



深入理解 Cisco 路由器内部操作

CCIE PROFESSIONAL DEVELOPMENT

CISCO IOS[®] 精髓

Vijay Bollapragada, Curtis Murphy, Russ White 著

邓劲生 白建军 齐宁 译

CISCO SYSTEMS

CISCO PRESS



中国电力出版社
www.cepp.com.cn

深入理解 Cisco 路由器内部操作

CCIE PROFESSIONAL DEVELOPMENT

CISCO IOS[®] 精髓

Vijay Bollapragada, Curtis Murphy, Russ White 著

邓劲生 白建军 齐宁 译

中国电力出版社

内 容 提 要

本书介绍了 IOS 及其操作的一些关键部分。主要分为三部分，第一部分讲述了 IOS 的基本体系结构，包括它的软件底层结构和交换体系结构。第二部分从第 3 章开始，讲述了在一些选定的 Cisco 硬件平台上应用 IOS。最后，在第 8 章描述了 IOS 如何实施其 QoS 机制。

本书适用于不同程度的网络设计者、实施者以及管理者阅读，也可供 Cisco 认证考试及高等院校网络专业师生参考。

图书在版编目 (CIP) 数据

Cisco IOS 精髓/ (美) 波拉普拉加达等著; 邓劲生, 白建军, 齐宁译.-北京: 中国电力出版社, 2000

ISBN 7-5083-0476-4

I .C… II.①波…②邓…③白…④齐… III.计算机网络-通信协议 IV.TP393

中国版本图书馆 CIP 数据核字 (2000) 第 75028 号

北京版权局著作权登记号 图字 01-2000-3953

本书英文版原名: Inside Cisco IOS Software Architecture

Authorized translation from the English language edition published by CISCO Press, an imprint of Prentice Hall Technical Reference.

Copyright © 2000.

All rights reserved including the right of reproduction in whole or in part in any form.

本书中文版由美国培生集团授权出版, 版权所有。

中国电力出版社出版、发行

(北京三里河路 6 号 100044 <http://www.infopower.com.cn>)

三河市实验小学印刷厂印刷

各地新华书店经售

*

2001 年 1 月第一版 2001 年 1 月北京第一次印刷

787 毫米×1092 毫米 16 开本 13.25 印张 303 千字

定价 29.00 元

版 权 所 有 翻 印 必 究

(本书如有印装质量问题, 我社发行部负责退换)

译 者 序

网络作为知识经济支柱产业——信息产业的主角，在国民经济发展过程中起着非常重要的作用。充分了解各种网络技术，也成为网络工程技术人员的迫切需求。

Cisco 公司作为全球最大的网络和电信设备提供商，为网络技术的发展做出了许多不可磨灭的贡献。其产品占全球 Internet 骨干路由器和交换机市场 65%~82% 的市场份额。而 IOS 技术解决方案，满足了广大客户的端到端的网络规划、设计及实施的需求。

本书集中于讲解 Cisco IOS 技术的一个领域，也是至关重要的一个领域——Cisco IOS 软件体系结构。对 IOS 的结构组成、报文交换，特别是几种特殊路由器平台，存储共享路由器、早期的 CBUS 路由器、基于存储片的系统以及 Cisco 7500、Cisco 12000 等的报文交换技术做了详细、清晰、高度精确的描述。最后还给出了 IOS 服务质量的概述及几种 QoS 方法示例。

在书中介绍了 IOS 软件体系结构及其操作的一些关键部分。主要分为三部分，第一部分讲述了 IOS 的基本体系结构，包括它的软件底层结构和交换体系结构。第二部分从第 3 章开始，讲述了在一些选定的 Cisco 硬件平台上应用 IOS。最后，在第 8 章描述了 IOS 如何实施其 QoS 机制。

本书内容翔实全面，广泛适用于不同程度的 Cisco 网络管理人员，可供网络设计者、实施者以及管理者阅读，也可供 Cisco 认证考试及大专院校网络专业师生参考。

本书由灯芯工作室组织翻译，由邓劲生、白建军、齐宁等完成主要章节的翻译，其他工作人员完成全书的录入、校对和审稿等工作，孙兆林统校全书。由于书中涉及的知识和技术范围较广，尽管我们为此付出了许多努力，但限于水平，书中难免还存在一些翻译不当之处，恳请广大读者批评指正。

译者
2000年9月

关于作者

Vijay Bollapragada, CCIE#1606, 现在是 Cisco 方案工程组的经理。这个方案工程组主要是用 Cisco 设备提供当今网络领域的解决方案并负责处理复杂的软件、硬件问题。Vijay 也为 Cisco 的工程师和客户做过几次讲座, 包括 Cisco 路由器体系结构、IP 组播、Internet 服务质量和 Internet 路由体系结构。他也是 Duke 大学电子工程系的一名副教授。

Curtis Murphy, CCIE#1521, 是 Cisco 在 Research Triangle Park 的适用设计部的一名工程师。他主要从事 Cisco IOS 软件产品的研制。Curtis 从 1989 年就开始在网络业界工作, 到 1994 年他进入 Cisco 公司。在 Cisco, 他既是技术支持中心的 IBM 协议专家, 又是中端和高端路由器的 IOS 软件开发部的一名工程师。

Russ White, CCIE#2635, 现在在 IOS 网络协议开发部和伸缩性研究组工作, 主要从事路由协议的设计、路由协议的应用和可伸缩网络的设计。Russ 也是 Cisco 出版社在 1999 年 6 月出版的 CCIE 专业开发: 高级 IP 网络设计(Advanced IP network design, ISBN:1-57870-097-3) 一书的作者之一。

献 辞

Vijay Bollapragada:

献给我最好的朋友和妻子，Leena。正是她的爱、给我的支持以及为我做出的牺牲才使我写了这本书。也献给我两个可爱的孩子，Amita 和 Abhishe。还献给我的父母和其他所有的朋友！

Curtis Murphy:

献给我亲爱的妻子，Ann，为她那欢快的笑容、充满爱意的支持以及情愿牺牲我们在一起的时光，让我写这本书。

Russ white:

献给我的妻子，Lori，在本书的写作过程中，她给了我许多支持。也献给我的女儿，Rebekah，她在我工作的时候也懂得了不再和我抢用电脑（我会很快为你装一些软件的）。也感谢上帝给了我这样一次机会、时间以及才能来完成这份鲁莽的冒险。

关于审稿者

Michael Brown, CCIE#2429, 是一名 Cisco 系统 ISP 支持工程师。他在北卡罗来纳州大学获得电子工程科学学士学位。在网络界工作的 11 年中, 他为客户提供了许多网络和技术方面的支持。加入 Cisco Research Triangle Park 的 ISP 支持小组以前, 他已经在 Cisco 布鲁塞尔、比利时技术支持中心工作了三年半时间。他曾为 Cisco 内部人员以及 Cisco 客户做过几次关于路由体系结构的讲座。

Jennifer Denaven Carroll, CCIE#1402, 是 Lucent 科技 Network care 专业服务的资深顾问。她在上福尼亚大学获得计算机科学学士学位。在过去的 11 年里, 利用各种 IOS 交换和报文排队策略, Jennifer 曾规划、设计并实施过多个基于 Cisco 的网络。

Ron Long, 是广域 Internet 开发领导者, 也是西南部完整通信的提供商。他负责规划、扩展、开发以及实施 Internet 服务和对第 2、3 层的访问。Ron 在 1995 年开始进入网络界, 当时他在德克萨斯的一个小镇上与别人共同创办了一家 ISP。从那时起, 他就为全球 1000 多家提供过咨询服务和实施 Cisco 网络方案。

Alexander Marrhold, CCIE # 3324, 拥有工业电子行业的 MSC 学位, 也有 MBA 学位。他是 PROIN 公司 (一家领导着欧洲的培训和咨询的公司) 的资深顾问和经理。他主要从事核心路由和提供服务的研究, 比如: MPLS、LocalLoop 网络设计和实施。除了是一名资深顾问, 他在服务提供领域还是一名 CCSI, 并且受过特殊的培训。Alexander 以前还在一所大学从事过无线通讯的教学工作, 也曾是化工学院 CIM 项目的负责人。

致 谢

没有很多人的帮助这本书是不能出版的，他们给了我许多建议和意见。首先我要感谢的是这本书的技术审阅者，他们是：Mike Brown、Chris Shaker、Derek Iverson、Alexander Marhold、Jeunifer Carroll、Ron Long。还要感谢 Bob Olsen、Eric Osborne、Siva Valliappan、Ruchi Kapoor 和 Lance MaCallum，他们给了我很多有价值的反馈。最后要感谢的是 Brett Bartow、Chris Cleveland 以及 Cisco 出版社的其他朋友，他们为这本书的出版付出了很多劳动。

前 言

现在，在书店里有许多关于互联网的书，从协议到网络设计技术，覆盖面特别广。毫无疑问互联网越来越成为人们欢迎的一个领域，比如说：Internet 的飞速发展以及传输语音、图像和数据所带来的极大的方便。

Cisco 在销售网络设备方面可以说取得了巨大的成功。其中网络建设的设备，Cisco 占整个市场销售额的 85% 还多。同时，Cisco IOS 软件也成了业界路由器操作系统的默认标准。虽然有许多关于 IOS 支持的网络设计和网络协议方面的资料，但很少不是由 Cisco 公司提供的。

这种关于 Cisco 信息的不足是可以理解的，毕竟，IOS 专属于 Cisco。但这绝不会给网络建设带来任何不便。在帮助设计和解决基于 IOS 的网络的过程中，我们发现很多情况下，要么是把 IOS 体系结构方面的技术局限于一个问题，要么是使之更难解决。另外我们也曾回答过无数个 Cisco 用户关于 IOS 各种版本的各式各样的问题。本书就是试图把 Cisco IOS 体系结构和操作方面的知识集中到一起，供大家使用参考。书中的信息有很多来源于论坛、Cisco 技术支持中心等，有很多信息是在 Cisco 文档中找不到的。

▣ 本书目的

《Cisco IOS 精髓》试图为网络设计者、建设者以及管理者提供一个方便的 IOS “操作手册”。本书主要介绍 IOS 内部体系结构及 IOS 操作的一些关键部分，同时也阐述了几种 Cisco 硬件平台的结构，因为每一版本的 IOS 都是与相应的 Cisco 硬件设备紧密的集成在一起，抛开硬件体系结构而单独讲述软件是比较困难的。但是这本书并不是有关 Cisco 硬件的详细操作手册，对硬件的描述只是为了更好地说明 Cisco IOS 软件的特性。读者可能会注意到本书并没有涉及太多的硬

件平台，特别是根本没有涉及 Catalyst 交换机产品，也没有提及许多“访问”路由器。在很多情况下，不提及那些平台要么是因为这些平台与书中提及的平台类似，要么是应该在与它们相配套的资料中提及（比如：Catalyst 系列交换机）。

▣ 本书结构

本书分为三部分。第一部分讲述 IOS 总体结构，包括它的软件底层结构和报文交换体系结构。第二部分，从第 3 章开始，介绍在一些 Cisco 硬件平台上应用 IOS，包括这些硬件平台的设计以及一些特定的操作。最后，第三部分（第 8 章）介绍 IOS 怎样实现 QoS（Quality of Service，服务质量）。主要由以下章节组成：

- **第 1 章：IOS 软件体系结构基础**——给出操作系统的一些基本概念，随后介绍了 IOS 软件的底层结构，包括：进程、存储管理、CPU 调度、报文缓冲以及设备驱动。
- **第 2 章：报文交换体系结构**——给出了交换结构的一个概览。主要描述了几个独立平台的交换方法的原理，包括：进程交换、快速交换、最优交换以及 Cisco 快速转发。
- **第 3 章：存储共享路由器**——用一个相对简单的存储共享的路由器来描述第 1 章和第 2 章所提到的技术在硬件平台上是如何应用的。
- **第 4 章：早期的 Cbus 路由器**——介绍早期 Cbus 路由器中 IOS 的特殊交换的特征。描述了 AGS+和 Cisco 7000 路由器的体系结构。
- **第 5 章：基于存储片的系统**——用 Cisco 7200 路由器的 IOS 作为例子，阐述了 Particle-based 报文缓冲机制的原理。也包括 Cisco 7200 路由器的报文交换应用。
- **第 6 章：Cisco 7500 路由器**——继续描述 Cisco 7500 路由器 IOS 的 Cbus 体系结构，同时也介绍了 IOS 分布交换在多功能接口处理器上的应用（Versatile Interface Processor, VIP）。

- **第 7 章：Cisco 吉比特交换路由器：12000**——介绍在 Cisco 12000 系列路由器上应用的 IOS。
- **第 8 章：服务质量 (QoS)**——给出 Cisco IOS QoS 的一个概览，同时也描述了几种 QoS 方法。包括：优先级排队、客户排队、加权公平排队以及改进的亏空轮环。
- **附录 A：网络流交换**——介绍 IOS 中网络流的特性（网络流用来监控网络流量和计费）。

目 录

译者序	I
关于作者	II
献辞	III
关于审稿者	IV
致谢	V
前言	VI
第 1 章 IOS 软件体系结构基础	1
1.1 操作系统基础	2
1.2 IOS 体系结构概述	6
1.3 存储器管理	7
1.4 IOS 进程	12
1.5 IOS 内核	19
1.6 报文缓冲管理	32
1.7 设备驱动程序	39
1.8 小结	40
第 2 章 报文交换体系结构	41
2.1 路由 101: 进程交换	43
2.2 快速交换: 向缓冲求助	47
2.3 最优交换	57
2.4 Cisco 快速转发	59
2.5 小结	66
第 3 章 存储共享路由器	67
3.1 存储共享路由器体系结构	68
3.2 存储共享路由器报文缓冲区	74
3.3 存储共享路由器的报文交换	78
3.4 小结	84
第 4 章 早期的 Cbus 路由器	85
4.1 AGS+硬件体系结构	86

4.2	Cbus 报文交换.....	88
4.3	Cisco 7000 系列路由器.....	91
1.4	小结.....	92
第 5 章	基于存储片的系统.....	93
5.1	用存储片进行缓冲区管理.....	94
5.2	Cisco 7200 系列路由器.....	98
5.3	Cisco 7200 系列路由器上的报文交换.....	105
5.4	小结.....	109
第 6 章	Cisco 7500 路由器.....	110
6.1	Cisco 7500 路由器的硬件结构.....	111
6.2	Cisco 7500 路由器的报文交换.....	121
6.3	VIP 结构.....	127
6.4	Cisco 7500 路由器疑难解答.....	138
6.5	小结.....	141
第 7 章	Cisco 吉比特交换路由器 12000.....	142
7.1	硬件结构.....	142
7.2	报文交换.....	157
7.3	小结.....	162
第 8 章	服务质量.....	163
8.1	QoS 概览.....	164
8.2	优先级排队.....	167
8.3	定制排队.....	169
8.4	加权公平排队.....	173
8.5	改进的亏空轮环.....	186
8.6	加权随机早检测.....	190
8.7	可选择报文废弃.....	193
8.8	其他 QoS 特性.....	194
8.9	小结.....	194
附录 A	网络流交换.....	195

IOS 软件体系结构基础

本章包括以下主要内容：

- 操作系统基础
- IOS 体系结构概述
- 存储器管理
- IOS 进程
- IOS 内核
- 报文缓冲管理
- 设备驱动

如果让读者选择最流行的、使用最广泛的计算机操作系统，你将会选择哪些呢？也许你会列出许多，诸如：UNIX、MS-DOS、Microsoft Windows 甚至是 IBM in MVS，这些都是很著名的操作系统，甚至家里的电脑正使用着其中的一种。现在，考虑一分钟：还有其他的吗？是否会想到 Cisco IOS？不，那是不可能的，虽然 IOS 是目前使用最广泛的操作系统之一。

不像上面提到的那些用于一般目的的操作系统，很多人根本就没有直接接触过

IOS。许多用计算机上网的人从来没有意识到 IOS 才是幕后英雄。即使那些知道 IOS，甚至直接使用 IOS 的人也常常不认为 IOS 是一种操作系统，他们认为 IOS 只不过是 Cisco 路由器上运行的一种软件而已。

确实，IOS 不运行字处理器或其他的应用程序，但实际上它仍旧是一种操作系统——一种特殊的用来交换数据报文的操作系统。正像我们即将看到的一样，IOS 体系结构的很大一部分都致力于使报文交换更加迅速、更加高效。

虽然组成 IOS 的组件中有许多也是组成通常所说的操作系统的组件，但这些组件都包括为 IOS 特别设计的一些关键的功能。那么本章就是要讨论那些 IOS 的基本组件——IOS 软件底层结构，以及设计它们的机制。

我们首先从一些基本的操作系统概念和术语开始。这些概念和术语，对理解 IOS 体系结构非常有用。如果读者对操作系统已经有一个整体的了解，也可以跳过这一部分，直接进入“IOS 体系结构概述”这一节。本章其余部分主要介绍 IOS 体系结构的一些主要组成。

1.1 操作系统基础

现代操作系统主要完成两个功能：硬件抽象和资源管理。

硬件抽象给软件开发者提供一个接口，这个接口在计算机硬件与他们的应用程序之间。有了这样的接口，软件开发者就不用再关心硬件的复杂性。实际上对硬件的编程只需一次（在操作系统内部），然后大家都可以共享。

操作系统完成的第二个功能是管理计算机的资源（包括：CPU 周期、存储器、磁盘空间等）。通过操作系统的有效管理，这些资源就可以被各种应用所共享。像硬件抽象功能一样，在操作系统内部构造一个资源管理体系，从而每一个应用的开发者都不必再为自己的程序编写资源管理代码。

1.1.1 CPU 资源管理和多任务

虽然一些操作系统在同一时刻只允许一个程序运行（例如：MS-DOS 的许多版本都是如此），现在也很容易发现许多操作系统也可以在同一时刻管理多个程序。在同一时刻运行多个程序称为多任务，而支持多任务的操作系统一般称之为多任务操作系

统。

为多任务操作系统编写的应用程序通常包括多个能在同一时刻运行的、独立的任务。这些小的子程序一般称之为线程，因为它们组成了一个程序内的单线指令序列。每个线程都含有自己的一个 CPU 寄存器的集合，称为上下文 (Context)，线程也能和同一程序内部的其他线程共享存储地址空间。一组共享相同的地址空间、为完成同一个任务且一起控制一组操作系统资源的线程称为一个进程 (Process)。在支持虚拟存储的操作系统和 CPU 中，每一个进程都在各自独立的地址空间运行，它们的地址并不交叉。

因为一个处理器在同一时刻只能运行一个程序的指令，操作系统必须判断某一时刻应该执行哪一个程序的指令集合 (哪一个线程)。决定究竟该执行哪个进程的机制称为调度 (Scheduling)，这通常由操作系统的一部分核心程序来完成，称之为内核 (kernel)。根据操作系统所采用的优化方式的不同，操作系统可以使用不同的线程调度策略。不同的应用 (批处理、交互、事务性、实时及其他类型的应用) 拥有不同的 CPU 使用特性，同时这些应用的整体性能也受所采用的调度策略影响。

最简单的调度策略就是根据每一个线程请求处理的先后顺序来分配处理器，先来的先用，直至执行完一个再执行下一个，这种方法也称为 FIFO (先进先出) 直至运行完调度。FIFO 的优点是：它比较容易实施，开销非常低，而且比较公平——所有的线程都是平等的，先来的先服务。

这种调度方法比较适合于批处理应用以及一些需要顺序处理，而且处理完马上就会结束的事务性应用。但是它不适用于交互性和实时性的应用。因为交互性应用需要很快的被调度，对 CPU 需求周期也比较短，只有这样它才能很快地把执行结果及时反馈给用户或服务于其他外部设备。

对于这些类型的应用，一个可行的解决方法就是赋予每一个线程不同的优先级。对于那些急需访问 CPU 的线程，比如实时应用的线程，赋予较高的优先级；而对于那些不太关键的线程，比如批处理应用的线程，赋予较低的优先级。具有高优先级的线程可以跳到任务队列的头部，迅速地访问 CPU。如果有几个相同优先级的线程同时到达，则按它们到达的顺序调度 (就像基本的 FIFO 策略)，这种调度机制就称为直至运行完带有优先级的调度。

虽然这种调度策略比基本的 FIFO 有所改进，但它有一个不足，这种方法很容易使某一个线程长期独占 CPU，从而不太适合于交叉的或实时的应用。具有高优先级的线程可能会一直等在一个已经开始运行、优先级比较低的线程后面，直至它运行完。要解决这个问题，一个可行的方法就是暂时挂起 (Suspend) 或抢占 (Preempt) 一个正在运行的线程，从而其他的线程可以使用 CPU。

☞ 线程抢先

有意的挂起一个线程而调度另外的线程称为抢占。采用抢占而不是直至运行完的调度策略称为抢占调度。采用这种调度策略的操作系统称为抢占式多任务操作系统。抢占主要是基于内核通过一个上下文转换（Context Switch）来周期性地改变现在运行的线程而实现的。上下文转换的触发可以是一个系统时钟（每一个线程都分到一个时钟长度），也可以是一个调用内核本身的函数。当一个上下文转换被触发时，内核就选择下一个线程访问 CPU，而被抢占的线程则被放于队列尾，并根据调度策略等待下一次被调度的机会。

📌 注意

当操作系统内核停止一个线程使用 CPU，而调度另一个线程访问 CPU 时，一个上下文转换就发生了。换句话说，当计算机改变它正执行的任务时，就发生了上下文转换。如果根据 CPU 占用时间来评价的话，上下文转换是非常费时的，因为所有正在运行的线程所使用的寄存器都得被保存，然后再重新为新的线程使用。上下文是被抢先的线程得知它在哪里被停止的基础，也是即将运行的线程得知它上次运行到哪里的基础。

多任务抢占有许多优点，包括下面这些：

- ◆ **它的可预测性：**在有限的范围内，一个线程可以知道它下次可能被调度的时间。例如：假设在内核设置受限的情况下，一个线程可以设置为 1 秒钟运行两次，因此编程者就能确定在哪个时间间隔里，该线程可能会被调度。
- ◆ **它的不易被打断性：**沿有任何单独的线程能长时间占用 CPU，一个陷入无限死循环的线程并不能阻止其他的线程被调度。

当然，抢占式多任务调度策略也有许多不足，比如：

- ◆ **它比运行直至结束调度策略的效率低：**总的来说，抢占式多任务调度系统总是倾向于经常进行上下文转换，这也意味着 CPU 将用更多的时间来进行线程调度和上下文转换。
- ◆ **它使应用软件变得更加复杂：**一个运行于抢占式系统内的线程可能随时都会