 数据存储概述篇。本篇主要针对计算机系统中与存储相关的技术进行介绍,包括存储器的组成结构、工作原理、存储管理系统以及不同存储

产品技术路线等内容。

(2) 数据灾备概述篇。本篇主要是针对数据灾备技术的不同方案类型以及在虚拟环境下如何实现灾备技术部署等内容进行介绍。

(3) 数据灾备应用篇。本篇主要针对整个行业典型的灾备技术解决方案进行介绍和模拟,以及对大型数据中心运维中有关数据安全的运维服务进行介绍。

(4) 数据灾备趋势篇。本篇主要讲述云计算的相关技术、设施框架类型以及各种云环境下的云灾备和数据迁移服务技术。

本书不仅适合作为本科院校及职业院校计算机类、通信类相关专业的教材使用,也可作为从事系统集成、数据容灾备份等领域工作的相关技术人员的参考用书。

本书由贾如春负责总体策划设计及统稿,并与企业专家周晓花共同担任本书主编,陈新华、王宏旭、吴粟担任本书的副编辑,全书由赵克林教授及企业专家康乐主审。感谢精容数安科技有限公司给予作者的支持与帮助,同时感谢其他高等院校中各位教师对本书提出的修改建议。

由于作者水平有限,本书涉及的知识面较广,虽然编者已经尽了最大努力,但书中难免存在错漏之处,欢迎广大读者批评指正。

编 者

2016 年 6 月

目 录

第一篇 数据存储概述

第 1 章 数据存储基础	3
1.1 数据存储发展历史	4
1.1.1 信息数据发展	4
1.1.2 存储的基本概念	5
1.2 数据存储原理	5
1.2.1 存储的基本原理	5
1.2.2 常见存储设备及其组件	6
1.2.3 存储网络的分类	9
1.2.4 存储网络的比较	12
1.2.5 RAID 存储技术	15
1.2.6 存储性能指标	15
1.3 数据存储技术	16
1.3.1 备份技术背景	16
1.3.2 数据备份原则	17
1.3.3 数据灾难基础	17
1.3.4 数据容灾与备份	18
1.3.5 数据容灾指标	19
1.3.6 数据容灾级别	19
1.4 数据存储与应用	21
1.4.1 数据存储与数据访问	21
1.4.2 存储网络对应用系统的影响	21
1.4.3 存储网络对应用系统的意义	23
任务拓展	24
第 2 章 存储应用环境	25
2.1 IT 基础设施	26
2.1.1 计算机系统组成	26
2.1.2 IT 系统简介	26
2.1.3 数据中心的概念	29

2.2 存储环境.....	31
2.2.1 主机的内部应用环境	31
2.2.2 传统内置存储遇到的问题	31
2.2.3 网络存储应用环境	31
2.3 存储介质.....	33
2.3.1 机械硬盘	33
2.3.2 SSD 硬盘	38
2.3.3 SSD 与 HDD 的比较	48
2.3.4 SSD 性能优势	50
2.3.5 SSD 在存储中的应用	51
2.4 存储主机应用系统.....	52
2.4.1 主机硬件系统	52
2.4.2 主机软件系统	54
2.4.3 主机的管理	56
2.5 存储中应用数据库.....	59
2.5.1 数据库的基本概念	59
2.5.2 数据库处理系统	60
2.5.3 数据库的基本结构	60
2.5.4 数据库的主要特点	60
2.5.5 数据库的结构模型	61
2.5.6 数据库的备份与恢复	62
任务拓展	65
第3章 数据存储技术	66
3.1 存储阵列系统的组成.....	67
3.1.1 存储阵列系统的基本概念	67
3.1.2 存储阵列在存储系统架构中的位置	67
3.1.3 存储阵列系统硬件的组成	67
3.2 传统的存储系统.....	68
3.2.1 传统的内置存储	68
3.2.2 外置存储网络的形态	69
3.2.3 DAS 存储的形态	70
3.2.4 DAS 存储的局限性	70
3.3 SAN 技术与应用	71
3.3.1 SAN 存储基础	71
3.3.2 FC 连接与协议	74
3.3.3 SCSI 协议介绍	80
3.3.4 iSCSI 协议介绍	85
3.3.5 IP SAN 的产生与发展	89

3.3.6 IP SAN 的组成和组网连接	92
3.3.7 FC 协议与 TCP 协议融合	95
3.4 NAS 技术与应用	104
3.4.1 NAS 存储基础	104
3.4.2 NAS 网络拓扑	104
3.4.3 NAS 的发展及特点	105
3.4.4 集群 NAS	105
3.4.5 NAS 与文件服务器的对比	105
3.4.6 NAS 系统的组成与部件	106
3.4.7 NAS 文件共享协议 CIFS 及 NFS	108
3.4.8 NAS 文件系统的 I/O 与性能	110
3.4.9 NAS 的实现与应用	111
任务拓展	117

第二篇 数据灾备概述

第 4 章 RAID 技术与应用	121
4.1 RAID 技术介绍	122
4.1.1 RAID 简介	122
4.1.2 RAID 的实现技术分类	122
4.1.3 RAID 的基本原理	122
4.1.4 RAID 的关键技术	123
4.1.5 RAID 的优势	125
4.2 RAID 级别分类	125
4.2.1 RAID0	126
4.2.2 RAID1	126
4.2.3 RAID2	126
4.2.4 RAID3	127
4.2.5 RAID4	128
4.2.6 RAID5	128
4.2.7 RAID6	129
4.2.8 RAID 组合等级	129
4.2.9 非标准 RAID 等级	132
4.2.10 RAID 的实现方式	136
4.3 RAID 的应用选择	137
任务拓展	138
第 5 章 虚拟磁带库技术	139
5.1 虚拟磁带库介绍及相关产品对比	140

5.1.1	架构	140
5.1.2	各种 VTL 产品间的重要差异	141
5.1.3	虚拟磁带库和磁带库的对比	143
5.2	虚拟磁带库的实现方式	144
5.2.1	备份软件型虚拟磁带库(第Ⅰ代 D2D)	144
5.2.2	应用服务器级虚拟磁带库方案(第Ⅱ代 D2D)	145
5.2.3	智能化专用型虚拟磁带库设备方案(第Ⅲ代 D2D)	145
5.3	虚拟磁带库功能介绍	147
5.3.1	新型环境与创新存储的应用	147
5.3.2	直接磁带输出	147
5.3.3	删除重复数据	148
5.3.4	I/O 负载平衡	148
5.3.5	硬件压缩	148
5.3.6	无服务器磁带备份	148
5.3.7	销毁虚拟磁带	148
5.3.8	磁带整合	149
5.3.9	磁带缓冲	149
5.3.10	按需扩容	149
5.4	虚拟磁带库的优势与劣势	149
5.4.1	VTL 的优点	150
5.4.2	VTL 的缺点	150
5.5	虚拟磁带库的管理方式	151
5.6	虚拟磁带库数据的迁移	151
5.7	虚拟磁带库与重复数据删除技术介绍	152
5.7.1	重复数据删除的概念	152
5.7.2	重复数据删除的技术基础	153
5.7.3	重复数据删除技术的分类	154
5.7.4	重复数据删除操作的基本原理	155
5.7.5	重复数据删除可选择的方式	156
5.7.6	重复数据删除的优势	157
5.8	虚拟磁带库的趋势介绍	159
5.8.1	变化中的虚拟磁带库市场	159
5.8.2	用户的需求	159
5.8.3	下一代产品的增强特性	160
	任务拓展	161
第 6 章	数据灾备与恢复技术	162
6.1	灾备技术介绍	163
6.1.1	数据备份技术	164

6.1.2	本地介质备份及异地介质存放方案	164
6.1.3	远程数据备份方案	165
6.2	基于目的端重复数据删除技术的备份方案介绍	167
6.2.1	技术描述	167
6.2.2	资源配置要求	168
6.2.3	技术特点	168
6.2.4	适用范围	168
6.2.5	可选技术	168
6.3	基于源端重复数据删除技术的备份方案介绍	168
6.3.1	技术描述	168
6.3.2	资源配置要求	169
6.3.3	技术特点	169
6.3.4	适用范围	170
6.3.5	可选技术	170
6.4	基于智能存储设备的数据复制技术	170
6.4.1	技术描述	170
6.4.2	技术架构及特点	170
6.4.3	生产—同城复制模式(即 A—B 模式)	171
6.4.4	资源配置要求	173
6.4.5	适用范围	174
6.4.6	主流技术	174
6.4.7	实施步骤	174
6.5	基于数据库的数据复制技术	174
6.5.1	技术描述	174
6.5.2	资源配置要求	175
6.5.3	技术特点	175
6.5.4	主要实施步骤	176
6.5.5	适用范围	176
6.5.6	主流技术	176
6.6	基于主机的数据复制方案介绍	176
6.6.1	数据卷镜像方案	176
6.6.2	数据卷复制方案	179
6.6.3	主要实施步骤	181
6.7	基于存储虚拟化的数据复制技术介绍	182
6.7.1	模式一:带外数据、带外管理模式	182
6.7.2	模式二:带内数据、带外管理模式	183
6.7.3	模式三:带内数据、带内管理模式	184
6.8	灾备技术对比表	186
	任务拓展	187

第7章 虚拟化技术	188
7.1 虚拟化技术概述	188
7.1.1 虚拟化的定义	188
7.1.2 虚拟化的发展历史	189
7.1.3 百花齐放的虚拟化技术	191
7.1.4 虚拟化的优势	192
7.1.5 虚拟化的目的	193
7.2 现有虚拟化技术的分析与对比	194
7.2.1 服务器虚拟化	194
7.2.2 网络虚拟化	214
7.2.3 存储虚拟化	217
7.2.4 应用虚拟化	220
7.2.5 虚拟化技术的比较	222
7.3 虚拟化架构对比分析	223
7.4 虚拟平台所面临的安全问题	227
7.4.1 使用虚拟化环境时存在的缺陷	227
7.4.2 保障虚拟服务器环境安全的措施	227
7.5 虚拟化未来发展趋势	229
7.5.1 被重构的 IT 图景	229
7.5.2 云计算的演进	229
任务拓展	231

第三篇 数据灾备应用

第8章 灾备系统设计与典型案例分析	235
8.1 灾备需求分析	236
8.1.1 信息可行性分析	236
8.1.2 故障分析	236
8.1.3 基础架构分析	237
8.2 系统的设计思路和设计原则	243
8.2.1 系统设计的理论依据和规范	243
8.2.2 系统设计方法论	243
8.2.3 技术路线	243
8.2.4 系统设计原则	244
8.3 备份系统建设的重要性	245
8.3.1 系统灾难分析	245
8.3.2 国家对数据安全的重视	246
8.3.3 备份系统保护数据安全	246

8.3.4 备份系统的保护场景	246
8.4 典型案例分析	246
8.4.1 基本数据保护及数据归档场景	246
8.4.2 应用保护场景	248
8.4.3 大型灾备场景	249
任务拓展	251
第9章 数据中心安全运维	252
9.1 运维体系介绍	252
9.1.1 运维管理服务体系	253
9.1.2 服务提升工具	255
9.1.3 运维监控平台的建设原则	255
9.1.4 系统集中监控方案	256
9.1.5 统一事件管理平台建设	259
9.1.6 报表管理系统	259
9.1.7 运维服务管理平台建设	260
9.1.8 应急机构与职责	261
9.1.9 突发事件分级	262
9.1.10 应急响应	265
9.2 运维制度管理	267
9.2.1 管理制度架构	267
9.2.2 管理制度说明	267
9.2.3 运维服务内容综述	269
9.2.4 基础设施保障服务	270
9.2.5 变更管理服务	272
9.2.6 问题管理服务	275
9.2.7 客户服务管理	277
9.2.8 运维服务质量管理	278
9.2.9 服务水平管理	279
9.2.10 灾难恢复服务	279
9.2.11 应急资源保障	280
9.2.12 应急宣传、培训与演练	281
9.3 安全管理服务	282
9.3.1 安全管理通则	282
9.3.2 人员的安全	282
9.3.3 物理的安全	283
9.3.4 安全审核	283
任务拓展	283

第四篇 数据灾备趋势

第 10 章 云计算应用与云灾备	287
10.1 云计算的应用	287
10.1.1 云物联	287
10.1.2 云安全	288
10.1.3 云存储	289
10.1.4 云游戏	289
10.1.5 云计算	289
10.1.6 私有云	289
10.1.7 云教育	290
10.1.8 云会议	290
10.1.9 云社交	290
10.2 云的三种服务模式和四种服务模型	290
10.2.1 云服务的模式	290
10.2.2 云服务的部署模型	291
10.3 云灾备介绍	293
10.3.1 云灾备的概念	293
10.3.2 云灾备服务详情	293
10.3.3 云灾备的服务类型	293
10.3.4 云灾备服务优势	294
10.3.5 基于云灾备的数据安全存储关键技术	294
10.4 Google 云计算原理介绍	295
10.4.1 Google 文件系统 GFS	295
10.4.2 系统架构	296
10.4.3 容错机制	298
10.4.4 系统管理技术	298
10.4.5 并行数据处理 MapReduce	299
10.4.6 实现机制	300
10.4.7 分布式锁服务 Chubby	303
10.4.8 通信协议	308
10.4.9 正确性与性能	310
10.4.10 分布式结构化数据表 Bigtable	311
任务拓展	319
第 11 章 大数据存储	320
11.1 大数据存储的概念	321
11.2 分布式存储系统介绍	324

11.2.1	分布式文件系统	324
11.2.2	典型架构	325
11.2.3	问题及解决方法	325
11.2.4	主控服务器	327
11.2.5	数据服务器	328
11.2.6	HDFS 介绍	329
11.2.7	分布式文件系统 HDFS 的特性	330
11.3	分布式数据库	331
11.3.1	分布式数据库结构	331
11.3.2	分布式并发控制技术	332
11.3.3	NoSQL 数据库介绍	334
11.3.4	HBase 介绍	337
11.4	关键技术分析	343
11.4.1	元数据管理	343
11.4.2	数据去重	344
11.4.3	数据分布和负载均衡	344
11.5	不同数据库公司的大数据主张	345
11.6	大数据时代的数据保护	347
11.6.1	HDFS	347
11.6.2	HBase	349
11.6.3	Zookeeper	351
11.6.4	OpenStack	352
	任务拓展	356
	参考文献	357



第 一 篇

数据存储概述

第 1 章 数据存储基础

第 2 章 存储应用环境

第 3 章 数据存储技术



第 1 章 数据存储基础



任务目标

- 了解数据存储的概念与发展历程；
- 了解数据存储的原理与物理结构；
- 了解数据存储的基本介质与技术；
- 了解数据灾难与容灾基础。



项目背景

存储技术作为信息技术的核心之一,一直伴随并推动着 IT 业各方面技术的协同发展,是当今 IT 领域中少数发展最为迅速的热点之一。纸的发明记载了人类的历史和文明,现代信息存储技术则大大超越了纸张记录的含义。如果说信息代表的是生存和生命、进取和发展,那么信息的价值是无可估量的,而存储作为信息的载体使信息的价值得到实现和增值,也就是说,存储的数据才是现代人类社会的真实财富所在。



项目描述

聚比特科技有限公司自成立以来一直从事互联网和电子商务等业务,但是随着信息化的发展,出现了很多瓶颈,越来越影响着企业的经营管理模式,这当中影响最大的便是企业的信息管理模式,随着应用互联网和电子商务业务应用的增长,企业信息数据呈爆炸性增长,一方面极大地促进了企业的发展,另一方面又对海量信息数据的存储和管理提出了新的挑战。如何有效地解决这些问题,成为企业及研究人员正在思考的问题。数据存储直接影响着行业的发展,也影响着企业自身的竞争力。



项目分析

如何高效、安全、可靠、完整地使用和保存宝贵的数据资料,又如何从这些浩瀚如海的信息中顺利地找到所需要的信息,成为聚比特科技有限公司的当务之急,只有使用高性能计算机网络存储系统,才能从根本上解决企业数据日益增长的需求。



项目实施

通过企业信息化专家分析形成一种自动化的存储管理手段,不仅可以解决现有企业关键数据的存储和管理需求,而且可以同时为网络上各种工作站用户提供数据备份的解决方案,减轻系统管理员的负担,有效地保护宝贵的数据及人力资源。并且当数据遇到灾难后,