

IT 先锋系列丛书

存储区域网络光纤通路技术

FIBRE CHANNEL FOR SANs

Alan F. Benner 著
胡先志 胡佳妮 等 译
艾 恕 巴继东 审校

**Mc
Graw
Hill** Education

人民邮电出版社
POSTS & TELECOMMUNICATIONS PRESS

IT 先锋系列丛书

存储区域网络光纤通路技术

Alan F. Benner 著

胡先志 胡佳妮 等译

艾 恕 巴继东 审校

人民邮电出版社

图书在版编目 (CIP) 数据

存储区域网络光纤通路技术 / (美) 本纳 (Benner, A. F.) 著; 胡先志等译.
—北京: 人民邮电出版社, 2003.4
(IT 先锋系列丛书)
ISBN 7-115-11050-6

I. 存… II. ①本… ②胡… III. 光纤通信 IV. TN929.11

中国版本图书馆 CIP 数据核字 (2003) 第 010736 号

IT 先锋系列丛书

存储区域网络光纤通路技术

- ◆ 著 Alan F. Benner
译 胡先志 胡佳妮 等
审校 艾 恕 巴继东
责任编辑 梁 凝
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
读者热线 010-67129258
北京汉魂图文设计有限公司制作
北京顺义振华印刷厂印刷
新华书店总店北京发行所经销
- ◆ 开本: 800×1000 1/16
印张: 15.75
字数: 332 千字 2003 年 4 月第 1 版
印数: 1-4 000 册 2003 年 4 月北京第 1 次印刷
著作权合同登记 图字: 01-2001-3979 号

ISBN 7-115-11050-6/TN · 2019

定价: 27.00 元

本书如有印装质量问题, 请与本社联系 电话: (010) 67129223

版 权 声 明

Alan F. Benner

Fibre Channel for SANs

ISBN: 0-07-137413-2

Copyright©2001 by the McGraw-Hill Companies, Inc.

Original language published by The McGraw-Hill Companies, Inc. All Rights reserved. No part of this publication may be reproduced or distributed in any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

Simplified Chinese translation edition jointly published by McGraw-Hill Education(Asia)Co. and Posts and Telecommunications Press.

本书中文简体字翻译版由人民邮电出版社和美国麦格劳-希尔教育(亚洲)出版公司合作出版。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有 McGraw-Hill 公司激光防伪标签,无标签者不得销售。

北京市版权局著作权合同登记 图字: 01-2001-3979 号

内 容 提 要

本书介绍了计算机通信领域中的最新技术之一的存储区域网络中的光纤通路。书中全面阐述了光纤通路特征、光纤通路协议、初始化和数据传送、FC-0 物理接口、有序集、注册和注销服务、服务的类型、检错与校对、流量控制、仲裁环、交换网络和光纤通路的未来展望等内容。本书的特点是内容新颖、介绍翔实、语言简炼、图表清晰。

本书可供从事计算机、通信等学科专业的师生阅读参考，也可供计算机通信网络建设与管理维护的工程技术人员学习参考，是一本介绍存储区域网络光纤通路技术的最新专著。

译 者 序

随着计算机通信技术的飞速发展，使得计算机网络应用越来越普及。为了帮助从事计算机通信工作的读者更快和更好地掌握网络技术，积极推动国内网络建设，受人民邮电出版社的委托，我们翻译了由美国麦格劳希尔出版公司 2001 年出版的这本介绍存储区域网络光纤通路的最新专著。

本书的目的是帮助读者学习和掌握计算机通信网络中的最新技术之一的存储区域网络光纤通路的特征、协议、应用及未来展望等有关内容。本书是一本介绍存储区域网络光纤通路基本理论、具体应用的专著。全书共分 19 章。书中对存储区域网络中的光纤通路特征、协议、初始化和数据传送、FC-0 物理接口、FC-1：8B/10B 转换机和接收机、有序集、FC-2 帧、FC-2：序列和交换、注册和注销服务、服务的类型、链路服务、检错与恢复、可选帧头和特殊功能、第 1 类连接管理、流量控制、仲裁环、交换网络、FC-4、SCSI、光纤通路上的 IP 和未来的工作等内容作了一一详细介绍。本书翻译分工为：胡先志译目录、第 1 章至第 5 章、附录；胡佳妮译第 6 章至第 8 章；熊斌译第 9 章至第 11 章；陆玉喜译第 12 章至第 14 章；熊伟芳译第 15 章、第 19 章；魏小萌译第 16 章；欧阳旭华译第 17 章至第 18 章。全书由巴继东，艾恕校对。

本书翻译工作始终得到人民邮电出版社的精心指导和热情帮助。在此，我们向为本书翻译出版付出辛勤劳动的所有人员表示诚挚的敬意和由衷的感谢。

由于光纤通路是计算机通信网络技术方面的最新技术之一，加之译者对该技术理解和翻译水平的局限，书中难免会出现一些谬误和不妥之处，敬请广大读者批评指正。

译者

2002 年 4 月 24 日于武汉南望山

前 言

在本书的第一版前言中，我简单表述了以一本书的形式来跟上信息潮流是很困难的，其首要原因是从书的开始撰写和实际出版的这段时间内，技术已发生了巨大的变化。其次，光纤通信、存储区域网络（SAN）和服务器组网如果说和以前有什么不同，那就是发展速度更快了。自本书第一版问世以来，至少通过了十多个新标准，而且另有一些标准或在研究之中，或已由更好的版本替代了。

幸运地是当今绝大多数的进步都集中在网络管理和数据管理层，这样，本书主要内容最好定位在网络的基本操作上。因此，我们期望本书的内容在今后一段相当的时间内仍将能够保持稳定。

自本书的第一版出版以来，光纤通路已在其得到使用的地方发生最根本的变化。光纤通路刚产生时，人们并不十分清晰它解决哪些问题最有效，因为是光纤通路是一种应用极为广泛的技术。其后，光纤通路已成为存储组网的事实上标准互连技术。实际上，术语“存储区域网络”也是自那以后新定义的。存储区域网络现在实际上已变成为与“光纤通路”的同义词。

自从光纤通路变得与存储器连网技术密不可分以来，光纤通路在解决存储器连网问题上已取得长足进展。不过本书中，我仍试图把重点始终放在整个光纤通路技术上，因为如果只讨论当今 SAN 的那些重要问题，就会忽略在未来极为重要的那部分技术风险。

本书有二个目的。首先，对光纤通路结构体系的概念，结构和目标作一详细的综述。专业读者在获得权威信息和查阅 ANSI 材料之前，应能通过本书理解光纤通路大部分技术细节。其次，也许更重要的是，本书力求证明：为什么网络设计要按照其自己的方式。组网技术继续变得越来越重要，且新技术也在不断地研发之中。随着新的网络化技术的发明，我希望本书将帮助读者对现有的网络发掘出许多有价值的东西。

关于表示方法这里也要简单说一下。在这类著作中确定一种统一的表示方法决非易事。因为本书内容跨越计算机和通信两个学科领域，但两学科惯用的表示方法稍有不同。例如，数据通信速率常用每秒兆比特度量，其中的“兆”意思是 10^6 ；而计算机数据则是以兆字节度量，其中的“兆”意思是 2^{20} 。本书则在某种程度上将两者结合起来，用“bit”来表示比特，如 Gbit/s，和用“Byte”来表示字节，如 MByte/s。本书的内容是面向通信的，所以前缀“兆（M）”和“吉（G）”的意思是 10^6 和 10^9 ，而不是 2^{20} 和 2^{30} 。本书中的所有数字都以二进制（‘01100101’），十六进制（x‘FF FFFD’）或十进制（65.532）形式书写。单个的比特则写成 1 或 0。

许多常用词，例如序列、交换和连接，在与它们的普通习惯用法完全不同的光纤通路中则具有特殊的含义。本书中具有特殊光纤通路意义的词通常都采用大字，以与它们的普通习惯用法相区别。这种大写通常与 ANSI 标准文件中所用的形式相一致。这里所提供的信息不存在版权问题，是通过普通的书、文章、ANSI 文件或其他参考材料得到的。

本书中所用的几个术语，例如 ATM 和 HIPPI，取自于其他的体系。书中用到的所有的商标都是其合法拥有的财产。以太网是施乐公司的商标。ESCON、FICON 和 SBCON 都是美国 IBM 公司在美国或其他国家（或同时在美国及国外）的商标或注册商标。InfiniBand 是无限区贸易协会的注册商标。

本书内容编排顺序如下。书中前几章给出光纤通路体系结构特点和目标的综述，同时也给出数据怎样在一个光纤通路网络中传送的例子。书中的中间几章涉及光纤通路的概念和结构的地方讲得比较详细，包括关于光纤通路所有的物理组成、逻辑结构、支持功能、流量控制和差错恢复等内容的章节。

书中的最后几章包括仲裁环的配置和操作，以及光纤通路结构映射到更上层协议，例如 SCSI 和 TCP/IP 的 IP 层的映射。这几章说明光纤通路是如何与现存的软件和操作系统层相适应。在书中的最后一章中，我斗胆地借此机会来对光纤通路、SAN 和服务器组网的未来作出某些预测。

我在写本书过程中，得到许多人士的大力帮助，但由于应感谢的人实在太多，在这里无法一一列举，不过我要尽力对他们一一致谢。对于遗漏未能对其致谢的人们，我只好事先向他们表示深深地歉意。非常感谢 Carl Zeitler、Kiwon Les、Mike Yang、Dan Eiscnhower、Ron Cash、Koger Weckly、Giles Erazier、Jerry Chapman、Jerry Rouse、Jonathan Thatcher、Bill George、Al Widmer、Tom McConathy、Casey Cannon、Gary Nutt、R.Bryan Cook、Paul Green、Dal Allan、Martin Sachs、Horst Truestedt、Richard Taborck、Schlto Van Doom and Roger Cummings，他们尤其对本书的第一版的写作给予许多帮助。感谢 Herman Presby、Lvan Kaminow 和 Jon Sauer 把我带入光网络世界，感谢 Frank Kampf、Bob Stucke、Harish Sethu、Doug Joseph 和 Bob Cypher 使我了解许多计算机通信方面的问题。同时也感谢 Steve Chapman 和 Caroline Lwvine 帮助我将本书各章组合为一体并完成本次写作计划，再次感谢 Steve 帮助我完成了第二版的工作。

我特别要感谢 Joe Mathis，他指导我如何成功地完成这本著作提供了许多指导。感谢 Renato Recio，他经常以示范的方式指导我写作。我也要感谢 InfiniBand 的同仁，你们明知道自己的身份，给了我与非凡人们一道工作的机会，我们有信心在今后做更多的事情。

我希望这本书能够清楚地表明，本人对构想和解释怎样建设更好的通信系统是非常关心。我期望本书将对广大读者有所帮助。

仅以本书献给 Micah，当我要做更重要的事情时，他代我花费了查找资料的时间，本书也献给 Ashley 和 Denali，他们正忙创造着自己的未来，即看看事情怎样发生将是令人开心的。

目 录

第 1 章 光纤通路和存储区域网络	1
1.1 光纤通路的特点	1
1.2 存储区域网络	3
1.3 SAN、LAN 和网络连接存储 (NAS)	5
1.3.1 LAN 与 SAN: 网络结构的差别	6
1.3.2 LAN 与 SAN: 协议结构的差别	6
1.3.3 网络连接存储 (NAS) 和存储区域网络 (SAN)	7
1.4 本书的目的	8
第 2 章 概述	9
2.1 FC-0 简介	11
2.2 FC-1 简介	12
2.3 FC-2 简介	13
2.3.1 物理模型: 节点、端口和拓扑	14
2.3.2 带宽和通信开销	15
2.3.3 基本组成部件及其层次结构	16
2.3.4 链路控制帧	19
2.3.5 一般交换结构模型	19
2.3.6 服务类别	21
2.3.7 基本的和扩展的链路服务命令	23
2.3.8 仲裁环功能	24
2.3.9 协议	24
2.3.10 分段和重组	25
2.3.11 数据压缩	25
2.3.12 检错和恢复	26
2.4 FC-3 简介	26
2.5 FC-4 简介	27
2.5.1 光纤通路上的 IP	27
2.5.2 光纤通路上的 SCSI	28
2.5.3 光纤通路上的 FICON/ESCON	28

第 3 章 初始化和数据传送	29
3.1 链路初始化	29
3.2 环初始化	31
3.3 交换结构与 N 端口注册	33
3.4 起始序列传送	34
3.5 交换完成	37
3.6 本例未涉及的协议	37
3.6.1 检错与恢复	38
3.6.2 第 1 类服务和交换结构的工作	38
3.6.3 其他协议工作	39
第 4 章 FC-0: 物理接口	40
4.1 一般特征	40
4.2 发射机和接收机	40
4.3 失效码的有意传送	41
4.4 FC-0 术语和技术方案	41
4.5 长波长激光器与单模光纤链路	43
4.5.1 光缆线路	44
4.5.2 发射机	44
4.5.3 接收机	45
4.6 短波长激光器与多模光纤链路	45
4.6.1 光缆线路	46
4.6.2 发射机	47
4.6.3 接收机	47
4.7 短波长激光器链路的开放式光纤控制安全系统	47
4.8 LED 多模光纤链路	48
4.9 75Ω 同轴电缆链路	48
4.10 150Ω 屏蔽双绞线电缆链路	50
第 5 章 FC-1: 8B/10B, 发射机和接收机	51
5.1 8B/10B 传送码概述	51
5.2 字符编码和解码	53
5.3 传送、接收和检错	57
5.4 发射机和接收机状态	59

第 6 章 有序集	61
6.1 有序集	62
6.1.1 帧起始定界符	62
6.1.2 帧终止定界符	62
6.1.3 原语信号——Idle 和 R_RDY	65
6.1.4 原语序列	65
6.2 非环状拓扑的端口状态	67
6.2.1 激活状态	67
6.2.2 链路恢复状态 (LR1、LR2 和 LR3)	68
6.2.3 链路失效状态 (LF1 与 LF2)	69
6.2.4 离线状态 (OL1、OL2 与 OL3)	70
6.3 原语序列协议	71
6.3.1 链路恢复协议	71
6.3.2 在线到离线协议	72
第 7 章 FC-2: 帧	73
7.1 帧格式	74
7.1.1 Idle 和 R_RDY 原语信号	74
7.1.2 帧起始定界符	74
7.1.3 帧头字段	74
7.1.4 可选帧头	74
7.1.5 净荷	75
7.1.6 循环冗余校验字段	75
7.1.7 帧终止定界符	75
7.2 帧头字段	75
7.2.1 路由控制字段: R_CTL	76
7.2.2 地址标识符: S_ID 和 D_ID	79
7.2.3 等级特别控制: CS_CTL	81
7.2.4 数据结构类型: TYPE	81
7.2.5 帧控制: F_CTL	82
7.2.6 序列 ID: SEQ_ID	87
7.2.7 数据字段控制: DF_CTL	87
7.2.8 序列计数: SEQ_CNT	88
7.2.9 发送端交换标识符: OX_ID	88
7.2.10 响应端交换标识符: RX_ID	88

7.2.11 参数	89
7.3 数据帧和链路控制帧	89
7.4 链路可靠性复原帧 (LCR)	91
7.5 对数据和链路控制帧的详细响应	92
7.5.1 R_RDY 响应	92
7.5.2 ACK 帧	92
7.5.3 F_BSY 帧	94
7.5.4 P_BSY 帧	95
7.5.5 F_RJT 和 P_RJT 帧	96
7.5.6 通报 (NTY) 和终止 (END) 帧	96
第 8 章 FC-2: 序列和交换	98
8.1 序列管理	98
8.1.1 序列初始化	99
8.1.2 序列处理有效性	99
8.1.3 序列计数管理	100
8.1.4 正常序列完成	100
8.2 交换管理	102
8.2.1 交换发起和 X_ID 的分配	102
8.2.2 交换管理	103
8.2.3 交换终止	103
8.3 序列和交换状态块	104
8.3.1 序列状态块规则	104
8.3.2 交换状态块规则	104
第 9 章 注册和注销服务	107
9.1 注册和注销概述	107
9.2 交换结构注册	110
9.3 N 端口注册	111
9.4 注销	112
9.5 N 端口和 F 端口服务参数	113
9.5.1 公用服务参数	113
9.5.2 N 端口名字和 F 端口名字	115
9.5.3 节点名字和交换结构名字	116
9.5.4 N 端口和 F 端口服务类型参数	116

9.5.5 提供商版本级别	119
9.6 进程注册和注销概述	119
9.7 估算端到端可靠性的过程	120
第 10 章 服务类型	124
10.1 第 1 类服务	126
10.2 第 2 类服务	128
10.3 第 3 类服务	129
10.4 混合服务	130
10.5 第 4 类服务	131
10.6 第 6 类服务	132
第 11 章 链路服务	133
11.1 基本链路服务命令	133
11.1.1 不可操作 (NOP)	134
11.1.2 拆除连接 (RMC)	134
11.1.3 基本接受 (BA_ACC)	134
11.1.4 基本拒绝 (BA_RJT)	135
11.1.5 中止序列 (ABTS) 命令	136
11.2 扩展链路服务命令概述	136
11.3 扩展链路服务应答	138
11.3.1 链路服务拒绝 (LS_RJT)	138
11.3.2 接受 (ACC)	138
11.4 ELS 请求: 注册、注销和中止交换	139
11.4.1 N 端口注册 (PLOGI)	139
11.4.2 F 端口注册 (FLOGI)	141
11.4.3 注销 (LOGO)	141
11.4.4 中止交换 (ABTX)	141
11.5 ELS 请求: 状态判断和发起请求	141
11.5.1 读连接状态 (RCS)	141
11.5.2 读交换状态块 (RES)	143
11.5.3 读序列状态块 (RSS)	143
11.5.4 请求序列发起 (RSI)	143
11.6 ELS 请求: 可靠性确定	143
11.6.1 建立数据流 (ESTS)	143

11.6.2	估计可靠性 (ESTC)	143
11.6.3	建议可靠性 (ADVC)	144
11.7	ELS 请求: 其他功能	144
11.7.1	读超时值 (RTV)	144
11.7.2	读链路差错状态块 (RLS)	144
11.7.3	响应 (ECHO)	144
11.7.4	测试 (TEST)	144
11.7.5	重设恢复条件 (RRQ)	145
11.8	ELS 请求别名标识符	146
11.8.1	获得别名_ID (GAID)	146
11.8.2	交换结构激活和去激活别名标识符 (FACT, FDACT)	146
11.8.3	N 端口激活和去激活别名标识符 (NACT, NDACT)	146
11.9	第 4 类服务 ELS 请求	146
11.9.1	服务质量请求 (QoSR)	146
11.9.2	读虚电路状态 (RVCS)	147
11.10	FC-4 链路服务	147
第 12 章	检错与恢复	148
12.1	概述	148
12.2	超时时限	149
12.2.1	R_T_TOV	149
12.2.2	E_D_TOV	149
12.2.3	R_A_TOV	149
12.2.4	CR_TOV	150
12.2.5	FC_AL 环超时	150
12.3	超时的作用	150
12.4	链路差错检验与恢复	152
12.5	帧接收与帧有效性	153
12.6	交换和序列的完整性	154
12.7	第 1 类和第 2 类服务的交换差错策略	155
12.7.1	所有丢弃策略的通用规则	155
12.7.2	中止、丢弃多序列差错策略	156
12.7.3	立刻重传的多序列丢弃	156
12.7.4	中止、丢弃单个序列差错策略	157
12.7.5	无限缓冲的差错策略	157

12.8	第3类序列差错处理	158
12.9	序列恢复	158
12.9.1	差错恢复策略	158
12.9.2	序列恢复用中止序列协议	159
12.9.3	第1类恢复的序列重传	161
12.10	停止序列协议简介	161
12.11	链路差错状态块规则	162
第13章	可选帧头和特殊功能	164
13.1	可选帧头	164
13.1.1	网络帧头	165
13.1.2	关联帧头	165
13.1.3	设备帧头	167
13.2	特殊功能	167
13.2.1	数据压缩	167
13.2.2	加密	167
第14章	第1类连接管理	168
14.1	简介	168
14.2	综述：正常过程	168
14.2.1	建立一个连接	169
14.2.2	连接过程	169
14.2.3	拆除一个连接	169
14.3	专线连接恢复	170
14.4	各种拓扑的专线连接	170
14.4.1	点到点和仲裁环拓扑	170
14.4.2	交换结构拓扑	171
14.5	第1类专线连接：详细操作	172
14.5.1	连接建立	172
14.5.2	单向与双向连接	174
14.5.3	入栈连接请求	174
14.5.4	拆除一个连接	175
14.5.5	呼叫延迟	175
14.5.6	抢先占有	176

第 15 章 流量控制	177
15.1 总的策略	177
15.2 光纤通路流量控制策略	178
15.3 端到端流量控制	179
15.3.1 用于端到端可靠性分配的接收缓冲区分配	180
15.3.2 影响端到端可靠性计算值的事件	180
15.3.3 序列接收端任务	181
15.4 缓冲区到缓冲区的流量控制	182
15.5 第 2 类服务的综合流量控制	184
第 16 章 仲裁环	185
16.1 概述	185
16.1.1 FC_AL 操作所需的额外功能	186
16.1.2 仲裁环集线器	187
16.2 AL_PA 物理地址	188
16.3 AL 原语信号和序列	189
16.3.1 原语信号	189
16.3.2 原语序列	191
16.4 环初始化	192
16.4.1 选择一个临时环主机	192
16.4.2 AL_PA 地址的选择	192
16.4.3 构建一个 AL_PA 地址映射	195
16.5 环端口状态机运行：一个示例	195
16.5.1 两端口环的初始化	196
16.5.2 通信的仲裁和初始化	197
16.5.3 仲裁和访问公平性	198
16.5.4 结束通信	198
16.5.5 交替缓冲区到缓冲区流量控制管理	198
16.5.6 连续通信	199
16.5.7 传送到另一个端口	199
16.6 性能和定时	200
16.7 仲裁环的优点和缺点	201
第 17 章 交换结构	203
17.1 网络特征	204

17.1.1	与其他网络的比较	204
17.1.2	交换机端口的类型	205
17.1.3	交换机和寻址体系	205
17.1.4	级别从属关系	206
17.2	F类业务	207
17.3	特殊交换功能	208
17.3.1	别名	208
17.3.2	组播	208
17.3.3	搜索组	209
17.3.4	循环组	209
17.4	众所周知的一般服务	209
17.4.1	广播——x'FF FFFF'	210
17.4.2	F端口——x'FF FFFE'	210
17.4.3	网络控制器——x'FF FFFD'	210
17.4.4	名字服务——x'FF FFFC'	210
17.4.5	时间服务——x'FF FFFB'	211
17.4.6	管理服务器——x'FFFFFA'	211
17.4.7	服务质量管理者——x'FFFF9'	211
17.4.8	别名服务——x'FF FFF8'	211
17.4.9	安全匙分配服务——x'FF FFF7'	211
17.4.10	时钟同步服务器——x'FF FFF6'	212
17.4.11	组播服务器——x'FF FFF5'	213
第 18 章	FC-4: SCSI 和光纤通路的 IP	214
18.1	光纤通路上的 FC-4 控制操作	214
18.2	光纤通路上的 IP	215
18.2.1	IP 和 ARP 数据报的传送	215
18.2.2	IP 操作	216
18.3	光纤通路上的 SCSI-3: FCP	219
18.3.1	光纤通路协议 (FCP) 信息单元格式	220
18.3.2	FCP 下的 I/O 工作举例	223
第 19 章	未来的工作: FC-PH-?和未来	226
19.1	光纤通路未来	227
19.2	未来“到处都是以太网”	227