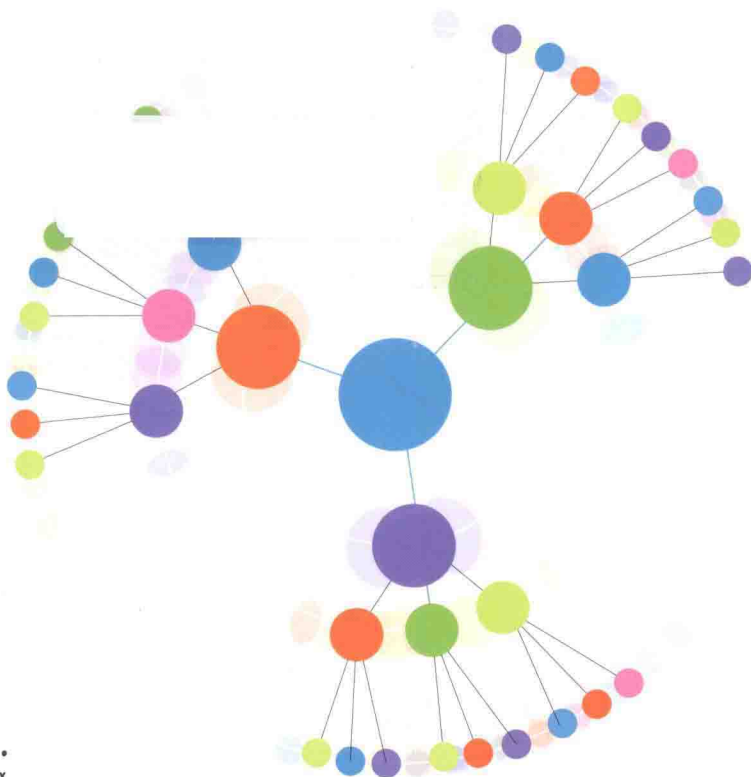


CDN技术详解

Dissecting Content Delivery Network

雷葆华 孙颖 王峰 陈晓益 等著

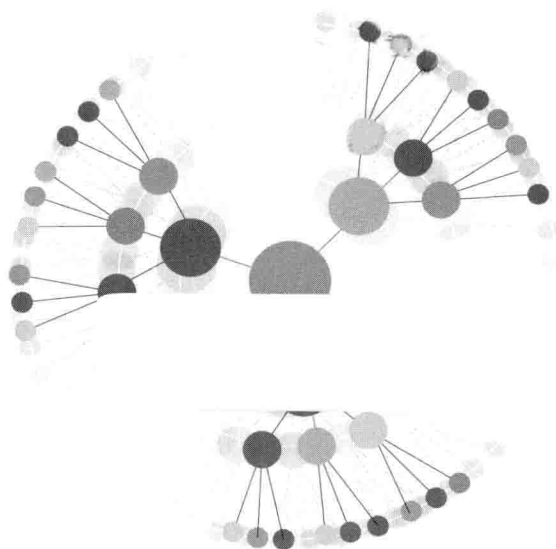


电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

转型时代丛书
中国电信北京研究院专家奉献

CDN技术详解

雷葆华 孙颖 王峰 陈晓益 蔡永顺 王志军 著



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书内容包括CDN技术的发展历程、关键技术、商业化服务现状,以及对未来的发展展望,对构成CDN系统的关键功能模块GSLB、SLB、Cache进行了重点讲解,除技术原理之外,还对实现这些功能模块所涉及的一些协议和开发工具进行了讲解,希望能帮助读者了解CDN这项技术,并对CDN系统的设计和开发有一些初步的体会。

本书适合从事互联网开发和运营工作的专业人士、电信运营服务从业人员,以及相关专业的高校学生。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

CDN 技术详解 / 雷葆华等著. —北京: 电子工业出版社, 2014.11

(转型时代丛书)

ISBN 978-7-121-24536-7

I. ①C... II. ①雷... III. ①计算机网络—网络结构—研究 IV. ①TP393.02

中国版本图书馆 CIP 数据核字 (2014) 第 239848 号

策划编辑: 刘 皎

责任编辑: 刘 舫

印 刷: 北京天来印务有限公司

装 订: 北京天来印务有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱

邮编: 100036

开 本: 720×1000 1/16

印张: 25.25 字数: 424 千字

版 次: 2014 年 11 月第 1 版

印 次: 2014 年 11 月第 1 次印刷

定 价: 69.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

转型时代丛书

指导委员会

主任委员：吴基传

副主任委员：杨杰

委 员：陈俊亮 李未 韦乐平

邬贺铨 张继平（按拼音顺序排序）

编委会

主 任：李志刚

副主任：侯春雨 赵慧玲

委 员：毕奇 朱健 野永东 谢朝阳

陈自清 杨峰义 王晓平 张成良

转型时代丛书

总序

“变化，无论是突如其来的，还是循序渐进的，有时都会淘汰你认为理所当然的一切”。
——《转型》

二十一世纪以来，信息化更加快速而深刻地改变着这个世界，大到全球经济社会的发展格局，小到每个人的日常工作生活。许多国家把数字化、信息化、智能化作为国家战略的关键主题，把信息基础设施建设作为后金融危机时代振兴经济的重要手段。同样，我国“十二五”规划也把全面提高信息化水平，特别是加快建设下一代国家信息基础设施、推动信息化和工业化深度融合、推进经济社会各领域信息化作为重要工作列入其中。

信息通信产业中新技术、新业务的不断快速发展不仅催生着新的经济增长点，造就了谷歌、Twitter、腾讯等一个又一个明星企业，引领整个行业及社会经济的发展方向，更重要的是它对人们生产、生活产生了深刻而久远的影响。我们的生产资料不仅仅是机器，还有电脑、手机、互联网；我们通过点击“百度”打开未知世界，通过“淘宝”购买商品，利用手机登录“Facebook”去了解彼此、评论时政，所有这一切都表明信息通信产业正在更广、更深地影响着我们每一个人，互联网/移动互联网已成为像水、电一样的生产、生活“必需品”。

环顾全球，整个信息通信产业正在朝着宽带化、移动化、智能化发展，特别是3G的普及和LTE的逐步成熟使得移动互联网一跃成为整个行业中最前沿、最具革命性的领域。智能管道、物联网、下一代互联网、云计算等一个个新的理念、新的信息服务模式正在席卷全球成为新热点。而这一切变化，都将对从事信息服务的企业，包括电信运营商，带来前所未有的机遇和挑战。适者生存法则同样适用于多变的企业生态系统。无论是百年老店，还是创业新秀，只有顺应信息化时代发展潮流，重新审视并及时调整企业的商业模式，抓住信息化带来的重大机遇，才能在变化中顺势前进。

鉴于此，这套“转型时代丛书”既有对智能宽带网络、移动互联网、云计算等新技术、新网络的研究和实践总结，也有对商业模式、营销变革等现代管理中关键问题的长期探索。相信此系列书籍能帮助您了解趋势，廓清谜团，抓住机会，与信息化时代共同成长。



2012年1月

推荐序

赵慧玲

中国电信北京研究院副院长

互联网已经成为今天的生活必需品，人们利用它获取信息和资讯、进行娱乐并相互沟通，在这背后有一个鲜为人知的幕后英雄——内容分发网络（CDN）。CDN 架设在 IP 网络之上，是互联网世界里的智能传送网络，连接了我们和互联网上无数色彩缤纷的应用。

CDN 的出现改变了互联网的生态面貌，促进了互联网的发展。在互联网商业化不久，由于缺乏必要的流量管理和疏通手段，骨干带宽被迅速消耗掉，IP 网络流量秩序濒于失衡。为解决这一难题，麻省理工学院的一批顶级科学家提出了 CDN 解决方案，从此以后，CDN 就成为互联网的基础设施，伴随着互联网的潮起潮落而发展。CDN 将内容推到网络的边缘，为用户提供就近性的边缘服务，从而保证服务的质量和整个网络上的访问秩序，解决了困扰内容提供商的内容“集中与分散”的两难选择，极大缓解了内容传送瓶颈。CDN 的发展促使整个互联网产业进一步分工合作，使网站和应用开发者们能够更关注于上层应用。

时间进入 2009 年，云计算和移动互联网时代到来，CDN 在这一背景下重新引起了各方的关注，蓝汛和网宿公司的成功上市，标志着资本市场对 CDN 的认可。我们看到，在国外很多公司的云计算服务中，CDN 成为一项标准配置，

而移动互联网中应用的多样性更需要 CDN 的分发能力。传统的 IP 网络以传输中立、协议简单为信条，而互联网，特别是在移动互联网时代，应用的丰富性要求网络能力能够随着应用不断拓展。中国电信正致力于打造“智能管道”，而 CDN 技术是打造智能管道的一项关键技术，是电信运营商未来服务运营的重要竞争力。CDN 与基础 IP 网络联系紧密，通过与各级网络之间的配合调度，在给用户提供优质服务的同时，也能降低骨干网的传输压力和峰谷差异。同时，基于 CDN 的海量信息感知能力，还能为业务开发者提供各种用户行为分析和预测数据，创造了新的商业模式。

中国电信从 2002 年开始进行 CDN 网络的建设，用于承载自营业务和对外服务。在这一过程中，我们经历了很多挫折，也积累了宝贵的经验。回顾这些年的工作，深感这一领域缺乏系统性的书籍和技术资料，使入门者的摸索过程显得更为困难。为帮助普及 CDN 的服务和应用概念，推动业界的技术沟通与交流，中国电信北京研究院组织从事 CDN 技术研究及相关专业技术的联合团队编写了本书，其中既有长期工作在 CDN 网络建设和服务一线，对现网运营支撑有丰富经验的工程师，又有互联网战略运营和业务的专家，还有部分从事云计算等信息技术领域研究的领军人才，这本书既是他们研究成果的智慧结晶，也是他们分享知识和经验的平台。相信他们的研究和心得会给广大读者带来思考、启迪和帮助。



2012 年 4 月，北京

序言

雷葆华

中国电信北京研究院云计算团队总监

CDN 是互联网服务背后重要的支撑者,它伴随着互联网商业化的出现而出现,一直可以看做互联网的卖水人。国内 CDN 产业的出现和国外基本在同一时间,但与国外不同,国内 CDN 发展的第一次浪潮从 2002 年左右开始,当时正值 ADSL 宽带业务大发展(从 64kb/s 的拨号、128kb/s 的 ISDN 升级到 512kb/s 的 ADSL),各省和地市级的运营商为填充提速后的宽带,增强宽带的吸引力,纷纷开始提供流媒体服务,CDN 作为流媒体系统的一个重要核心组成部分开始被运营商接受和部署,并逐步在国内市场普及。通过 CDN 的支撑,运营商们为用户提供了有服务质量保证的流媒体服务,有力地推动了宽带的发展,成为国内 CDN 发展第一个高潮的最主要的推动者。可以说,CDN 为中国第一次带宽升级起到了重要的推动和支撑作用。

从 2008 年开始,3G 牌照的发放和新一轮宽带提速的推进,为移动互联网和云计算时代的到来奠定了基础条件。在这一新的发展阶段,CDN 作为互联网服务的重要技术保证方式和一种基础资源服务,再一次受到了整个业界的关注,ChinaCache 和网宿公司的成功上市,也标志着资本市场对 CDN 的认可。但是,客观地说,从 CDN 服务的普及率、服务的附加值和产品的核心竞争力几方面看,我国的 CDN 产业与欧美和日韩等国相比仍有不小的差距,整个行业还属

于粗放型的发展阶段。其中一个比较重要的原因是国内这方面的人才和技术积累比较少,相关的书籍、资料很少,特别是能够系统性地介绍 CDN 技术的书籍在国内仍是一个空白。为此,我组织我们的研发团队利用业余时间编写了这本书,希望能够帮助读者系统地了解 CDN 的基本原理和基础知识,掌握其中的一些关键技术,并对整个 CDN 产业的发展有一个比较全面的认识,通过书籍为推动这个产业的发展尽一份绵薄之力。

为更好地让读者理解,我们采用循序渐进的方式介绍相关知识。

第 1 章是引言,通过对 CDN 的基本概念、发展历史、作用 and 价值的介绍,让读者了解为什么需要 CDN,它能做什么,对 CDN 形成一个初步的认识。

第 2 章是 CDN 技术概述,对 CDN 的系统架构进行整体阐述,同时通过对 CDN 不同服务类型的分类描述,让读者对 CDN 的技术框架有一个比较全面的了解。

第 3 章至第 5 章,从对静态网页加速技术的实现入手,分别对 Web 缓存、集群和负载均衡、全局负载均衡和内容路由技术这三大关键技术的技术原理、使用方法和实现实例进行介绍,使读者能够了解 CDN 技术的发展脉络,同时对 CDN 系统的部署从单台 Cache 缓存技术到集群的实现再到整个 CDN 系统的分布扩展有一个全面的认识。

第 6 章重点介绍流媒体 CDN 系统的实现,通过对比流媒体加速和网页加速的区别,阐述流媒体 CDN 的关键技术,在该章的最后以 IPTV CDN 服务系统为例,帮助读者理解相关内容。

第 7 章分别对动态网页加速技术和应用交付技术进行介绍。它们分别是 Web 2.0 网站加速和企业内部应用加速的关键技术。

第 8 章从商业应用价值角度对 CDN 的产业现状和商业价值进行分析。

第9章对CDN和云计算、P2P等其他相关技术的关系进行阐述，同时作为总结，对CDN的下一步发展进行了分析和展望，提出进一步的发展方向。

本书部分内容和案例来自于我们的实践经验和成果，同时本书的编写得到了中国电信集团北京研究院的支持，参考了大量业界的研究成果和相关技术材料，在此一并感谢。最后，由于成稿仓促，知识水平和认识的局限，书中难免有纰漏之处，请各位专家和读者不吝赐教。

雷葆华

2012年4月，北京

致谢

我是在 2003 年开始接触和了解 CDN 的，后来长期负责中国电信 CDN 骨干网的技术方案制订、工程实施和网络规划等工作。回头看我对 CDN 的学习和认识过程，深感当时国内在这方面资料和信息之匮乏，我自己是通过两年多的研究、测试、工程设计和实施工作，在经过长期的实践摸索和经验总结之后，才开始对 CDN 技术有一个比较系统的理解和认识，通过进一步的研究和工作积累，才逐步形成了一套 CDN 技术体系和研究方法，并将 CDN 与其他领域的研究结合起来，取得了一些成绩。在这一过程中得到了很多人的帮助和支持，在此表示真诚的感谢。

首先，要感谢中国电信集团北京研究院的各级领导，没有院领导和部门的长期支持，就不会有系统地研究 CDN 技术的机会，也不会有今天的这些成绩。

其次，要感谢中国电信集团公司的相关领导对我们的理解和信任，对我们研究工作的大力支持和指导，使得我们的技术和业务研究成果能够在现网中得到检验和应用，积累了宝贵的经验。

另外，还要感谢业界的专家和合作伙伴的技术人员，他们无私地奉献了专业知识和经验，特别是在 CDN 技术研究初期，资料匮乏，通过与他们的沟

通、交流，获得了很多知识。

同时，还要感谢电子工业出版社的领导、刘皎编辑以及为本书付出辛劳的出版社的朋友。

最后，要感谢我们的家人，本书的几位作者日常的科研工作和任务都很繁重，主要利用业余时间完成本书，没有家人的理解和支持，本书很难顺利付印。

目录

第1章 引言	1
1.1 CDN 的基本概念和产生背景	2
1.2 CDN 的基本工作过程	5
1.3 CDN 的发展历史	8
1.4 CDN 对互联网产业的价值和作用	13
第2章 CDN 技术概述	16
2.1 CDN 的系统架构	17
2.1.1 功能架构	17
2.1.2 部署架构	21
2.2 CDN 系统分类	23
2.2.1 基于不同内容承载类型的分类	24
2.2.2 基于内容生成机制的分类和分层加速服务	27
2.3 小结	29
第3章 内容缓存工作原理及实现	32
3.1 内容缓存技术的发展背景	33
3.1.1 网站的问题和需求	33
3.1.2 CDN 出现前的网站服务技术	35

3.2	Cache 设备的工作方式和设计要求.....	38
3.2.1	正向代理	39
3.2.2	反向代理	41
3.2.3	透明代理	42
3.2.4	Web Cache 产品实现关键要素分析	44
3.3	Web Cache 的实现基础 ——基于 HTTP 协议的 Web 缓存技术	45
3.3.1	Web 与 HTTP.....	45
3.3.2	HTTP 协议工作原理	48
3.3.3	HTTP 中的 Cookie 和 Session.....	68
3.3.4	HTTPS 安全协议.....	74
3.3.5	HTTP 协议中的缓存技术	76
3.4	Web Cache 技术实现关键点分析	82
3.4.1	Web Cache 关键性能指标说明	82
3.4.2	内容存储机制.....	85
3.4.3	内容更新机制.....	86
3.4.4	Web Cache 协议优化.....	90
3.4.5	Web Cache 安全实现机制	92
3.5	开源 Web 缓存代理软件——Squid	94

第 4 章 集群服务与负载均衡技术 97

4.1	服务器集群技术	98
4.1.1	集群的基本概念.....	98
4.1.2	集群的分类	99
4.1.3	集群的系统结构.....	101
4.1.4	CDN 负载均衡集群.....	102
4.2	Cache 集群协同交互方法	103
4.2.1	ICP.....	104
4.2.2	HTCP	105
4.2.3	Cache Digest.....	106

4.2.4	Cache Pre-filling	106
4.2.5	CARP	107
4.3	负载均衡技术的实现	108
4.3.1	负载均衡关键技术	110
4.3.2	负载均衡部署方式	115
4.3.3	服务器负载均衡	118
4.3.4	链路负载均衡	125
4.4	开源负载均衡软件	130
4.4.1	LVS	130
4.4.2	Nginx	132
 第 5 章 全局负载均衡工作原理及实现.....		134
5.1	全局负载均衡在 CDN 系统中的作用	135
5.2	基于 DNS 解析的 GSLB 实现机制	136
5.2.1	DNS 的产生背景	136
5.2.2	DNS 基本工作原理	137
5.2.3	基于 DNS 解析的 GSLB 工作方式	147
5.2.4	负载均衡的策略判断条件信息	150
5.2.5	开源 DNS 服务软件——BIND	153
5.3	基于 DNS 的 GSLB 应用部署方法	155
5.3.1	GSLB 应用部署时的一些基本概念	155
5.3.2	负载均衡策略	160
5.3.3	GSLB 部署中的关键问题	171
5.4	基于应用层协议重定向的 GSLB	177
5.4.1	HTTP 重定向基本原理	177
5.4.2	基于 HTTP 重定向的 GSLB 工作流程	179
5.5	基于 IP 路由的 GSLB	181
5.6	小结	184

第 6 章 流媒体 CDN 系统的组成和关键技术188

6.1	流媒体系统工作原理概述	191
6.2	流媒体传送协议体系	194
6.2.1	RTP 和 RTCP	196
6.2.2	RTSP	200
6.2.3	RTMP	206
6.2.4	HTTP Streaming	215
6.2.5	MPEG-2 TS	225
6.3	流媒体业务对 CDN 提出的要求和挑战	227
6.3.1	流媒体加速与 Web 加速之间的业务差异	227
6.3.2	流媒体 CDN 系统架构描述	229
6.3.3	小结	231
6.4	流媒体 CDN 系统的关键技术实现	232
6.4.1	Cache 的设计实现	232
6.4.2	负载均衡系统设计实现	237
6.4.3	内容分发机制设计实现	239
6.4.4	组网模式	241
6.4.5	内容文件预处理技术	242
6.4.6	防盗链机制和实现	245

第 7 章 动态内容加速服务的实现249

7.1	动态内容加速技术	250
7.1.1	业务逻辑层加速技术：边缘计算	254
7.1.2	数据访问层加速技术：数据库复制	256
7.1.3	用户数据层加速技术：用户数据复制	260
7.2	应用加速技术	262
7.2.1	应用加速技术概述	262
7.2.2	广域网加速技术	263
7.2.3	SSL 加速技术介绍	273