

智慧协同网络 资源分配和业务管理

**Resource Allocation and Traffic Management
for Smart and Cooperative Networks**

王雄 任婧 王晟 徐世中 赵阳明 



国家重点图书出版规划项目



国之重器出版工程

网络强国建设

智慧协同标识网络系列

智慧协同网络 资源分配和业务管理

Resource Allocation and Traffic Management
for Smart and Cooperative Networks

王雄 任婧 王晟 徐世中 赵阳明 著

人民邮电出版社
北京

图书在版编目 (C I P) 数据

智慧协同网络资源分配和业务管理 / 王雄等著. --
北京 : 人民邮电出版社, 2018. 7
国之重器出版工程. 智慧协同标识网络系列
ISBN 978-7-115-48560-1

I. ①智… II. ①王… III. ①计算机网络—资源分配—研究②计算机网络—业务管理—研究 IV. ①TP393

中国版本图书馆CIP数据核字(2018)第102678号

内 容 提 要

本书主要讲解智慧协同网络的特点和架构原理,并重点讲解智慧协同网络的资源分配和业务管理方法。本书分成3个部分:第1部分(第1章)主要介绍智慧协同网络的特点和架构原理,包括信息中心网络、软件定义网络和数据中心网络;第2部分(第2~4章)主要介绍网络感知方法和基于网络感知的资源分配和业务管理方法;第3部分(第5~7章)主要介绍基于协同的网络资源分配和业务管理方法。

本书内容是作者近3年来最新的研究成果,主要涉及主要的智慧协同网络资源分配和业务管理问题,对下一代网络技术感兴趣的读者能从中获取有参考价值的信息。初次接触下一代网络技术的读者可以了解智慧协同网络涉及的主要资源分配和业务管理问题,并对智慧协同网络的特性有初步认识。本书可供通信、计算机、网络工程技术人员和科研人员阅读,也可供高等院校和院所相关专业师生参考。

◆ 著 王 雄 任 婧 王 晟 徐世中 赵阳明

责任编辑 代晓丽

责任印制 杨林杰

◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号

邮编 100164 电子邮件 315@ptpress.com.cn

网址 <http://www.ptpress.com.cn>

固安县铭成印刷有限公司印刷

◆ 开本: 16

印张: 11.75

2018年7月第1版

字数: 217千字

2018年7月河北第1次印刷

定价: 88.00元

读者服务热线: (010)81055488 印装质量热线: (010)81055316

反盗版热线: (010)81055315

《国之重器出版工程》

编辑委员会

编辑委员会主任：苗 圩

编辑委员会副主任：刘利华 辛国斌

编辑委员会委员：

冯长辉	梁志峰	高东升	姜子琨	许科敏
陈 因	郑立新	马向晖	高云虎	金 鑫
李 巍	李 东	高延敏	何 琼	刁石京
谢少锋	闻 库	韩 夏	赵志国	谢远生
赵永红	韩占武	刘 多	尹丽波	赵 波
卢 山	徐惠彬	赵长禄	周 玉	姚 郁
张 炜	聂 宏	付梦印	季仲华	

专家委员会委员（按姓氏笔画排列）：

于 全 中国工程院院士

王少萍 “长江学者奖励计划”特聘教授

王建民 清华大学软件学院院长

王哲荣 中国工程院院士

王 越 中国科学院院士、中国工程院院士

尤肖虎 “长江学者奖励计划”特聘教授

邓宗全 中国工程院院士

甘晓华 中国工程院院士

叶培建 中国科学院院士

朱英富 中国工程院院士

朵英贤 中国工程院院士

邬贺铨 中国工程院院士

刘大响 中国工程院院士

刘怡昕 中国工程院院士

刘韵洁 中国工程院院士

孙逢春 中国工程院院士

苏彦庆 “长江学者奖励计划”特聘教授



- 苏哲子 中国工程院院士
- 李伯虎 中国工程院院士
- 李应红 中国科学院院士
- 李新亚 国家制造强国建设战略咨询委员会委员、
中国机械工业联合会副会长
- 杨德森 中国工程院院士
- 张宏科 北京交通大学下一代互联网互联设备国家
工程实验室主任
- 陆建勋 中国工程院院士
- 陆燕荪 国家制造强国建设战略咨询委员会委员、原
机械工业部副部长
- 陈一坚 中国工程院院士
- 陈懋章 中国工程院院士
- 金东寒 中国工程院院士
- 周立伟 中国工程院院士
- 郑纬民 中国计算机学会原理事长
- 郑建华 中国科学院院士



屈贤明 国家制造强国建设战略咨询委员会委员、工业和信息化部智能制造专家咨询委员会副主任

项昌乐 “长江学者奖励计划”特聘教授，中国科协书记处书记，北京理工大学党委副书记、副校长

柳百成 中国工程院院士

闻雪友 中国工程院院士

徐德民 中国工程院院士

唐长红 中国工程院院士

黄卫东 “长江学者奖励计划”特聘教授

黄先祥 中国工程院院士

黄 维 中国科学院院士、西北工业大学常务副校长

董景辰 工业和信息化部智能制造专家咨询委员会委员

焦宗夏 “长江学者奖励计划”特聘教授



前言

互联网在设计之初其目的是为了提供端到端主机间的信息共享服务，因此，互联网络采用了以地址为中心的网络体系架构。因互联网以地址为中心的网络架构具有开放、简单、易扩展等优点，在过去的三十几年中，互联网络取得了巨大的成功。如今，互联网已成为支撑现代社会经济、科技、教育、医疗、文化等方面最重要的信息基础设施，是衡量一个国家科技和经济水平的重要标志之一。

随着宽带无线移动通信、高速光传输和云计算技术的迅速发展，各种高带宽、低时延、高安全和高可靠性要求的互联网络创新应用不断涌现。然而，面对海量的高要求互联网应用，30多年前发明的以IPv4为核心技术的互联网面临着越来越严重的技术挑战，主要包括：内容分发效率低下、地址空间不足、安全性低、可移动性差等。在维持现有互联网体系架构不变的前提下，要彻底地解决这些问题是非常困难的。为了解决这些问题，工业界和学术界认为下一代互联网络应该采用 Clean State 的网络架构。

为了应对现有互联网络的技术挑战，美国、欧洲等发达国家和地区从20世纪90年代中期就先后开始下一代互联网研究。在国家重点基础研究发展计划（“973”计划）、国家高技术研究发展计划（“863”计划）、国家自然科学基金的大力支持下，中国从20世纪90年代后期就开始了下一代互联网络体系架构和关键技术的研究。在全球研究人员的共同努力和众多项目的资助下，下一代互联网络体系架构和关键技术的研究取得了很多有价值的研究成果，其中，有代表性的包括信息中心网络、软件定义网络、名址分离网络、标识网络等。

通过对目前提出的下一代互联网络架构的分析，我们发现下一代互联网应



具有智慧协同的能力。所谓智慧,是指网络应该具有感知传输的内容、及时准确感知网络状态的动态变化和感知不同网络应用特点的能力,并能根据感知到的信息做出优化的决策。而所谓协同,是指网络能够实现异构体系架构间的协同互通,能够协同各种可用手段来优化网络性能,以及支持多种网络优化目标的协同决策。在本书中,我们将具备智慧协同能力的网络称为智慧协同网络。智慧协同网络不是指某一种特定类型的网络,它覆盖的范围较为宽泛,凡是具备智慧和协同特性的网络,我们都可以视为智慧协同网络。

资源分配和业务管理是智慧协同网络的关键问题之一,该问题关系到网络业务的需求是否能得到有效满足以及网络资源是否能得到有效利用。本书主要讨论智慧协同网络中的资源管理方法,具体而言,主要包括信息中心网络、SDN 和数据中心网络中的资源分配和业务管理方法。全书分成 3 个部分:第 1 部分(第 1 章)主要介绍智慧协同网络的特点和架构原理;第 2 部分(第 2 ~ 4 章)分别介绍信息中心网络中的缓存管理方法、数据中心网络 Coflow 的调度和路由问题,讨论 SDN 中流量矩阵测量问题;第 3 部分(第 5 ~ 7 章)讨论多种不同架构信息中心网络协同互通问题,介绍网络资源管理中不同决策目标的联合优化问题和不同决策手段联合优化的网络资源管理方法。本书不能完全罗列智慧协同网络中的所有资源管理方法,本书各个章节的网络资源管理方法都是作者团队近年来的研究成果,是研究团队集体研究的成果总结。

在此,特别感谢科学技术部国家重点基础研究发展计划(“973”计划)项目“资源动态适配机制与理论”(2013CB329103)、国家自然科学基金项目“面向 ICN 的网络级内嵌式缓存架构与配置管理方法研究”(61301153)和“全光网中基于监测迹的链路失效定位技术的研究”(61271165)、长江学者和创新团队发展计划和高等学校学科创新引智计划(又称“111 计划”)(B14039)的大力资助。本书的所有成果都是在这些项目的资助下完成的。

此外,由于编著者水平有限,书中难免有错漏,希望读者批评指正。

作 者



目 录

第 1 章 概述	1
1.1 现有互联网存在的问题	2
1.2 未来互联网与智慧协同网络	4
1.3 信息中心网络	7
1.4 软件定义网络	11
1.5 数据中心网络	13
参考文献	14
第 2 章 信息中心网络缓存管理	17
2.1 缓存感知的路由策略	19
2.1.1 研究动机	19
2.1.2 缓存感知的 K-Anycast 路由策略	20
2.1.3 实验设计及结果	26
2.2 路径上协作缓存管理策略	34
2.2.1 研究动机	34
2.2.2 MAGIC 缓存机制的设计	36
2.2.3 实验设计及结果	38
参考文献	43

第 3 章 SDN 中在线流量矩阵估计方法 45

3.1 背景介绍 48

3.2 流量测量的整体量框架 48

3.3 问题描述 50

3.4 流量测量规则设计方法 51

3.4.1 MLRF 51

3.4.2 LFF 54

3.4.3 MLRF 与 LFF 的讨论分析 59

3.5 仿真实验分析 61

3.5.1 仿真设置 61

3.5.2 仿真结果与分析 63

参考文献 69

第 4 章 数据中心网络流量的路由与调度联合优化 71

4.1 背景介绍 72

4.2 路由和调度对流量优化的意义 73

4.3 RAPIER 总体设计方案 76

4.3.1 RAPIER 应该具有的特性 76

4.3.2 RAPIER 的总体设计 77

4.3.3 RAPIER 的关键算法设计 79

4.3.4 仿真结果与分析 85

参考文献 93

第 5 章 数据中心网络虚拟机放置与拓扑控制联合优化 95

5.1 背景介绍 96

5.2 数据中心动态拓扑与虚拟机放置 99

5.2.1 OSA 交换结构 99

5.2.2 虚拟机放置问题 101

5.2.3 联合优化网络拓扑和虚拟机放置的原因 101

5.3 问题建模 102

5.3.1 网络模型 102

5.3.2 代价模型 103



5.3.3	模型约束	104
5.4	离线算法设计	105
5.4.1	模型分析	106
5.4.2	子问题 1 的求解——虚拟机分组	107
5.4.3	子问题 2 的求解——拓扑设计与路由	109
5.4.4	算法分析	111
5.4.5	算法应用讨论	114
5.5	在线算法设计	114
5.5.1	在线算法设计分析	114
5.5.2	虚拟机放置在线优化	115
5.5.3	离线算法的使用	116
5.5.4	租户的进入和退出	117
5.6	仿真及实验结论	117
5.6.1	离线算法性能受 ToR 最大度数的影响研究	118
5.6.2	离线算法性能受网络大小的影响研究	121
5.6.3	在线算法性能研究	123
	参考文献	124
第 6 章	数据中心网络业务量工程中的多目标联合优化	127
6.1	背景介绍	128
6.2	负载均衡模型	130
6.3	能量效率模型	131
6.4	纳什议价模型	132
6.5	问题分析	134
6.5.1	方案应具有的特性	134
6.5.2	纳什议价的优势	136
6.6	基于纳什议价的多目标优化方案	136
6.6.1	纳什议价模型与威胁值博弈	137
6.6.2	得到具有公平性的解的方法	139
6.6.3	求解纳什均衡模型的方法	145
6.7	仿真及结果分析	146
6.7.1	简单平行链路网络中的应用	147
6.7.2	Fattree 中的应用	148



6.7.3 NSFNet 中的应用	149
参考文献	151
第 7 章 信息中心网络的部署及协同互通	153
7.1 背景介绍	155
7.2 研究现状	157
7.3 信息中心网络可能的部署场景及挑战	159
7.4 通用信息中心网络部署架构的设计和实现	161
7.4.1 VICN 的设计思路	161
7.4.2 VICN 架构设计	162
7.4.3 VICN 原型系统的实现	164
7.5 实验设计及结果	169
7.5.1 实验部署	170
7.5.2 实验结果	171
参考文献	172
中英文对照表	173
名词索引	175



第1章 概 述

本章首先介绍了现有互联网起源及主要使用的技术，并分析了现有互联网的主要缺陷。在此基础上，介绍了未来网络业务对互联网的主要要求——智慧协同的能力。最后，重点介绍了业界公认的三种典型的智慧协同网络基本架构和原理，包括信息中心网络、软件定义网络和数据中心网络。



| 1.1 现有互联网存在的问题 |

互联网自 20 世纪 70 年代诞生以来，已经全方位地融入人类生活的各个方面，极大地改变了人类沟通、工作和生活的方式。在互联网体系结构设计之初，其目的仅是为了在两台计算机之间实现数据传递。因此，诞生于 20 世纪 70 年代的互联网体系架构采用主机到主机（Host-to-Host）的设计原则。在该原则中，每个主机都需要分配一个全网唯一的标识符（IP 地址），网络节点根据数据分组中携带的目的 IP 地址来转发数据分组。由于网络只需要根据目的 IP 采用“尽力而为”的方式来转发数据分组，因此，如何保证数据传输的可靠性和安全性等复杂功能均由网络终端完成。互联网的这种设计思路非常简单，并具有较好的健壮性、开放性和可扩展性，这为互联网的快速发展奠定了非常坚实的基础。

由于采用主机到主机的设计原则，所以现有互联网的通信过程主要解决网络中主机“在哪里”的问题，网络中的所有主机（计算机主机、交换设备以及移动终端等）都具有一个 IP 地址，以表示其位置信息。主机上的一切实体对象（内容、应用程序等）都与该主机地址绑定。当应用程序想要获取某个特定内容对象时，必须首先指定拥有该内容对象的某个特定主机作为内容提供者，并给出其地址，然后网络负责将内容对象从指定的源主机地址递送到指定的目



的地址。

随着互联网和网络应用的飞速发展,网络中需要递送的内容呈爆发式增长的态势。根据中国互联网信息中心(CNNIC)的报告,截至2017年12月,中国互联网用户规模达7.72亿,普及率达55.8%。随着大量新型网络应用的兴起(如在线视频和在线游戏等),互联网流量将进一步显著增长^[1]。根据思科公司的预测,到2018年,将有超过40亿的互联网用户,平均每月将产生132 EB的网络流量,其中79%的流量将由与内容递送相关的业务产生^[2]。

巨量的用户需求对互联网架构的性能也提出了新的挑战,比较突出的问题有以下几个方面。

(1) 无法保证用户自行指定的内容拥有者是最优的内容提供者

网络中往往有多个内容拥有者可以提供同一内容对象,但用户指定的内容拥有者可能并不是最优的提供者。如果用户指定的内容拥有者与其距离较远,用户将经历较大的服务时延,且需要较多的网络链路资源来递送内容;如果用户指定的内容拥有者负载较重,用户的内容请求可能无法获得响应。出现这一问题的根本原因在于,即便有很多主机拥有同一内容对象,用户往往仅能从有限的渠道(如搜索引擎、网站链接等)获得部分主机地址信息。同时,用户也无法获得底层网络拓扑、时延等信息,因此,无法判断所知晓的内容拥有者是否在距离上或时延上是最优的。

(2) 无法避免在相同链路上多次传输重复的内容对象

当多个用户请求同一内容对象时,在以主机为中心的通信模式下,需要在每个用户和其指定的内容拥有者之间通过独立的端到端通信递送所需内容。而网络设备仅根据主机地址递交分组,因此,在多条端到端通信所经过的相同链路上,可能反复传输相同的内容对象。

(3) 网络的安全性差

由于传统互联网将网络安全作为网络应用交给端系统完成,因此,最常采用的保障内容安全的方法是在端系统进行加密/解密。这意味着内容对象的真实性取决于主机的可信度,往往需要一种额外的安全手段来识别、定位和连接此可信主机。网络本身无法保证其递送内容的真实性。此外,由于所有的通信都是基于地址来进行的,网络的攻击者可以有针对性地对单个主机/地址发动DDoS攻击。

(4) 网络的可控可管性差

为了保证网络的可扩展性,传统互联网通常采用分布式的控制协议,如OSP、IS-IS、BGP等,然而这些分布式控制协议基本不能实现对流的精细化控制,并且任何细微的网络参数调整都有可能网络状态的震荡,引起网络的不稳定性。因此,如何保证业务的QoS和精细化的网络控制及管理一直都是传统互



联网面临的主要问题之一。虽然为了增强网络的可控可管能力, IETF 也提出了一些新的控制, 如 MPLS, 但是由于网络架构的缺陷, 这些协议仍然面临配置和管理复杂的问题。

(5) 网络不支持节点移动

在传统互联网中, 每个通信终端都有一个 IP 地址, 所有的通信都是将数据从一个地址发往另一个地址。为了管理方便, 分配给每个机构的 IP 地址都是固定不变的。因此, 当一个终端从一个地方移动到另一个地方时, 它的 IP 地址也必须重新分配, 这必然导致通信的中断。为了解决这一问题, 通常的做法是在网络中增加移动代理, 通过移动代理来管理用户位置的变化。

综上所述, 传统互联网体系架构本身的局限性, 导致了一系列很难解决的问题。虽然, 目前这些问题都有一些方法可以解决, 但这些方法都是通过现在网络体系架构上打补丁的手段, 其带来的问题是网络管理实体越来越复杂, 网络管理非常困难。为了从根本上解决现有互联网的这些问题, 更好的方法是根据现有网络需求, 重新设计互联网体系架构。

| 1.2 未来互联网与智慧协同网络 |

近年来, 世界各国都积极开展了未来互联网体系的研究工作。如美国国家自然科学基金会于 2005 年和 2006 年先后启动了全球网络创新环境 (Global Environment for Networking Innovations, GENI) 计划和未来互联网设计 (Future Internet Network Design, FIND) 计划。日本于 2006 年启动了 AKARI 计划, 目标是在 2015 年之前研究出一个全新的网络架构。欧盟于 2008 年启动了未来互联网研究和实验 (Future Internet Research and Experimentation, FIRE) 计划, 拟对未来互联网架构和服务机制进行研究, 并建立相应的测试平台。在 GENI 和 FIND 计划的基础上, 美国国家自然科学基金会于 2010 年发布了未来互联网架构 (Future Internet Architecture, FIA) 计划, 先后资助了命名数据网络 (Named Data Networking, NDN)、MobilityFirst、NEBULA、表达式互联网架构 (eXpressive Internet Architecture, XIA)、ChoiceNet 这 5 个重大项目, 分别从不同侧面研究未来互联网的体系架构。2012 年 6 月, 美国前总统奥巴马签署行政命令并启动 US IGNITE 计划, 进一步加强美国在未来互联网体系与应用方面的基础研究, 以巩固美国在信息网络领域的领导地位。

我国也非常重视对未来信息网络体系结构和关键理论及技术的研究。国家