

# 基于领域专用软硬件协同的多模态网络环境构造技术

胡宇翔, 崔子熙, 李子勇, 董永吉, 崔鹏帅, 邬江兴

(信息工程大学信息技术研究所, 河南 郑州 450001)

**摘 要:** 信息网络与经济社会深度融合发展带来了网络模态共生共存、动态并发、演进变革融合态势。首先, 从基于多模态网络环境的垂直行业应用需求出发, 提出了一种基于领域专用软硬件协同的多模态网元设计方案, 讨论了多模态网络环境构造关键技术; 然后, 针对多样化网络模态对网元内异构资源竞争的问题, 提出了一种存算转一体的异构软硬件资源分配方法, 根据模态业务量动态调整资源分配和置换策略, 实现异构资源的高效利用。实验结果表明, 所提方案可为构建技术体系与支撑环境分离的多模态网络提供基础平台支撑。

**关键词:** 多模态网络环境; 多模态网元; 领域专用软硬件协同

**中图分类号:** TN934

**文献标志码:** A

**DOI:** 10.11959/j.issn.1000-436x.2022086

## Construction technologies of polymorphic network environment based on codesign of domain-specific software/hardware

HU Yuxiang, CUI Zixi, LI Ziyong, DONG Yongji, CUI Pengshuai, WU Jiangxing

Institute of Information Technology, Information Engineering University, Zhengzhou 450001, China

**Abstract:** The deeply integration of information network and economic society has brought the trend of symbiotic coexistence, dynamic concurrency, evolution and revolution of network modals. Firstly, in order to satisfy the requirements of vertical industry applications, a scheme of the polymorphic network element based on codesign of domain-specific hardware/software was proposed, the key technologies to construct polymorphic network environment were discussed as well. Then, to solve the resource competition problem between diversified network modals, a heterogeneous resource allocation and replacement method was proposed by integration of computing, storage and forwarding, which dynamically adjusted the resource-allocation strategy according to the traffic volume to realize the efficient resource utilization. Experimental results show that the proposed scheme can provide the basic platform support for the construction of polymorphic network by separation of technical systems and supporting environment.

**Keywords:** polymorphic network environment, polymorphic network element, codesign of domain-specific software/hardware

## 0 引言

近年来, 信息网络已逐渐成为与国家发展、人民生活息息相关的重要基础设施, 海量内容获取、高清视频直播、低时延控制等新兴业务的不断涌现, 现有信息网络以变长分组交换、尽力而为传送

和 TCP/IP 等为基础, 网络功能单一僵化, 且对网络体系的演进式修补使系统工程复杂度失控, 广义安全危机凸显, 在应对人-机-物大规模智能组网、全时空随遇接入、服务性能可确保、安全性能可量化设计与验证度量、网络大数据或分布式网络计算等多元化应用的灵活性/扩展性、差异化/定制化服

收稿日期: 2022-01-05; 修回日期: 2022-03-21

基金项目: 国家重点研发计划基金资助项目 (No.2019YFB1802502); 国家自然科学基金资助项目 (No.61872382)

**Foundation Items:** The National Key Research and Development Program of China (No.2019YFB1802502), The National Natural Science Foundation of China (No.61872382)

务、移动性支持、全域运维管理、安全可信等方面面临诸多困境<sup>[1]</sup>。

为此,国内外研究人员提出了以信息/内容中心网络、地球剖分网络、身份标识网络等为代表的多种新型路由寻址技术及网络架构,并已经在不同的网络场景中得到部署,显示出强大的生命力和应用价值。面向未来典型应用场景需求,郑秀丽等<sup>[2]</sup>对 IP 技术进行继承式发展,提出了一种新型的网络协议体系——NewIP,构造未来网络“新腰”。命名数据网络(NDN, named data networking)<sup>[3]</sup>是一种基于内容标识的新型网络架构,以“生产者-消费者”的通信模式构建去 IP 化的内容分发逻辑。文献[4]提出了一种面向多源信息采集的地理位置路由寻址机制,验证了位置信息编码路由机制在物联网应用场景中的有效性。文献[5]提出了一种软件定义网络、存储和计算新融合架构,引入全网能耗度量解决多维资源容量分配问题,最终实现以信息内容为中心的“生产者-消费者”通信模式。

面对“人-机-物”万物智联时代的多元化、个性化网络能力需求,我国邬江兴院士<sup>[6]</sup>自主创新提出了“技术体系与支撑环境分离”的多模态网络发展范式,将现有或将来的各种网络技术体系以“网络模态”的形态在多模态网络环境(PINE, polymorphic network environment)中动态加载和运行,按照模态自定义的软硬件组态、报文格式、路由协议、交换方式、转发逻辑、业务特点、运维规范和安全策略等进行处理,实现多种网络模态在同一技术物理环境内的共生共存、独立演进与变革,由多元化技术体系满足垂直行业的多样化需求。文献[7]提出了多模态智慧网络架构及其运行逻辑,支持网络结构的全维可定义和各层功能的多模态呈现,为网络智能化、多样化、个性化、高稳健性和高效率发展提供了一种解决方案。

在多模态网络发展范式中,PINE 是支撑网络模态共生共存、动态并发、演进变革融合的运行逻辑和资源实体,它通过网络全维可定义等技术手段和标准化的网络功能即插即用接口,基于面向各种领域专用的软硬件资源如多核、专用集成电路(ASIC, application specific integrated circuit)、现场可编程门阵列(FPGA, field programmable gate array)等,构建出可灵活定义和编排的网络基线功能,并通过网络基线元素再定义和服务自动编排来灵活

适配多元化网络模态的需求,从而为多模态的网络技术体系提供柔性化组织的可靠“底座网络”支撑。

本文从多模态网络环境的上述目标出发,提出基于领域专用软硬件协同的多模态网络环境构造技术。具体来说,本文工作主要包括以下 2 个方面。

1) 在架构设计层面,结合基于多模态网络环境的垂直行业应用方案,提出了一种基于领域专用软硬件协同的多模态网元设计,讨论了多模态网络环境构造的关键技术,为多模态网络环境技术研究提供架构支撑。

2) 在资源分配层面,针对多样化网络模态对网元内异构资源竞争的问题,提出了一种存算转一体的异构软硬件资源分配方法,可根据模态业务量动态调整资源分配和置换策略,实现异构资源的高效利用。

## 1 多模态网络环境设计

### 1.1 基于多模态网络环境的垂直行业应用

面对确定性时延、低功耗等更加精细化的垂直行业应用场景,传统网络技术体系表现出无所适从的窘境,只能通过虚拟专网、网络切片、边缘计算等方案作为过渡或者补充,很难从网络基础能力上满足行业应用需求。例如,在以微秒为量化单位的金融交易市场中,网络通信的低时延意味着行情商能够获得更多的交易机会与更大的决策空间,需要设计领域专用的 TCP/IP 引擎<sup>[8]</sup>。特别是在“摩尔定律”和“登纳德缩比定律”陷入困境的今天,源于消费互联网时代的 IP 网络无法单纯地依靠增加各种传输资源扩展传输带宽、无节制堆砌软硬件代码量、提升 CPU 性能来满足产业互联网时代新的发展要求<sup>[6]</sup>。

面对上述挑战,多模态网络环境提出构建多样化新型寻址路由技术及其工程实现的协议栈,支持在各个协议栈“通道”内的数据协议格式、查表转发方式、运行组织逻辑等由网络模态自定义处理,由多元化网络模态承载复杂、定制化的网络业务处理,支撑起网际互连网络、时延敏感网络、位置敏感网络、移动通信网络等复杂的应用场景,从根本上满足网络智慧化、多元化、个性化、高稳健、高效能的业务需求。

基于多模态网络环境的垂直行业应用如图 1 所示<sup>[6]</sup>。在该方案中,多模态网络环境由多个多模态网元互联而成,使各种网络模态屏蔽底层异构平台

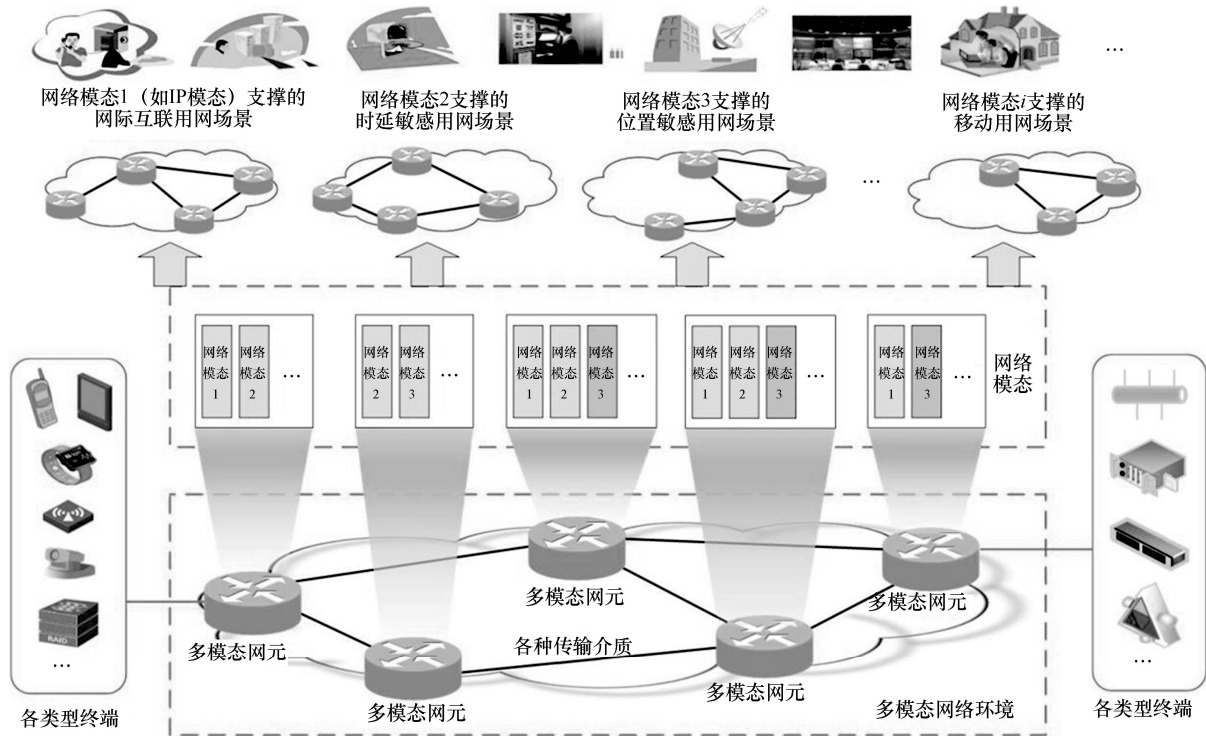


图1 基于多模态网络环境的垂直行业应用

的技术细节，向上支持多种模态在运行时共生共存，既能实现多元网络的智能、高效、安全一体化部署和管理，又能保障各类业务及网络技术体制自持发展、相互隔离。其中，多模态网元专注于网络基线能力的提供，对网元设备内部的计算、存储、转发等异构软硬件资源进行统一抽象，将不同的资源组合映射到网络模态对应的特定网络功能。在初始化阶段，多模态网络环境依据可用资源总量与模态能力请求求解最佳的资源分配，保障网络多元化性能指标；在运行过程中，各模态之间的处理逻辑相互独立、分配资源相互隔离，同时模态的资源配比依据网络环境与业务流负载率的变化进行动态调整。

为了实现网络模态的智能加载和部署，多模态网络环境从协议处理灵活性和资源属性 2 个方面进行支撑。

一方面，针对传统网元设备的功能固化、协议不可定制等问题，面向各种网络模态对协议解析和处理逻辑的深度可编程需求，进行面向数据报文全域的可编程解析与处理，并通过规范的编程接口实现自定义的报文处理逻辑配置。在该方面，业界已经开展了诸多工作，例如，Pat 等<sup>[9]</sup>推出可编程协议无关报文处理(P4, programming protocol-independent packet processors)语言，进一步开放底层转发流水

线，通过统一的编程语言控制报文处理逻辑，支持网络功能按需重构，实现了网络协议与物理设备的解耦。Broadcom 公司开发的网络编程语言同样可用于定义和编译其产品线中具有不同硬件架构（如 Trident 4 和 Jericho 2）的可编程网络设备<sup>[10]</sup>，提供一个编译器驱动的数据平面，具有高功能容量和功能并行性。

另一方面，扩展网络数据平面的资源空间，将计算、存储甚至安全相关处理资源引入多模态网络环境中，面向各种网络模态的定制化处理逻辑，协同调度各种软硬件异构资源实现复杂网络功能处理，将面向特定网络模态的流量计算处理/加密、内容缓存等操作在多模态网络环境中卸载实现。在该方面，业界也已经开展了诸多工作。Scholz 等<sup>[11]</sup>分别在 3 个不同的 P4 目标平台（CPU、SmartNIC、NetFPGA）中实现了一种加密哈希函数的原型方案，以减少针对哈希冲突的潜在攻击。Mahmood 等<sup>[12]</sup>基于软件定义网络架构完成一种面向 NDN 的实施方案，通过网内缓存机制实现交换节点的带状状态转发功能。

## 1.2 基于领域专用软硬件协同的多模态网元设计

近年来，各种领域专用软硬件层出不穷，如面向高速报文处理的 FPGA、面向可编程报文交换的

ASIC、面向复杂网络功能逻辑的网络处理单元 (NPU, network processing unit), 以及面向复杂报文处理的数据平面开发套件 (DPDK, data plane development kit) 等。这些面向特定领域专用的软硬件功能特色明显, 适配领域聚焦, 为面向特定应用场景和网络模态需求提供了更丰富的资源支撑。为了快速加载各种网络模态、有效释放端系统和应用处理资源, 需要将面向领域专用的各种算力和缓存等资源部署至多模态网络环境中, 形成计算存储转发一体化的多模态网元。例如, 以内容为中心的 NDN 依赖一定的存储空间对数据资源进行缓存, 以便针对后续相同兴趣报文进行快速响应<sup>[12]</sup>。因此, 多模态网络环境构造的核心就是设计基于领域专用软硬件协同的多模态网元, 完成网元系统内部的异质资源的统一表征与置换, 支持快速构建面向给定网络模态的处理逻辑, 为实现“支撑环境与技术体制分离”奠定基础。

本文提出一种基于领域专用软硬件协同的多模态网元设计方案, 如图 2 所示。该网元由高性能转发功能模组、基于异构软硬加资源协同的软件定义功能模组和网元逻辑配置模组组成。

1) 高性能转发功能模组构建在基于 ASIC 或网络处理机 (NP, network processor) 的高性能可编程交换芯片之上, 具备可编程的模态识别、协议解析和报文处理等功能, 并对指定类型的网络模态进行标记以及向后端软件定义功能模组的数据分发, 完成相应的协议报文输入/输出功能。

2) 软件定义功能模组由一系列面向网络处理领域专用的异构软硬件资源组成, 如各种 CPU、

FPGA、ASIC 器件等。这些异构软硬件资源经内部物理链路与硬件转发功能模组互通。各功能实例是由网络模态依据特定策略定义的模态处理逻辑, 能够灵活组合平台资源以完成特定的流处理机制。该模组中, 调度模块用于统一管理计算存储转发资源, 通过资源隔离技术使同一功能实例服务于多种模态流水线。

3) 网元逻辑配置模组用于承载多模态网络编译套件, 根据网络模态需求完成网元系统内部异构资源的协同编译, 并将资源处理逻辑配置到软硬件功能模组中。在网元运行过程中, 该模组监视各功能实体的运行状态和资源利用率, 并对资源协同处理逻辑进行动态编译和配置。

可以看出, 多模态网元中各种网络模态的构建过程是一种面向异构软硬件资源的协同调度和组合过程, 所有网络基线能力都已融入模态处理流水线逻辑中, 它与目前网络功能虚拟化技术中的服务功能链编排有着本质不同。

## 2 多模态网络环境构造关键技术

### 2.1 基于高级抽象模型的网络功能重构

基于高级抽象模型的网络开放可编程技术及多元化异构平台技术的发展, 为构造异构资源一体化编译的多模态网络环境提供了重要支撑。现有可编程的各种处理平台包括 CPU、ASIC、FPGA、NP 等, 在成本、功耗、吞吐量、灵活度、功能性等方面各有优势, 丰富了网络功能向数据平面卸载的载体对象。为了实现编程标准化, P4 仅仅定义出“匹配+动作”的流水线模型, 允许各平台在抽象的协

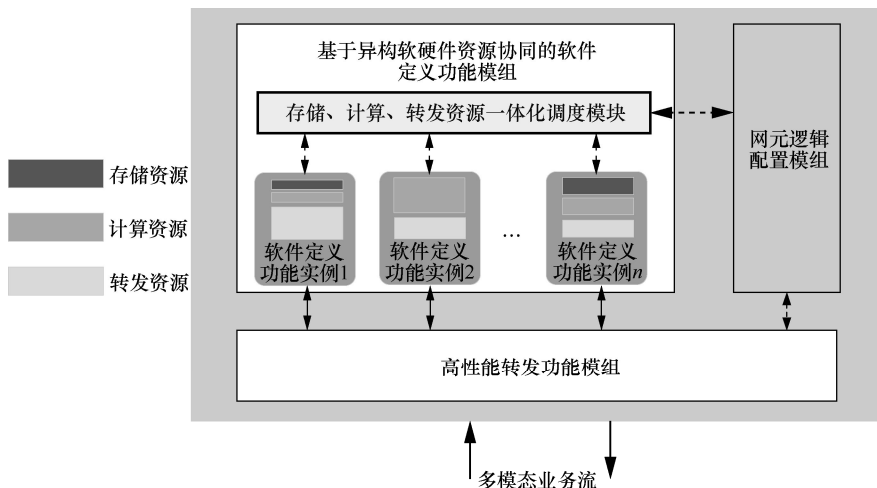


图 2 基于领域专用软硬件协同的多模态网元设计方案

议无关交换架构（PISA, protocol independent switch architecture）概念之上构建具体的、与底层架构耦合的单元布局<sup>[13]</sup>。虽然这种语言设计可以灵活地适应一系列目标平台，但也使数据平面针对报文的操作原语限于字段修改、协议重组等基本转发功能，同时增加了编写、组织和重构多模态功能程序的难度<sup>[14]</sup>。

多模态网元将用户编程的维度由“匹配+动作”扩展至计算、存储、转发等多个维度，本质上是一种面向模态的新型轻量级的“异质编程架构”，将该架构抽象出通用的高级语言表示，对异构的多模态网络资源组合、重构和编排等提供了细粒度的操作原语和交互接口。其核心思想是将报文处理流程提炼为功能逻辑本质，并从硬件层结构中抽象出来。在可定义的粒度方面，既支持用户定义完整的控制模块实现特定网络功能，又支持各功能模块之间通过良好的接口交互元数据信息；在可定义的灵活性方面，既支持常用的、平台提前定义的函数功能和代码库，又支持将用户自定义的计算/存储/安全操作以外部扩展对象进行调用。

### 2.2 网络模态的智能生成与部署

在多资源共存的异构网络环境中，可移植性、可扩展性和组合性问题从根本上破坏了大规模可

编程网络的可管理性，网络模态加载容易陷入无休止的程序重构和硬件配置<sup>[15]</sup>。针对网络模态的生成、部署与隔离，需要形成标准的多平台协同编译 workflow，向上屏蔽冗余的底层硬件细节，降低网络管理复杂度。

为了构建高效的计算存储转发一体化编译工作环境，PINE 全面推进异构资源平台的硬件抽象描述及配套的编译套件，支持网络模态功能快速、按需加载至异构平台。面向异构软硬件资源的网络模态编译流程如图 3 所示，编译套件接受以下 3 个方面的信息作为输入：1) 基于高级网络编程的模态程序（2.1 节）；2) 异构平台配置信息；3) 特定网络功能与目标平台的映射关系。前端编译模组由输入生成一个上下文感知的中间表示，其中包括指令依赖性、部署约束等重要信息。后端编译模组将上述中间表示编码为可满足性模理论（SMT, satisfiability modulo theory）下的各种约束，通过对一组约束进行求解，最终分别获得各模态在异构功能平台的流水线配置。

在根据平台相关的后端编译器生成初始化配置之后，多模态网元操作系统通过容器化技术将自定义流水线加载至对应的承载平台。因此，单个模态功能的重构不会对系统整体的稳定性造成影响，

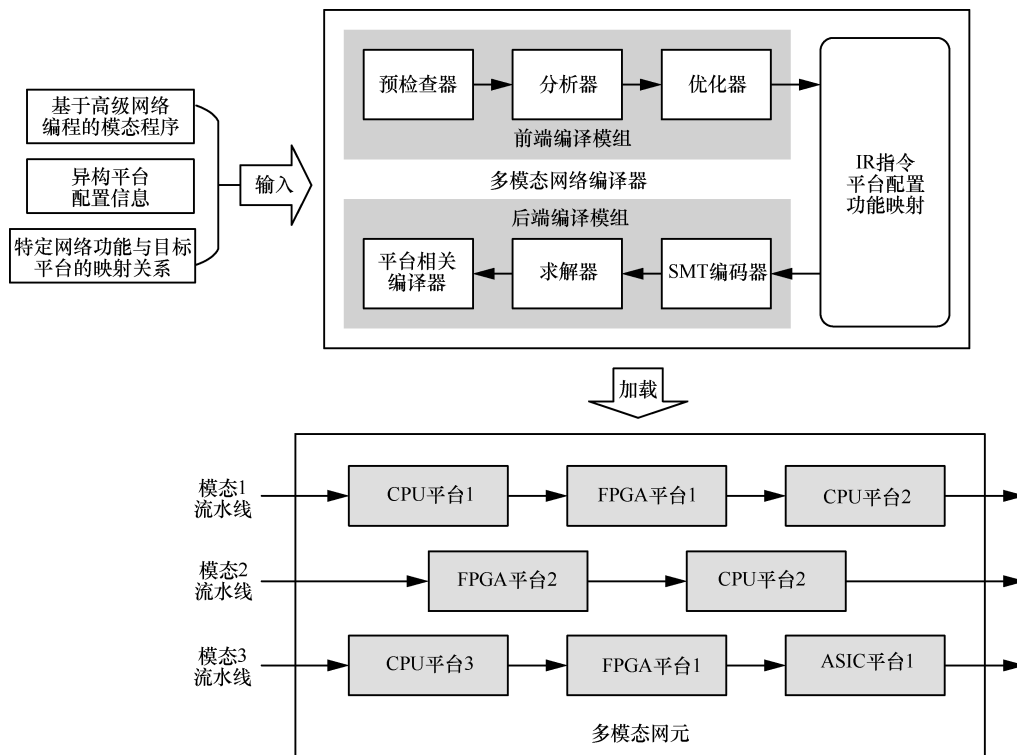


图3 面向异构软硬件资源的网络模态编译流程

服务维护与升级能够在系统运行时进行,不影响底层网络流量的转发。

### 2.3 多样化异构软硬件资源的可编程调度与互联

在多模态网络环境中,多模态网元设备可通过动态编程、重构等构造方式,实现特定功能的模态实现。在此过程中,多模态网元设备需协同调度内部的计算、存储、转发甚至是安全处理等资源,通过对资源的状态进行细粒度感知,在异构资源间形成可编程的调度和互联通道,从而构建出满足模态需要的数据处理逻辑。构建上述可编程互联通道的基础,不是各种异构资源间使用传统的固定线路连接方式,而是采用可编程的方式构建灵活、动态的连接通道,通过在资源间流转带有标识的报文来满足消息交互需求。

为此,多模态网元配置模组向硬件转发功能模组配置路由,通过软件定义功能模组的互联通道完成复杂的报文处理逻辑。如图 4 所示,各异构资源之间可编程互联机制通过封装数据报头的方式体现。当报文到达后,各资源实体根据报头语义进行相应的处理,完成后更新报头内容传递给下一个资源实体进行后续处理。该机制的优势在于可以通过定义报头格式引入新的控制语义,以适应资源实体的替换,支撑节点功能的改变。

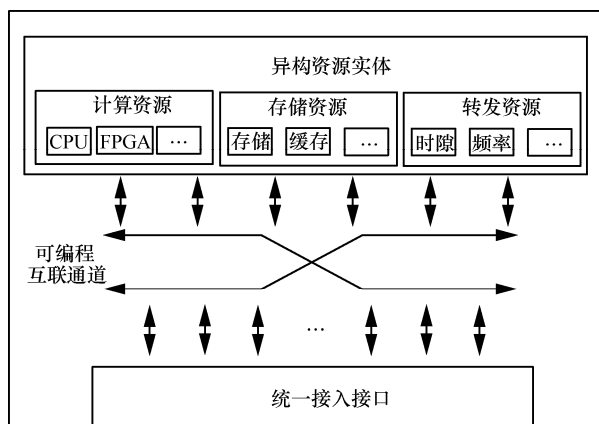


图 4 多模态网元内异构资源间的可编程调度和互联机制

多模态网元设备通过构建灵活可变的异构资源连接关系,使网络模态在安装部署时只需要关注所需资源的类型、能力等,而不再受平台灵活性限制,甚至可以动态调整资源的配置状态,增加或删除平台内部资源,从而实现模态的动态、灵活部署。

### 2.4 融合异构资源的内生安全数据流处理

PINE 充分实践“软件定义”的思想来实现数据平面功能的全维可定义,但数控分离与网络可编

程给网络运行环境带来了新的安全问题——开放的技术组件堆叠极易导致网元设备遭到软硬件漏洞、故障导致的蓄意攻击,无法提供正常的网络服务。为此,多模态网元从报文流水线层面引入内生安全构造技术<sup>[16]</sup>,针对单个模态流水线任务的关键功能,在保持功能等价的条件下分别在多个异构平台上构建冗余执行体;在此基础上,通过负反馈控制机制赋予执行体结构表征的不确定性,以抵御开放式产业生态环境下的已知/未知漏洞、后门,为多模态网络业务的长效运行提供稳固保障。

图 5 展示了融合异构资源的内生安全数据流处理模型,主要包括解析分发器、逆解析裁决器、单模态流水线结构、负反馈控制器等模块。其中,解析分发器、逆解析裁决器均对应于多模态网元的高性能转发功能模组,而负反馈控制器对应于多模态网元逻辑配置模组。该模型将网络模态的功能部署到不同的软硬件资源上,从而构建出多条处理逻辑完全相同的模态处理流水线。在数据流进入多模态网元后,解析分发器根据负反馈控制指令(即内部处理逻辑控制)将解析后的数据分发至多个流水线。每个流水线完成对报文处理流程后,逆解析裁决器再次对报文进行整合,依出端口配置转发流量。

在上述过程中,若逆解析裁决器发现某一流水线发生非正常扰动,则报告至负反馈控制器。负反馈控制器根据裁决结果对流水线资源池中输出结果异常的流水线进行清洗,并上报至控制器完成进一步的模态重构机制。因此,基于动态异构冗余的融合异构资源的内生安全数据流处理机制,从根本上抑制了抑制随机性失效和人为蓄意扰动,构建了“高可靠、高可信、高可用”的多模态网络环境。

## 3 异构软硬件资源分配方法

在多模态网络环境中,针对多元化业务构建的模态流水线,其资源需求呈现差异化特征。具体到单个多模态网元中,相同的物理资源在多种模态配置之间可能存在竞争关系。例如,多核 CPU 平台执行基于哈希加密函数的计算任务和基于深度报文检测的字符搜索任务时均消耗大量的内存资源。对于任意软件功能实例,构建不同模态流水线的所需计算、存储、转发功能容量均不能超越片上资源总和。因此,给定资源总量,多模态网元需针对网络模态对异质资源的需求进行动态、高效的资源分配。

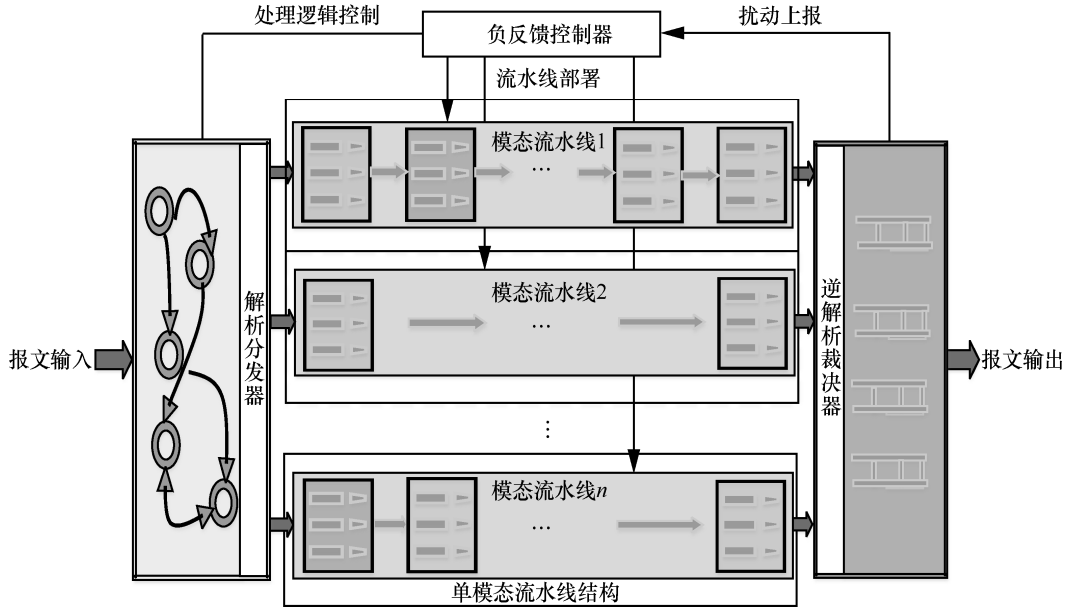


图5 融合异构资源的内生安全数据流处理模型

### 3.1 多模态网元资源分配建模

多模态网元  $v$  所能提供的物理资源集合可表示为  $A_v$ ，其中，在多模态网元  $v$  中内嵌类型  $r$  的资源集合为  $R_v^r, \forall r \in A_v$ 。以计算、存储、转发（网络）资源为例，为建模方便，这里将异质异构的同类型资源视作等效的，则多模态网元  $v$  中资源图谱可以抽象为  $A_v = (R_v^c, R_v^s, R_v^n)$ ，其中， $R_v^c$ 、 $R_v^s$ 、 $R_v^n$  分别表示计算资源、存储资源、转发资源。不同模态间的资源保障根据不同类型资源（计算、存储、转发等）提供多种资源组合调度方案，这里假设模态  $i$  存在  $m$  种资源决策  $\Gamma_v^i = \{\phi_m^i, \dots, \phi_m^i\}$ ，其中， $\phi_m^i = f(r_v^c, r_v^s, r_v^n)$  表示每种资源决策均可保障模态正常提供服务所需要的计算、存储、转发等组合资源。 $\theta_m^i$  表示模态  $i$  是否选择第  $m$  个资源保障方案  $\phi_m^i = f(r_v^c, r_v^s, r_v^n)$

$$\theta_m^i = \begin{cases} 1, & \text{模态 } i \text{ 选择方案 } m \\ 0, & \text{其他} \end{cases} \quad (1)$$

针对多模态网元的资源约束条件如下。

1) 所有模态的计算资源不超过网元的最大计算能力，包括运行模态运维逻辑需要的计算资源  $\phi_m^i = f(r_v^c)$  和处理对应模态业务所需要的计算资源。

2) 所有模态的存储资源不超过网元的存储资源，包括运行模态运维逻辑需要的存储资源和处理对应模态业务所需要的存储资源。

3) 所有模态所分配的带宽资源不超过网元的最大交换带宽  $\phi_m^i = f(R_v^n)$ 。

上述约束条件具体可表示为

$$\sum_{\forall i} \sum_{\forall m} \theta_m^i \phi_m^i(r_v^c) + \sum_{\forall i} \sum_{\forall m} \sum_{\forall k} \theta_m^i y_{i,k}^c \lambda_i^k(t) \leq R_v^c \quad (2)$$

$$\sum_{\forall i} \sum_{\forall m} \theta_m^i \phi_m^i(r_v^s) + \sum_{\forall i} \sum_{\forall m} \sum_{\forall k} \theta_m^i y_{i,k}^s \lambda_i^k(t) \leq R_v^s \quad (3)$$

$$\sum_{\forall i} \sum_{\forall m} \theta_m^i \phi_m^i(r_v^n) \leq R_v^n \quad (4)$$

其中， $\lambda_i^k(t)$  表示模态服务请求到达率， $k$  表示相应的业务， $y_{i,k}^c$  表示模态  $i$  服务业务请求  $k$  所需要的计算资源， $y_{i,k}^s$  表示模态  $i$  服务业务请求  $k$  所需要的存储资源。式(2)和式(3)相加的因子中前者是运行模态运维逻辑需要的计算资源和存储资源，后者是处理对应模态业务所需要的计算资源和存储资源。

这里，以模态的处理时延作为衡量模态服务性能的一个指标，每个模态的服务请求到达率为  $\lambda_i(t)$ ，在多模态网元中，各模态的服务请求相互独立。应用 Little 原理<sup>[17]</sup>，模态  $i$  的服务处理时延为

$$\mathcal{Q}_i(t) = \frac{1}{\alpha_i(\bar{r}_i) - \lambda_i(t)} \quad (5)$$

其中， $\alpha_i(\bar{r}_i)$  为模态  $i$  的处理能力，与其所分配的计算资源  $\bar{r}_i$  有关。则多模态网元  $v$  的平均处理时延为  $\zeta_n(t)$ ，与其所承载的每个模态的处理时延相关。

$$\zeta_v(t) = \frac{\sum_{\forall i} g_i(t) \lambda_i(t)}{\sum_{\forall i} \lambda_i(t)} \quad (6)$$

本文将多模态网元看作一个计算-存储-转发融合的系统，其运维成本和服务性能由其内嵌的异质异构的计算资源、存储资源、转发资源的融合协调能力综合体现。需要特别说明的是，网络中的存储、计算与转发资源具备能力补偿的关系，例如，缓存在不同网络节点或位置的部署策略，以及缓存的存取速率和容量，均会对相关节点的转发资源开销产生影响；同样，计算资源的部署位置和能力也会对转发资源产生影响，转发策略的不同也会导致不同带宽的流量被分配到具备差异化计算能力的资源上。因此，网络模态的资源分配情况不仅会在业务流量变化时调整，也会在异构资源间置换时进行调整。

综上，本文通过分析计算-存储-转发资源的内在转化关系，通过计算、存储、转发资源的一体化调度研究多模态网元系统的效用函数和服务性能，其优化目标旨在确定多模态网元中所有模态的资源保障决策变量  $\{P | \theta_m^i = 1, \forall i, m\}$ ，在保证确保模态业务时延约束的同时尽可能减少网络资源消耗。其中，资源消耗和计算、存储、转发的资源使用量有关，可表示为  $\Psi_v(\tilde{r}_v^c, \tilde{r}_v^s, \tilde{r}_v^n)$ 。因此，优化目标函数可表示为

$$\begin{aligned} & \min \tau \Psi_v(\tilde{r}_v^c, \tilde{r}_v^s, \tilde{r}_v^n) + (1 - \tau) \zeta_v(t) \\ & \text{s.t.} \\ & \sum_{\forall i} \sum_{\forall m} \theta_m^i \phi_m^i(r_v^c) + \sum_{\forall i} \sum_{\forall m} \sum_{\forall k} \theta_m^i y_{i,k}^c \lambda_i^k(t) \leq R_v^c \\ & \sum_{\forall i} \sum_{\forall m} \theta_m^i \phi_m^i(r_v^s) + \sum_{\forall i} \sum_{\forall m} \sum_{\forall k} \theta_m^i y_{i,k}^s \lambda_i^k(t) \leq R_v^s \\ & \sum_{\forall i} \sum_{\forall m} \theta_m^i \phi_m^i(r_v^n) \leq R_v^n \end{aligned} \quad (7)$$

### 3.2 算法描述

多样化网络模态的共生共存，本质上是 NP 难问题。随着模态业务规模的增加，如何在较短时间内得到最优解是需要考虑的问题。本文设计面向多模态业务流的计算-存储-转发一体化资源优化算法 Integrated-CSN，该算法借助模拟退火算法<sup>[18]</sup>进行模型求解，相关参数设置如下：初始温度  $T_0$  为 900 K，终止温度  $T_{\text{end}}$  为 1 K，迭代次数  $N$  为 80，降温速度  $R$  为每 80 次降温 0.95 K。需要说明的是，本文重点关注的是如何对多模态网络环境中的异构软硬件资源进行分配建模，求解算法并不局限于

本文所采用的算法，也可采用其他求解逻辑或数据空间不同的算法，如粒子群算法、增强学习算法等。

Integrated-CSN 算法流程如下。

#### 1) 初始资源分配策略

Integrated-CSN 算法首先产生初始的资源分配策略  $\phi_{\text{init}}$ 。Integrated-CSN 算法的求解速率与初始状态有关，部署时的初始解设置为针对每种模态运维逻辑随机选择一种资源保障方案，每种模态的业务处理资源初始权重设为平均分配。当网络状态再次改变时，Integrated-CSN 算法的初始资源分配策略即上一个网络状态的稳定解。

#### 2) 新的资源分配策略的产生

Integrated-CSN 算法的求解迭代速率取决于新的资源分配策略的产生方式，由于解空间的维度较多，传统的完全随机产生新解方式效率较低。为了提高求解效率，这里针对 Integrated-CSN 算法的优化目标，设计一种新的搜索策略，每次产生新解仅针对一种模态随机选择一种保障方案，同时，为了保障每种模态的服务质量，根据模态业务流的处理时延为每种模态分配相应的权重值，每个模态的权重是相对，每个模态权重的比值作为模态处理业务的计算资源的分配方案。

#### 3) 按 Metropolis 准则接受新解

Integrated-CSN 算法基于模拟退火算法，按照 Metropolis 准则<sup>[18]</sup>判断新解是否应该接受。该解的接受概率为

$$P(T, \Omega, \Omega') = \begin{cases} 1, & \Omega' < \Omega \\ e^{-\frac{\Omega' - \Omega}{T}}, & \Omega' \geq \Omega \end{cases} \quad (8)$$

Integrated-CSN 算法通过改进模拟退过算法来求解最优的资源调度策略，算法运行过程中，在每个温度下运行  $N$  次，根据步骤 2) 中设计的搜索策略产生随机解，并按 Metropolis 准则以一定概率接受新解，防止陷入局部最优解，当算法温度降至  $T_{\text{end}}$  以下时，输出搜索的最优的资源分配策略。

## 4 实验验证

### 4.1 仿真参数设置

网元节点的计算资源量设为 2 000 MIPS（每秒执行百万指令数），存储资源量设为 10 TB，交换带宽容量设为 1 Gbit/s。网元内部支持 3 种模态数据处理逻辑，参照网络中一些典型的服务功能参数<sup>[19]</sup>对每种模态数据处理逻辑的参数进行设置：每种模

态数据处理逻辑所需计算资源为 20~80 MIPS, 存储资源设置为 1~100 GB。为了体现算法灵活调度资源的能力, 实验过程中, 随机产生服务请求强度范围为 200~2 200 个, 各模态业务量比例按 3:2:1 注入网元, 每个模态业务请求的计算资源在 2~8 MIPS 服从均匀分布, 请求带宽在 0.4~0.7 Mbit/s 服从均匀分布, 请求存储资源在 0.5~5 GB 服从均匀分布。

#### 4.2 对比方案

Static-CSN<sup>[20]</sup>是一种静态的资源分配算法, 参照基于容器的操作系统虚拟化技术预先分配固定资源给每种模态以实现物理资源隔离, 这里网元内部计算、存储、转发资源均平均分配, 不会随网络状态改变而改变。

DRF-CSN<sup>[21]</sup>针对每种模态的主导资源执行 DRF (dominant resource fairness) 算法 (DRF 是一种通用的多资源的 max-min fairness 分配策略), 支持异构环境下多种类型资源的调度。其核心思想是一个模态的资源分配由该模态的主导资源 (或计算或存储或带宽) 决定。

Integrated-CSN 是本文提出的计算存储转发资源的一体化调度算法, 同时对计算、存储、转发资源进行联合分配和置换优化。

#### 4.3 仿真结果

本文针对异构资源融合系统中的多资源分配算法进行对比, 记录网元的平均处理时延和网元资源开销, 仿真结果如下。

##### 1) 不同服务请求强度测试

不同服务请求强度下 3 种算法性能对比如图 6 所示。从图 6(a)可以看出, 所提 Integrated-CSN 算法的网元平均处理时延最小, DRF-CSN 次之, Static-CSN 的平均处理时延最大。当服务请求强度较小时, 3 种资源保障方案的平均处理时延差距较小。随着服务请求强度的增加, Static-CSN 由于其采用资源平均分配, 不能根据网络服务请求强度的改变而变化, 当每种模态达到资源瓶颈时, 其平均处理时延随着服务强度增大而快速增加。DRF-CSN 算法对每种模态的主导资源 (计算资源或存储资源或转发资源) 分别展开资源优化分配, 相较于 Static-CSN, 减少了处理时延, 但其难以更好地适应不同模态的业务强度从而有针对性地部署资源。Integrated-CSN 通过计算、存储、转发资源的联合优化, 通过不同资源的融合弥补, 可实现各模态的最佳资源保障, 降低了平均处理时延。

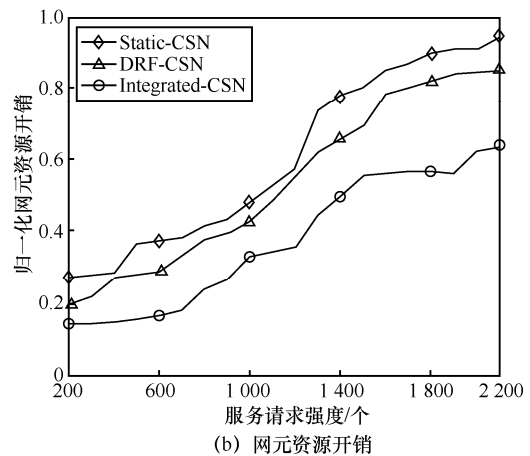
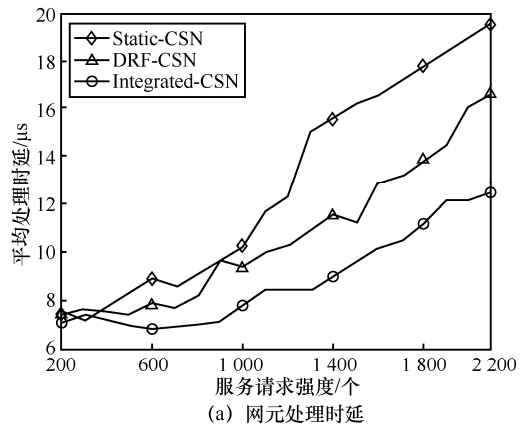


图 6 不同服务请求强度下 3 种算法性能对比

从图 6(b)中可以看出, 随着服务请求强度的增加, 3 种算法的资源开销都在升高。Static-CSN 采用固定的资源保障策略, 不能随模态业务负载的变化动态调整资源, 其资源开销较高。DRF-CSN 虽然能较大幅度地满足模态业务主导资源的需求, 但仍会造成其他资源的利用率较低, 其资源开销也偏高。Integrated-CSN 可根据模态业务量动态选择多种资源保障策略, 实现计算资源、存储资源和转发资源的动态转化, 其资源利用率最高, 因此, 相同的业务量其资源开销最小。

##### 2) 相同服务请求强度对比

相同服务请求强度下 3 种算法性能对比如图 7 所示。实验过程中, 在每一个时隙内, 向网元注入的服务请求总数相等, 但各模态业务比例随机变化。从图 7 中可以看出, 相同的业务强度下, Integrated-CSN 算法的平均处理时延和资源开销最低。Static-CSN 和 DRF-CSN 算法的时延较高且抖动较大, 这是由于 Static-CSN 采用固定的资源分配难以保障不同模态业务的资源需求, 导致网元性能不稳定, 造成较大的资源开销; DRF-CSN 以模态

主导资源为依据在一定程度上可均衡网元资源, 适应模态业务变化, 但当某一模态业务量占比较大时, 其仍会导致其他资源的瓶颈效应。Integrated-CSN 算法可以根据不同模态的业务量对计算、存储、转发资源联合优化, 能更充分地利用资源, 保障了各模态处理时延需求的同时, 降低了资源开销。

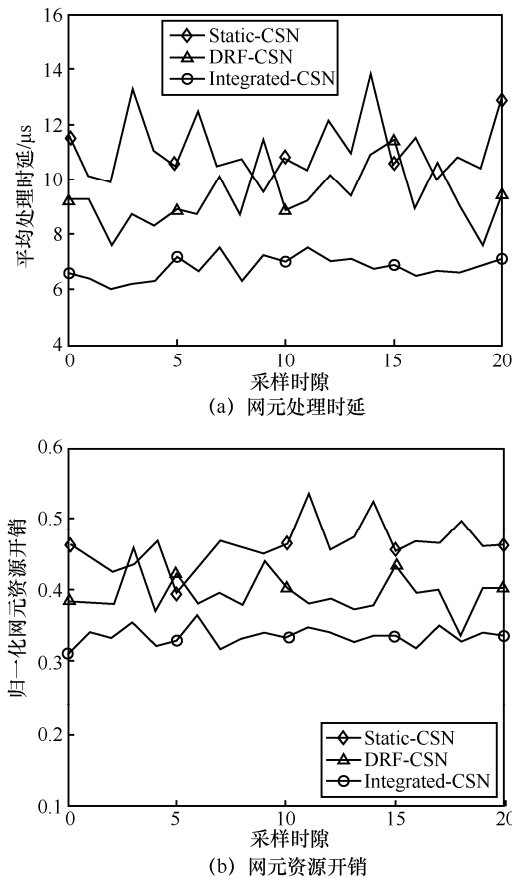


图 7 相同服务请求强度下 3 种算法性能对比

## 5 结束语

以“技术体系与支撑环境分离”的多模态网络发展范式为指导, 本文围绕基于领域专用软硬件协同的多模态网络环境构造技术开展研究。首先, 对面向领域专用的各种异构存储、计算、转发资源进行统一表达、利用和组合, 从基于多模态网络环境的垂直行业应用需求出发, 提出了一种基于领域专用软硬件协同的多模态网元设计方案, 讨论了多模态网络环境构造关键技术, 为多元化网络模态共生共存提供环境架构设计支撑。之后, 针对多样化网络模态对网元内异构资源竞争的问题, 提出了一种

存算转一体的异构软硬件资源分配方法, 可根据模态业务量动态调整资源分配和置换策略, 实现异构资源的高效利用。

## 参考文献:

- [1] 胡宇翔, 伊鹏, 孙鹏浩, 等. 全维可定义的多模态智慧网络体系研究[J]. 通信学报, 2019, 40(8): 1-12.  
HU Y X, YI P, SUN P H, et al. Research on the full-dimensional defined polymorphic smart network[J]. Journal on Communications, 2019, 40(8): 1-12.
- [2] 郑秀丽, 蒋胜, 王闯. NewIP: 开拓未来数据网络的新连接和新能力[J]. 电信科学, 2019, 35(9): 2-11.  
ZHENG X L, JIANG S, WANG C. NewIP: new connectivity and capabilities of upgrading future data network[J]. Telecommunications Science, 2019, 35(9): 2-11.
- [3] 杨仁宇, 韩毅刚, 章帆, 等. 基于 SDN 的内容中心网络研究综述[J]. 计算机科学, 2019, 46(1): 13-20.  
YANG R Y, HAN Y G, ZHANG F, et al. Survey of content centric network based on SDN[J]. Computer Science, 2019, 46(1): 13-20.
- [4] 侯赛凤, 田乐, 胡宇翔, 等. 面向多源信息采集的地理位置路由寻址机制[J]. 北京邮电大学学报, 2021, 44(2): 1-7.  
HOU S F, TIAN L, HU Y X, et al. Geographic routing and addressing scheme for multi-source data retrieval[J]. Journal of Beijing University of Posts and Telecommunications, 2021, 44(2): 1-7.
- [5] CHEN Q X, YU F R, HUANG T, et al. An integrated framework for software defined networking, caching, and computing[J]. IEEE Network, 2017, 31(3): 46-55.
- [6] 邬江兴, 胡宇翔. 网络技术体系与支撑环境分离的发展范式[J]. 信息通信技术与政策, 2021, 47(8): 1-11.  
WU J X, HU Y X. The development paradigm of separation between network technical system and supporting environment[J]. Information and Communications Technology and Policy, 2021, 47(8): 1-11.
- [7] HU Y X, LI D, SUN P H, et al. Polymorphic smart network: an open, flexible and universal architecture for future heterogeneous networks[J]. IEEE Transactions on Network Science and Engineering, 2020, 7(4): 2515-2525.
- [8] 冯一飞, 丁楠, 叶钧超, 等. 领域专用低延迟高带宽 TCP/IP 引擎设计与实现[J]. 计算机工程, 2021: doi.org/10.19678/j.issn.1000-3428.0063200.  
FENG Y F, DING N, YE J C, et al. Design and implementation of domain-specific low-latency and high-bandwidth TCP/IP offload engine[J]. Computer Engineering, 2021: doi.org/10.19678/j.issn.1000-3428.0063200.
- [9] PAT B, DAN D, GLEN G, et al. P4: programming protocol-independent packet processors[J]. Computer Communication Review, 2014, 44(3):87-95.
- [10] Broadcom. Broadcom's new Trident 4 and Jericho 2 switch devices offer programmability at scale[R]. 2020-06-18.
- [11] SCHOLZ D, OELDEMANN A, GEYER F, et al. Cryptographic hashing in P4 data planes[C]//Proceedings of 2019 ACM/IEEE Symposium on Architectures for Networking and Communications Systems. Piscataway: IEEE Press, 2019: 1-6.
- [12] MAHMOOD A, CASETTI C, CHIASSERINI C F, et al. Efficient caching through stateful SDN in named data networking[J]. Transac-

- tions on Emerging Telecommunications Technologies, 2018, 29(1): e3271.
- [13] SONI H, RIFAI M, KUMAR P, et al. Composing dataplane programs with  $\mu$ P4[C]//Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication. New York: ACM Press, 2020: 329-343.
- [14] YU Z, JUN B. ClickP4: towards modular programming of P4[C]//Proceedings of the SIGCOMM Posters and Demos. New York: ACM Press, 2017: 22-24.
- [15] GAO J Q, ZHAI E N, LIU H H, et al. Lyra: a cross-platform language and compiler for data plane programming on heterogeneous ASICs[C]//Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication. New York: ACM Press, 2020: 435-450.
- [16] 邬江兴. 网络空间拟态防御导论[M]. 北京: 科学出版社, 2017.  
WU J X. Introduction to cyberspace mimic defense[M]. Beijing: Science Press, 2017.
- [17] LITTLE J D C, GRAVES S C. Little's law[M]. Berlin: Springer, 2008.
- [18] BOULEIMEN K, LECOCQ H. A new efficient simulated annealing algorithm for the resource-constrained project scheduling problem and its multiple mode version[J]. European Journal of Operational Research, 2003, 149(2): 268-281.
- [19] PEI J N, HONG P L, XUE K P, et al. Efficiently embedding service function chains with dynamic virtual network function placement in geo-distributed cloud system[J]. IEEE Transactions on Parallel and Distributed Systems, 2019, 30(10): 2179-2192.
- [20] SOLTESZ S, PÖTZL H, FIUCZYNSKI M E, et al. Container-based operating system virtualization: a scalable, high-performance alternative to hypervisors[C]//Proceedings of the 2nd ACM SIGOPS/EuroSys European Conference on Computer Systems 2007. New York: ACM Press, 2007: 275-287.
- [21] BIAN S M, HUANG X, SHAO Z Y. Online task scheduling for fog computing with multi-resource fairness[C]//Proceedings of 2019 IEEE 90th Vehicular Technology Conference. Piscataway: IEEE Press, 2019: 1-5.

## [作者简介]



胡宇翔(1982-),男,河南周口人,博士,信息工程大学教授、博士生导师,主要研究方向为新型网络体系结构、路由与交换技术。



崔子熙(1996-),男,河南焦作人,信息工程大学博士生,主要研究方向为新型网络体系结构、路由与交换技术。



李子勇(1995-),男,安徽蚌埠人,信息工程大学博士生,主要研究方向为新型网络体系结构、路由与交换技术。

董永吉(1983-),男,吉林汪清人,博士,信息工程大学副研究员,主要研究方向为可编程网络数据平面、协议无关转发技术。

崔鹏帅(1990-),男,河南安阳人,博士,信息工程大学副研究员,主要研究方向为可编程网络数据平面、协议无关转发技术。

邬江兴(1953-),男,安徽金寨人,博士,中国工程院院士,信息工程大学教授、博士生导师,主要研究方向为网络通信与安全。