

大规模网络向IPv6单栈演进的技术方案



Technical Solution of Transition to Large-Scale IPv6-Only Networks

解冲锋/XIE Chongfeng¹, 李星/LI Xing², 李震/LI Zhen³,
余勇志/YU Yongzhi⁴

(1. 中国电信集团有限公司, 中国 北京 102209;
2. 清华大学, 中国 北京 100084;
3. 下一代互联网国家工程中心, 中国 北京 100176;
4. 中国电信股份有限公司重庆分公司, 中国 重庆 401121)
(1. China Telecom Group Co., Ltd., Beijing 102209, China;
2. Tsinghua University, Beijing 100084, China;
3. China Future Internet Engineering Center, Beijing 100176, China;
4. China Telecom Chongqing Branch, Chongqing 401121, China)

DOI: 10.12142/ZTETJ.202201012

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20220217.1716.004.html>

网络出版日期: 2022-02-18

收稿日期: 2021-12-15

摘要: 中国互联网全面进入了互联网协议第4版/互联网协议第6版(IPv4/IPv6)双栈运行阶段, 未来如何向IPv6单栈网络演进是新的技术和产业挑战。结合中国的最新政策要求, 在分析已有IPv6单栈过渡技术的基础上, 提出了面向大规模网络的多域纯IPv6网络的总体框架、演进路线和相关方案, 并给出发展IPv6单栈网络的策略建议。

关键词: IPv6单栈; 全局映射规则; 多域; IPv6段路由(SRv6); 边界网关协议4.0扩展(BGP4+)

Abstract: With the Internet in China fully entering the Internet protocol version 4/Internet protocol version 6 (IPv4/IPv6) dual-stack operation stage, how to evolve into IPv6 single-stack network in the future is a new technical and industrial challenge. Combined with the latest policy requirements of China and based on analyzing the existing IPv6 single-stack transition technology, the overall framework, evolution route, and related solutions of multi-domain IPv6-only network for large-scale networks are put forward, and the strategic suggestions for the development of IPv6 single-stack networks are proposed.

Keywords: IPv6-only network; global mapping rule; multi-domain; SRv6; BGP4+

作为新一代互联网网络协议, 互联网协议第6版(IPv6)是全球互联网升级演进的必然趋势和网络技术创新的重要方向, 其地位和趋势在全球已基本无争议。IPv6的发展在中国一直受到高度重视, 特别是自从2017年中共中央办公厅、国务院办公厅印发《推进互联网协议第六版(IPv6)规模部署行动计划》以后, 中国的IPv6部署大大加速, 并已实现全面部署, 网络基础设施也已进入互联网协议第4版(IPv4)/IPv6双栈运行阶段。

随着IPv6的规模应用, IPv6单栈网络的部署已经成为新的阶段性重要目标。从双栈向IPv6单栈演进的趋势日益明显, 并成为IPv6与经济社会深度融合的先决条件。在国际上, 发达国家也在积极推进网络和业务的IPv6单栈化。例如, 2020年11月美国白宫管理和预算办公室发布指南, 要求“美国各机构尽快完成向IPv6的过渡, 确保到2025财年

末, 联邦网络上超过80%的IP资源是IPv6单栈”。2021年7月, 中国中央网络安全和信息化委员会办公室、国家发展和改革委员会、工业和信息化部联合发布的《关于加快推进互联网协议第六版(IPv6)规模部署和应用工作的通知》指出, 增强IPv6网络互联互通能力, 积极推进IPv6单栈网络部署, 是我国未来推进IPv6工作的重点任务之一^[1]。

网络协议栈是互联网运行的基石。单一网络协议栈是互联网的最基本要求, 是互联互通的前提。IPv4/IPv6双栈网络维护成本高, 安全风险大。这不仅给设备提出很高的要求, 还给业务发展带来困惑甚至阻碍。从双栈向IPv6单栈演进有利于网络的运营维护, 并且具有多个优点: 单协议栈的运营工作量和成本比双栈显著降低; 单栈网络的风险暴露面减少, 使得网络更加安全; 网络架构更简单, 路由表数量减少, 对设备的要求降低, 转发能力得到提升。对于互联网

业务来说, IPv6单栈运营更为容易, 减少了在支持IPv4能力方面的不必要成本投入。

本文围绕中国IPv6发展的总体规划, 分析了大规模网络向IPv6单栈演进的内涵和思路, 然后面向大规模、多域、多场景网络提出了向IPv6单栈演进的网络总体架构和实施策略建议。

1 IPv6单栈的概念分析

通俗地说, IPv6单栈网络就是在网络中关闭IPv4协议栈并以IPv6协议为核心进行编址、路由和转发的网络, 它支持基于SRv6^[2]的能力创新。IPv6单栈是网络架构和能力的全新变化。IPv6单栈化的目的是构建极简、智能、安全、绿色的新型网络, 它是IPv6发展的最终方向。作为新型的网络基础设施, IPv6单栈网络主要承载各种互联网业务。按照互联网业务对于协议支持的程度不同, IPv6单栈网络呈现两种不同的形态: 综合态和终极态。

(1) 综合态。网络不仅要承载原生IPv6业务流量, 还要承载现存的IPv4业务流量, 并在边缘支持IPv4业务流量的接入和穿越。如图1所示, 在T2~T3阶段, SRv6等IPv6创新技术逐渐得到广泛应用。

(2) 终极态。在互联网中的业务全部IPv6化的情况下, 网络本身不再接入IPv4业务流。在图1T3以后的状态中, 网络将关闭边缘的IPv4属性。T3以后的状态是向IPv6网络过渡的终极态。

从双栈进入到IPv6单栈的阶段, 需要经过T1~T2之间的IPv6单栈部署阶段。这一阶段所需的时间一般为5~6年(时间长短受到实际IPv6流量占比大小的影响)。网络向IPv6单栈的推进也会促进互联网业务向IPv6迁移, 进而对网络中IPv6流量的占比产生影响, 因此单栈部署期是网络和业务互动的过程。

目前来看, 互联网业务要全部实现IPv6化还需要较长时间, 因此短期内实施终极态IPv6单栈是不现实的; 但如

果不推进IPv6单栈化而让双栈长久持续下去, 那么业务的IPv6改造动力就会更加不足, 这会影响IPv6流量的提升和生态的形成。综合以上因素, 当前推动综合态的IPv6单栈方案更具有现实意义和紧迫性。

综合态IPv6单栈的关键问题是, 在逐渐关闭IPv4协议栈以后, 网络中设备如何有效地支持剩余IPv4业务的承载, 以确保用户体验不降低。

截至目前, 互联网工程任务组(IETF)针对不同的场景设计了多种纯IPv6过渡方案, 如双栈(DS)-Lite^[3]、轻量级4over6^[4]、翻译型地址+端口映射转换机制(MAP-T)^[5]、封装型地址+端口映射转换机制(MAT-E)^[6]、NAT64/464XLAT^[7]等。这些方案均将IPv4作为IPv6网络的一种业务, 因此属于综合态的IPv6单栈方案, 并已获得商用。464XLAT是移动网中最成熟的IPv6单栈方案, 该方案已获得苹果iOS和安卓等操作系统的支持, 并得到了广泛应用。在实践方面, 中国分别在不同场景进行了IPv6单栈的试点研究。例如, CERNET是国际上首次采用IPv6单栈的网络, 也是迄今为止全球规模最大的IPv6单栈主干网。此外, 中国有些运营商也开始在4G和5G移动网络上进行IPv6单栈试点。

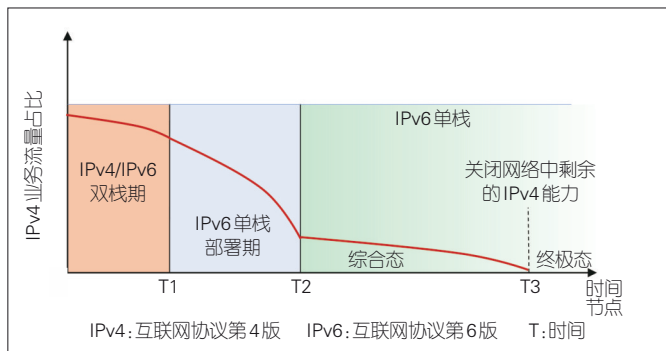
大型网络的IPv6单栈化必须具有如下几个特性。

- 开放性: 支持与外部各种网络(IPv4单栈、IPv4/IPv6双栈和IPv6单栈)的互通, 确保对网络内外业务的正常访问, 使用户体验不降低;
- 协议单纯性: 以IPv6协议为唯一基础协议进行编址、路由和转发等;
- 支持IPv6基础上的技术创新: 与智能化网络架构协同, 可根据业务需求和网络维护需求实现基于SRv6的流量调度和网络服务化;
- 地址真实可信: 支持互联网真实源地址的精确定位和地址溯源, 并可在地址中嵌入用户或者终端身份信息, 增强对网络和用户的安全管控能力。

目前, 中国数字化转型方兴未艾, 云计算、边缘计算和工业互联网等业务都需要一个先进的网络基础设施来支撑。这个基础设施应该是以IPv6为核心协议的新一代基础网络, 应避免进入IPv4误区。因此, 网络的IPv6单栈化也和数字化转型密切相关, 其在中国的部署具有必要性和紧迫性。另外, 中国推动互联网的单栈化有利于引领国际IPv6发展潮流, 可为全球互联网发展贡献力量。

2 多域网络引入IPv6单栈时的主要问题

如前所述, 目前业界已有多种IPv6单栈过渡技术方案。



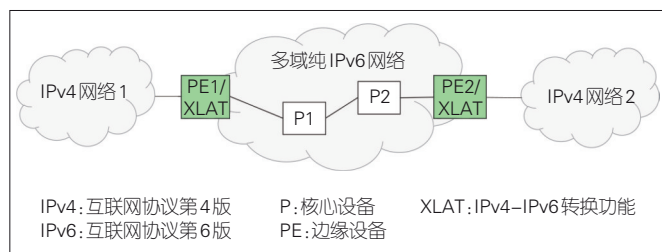
▲图1 网络向IPv6单栈演进的路径

这些方案在支持IPv4业务的承载时,需要不同类型的IPv4和IPv6转换技术。例如,464XLAT采用有状态NAT64的翻译技术,IVI采用无状态翻译NAT64技术,DS-Lite采用基于地址族转换路由器(AFTR)的4over6隧道技术,而主干网则采用GRE隧道或无状态翻译技术等。对于运营商来说,在引入IPv6单栈时面临的首要问题是采用何种技术。通常情况下,大规模IP网络是由多个自治系统组成的。各自治系统服务不同的场景,而且经常由不同的组织来管理,并采用不同的路由和安全策略。即便是在同一个运营商内,每个自治系统都由不同的机构或部门来管理和运营。在引入IPv6单栈方案时,如果自治系统之间缺乏协同,就需要在数据路径上将IPv6数据包转回到IPv4,然后在下一个域又转回IPv6。这样网络中就会出现较多功能不同的IPv4-IPv6数据包转换网关,而且转换次数会随着自制系统数量的增加而增加。如图2所示,在含3个域的网络上,各自治系统采用不同的IPv6单栈过渡技术。IPv4业务数据在跨域时需要恢复出IPv4数据包,这会导致网络中出现5次基于网关的IPv4和IPv6转换。过多的IPv4和IPv6转换使网络变得复杂,也使设备投资相应增大。因此,我们迫切需要协同域间的单栈方案,以消除不必要的转换功能,提高数据转发效率。

3 多域纯IPv6技术总体方案

针对以上问题,本文提出了面向大规模、多域和多场景的纯IPv6组网架构及相关技术方案,为不同自治系统间的IPv6单栈方案协同组网提供框架支持。

多域纯IPv6网络的目标是以IPv6为基础协议构建多域互连的网络基础设施,在进行IPv4业务的承载时遵循“IPv4 As A Service”的思路,即把IPv4作为一种业务。对于客户侧发出的IPv4数据包,为了不让IPv4数据包进入网络,网络边缘设备会将IPv4数据包适配成IPv6数据包,如图3所



▲图3 基于XLAT传输IPv4业务数据的过程

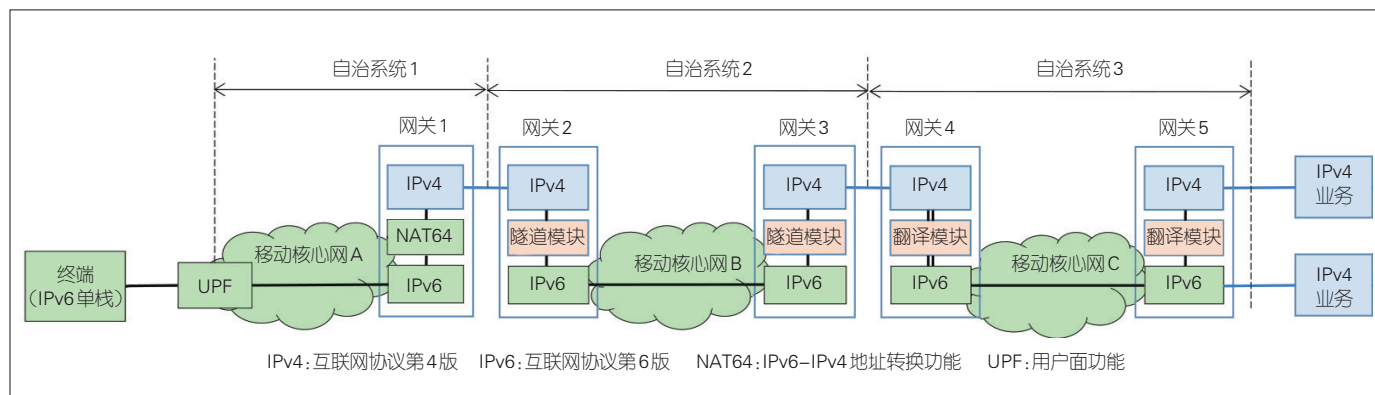
示。为支持在多域纯IPv6网络中传送IPv4业务数据包,在入口边缘设备(PE)上运行的XLAT需要将IPv4业务数据包转换成IPv6数据包,并通过IPv6路由系统将该数据包送到正确的出口PE,然后IPv6数据包被恢复成IPv4数据包;而在域间IPv6数据包不需要被恢复成IPv4数据包,这样系统最多需要两次IPv4和IPv6的转换。为此,各域间需要“达成转换共识”,即采用全局映射规则。具体来说,多域纯IPv6技术方案包括如下几个部分:

(1) 基于全局映射规则的IPv4-IPv6地址映射

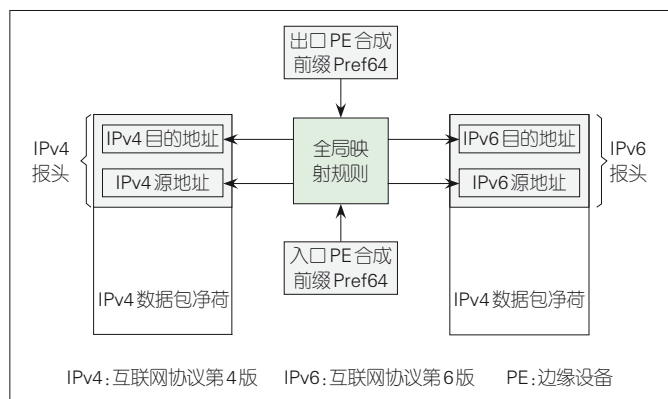
为了提升转化效率,本方案将整个IPv4地址空间“映射”到IPv6地址空间中,即把IPv4看作IPv6的一个“子集”。具体实现方式是通过添加IPv6合成前缀将IPv4地址映射到IPv6地址。全局映射规则就是所有IPv4地址块与IPv6合成前缀的映射关系: {IPv4 address block: Pref64}。对于需要通过纯IPv6网络传送的IPv4数据包,入口PE设备可基于统一定义的全局映射规则对其IPv4源和目的地址映射生成对应的IPv6源和目的地址。同时该IPv4数据包被转换成IPv6数据包,如图4所示,然后在纯IPv6的网络中进行域内和跨域传送。为了支持地址转换,全网PE都配置IPv4地址块对应的IPv6合成前缀。

(2) 基于BGP4+协议扩展跨域交换映射规则

为了将IPv4业务数据包正确地传送到网络出口,需要在PE中将IPv4地址块对应的路由映射合成IPv6路由。对于



▲图2 域间独立采用不同的IPv6单栈技术方案



▲图4 基于全局映射规则的IPv4域IPv6地址转换

特定的IPv4地址块来说，在纯IPv6网络中，可发布IPv6合成路由的PE就是该IPv4地址块的关联PE。关联PE也同样都有相应的Pref64。本方案采用PE的前缀Pref64来生成IPv4地址块所对应的IPv6合成路由，然后在全网交换IPv4地址块在纯IPv6网络中关联的边缘PE（出口PE）的位置，即IPv4地址块的IPv4-IPv6映射规则。为此，可通过BGP4+协议扩展^[8]在域内和域间的设备间交换IPv6合成路由和IPv4-IPv6映射规则。

(3) 兼容隧道和翻译的灵活转发方式

当用户发出的IPv4数据包达到IPv6网络边缘时，入口PE的XLAT模块根据本地的IPv4-IPv6映射规则将其转换为IPv6数据包，然后该数据包被转发到对应的网络出口PE。要说明的是，XLAT在转发面既可以支持翻译方式，也可以支持4over6的封装方式。

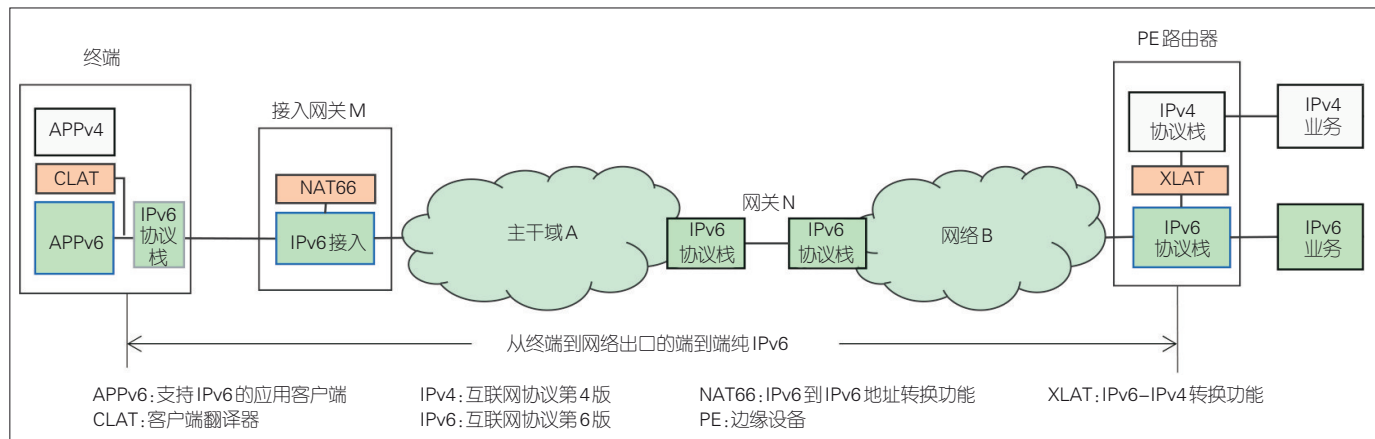
本方案支持域间IPv6单栈能力协同，包括接入域和主干域的协同、主干域间的协同、运营商间的协同，通过协同降低IPv4业务承载中的转换次数，来确保数据传输的效率。如图5所示，接入域为某个移动核心网，可通过464XLAT为

移动终端提供IPv6单栈接入服务。在464XLAT方案中，为了在IPv6网络中支持对IPv4业务的访问，客户端翻译器（CLAT）^[9]需要将IPv4客户端（APPv4）发出的IPv4数据包转换成IPv6数据包。主干域A和B分别为提供IPv6转发服务的主干网。本方案可以融合接入域、主干域A、主干域B的IPv6单栈能力，形成从终端到网络主干域B出口的端到端IPv6单栈能力，并且只需要在出口PE路由器进行一次IPv4-IPv6转换，不需要在域间互通的位置（M和N处）将IPv6数据包恢复成IPv4数据包。

4 向IPv6单栈过渡的策略建议

为了在中国尽早推动IPv6单栈网络的产业成熟和网络部署，需要做好如下几方面的工作：

- 从政策方面对IPv6单栈给予引导支持，并在必要时给出关闭网络中IPv4协议栈的实施路线、推进策略和时间点。
- 网络边缘的设备须满足多域纯IPv6标准的能力要求。新入网操作系统和终端（4G/5G移动终端、固网CPE终端、物联网终端和其他新型终端等）都应当具有标准要求的IPv6单栈能力，并优先以纯IPv6方式接入网络。
- 推动互联网应用向IPv6迁移，提升IPv6网络的流量占比，为IPv6单栈化创造良好的流量条件。新上线的云计算、工业互联网、物联网等应按照纯IPv6要求来建设和运营，能在IPv6单栈网络环境下运行。
- 网络运营商和大型互联网公司应按照标准尽早进行IPv6单栈商用试点，并加快在IPv6单栈网络方面的相互协作；鼓励并支持在中国开展IPv6单栈示范区的建设和实践。
- 强化IPv6对网络安全的提升，确保网络安全领域内新的系统、软硬件、安全策略能够在纯IPv6的基础上进行规划设计，加速中国网络信息安全体系向IPv6单栈的同步



▲图5 基于多域纯IPv6实现的端到端IPv6数据通信

过渡。

5 结束语

随着IPv6在中国的规模应用,IPv6单栈网络已经成为新的重要目标,也是中国IPv6部署迈上新台阶的必由之路。IPv6单栈的本质是做“减法”,即消除网络中不必要的功能和协议冗余,优化网络架构。网络的IPv6单栈化和中国的数字化转型密切相关,其部署具有必要性和紧迫性。在实施层面,IPv6单栈化网络在较长时期内仍需要支持剩余IPv4业务的综合承载,因此网络的边缘还须维持IPv4特性。只有等IPv6流量增长大大超过IPv4后,IPv4协议才可彻底退网。本文提出了大规模、多域、多场景下的纯IPv6组网架构,在支持域间和运营商间协同的基础上,整合不同域间的IPv6单栈能力,提高了数据传输效率,并在最后提出了IPv6单栈部署的策略建议。

致谢

阿里巴巴集团IPv6首席架构师宋林健、中国电信股份有限公司研究院马晨昊和李聪为本文的方案研究和现网试验做了大量工作,对他们表示特别感谢!

参考文献

- [1] 中央网络安全和信息化委员会办公室,国家发展和改革委员会,工业和信息化部. 关于加快推进互联网协议第六版(IPv6)规模部署和应用工作的通知 [EB/OL]. (2021-07-12) [2021-12-10]. http://www.cac.gov.cn/2021-07/23/c_1628629122784001.htm
- [2] IETF. SRv6 network programming: IETF RFC 8996 [S]. 2019
- [3] IETF. Dual-stack lite broadband deployments following IPv4 exhaustion: IETF RFC 6333 [S]. 2011
- [4] IETF. Lightweight 4over6: an extension to the dual-stack lite architecture: IETF RFC 7596 [S]. 2015
- [5] IETF. Mapping of address and port using translation (MAP-T): IETF RFC 7599 [S]. 2015
- [6] IETF. Mapping of address and port with encapsulation (MAP-E): IETF RFC 7597 [S]. 2015
- [7] IETF. 464XLAT: combination of stateful and stateless translation: IETF RFC 6877 [S]. 2013
- [8] IETF. Multiprotocol extensions for BGP-4: RFC 4760 [S]. 2007
- [9] IETF. IP/ICMP translation algorithm: RFC 6145 [S]. 2011

作者简介



解冲锋, 中国电信集团公司科技委委员、教授级高工, 欧洲电信标准化协会(ETSI) IPE 工组副主席, 北京市IPv6重点实验室主任; 主要研究方向为下一代互联网、IPv6、网络架构、SDN/NFV、物联网、云网融合等; 负责中国电信的多项科研项目, 以及多项国家“863”计划、CNGI、国家重大专项等科研项目; 在IETF发布RFC 5项, 发表论文10余篇, 拥有授权专利50项。



李星, 清华大学教授、博士生导师, CERNET网络中心副主任, 中国互联网协会学术工作委员会主任, 曾担任亚太网络工作组(APNG)主席、亚太网络信息中心(APNIC)理事会委员, 中国首位在互联网架构董事会(IAB)的成员, IETF Softwire工作组的发起人和技术顾问, 中国首个非中文相关标准RFC4925的第一作者, 国际互联网协会“互联网名人堂”入选者; 发明了无状态、基于运营商前缀的翻译技术IVI, 该技术已被列为IETF核心系列标准(RFC6052、RFC6144、RFC6145、RFC6219等); 在IETF发布RFC 11项。



李震, 下一代互联网国家工程中心总工程师, IPv6 Forum Fellow; 研究方向为IPv6、SDN/NFV、DNS、IOT等; 主持和参与多个重大科研项目, 包括GO4IT中欧合作项目、国家发改委CNGI重大专项、北京市重大科技专项等, 在下一代互联网标准开发设计、物联网应用以及网络、应用、DNS等基础设施IPv6过渡方面拥有丰富的研究和开发实践经验。



余勇志, 中国电信股份有限公司重庆分公司高级工程师、中国电信股份有限公司高级技术专家(IP专业); 主要研究方向为新型IP网络技术、移动互联网技术、云网融合等。