智慧协同标识网络系列

RESOURCE ALLOCATION AND TRAFFIC MANAGEMENT FOR SMART AND COOPERATIVE NETWORKS

智慧协同网络

资源分配和业务管理

■ 王雄 任婧 王晟 徐世中 赵阳明 著

中国工信出版集团

人民邮电出版社
POSTS & TELECOM PRESS

智慧协同标识网络系列

RESOURCE ALLOCATION AND TRAFFIC MANAGEMENT FOR SMART AND COOPERATIVE NETWORKS

智慧协同网络

资源分配和业务管理

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（CIP）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, 2016.7
（智慧协同标识网络系列）
ISBN 978-7-115-41892-0

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(2016)第057391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以及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。

本书内容都是作者近 3 年来最新的研究成果，主要内容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。对于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，也可以供高等院校和研究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　彭志环

◆ 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：700×1000　1/16
印张：12.5　　2016 年 7 月第 1 版
字数：203 千字　　2016 年 7 月河北第 1 次印刷

定价：68.00 元

读者服务热线：(010)81055488　印装质量热线：(010)81055316
反盗版热线：(010)81055315

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。互联网以地址为中心的网络架构因其具有开放、简单、易扩展等优点，在过去的 30 年中，互联网络取得的了巨大的成功。如今，互联网已成为支撑现代社会经济、科技、教育、医疗、文化等方面的最重要信息基础设施，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的不断涌现。然而，面对海量的高要求互联网应用，30 多年前发明的以 IPv4 协议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。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“973”计划、国家高技术研究发展“863”计划、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，凡是能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的利用。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。全书可以分成 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）分别介绍了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，讨论了 SDN 中流量矩阵测量问题。第 3 部分（第 5～7 章）讨论了多种不同架构信息中心网络协同互通问题，介绍了网络资源管理中不同决策目标的协同优化问题和不同决策手段协同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，还特别感谢国家科技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）、长江学者创新团队发展计划和高等学校学科创新引智计划“111 引智计划”（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

2015 年 12 月

第1章 概述

第2章 信息中心网络缓存管理

第3章 SDN 中在线流量矩阵估计方法

第1章 概 述

1.1 现有互联网络存在的问题

自 20 世纪 70 年代互联网发明以来，互联网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 年代的互联网络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都由网络终端来完成。互联网络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，这为互联网络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，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（电脑主机、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，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炸式增长的态势。根据思科公司的统计，在 2013 年，全球超过 36 亿互联网用户每月平均产生 51.2 EB 的网络流量，包括从网络中获取内容对象或通过网络进行通信[1]。根据中国互联网信息中心（CNNIC）的报告，截至 2014 年 6 月，中国互联网用户规模达 6.32 亿，普及率达 46.9%[2]。随着大量新型网络应用的兴起（比如在线视频和在线游戏等），互联网流量将进一步显著增长。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[1]。

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（比如通过搜索引擎、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/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，网络的攻击者可以有针对性地对单个主机/地址来发动 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，比如 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导致网络状态的震荡，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络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，比如 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，这必然导致了通信的中断。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在现有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，造成网络管理非常困难。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 未来互联网络与智慧协同网络

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，如美国自然科学基金委于 2005 年和 2006 年先后启动了 GENI（Global Environment for Networking Innovations，全球网络创新环境）计划[3]和 FIND（Future Internet Network Design，未来互联网络设计）[4]。日本于 2006 年启动了 AKARI 计划[5]，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 FIRE（Future Internet Research and Experimentation，未来互联网研究和实验）计划[6]，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，美国自然科学基金委于 2010 年发布了 FIA（Future Internet Architecture，未来互联网架构）计划，先后资助了 NDN（Named Data Networking，命名数据网络）[7]、MobilityFirst[8]、NEBULA[9]、XIA（eXpressive Internet Architecture）[10]、ChoiceNet[11]等 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，美国总统奥巴马签署行政命令并启动 US IGNITE 计划[12]，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“973”计划先后启动了“一体化可信网络与普适服务体系基础研究”[13]“可测可控可管的 IP 网的基础研究”“新一代互联网体系结构和协议基础研究”“面向服务的未来互联网体系结构与机制研究”和“可重构信息通信基础网络体系研究”等项目。国家“863”

计划先后启动了“身份与位置分离的新型路由关键技术与实验系统”“三网融合演进技术与系统研究”等项目。国家自然科学基金委先后启动了“未来互联网体系理论及关键技术研究”“后IP网络体系结构及其机理探索”“未来网络体系结构与关键技术”等重点项目。

在全球研究人员的共同努力和众多项目的资助下，未来互联网体系架构的研究取得了很多耀眼的成果，比如，在GENI的资助下，美国斯坦福大学Nick McKeown带领的Clean Slate课题组于2006年提出了软件定义网络（Software Defined Networking，SDN）的概念[14]。SDN将网络的控制面和数据面进行分离，控制面单独地运行于逻辑上集中的网络控制器上，而数据平面则分布在各个通用的网络转发设备上，这种创新性的设计大大地增加了网络的可控可管性，并降低了网络设备的复杂性。在美国自然基金重点项目的支持下，美国加州大学洛杉矶分校的张丽霞教授带领的研究团队提出了NDN体系架构[15]。NDN直接对传输的内容进行命名，网络用户根据内容的名字来获取内容和服务，而路由器根据内容名来转发内容；在欧盟PSIRP（Publish-Subscribe Internet Routing Paradigm，发布订阅互联网路由范例）项目的支持下，P. Jokela等人提出了PSIRP网络体系架构[16]，与NDN类似，PSIRP也是以信息为中心的网络体系架构。在我国“973”项目的支持下，北京交通大学张宏科教授团队提出了智慧协同标识网络[17]。

除了网络体系架构上的支持外，为了提高网络的数据分发能力，适应快速变化的网络新业务，增强网络的可管控能力，未来互联网还必须具有以下能力。

（1）感知传输内容

在传统互联网中，网络节点只关心数据分组中的地址信息，而不关心数据分组中携带的传输内容，从而网络无法缓存传输的内容，导致网络的带宽浪费和内容分发效率低下。为了提高网络对内容的分发效率，网络节点需要辨识经过的内容，并根据内容的特性对内容进行缓存和处理，从而节约网络带宽资源，提高内容的分发效率。

（2）及时准确地感知网络状态

要有效地管理和优化网络，网络必须首先实时地感知网络的当前状态，比如网络流量的大小、网络链路的性能指标、网络节点和链路的负载情况等。传统互联网由于缺乏精细化的管理手段，网络状态的感知通常存在很大的误差，从而引起网络决策的失效。

（3）感知和区分不同网络应用

网络中不同的应用有不同的特点，不同的应用对网络的需求也可能截然不同。因此，为了给各种网络应用提供最好的服务，网络首先需要感知不同网络应用的特点和需求，以便为网络的决策提供精确的输入。

（4）不同网络架构间的协同互通

从前面的论述中可以发现，目前研究界已经提出了若干种未来互联网的体系架构，每种架构都有各自的优点和缺点。因此，未来互联网很可能将是由采用不同体系架构的网络组成。在这种情况下，不同网络架构间的协同互通就变得非常重要。

（5）不同网络优化目标协同决策

网络中通常存在若干不同的优化目标，比如负载均衡、网络节能、网络使用代价等。这些优化目标通常还是相互矛盾的，即优化一个目标可能会损害另一个目标，比如，优化网络负载均衡就需要将流尽量均衡地分配到网络的各条链路上，然而这种优化方案却不利于网络的节能（因为网络节能需要尽量多地关掉网络节点和链路）。因此，为了达到一个理想的决策方案，网络的资源管理和调度方案需要考虑不同优化目标间的协同决策。

（6）不同决策手段的协同决策

为了提高网络的资源利用和网络的服务性能，通常可以采用多种不同的手段，比如可以改变业务路由、改变网络的物理或逻辑拓扑等。然而，在实际的网络环境中，只通过单一的手段往往很难达到满意的优化效果。因此，为了增强网络的服务性能，网络必须能协同多种手段来进行决策优化。

综上所述，未来互联网需要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。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未来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，凡是能具备智慧和协同特性的网络我们都可以视之为智慧协同网络。

目前，研究界普遍认为未来智慧协同网络主要包含 3 种类型：信息中心网络、软件定义网络和数据中心网络。因此，在接下来的小节中，我们将分别介绍这几种网络

的基本架构和原理。

1.3 信息中心网络

作为一种重要的未来网络架构，内容中心网络（Content Centric Networking，CCN）可以解决传统互联网以主机为中心的通信模式与快速增长的内容递送需求之间的矛盾。内容中心网络是从网络架构的角度出发，提出将网络中的主机位置和内容对象进行解耦。内容对象具有唯一的标识，不再与主机地址进行绑定。用户直接使用内容名请求所需要的内容对象。网络设备可以识别内容名，并能够根据内容名、用户位置等信息找到适当的内容拥有者作为内容提供者。同时，网络可以利用多播和缓存等技术，避免在网络中重复递送相同的内容对象，也可以利用内容名的唯一性来验证内容对象的真实性。

目前已有大量内容中心网络架构被提出，在这里我们选择最具代表性的 3 种架构进行简述，即 DONA（Data-Oriented Network Architecture，面向数据的网络架构）[18]、NDN[15]和 PURSUIT[19]。

（1）DONA

DONA 是较早出现的具有典型代表意义的网络架构[18]，如图 1-1 所示。相对于传统互联网设计对主机的关注，用户实际上更关心内容对象本身，DONA 认为新的网络架构需要满足如下要求。

① 内容名的永久性：只要内容对象本身是可用的，其对应的内容名也就是一直有效不变的，即内容名不会随内容对象所在位置的变化而变化。

② 内容对象的可用性：只要内容对象还存在于网络中，无需依赖特殊的应用层机制，网络可以为用户找到最合适的内容提供者。

③ 内容对象的真实性：网络需要为用户提供认证内容对象的机制，使其能够直接对内容对象的真实性进行认证，而不依赖于对获取内容对象的方式的认证。

基于以上 3 点，DONA 对网络中内容对象的命名机制和名字解析系统进行了重新设计。在 DONA 中，内容对象被赋予永久的名字。为了满足永久性和真实性的要求，DONA 采用了扁平自认证（Flat Self-Certification）的名字。DONA 中的内容名

根据主成分（Principal）构造，每个主成分都被赋予一对公私密钥（Public-Private Key）。每个内容对象都和一个主成分关联。内容名的格式为 P:L，其中，P 为主成分的公钥的加密散列值，用户可利用 P 验证主成分的身份；L 为主成分为此内容对象所选的标签，该标签可为内容对象本身的散列值，因此，用户可以根据内容名校验内容对象本身。

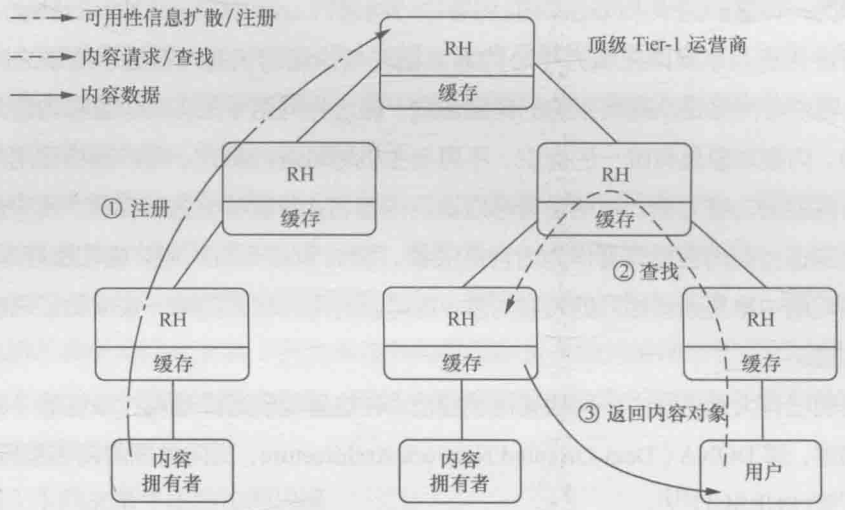

图 1-1　DONA 架构示意图

为了实现内容对象的高可用性，DONA 重新构建了名字解析系统。该名字解析系统由多个解析处理器（Resolution Handler，RH）构成。各自治域（Autonomous System，AS）自行决定本自治域内的 RH 组织结构，逻辑上对外呈现单个 RH，这些 RH 按照 AS 间拓扑关系进行连接。如图 1-1 所示，内容拥有者需要向名字解析系统发送注册信息完成注册（图 1-1 步骤①）。用户向名字解析系统查询能够提供所需内容对象的内容拥有者。名字解析系统根据内容名找到最合适的内容提供者，该内容提供者可能是已注册的内容拥有者或 RH 上的缓存（图 1-1 步骤②），内容对象则根据网络地址递送给用户（图 1-1 步骤③）。

（2）NDN

NDN[15]及其前续项目 CCN 重新设计了网络协议栈，使用内容层取代传统 IP 层作为网络新的“细腰”，将 IP 技术作为内容层的一种底层技术。相对于 DONA 采用的扁平自认证名字，NDN 采用了类似于统一资源定位符（Uniform/Universal