中国科学: 信息科学 2019 年 第 49 卷 第 9 期: 1186–1204
SCIENTIA SINICA Informationis
论文

多边共管的多模态网络标识域名生成管理解析原型系统

李挥[1*], 邬江兴[2], 邢凯轩[1], 伊鹏[2], 陈世胜[3], 梁伟[3], 魏进武[4], 李卫[4], 朱伏生[5], 田开颜[6], 朱江[6], 陆以勤[7], 徐恪[8], 宋佳兴[8], 刘怡俊[9], 董永吉[2], 韩永祥[10], 侯韩旭[10], 马军锋[11], 徐睿[1], 阙建明[1], 杨伟豪[12], 缪伟豪[13], 郑泽峰[14], 孙涛[15], 韦国华[1], 綦九华[1], 刘吉[1], 白永杰[1], 宁崇辉[1], 王菡[1], 张昕淳[1], 胡嘉伟[1], 黄健森[1], 吕赛[1], 刘馨蔚[1], 李更新[1]

1. 北京大学深圳研究生院, 深圳 518055
2. 国家数字交换系统工程技术研究中心, 郑州 450002
3. 中国电信股份有限公司, 北京 100033
4. 中国联合网络通信有限公司, 北京 100032
5. 广东省新一代通信与网络创新研究院, 广州 510670
6. 北京金山云网络技术有限公司, 北京 100085
7. 华南理工大学电子与信息学院, 广州 510006
8. 清华大学信息科学技术学院, 北京 100084
9. 广东工业大学计算机学院, 广州 510006
10. 东莞理工学院电子工程与智能化学院, 东莞 523808
11. 中国信息通信研究院, 北京 100191
12. 香港中文大学信息工程系, 香港
13. 香港科技大学电子与计算机工程系, 香港
14. 澳门科技大学资讯科技学院, 澳门
15. 深圳大学城网络信息中心, 深圳 518055

* 通信作者. E-mail: lih64@pkusz.edu.cn

收稿日期: 2019–04–12; 接受日期: 2019–06–27; 网络出版日期: 2019–08–29

鹏城实验室 – 大湾区未来网络试验与应用环境 (批准号: PCL2018KP001)、国家自然科学基金 (批准号: 61671001)、国家自然科学基金创新研究群体项目 (批准号: 61521003)、国家重点研发计划 (批准号: 2016YFB0800101, 2017YFB0803204)、中国工程科技中长期发展战略研究项目 (批准号: 2016-ZCQ-04)、深圳市学科建设项目 “数据科学与智能计算” 和深圳市基础研究课题 (批准号: JCYJ20170306092030521) 资助

摘要 面对新形势下互联网蓬勃发展给现有网络体系带来的挑战, 现有单边主义下的域名管理系统 (DNS) 在专业化服务质量和安全管控方面已力不从心. 本文从分析现有网络体系问题出发, 结合世界网络发展趋势, 明确新型网络基本技术特征, 以高安全、高鲁棒、高效能和高可用性为导向, 提

引用格式: 李挥, 邬江兴, 邢凯轩, 等. 多边共管的多模态网络标识域名生成管理解析原型系统. 中国科学: 信息科学, 2019, 49: 1186–1204, doi: 10.1360/N112019-00070
Li H, Wu J X, Xing K X, et al. Prototype and testing report of a multi-identifier system for reconfigurable network architecture under co-governing (in Chinese). Sci Sin Inform, 2019, 49: 1186–1204, doi: 10.1360/N112019-00070

出新型多模态域名管理技术框架, 实现融合多边共管的投票管理和多模态网络下标识域名互访和互隧道功能的原型系统. 原型系统通过现网测试验证了理论的正确性和可行性. 该系统可以方便地应用于自主多模态标识空间下的主权网及高安全专网, 为全世界对网络共管共治的诉求提供了中国方案.

关键词 多边共管, 域名解析, 多模态网络, 联盟链共识 PoV, 多模标识互相隧道/互译, 后 IP 高安全专网

1 引言

互联网在近半个世纪里, 经历了从无到有, 从简单到复杂的高速发展. 以互联网协议 (IP) 为主的网络在人类生活中拥有举足轻重的地位. IP 网络在设计之初是为了实现简单的端到端通信, 随着大数据、云计算、移动互联网、物联网的大规模部署和应用, 网络信息量和标识量以成倍的速度增长. 面对新形势下网络发展, IP 网络出现资源枯竭、业务适应能力差等严重问题, 再加上现有体系下, 美国垄断顶级 IP 域名根服务器等重要网络资源, 其霸权行为已经对全球网络空间构成了极大的威胁. 传统 IP 体系下的安全性差、可管控能力弱的问题亟待解决. 近些年, 全球先进国家都开始研究独立于 IP 的新型网络体系. 大致主流共识是, 后 IP 时代的网络应当是支持包含内容、身份、IP 地址, 及地理空间位置等多标识, 即多模态域名标识的新型网络体系. 多模态域名标识是全人类共同拥有的网络空间, 必须由各国共管、共治、共享. 本文从共管共治共享的理念出发, 在分析现有网络体系结构问题的基础上, 基于多边共管、多模寻址、内生安全、高效可用和隐私保护等特点, 解决全球面对的 IP 网络两大缺陷问题, 设计了多边共管共治的新型域名管理系统架构, 掌握了多模态网络标识互访互隧道等关键技术, 完成新型域名标识管理原型系统的开发. 目前, 该原型系统实现了用户内容与其私钥签名的绑定机制和多边共同管理网络标识空间的生成管理解析功能, 以及包括 IP-CCN-IP, IP-CCN, CCN-IP, CCN-IP-CCN, CCN-CCN 多种隧道传输和标识互访场景, 实现多种传输信道异构环境下高清视频播放. 原型系统已在由北京大学深圳研究生院、中国电信、中国联通、广东省新一代通信与网络创新研究院、金山云、华南理工大学、广东工业大学、香港科技大学、香港中文大学、澳门科技大学组成的实际运营网络上测试了多边共管网络标识的生成、管理、解析功能, 以及多模标识的互访和多模态网络互隧道场景. 实际运营商级别的网络部署实验结果说明系统方案完全可行. 此外, 通过实时高清视频传输播放测试, 结果表明系统性能良好, 符合预期, 进一步开发后可以进入实际应用.

2 现有体系问题及应对措施

2000 年, ICANN (The Internet Corporation for Assigned Names and Numbers) 在全球部署 13 个根服务器, RFC2535[1)] 宣称受字节数限制, 根服务器数量无法进一步扩展, 此后几年各根服务器开始在全球广泛设立镜像服务器, 通过任播技术响应域名解析请求. 而根服务器、域名、AS 号等关键互联网资源管理权仍属于美国商务部下属国家电信和信息管理局 NTIA (National Telecommunications and Information Administration). 单一国家管理、中心化架构的 DNS (domain name system) 给全球互联网安全带来巨大威胁. 2002 年 DNS 根服务器遭受大规模 DDoS 攻击, 导致全球域名解析服务受到严重影响. 2014 年中国 DNS 解析发生故障, 所有通用域遭到不同程度的 DNS 污染. 2015 年土耳其国家

1) http://www.rfc-editor.org/info/rfc2535.

顶级域遭到攻击, 几乎所有 (.TR) 域无法访问. IP 体系的先天缺陷不能适应以内容为中心、高速移动、物联网和工业互联网的业务需求. 研发替换 IP 体系的新型网络体系势在必行. 2010 年美国国家科学基金资助了 4 项未来网络架构 FIA 计划, 于 2015 年进入第 2 阶段. 其中 Named Data Network 项目旨在研发建立以内容为中心的新型网络架构. 但由于其颠覆式的体系架构, 导致运营商在实际部署时仍存在诸多难题. 就国内发展趋势而言, 国家非常重视对未来网络体系架构及域名解析的关键理论和技术的研究. “十一五” 以来, 国家对新一代信息网络基础理论研究进行了大规模的部署和支持. 2007 年, 国家重点基础研究发展计划 (973 计划) 启动了 “可测可控可管的 IP 网的基础研究” 项目, 主要针对现有 IP 网络的可测可控可管性进行了深入的研究. 2008 年, 国家 973 计划资助了 “新一代互联网体系结构和协议基础研究” 项目, 研究面向未来的新一代互联网体系结构与协议. 2011 年国家 973 计划支持了 “面向服务的未来互联网体系结构和机制研究” 和 “可重构信息通信基础网络体系研究” 两个项目 [1]. 前者从 “演进式” 与 “革命式” 两个研究思路对以服务为中心的未来互联网体系结构开展研究; 后者提出了 “可重构网络” 思想并建立可重构信息通信基础网络体系. 这些项目的开展, 意味着现有网络体系迈向未来网络体系的不可抗拒的发展趋势 [2,3].

面对中心化的技术和管控风险, 去中心化的多边共管共治便成为全球对域名空间管理的诉求. 国际上 Namecoin 项目首先提出了基于比特币区块链网络的分布式域名存储及合并挖矿等若干解决方案. Blockstack 项目提出虚拟链技术以支持逻辑层在不同底层链之间的移植 [4], 并对区块链网络架构、分布式数据存储, 以及无限分类账本等技术进行了深入的研究, 从而有效地增强了区块链域名系统的整体鲁棒性以及可重用性. 但是上述两个项目由于依赖的比特币底层技术, 域名解析系统均是对现有 DNS 系统的补充和替换, 无法从根本上解决现有网络架构中安全性和 IP 层所存在的 “细腰” 结构问题, 成为了制约网络总体功能的瓶颈.

我国北京邮电大学提出的一种基于区块链的开放数据索引命名 ODIN (open data index name) 模型, 该模型使用比特币侧链, 并直接将数据存于底层区块链中, 导致其存在数据存储限制大、读写内容缓慢等问题, 无法很好适应海量数据标识的处理需求.

上述主流应对措施虽然可以解决现有体系下若干问题, 但设计的系统仍存在性能低下、应用部署困难的问题, 无法满足新时代我国对互联网安全的需求 [5]. 基于上述认知, 我们认为新型网络应该具备全维可定义 [6]、多样化寻址路由、智慧化和广义鲁棒性等基本技术特征. 在开放的网络架构中, 实现对网络拓扑、协议、软硬件、接口等进行全维度可定义, 突破多样化寻址和路由技术的难题, 实现高强度鲁棒性, 才能充分满足新形势下人类对网络多元化的需求.

3 共管共治的多模态域名管理系统架构与关键技术

3.1 系统架构

本文提出的系统结合区块链技术, 对无中心化管理、各方参与、多边共管、平等开放的多模态新型域名解析系统进行探索研究, 实现安全可靠、高效传输、可大规模部署的特性, 推动现有网络体系向新型网络体系平稳过渡 [7~9].

共管共治的多模态域名管理系统框架如图 1 所示, 互联网管控的权利交由全世界互联网参与者, 不再是某个独立的机构垄断, 实现后 IP 时代网络空间的多边共管共治共享, 平等开放. 整个新型多标识的网络系统采用自上到下层级化网络域进行划分. 其中网络的顶级域由各个国家的政府机构作为顶级域名节点, 共同维持一条联盟链来达成全网共识, 实现互联网共管共治的本愿. 网络内所有的网络

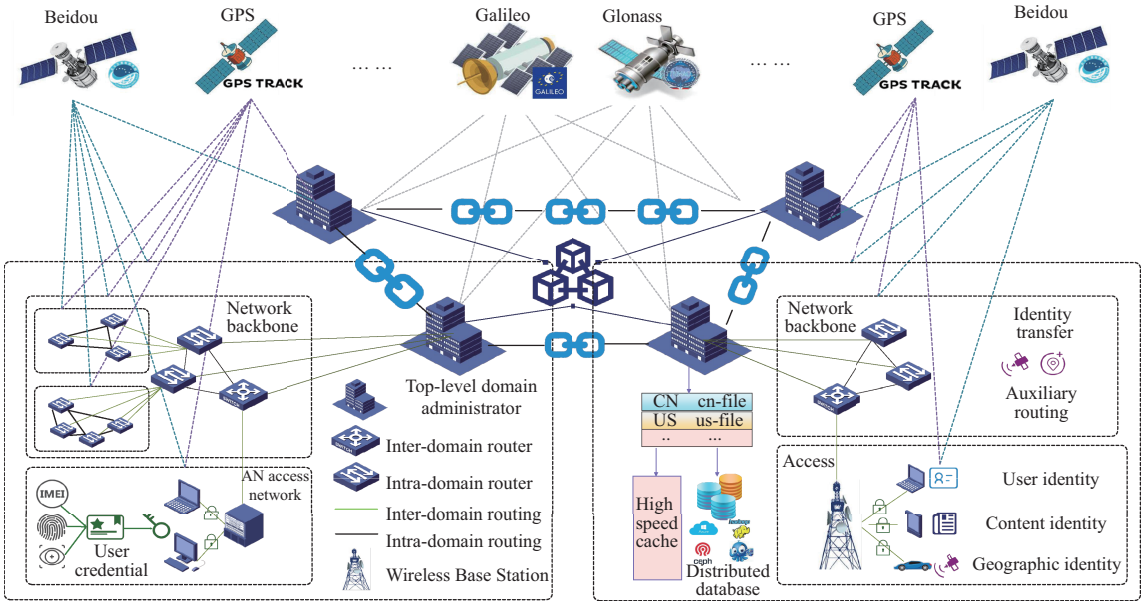

图 1 (网络版彩图) 新型域名系统整体系统架构模型

Figure 1 (Color online) Overall architecture of new domain name system

资源都将锁存在区块链上, 保证网络资源真实可信、不被篡改. 一级域及其他域由相应国家及专业机构管理, 其域内的标识管理方式、标识注册方案, 及共识算法可以不同, 其具体实现细节也可不同, 通过低耦合的方式来保证系统之间的安全性, 实现各层级之间的特殊性以及可定制性. 上下域之间通过网络监管节点作为数据访问接口以实现层级化的数据传输 [7~9].

正如图 2 所示, 系统中存在监管节点、个人用户, 以及企业用户等网络节点. 网络监管节点存在于顶级及各层级域内, 主要负责域内用户管理、标识注册、标识转换, 以及标识路由等服务, 同时每个网络监管节点存有面向内容网络标识, 空间信息标识、身份信息, 及 IP 地址等多模态标识. 新型网络支持包括身份、内容、空间信息, 及 IP 地址标识等多种标识共存的网络层路由寻址. 其网络中的所有的资源的内容标识均会和发布者的身份标识相互绑定, 用户登陆网络时的空间信息标识及访问的网络资源将记录在所在域的网络监管节点区块链上用于安全监管及数据保护. 自上而下共分为控制层、路由层和数据层.

最上层控制层负责域名管理、权限管理等更多与线下相结合的事务, 完成对区块信息的校验并在达成区块链共识后将记录域内的路由状态以及域内请求的认证. 使得系统全网内容统一, 具有极强的不易篡改性及可追溯性.

中间层路由层则完成对地址标识、内容标识、身份标识等多种网络标识的注册、解析等操作, 并负责数据包的转发及过滤. 各级节点依靠投票共识算法完成各层级区块链的数据一致性.

从底层数据层面来看, 系统具有一个高效的、分布式关系数据库. 其中包括区块链数据子层和云存储子层. 区块链数据子层存放标识解析的最小必需的数据, 称之为链上数据, 其数据存储格式采用区块链链式存储. 云存储子层存放网络标识的全部信息, 称之为链下数据, 其数据采用本地数据库存储.

标识的注册和网络资源请求步骤如图 2 所示, 资源注册步骤如下.

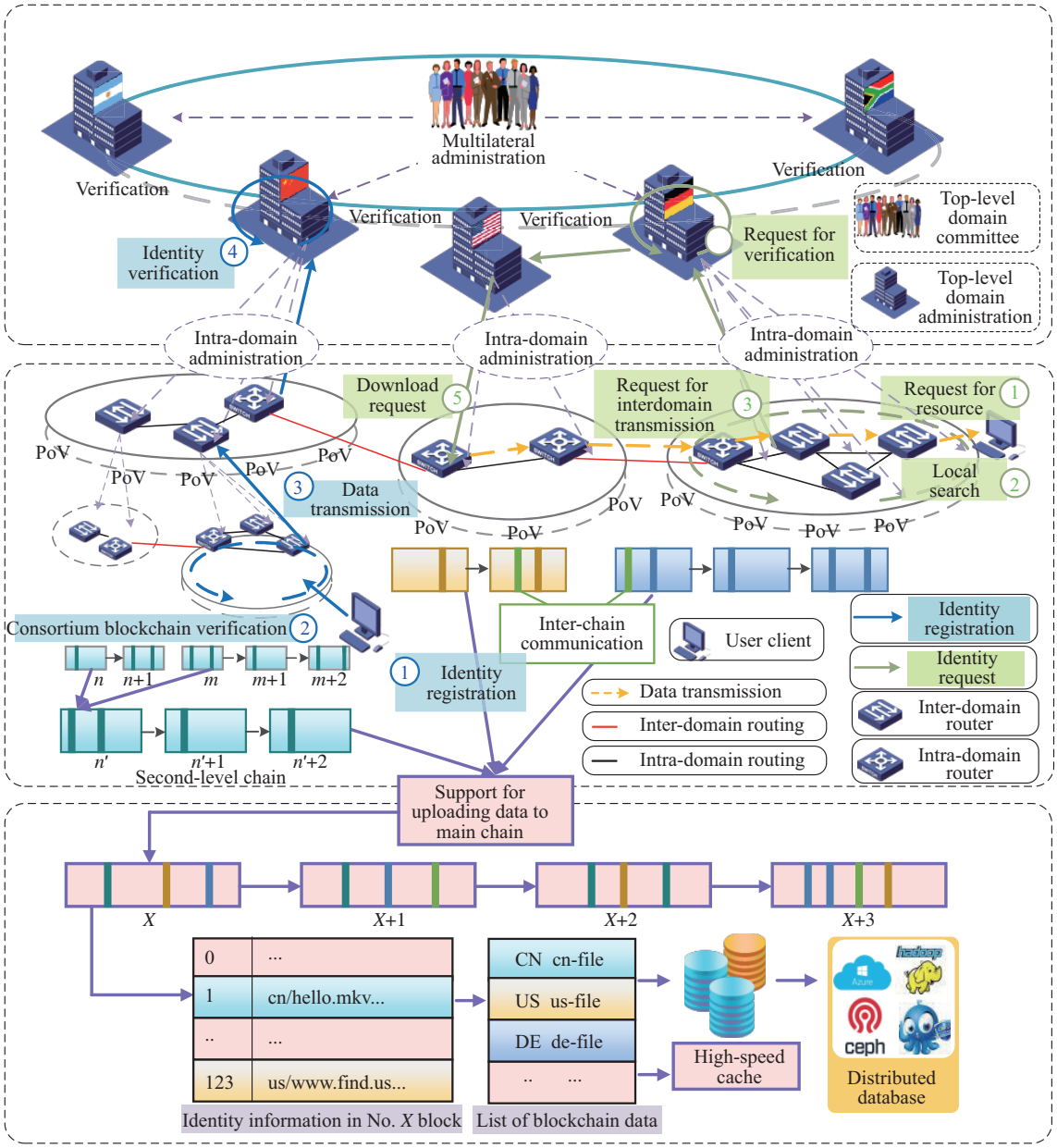

图 2 (网络版彩图) 多标识网络管理架构及标识注册和资源传输示意图

Figure 2 (Color online) Multi-mode network identification management architecture and operation flow

步骤 1. 注册资源标识. 在新构建的基于区块链的新型未来网络模型中, 任何能够被路由寻址的资源都要求先向联盟链中的顶级域名服务器注册, 只有当本资源标识通过联盟链认可并成功分配以后, 该资源才能被其他网络设备访问.

步骤 2. 数据路由转发. 当路由器接收到用户的注册请求之后, 将按照路由协议将其注册数据报文传输到其所在域的控制器进行后续认证及注册操作.

步骤 3. 联盟链节点认证. 联盟链节点在收到用户传输的资源标识请求之后, 联盟链节点将对其内容进行相关审查, 随后通过 PoV 共识算法对该通过的资源标识进行注册. 随后将返回给原申请节

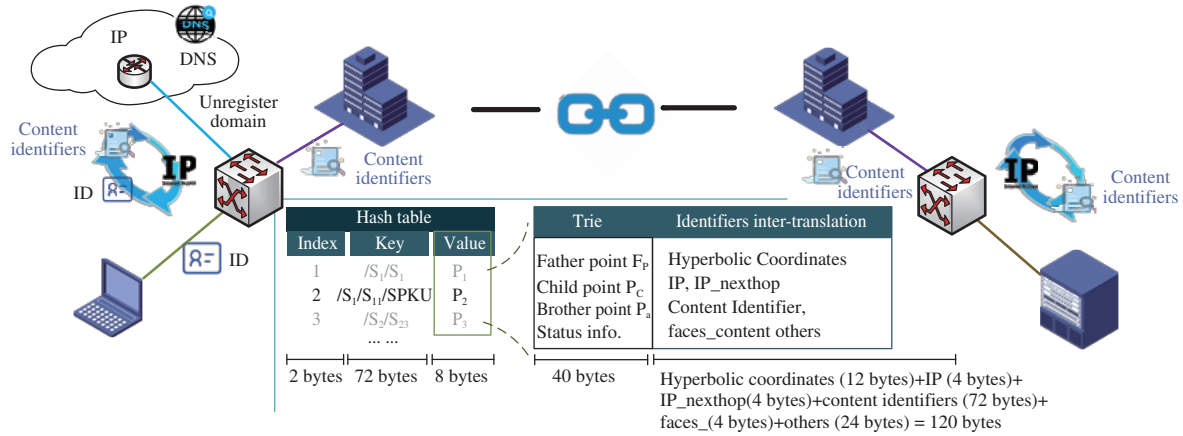

图 3　(网络版彩图) 多标识互译

Figure 3　(Color online) Multi-identifiers inter-translation

点相应的确认信号. 由于采用区块链数据库与区块内容相分离的原则, 原标识信息将存储在顶级域名的区块链数据库之上, 每当有一个记账完成之时, 全网将进行相应的区块链数据库同步工作以确认各个顶级域名之间的资源标识信息对等且统一.

网络资源请求流程如下.

步骤 1. 向最近的路由器传送兴趣 (interest) 数据包. 当请求的内容已获得网络注册的时候, 客户端即可使用相应的统一资源标识符 (uniform resource identifier, URI) 获取所需要的资源数据 (data) 包. 当最近的路由器接收到用户所发出的请求之后, 将通过查询转发表来确定是否要向上级域名服务器转发请求.

步骤 2. 数据路由转发. 当路由器接收到用户的查询请求之后, 将按照一定的路由协议将其 interest 数据报文传输到其所在域的控制器进行后续认证及查询操作.

步骤 3. 当联盟链节点接收到 interest 包请求后将进行请求验证. 其验证内容包括资源标识是否存在、请求是否合法等.

步骤 4. 联盟链节点下发转发路径表. 若顶级域名服务器在查询到相关已被注册标识, 将根据现有网络的动态拓扑结构来下发相应的路由信息. 网络中的转发线路上的相关路由器将收到新的转发路径表, 通过多跳路由建立数据传输通路.

在系统中, 每个域内的路由器维护一张拥有多种标识的路由信息表. 表中记录了资源所拥有的各种标识信息, 例如 IP、内容标识、身份标识和地理空间标识等. 用户在查询资源时, 可能在路由转发过程中需要转换标识. 标识间的互译转发过程如图 3 所示. 目前对我们所提出的基于 Hash 表的一种特殊的数据结构已经在普通服务器上实现支持超过 35 亿条 FIB (forward interest base) 条目的能力. 我们以内容标识和 IP 标识为例, 具体的用户进行资源的访问时, 会有以下 3 步.

步骤 1. 多标识路由器进行查询. (1) 传统域名, 则直接调用 DNS 查询. (2) IP 地址, 若标识互译转发表中存在, 则进行互译或转发, 否则代理访问传统 IP 网络. (3) CCN、身份 + 内容等其他标识, 先在 CS (content store)、PIT (pending interest table), 以及互译转发表中进行查询, 若存在则进行互译或转发, 否则转步骤 2.

步骤 2. 若在当前域中不存在该标识, 多标识路由器将递归的向上查询, 直到顶层域.

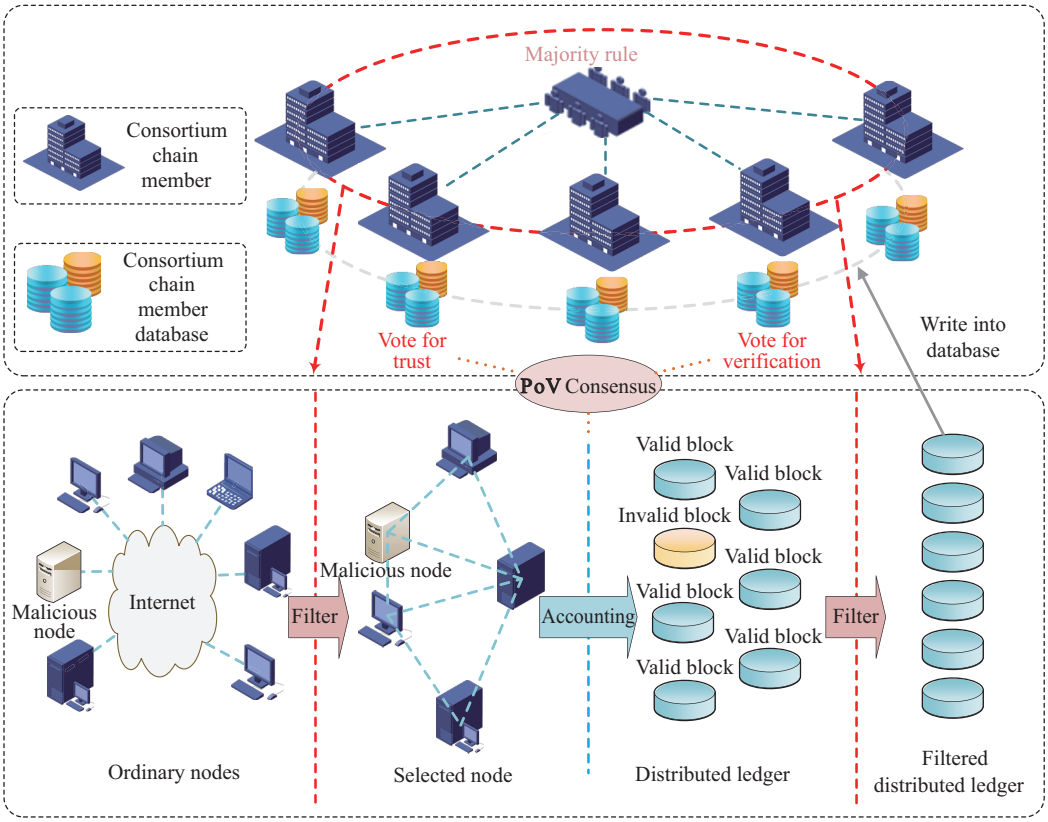

图 4　(网络版彩图) 分离投票权和记账权的共识设计思想

Figure 4　(Color online) A consensus design for separating voting and accounting

步骤 3. 若在顶层域中无该标识信息, 将按照标识信息特定的下级域查询, 直到标识指明的最底层域, 如存在则返回相应结果. 否则返回查询错误信息.

由于新型网络的所有各方发布内容和访问行为都受到有效的保护及管理, 其接入网络所产生的行为不可抵赖. 任何网络攻击或非法行为也将被域内区块链记录下来, 因此以这些标识的使用将使得网络空间处于有序与安全的状态, 将引导用户的各种流量承载到与身份绑定的新型标识网络如面向内容标识、身份标识上来. 而自然地减少没有任何安全保障的 IP 网络流量. 追求高可信服务的信息发布方将把他们的数据发布到新型标识上, 从而自然引导网络流量及体系的变革, 逐步实现去 IP 化.

3.2　关键技术

3.2.1　高效能无分叉的适应联盟协同运作的共识算法

基于网络标识管理系统框架的理论基础, 我们在系统中引入团队自主设计提出的新型共识算法 “proof of vote (PoV)” [10]. 其核心在于分离投票权与记账权, 由联盟成员共同投票进行 “去中心化” 仲裁, 无需集中式地信任某一成员机构. 利用联盟链模型中节点的特殊身份, 遵从 “少数服从多数” 的原则, 将投票结果作为系统生成有效区块的合法证明. 投票证明的思想在 PoV 共识的设计中由两种投票体现, 如图 4 所示.

Proof of vote on butlers (管家的信任投票). 用于对管家的信任投票, 委员在每一个任期结束

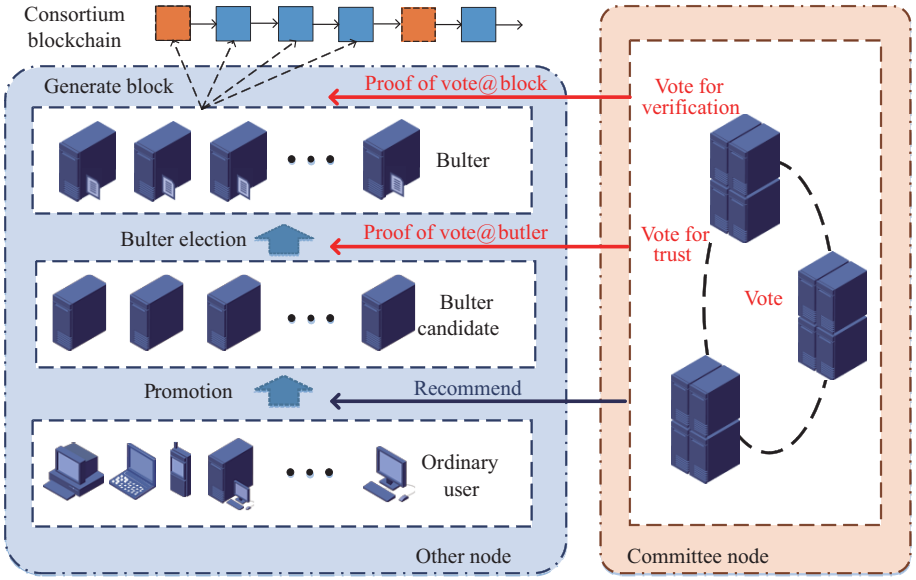

图 5 (网络版彩图) PoV 两种投票的思想体现

Figure 5 (Color online) Concept of two vote in PoV

的环节给管家投的总票数表示全体委员对此管家的信任度, 每个管家的可靠度通过委员对管家的投票结果来证明.

Proof of vote on blocks (区块的验证投票). 用于对区块产生的合法性进行验证投票, 每个区块必须获得超过半数以上的委员验证通过才能被认为是有效的合法区块. 若系统需要修正某次结果, 在超过半数委员的赞成后, 便可以修改该结果, 每个区块的合法性通过联盟的投票结果来证明.

同时 PoV 算法设立了管家、管家候选人角色, 如图 5 所示. 联盟链的决策结果由管家团队执行. 管家团队通过去中心化投票进行动态更换, 无形中在联盟节点间形成了一种投票共识. "少数服从多数的投票结果" 可以唯一确认一个最终决策, 使得联盟链系统在联盟节点的去中心化管理下, 稳定的运行.

投票证明的思想在共识机制的设计中通过两种投票体现, 第 1 种是委员对是否赞同区块产生的投票, 第 2 种是委员对管家候选人的投票. 委员通过返回签名的方式投票. 第 1 种情形是委员对区块的合法性的进行表决. 由值班管家 x 产生区块, 并发送给所有委员. 若委员同意此区块的产生, 则加密签名区块头和当前时间戳, 将签名和时间戳返回给管家 x. 当管家 x 在规定时间内收到超过半数委员节点以上签名的时候, 区块有效; 否则, 区块作废, 由管家 $x+1$ 重新封装区块. 第 2 种情形是委员对管家候选人的投票. 在任期中最后一个值班周期, 委员向值班管家 y 发送已签名的投票事务. 管家 y 收集并计算票数后, 用一个特殊区块封装所有的投票事务与结果, 发送给所有委员用于表决区块的合法性. 委员发送的投票事务中包含两种票的组合: (1) 正常票. 委员根据自己维护的管家候选人列表中的评分, 按分数高低给出分数较高的候选人序列. (2) 指定票. 考虑人为因素, 委员可以设置一组指定的候选人序列, 或者是随机的候选人序列, 增加管家的流动性. 委员对管家的投票体现委员对此管家的信任程度, 而管家的随机轮流记账增加管家的流动性, 避免某个机构控制大部分优秀的管家持久的占据管家行列, 也避免了部分经常被选举上的管家被大范围收买的可能性, 使得系统更加的安全可靠.

3.2.2 新型网络标识互通及渐进式部署方案

大数据时代未来网络体系架构的设计与发展是当今互联网领域的重要研究课题. 近年来, 人们提出了许多革命式网络体系结构方案, 在这些方案中, 信息中心网络 (ICN) 被认为是一种能够较好满足用户对信息传递需求的新型网络体系结构. 在 ICN 网络中研究最多的是, 内容中心网络 (CCN), 因为其以内容为中心进行数据的传输, 数据传输过程中利用数据本身进行命名, 彻底改变了以 TCP/IP 协议为基础以地址为中心的网络传输体系.

CCN 由 TCP/IP 拥塞控制发明人 van Jacobson 提出 [11]. 2016 年为了推动 CCN 网络的应用, 提出用 CCN 来做隧道技术传输 IP 数据包, 该理论的主要创新点是提出了 “转换代理的思想”, 在 IP 和 CCN 网络之间用代理技术进行数据包的转换 [12]. 然而, 实验还是在 Overlay IP 的环境下进行, 本质上还是两个 IP 网络进行数据交互, 没有实现真实意义上的 “隧道”, 2017 年首次出现了 “双栈” 的思想来实验渐进部署 CCN 网络 [13], 但是该思想停留在理论阶段, 只是提出了双栈的概念, 没有实现该双栈. 同年, 为了充分理由底层资源, 首次提出将 face 接口和 MAC 地址作动态映射 [14], 将内容层和以太网链路层作为一个连接, 同样该理论也只是停留在概念阶段, 并没有设计出部署方案. 2018 年, 思科公司基于以太网的 CCN 设计了一种 socket, 可以直接实现数据包的传输. 同样在 2018 年, 出现了基于以太网传输的 CCN 来做隧道技术传输 IP 数据包的思想, 最核心的创新点是将兴趣包进行了载荷传输, 将请求兴趣包中装了 IP 数据包. 不过该方案是在 Overlay IP 环境下做的仿真实验, 并不是真正意义上的部署.

目前, 国际上以 IP 为隧道承载内容中心网络技术已经成熟, 而反向隧道 (IP-CCN-IP) 承载未见实质进展, 完善该功能是实现可渐进部署的技术基础. 国际上针对 IP 网络和 CCN 网络的融合还没有定论 [15].

本系统设计是通过直接将 CCN 网络部署在以太网 MAC 层之上, 彻底摆脱 IP 依赖. 融合以太网传输、TCP 传输、UDP 传输网络的架构如图 6 所示. 系统在真实的网络环境中实现了 CCN 网络通信, 实现多模态标识互访和多模态网络互隧道过程.

为了使得两个 TCP 端通过中间的 CCN 网络进行通信, 在 IP 网与 CCN 网的边界处需要设置一对转换节点, 即 “发送多栈” 和 “接收多栈”, 分别连接到 TCP 的发送端与接收端. 不同网络寻址标识的转换与包封装都在多标识路由器中进行. 每个多标识路由器都有一个名字作为 CCN 网中的路由前缀, 该名字和其 IP 地址之间存在着映射关系, 从而使得 CCN 数据包顺利地传到指定的多标识路由器进行下一步处理.

建立连接. 两个 IP 网络的 TCP 端建立连接的 3 次握手过程等价于本 TCP/CCN 系统的内容网络 3 次兴趣包交换过程. 如图 7 所示, TCP/CCN 系统中通过 CCN 网络发送带有 SYN, SYN+ACK, ACK 控制信令的兴趣包的包头完成两个 IP 网络 TCP 端 3 次握手过程. IP 网络 TCP 客户端向多标识路由器发送 SYN, SYN+ACK, ACK 连接控制信令, CCN 网络通过将连接控制信令封装在 CCN 网络兴趣包的包头中完成对 TCP 连接控制信令的传输, 最后多标识路由器完成连接控制信令转发到 IP 网络的 TCP 服务端, 从而完成 TCP/CCN 系统两个 IP 网络 TCP 端连接的建立.

拆除连接. 两个 TCP 端拆除连接的 4 次挥手过程等效于本系统的 4 次兴趣包交换过程, 兴趣包交换逻辑与 TCP/CCN 建立连接过程相似.

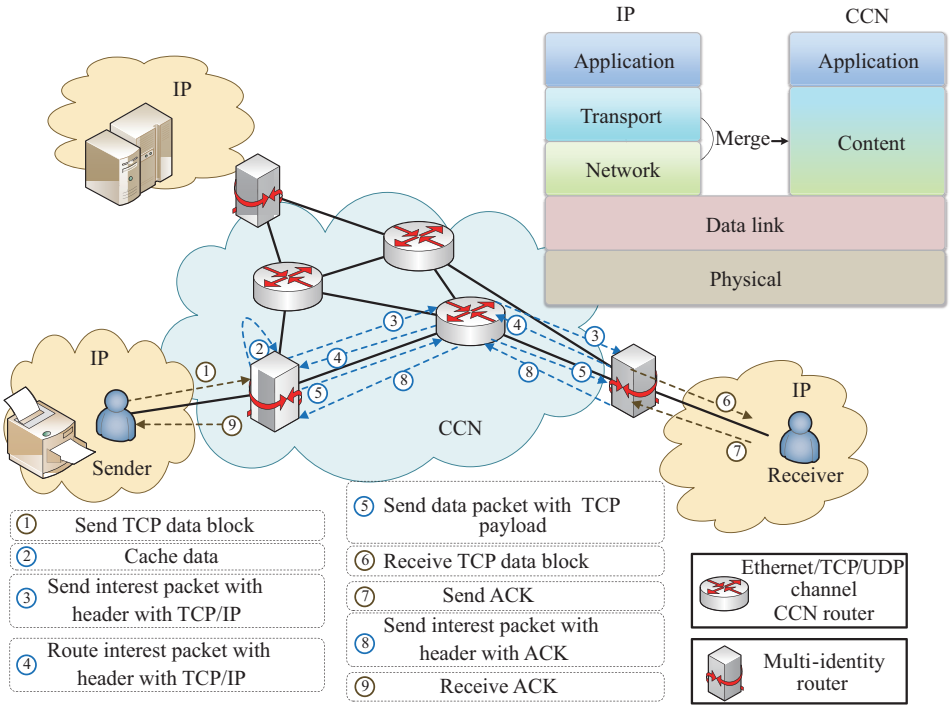

图 6　(网络版彩图) CCN 承载 IP 网络示意图

Figure 6　(Color online) Communication between CCN and IP network

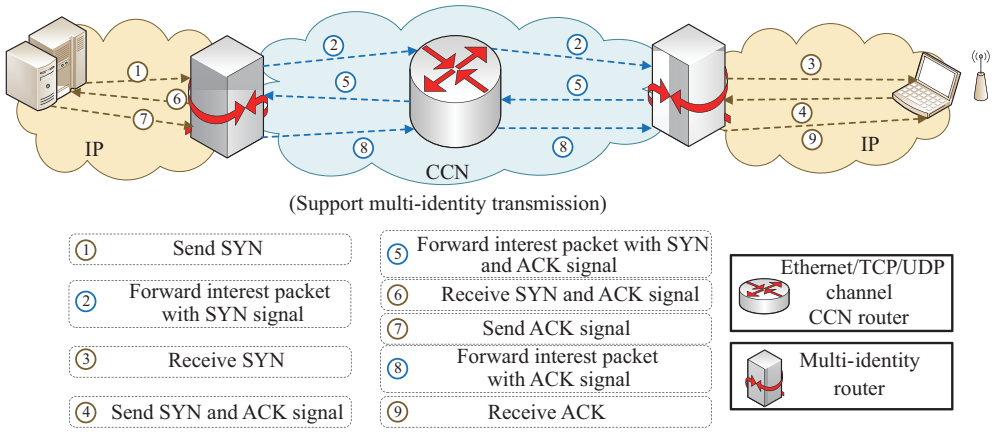

图 7　(网络版彩图) TCP/CCN 建立连接

Figure 7　(Color online) TCP/CCN connection establish

4　新型多模态域名管理系统原型机及现网测试验证

4.1　管理员端和用户端功能描述

4.1.1　区块链管理员功能描述

目前, 该原型系统实现了用户内容与其私钥签名的绑定机制, 完成包括区块链管理员模块和客户

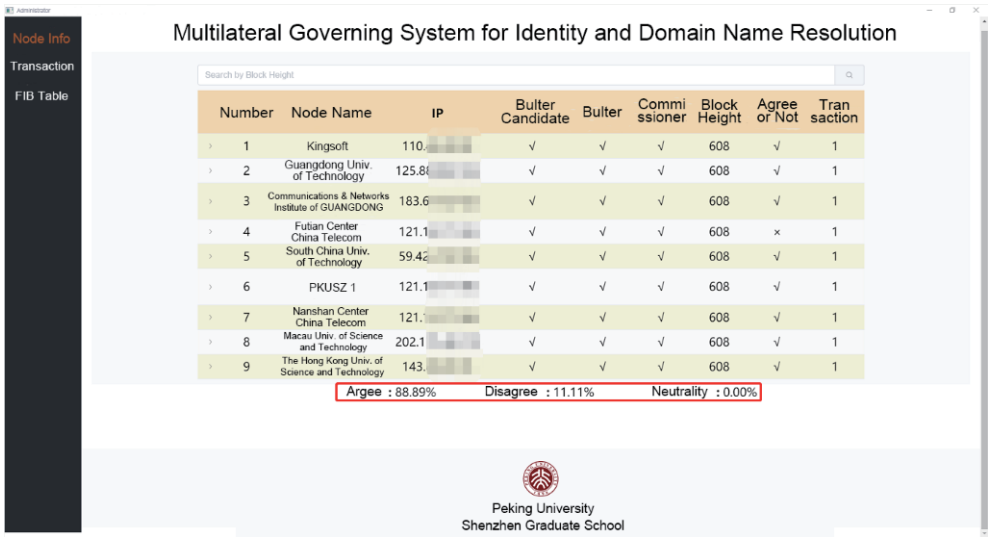

图 8　(网络版彩图) 区块链节点信息实时监控

Figure 8　(Color online) Endogenous safety architecture

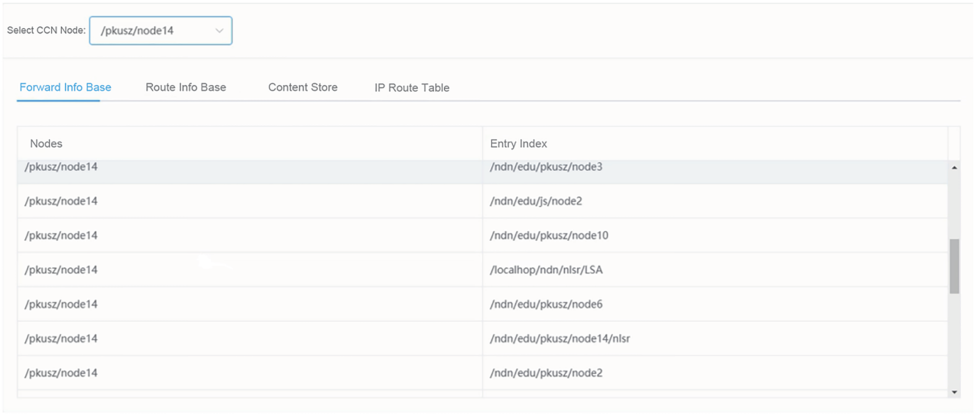

图 9　(网络版彩图) 节点路由状态信息

Figure 9　(Color online) Node routing status information

端模块在内的原型系统开发.

管理员对区块链的管理包括对节点运行状态的实时展示、对区块链数据的查询, 以及对区块链节点的操作和配置. 区块链节点将生成标识数据、注册用户等用户、标识管理操作以交易的形式存储在各个区块中, 而具体的用户和标识数据则在链外存储, 因此客户端不仅需要提供对区块数据的查询功能, 还需要提供用户数据和标识数据的查询功能. 用户端每一次注册和资源发布行为都对在全网进行一次共识, 体现在管理员界面交易数量一列. 在系统中实现了以下功能: 区块链状态实时监控、区块链数据查询、区块链网络管理功能和区块链近期交易查询, 如图 8 所示.

多标识路由节点提供了不同标识之间的转换和路由导航功能, 每个节点内部都维护了 FIB 表、路由表等路由信息, 实现多标识路由节点状态监控和多标识路由节点数据展示功能, 节点路由状态信息如图 9 所示.

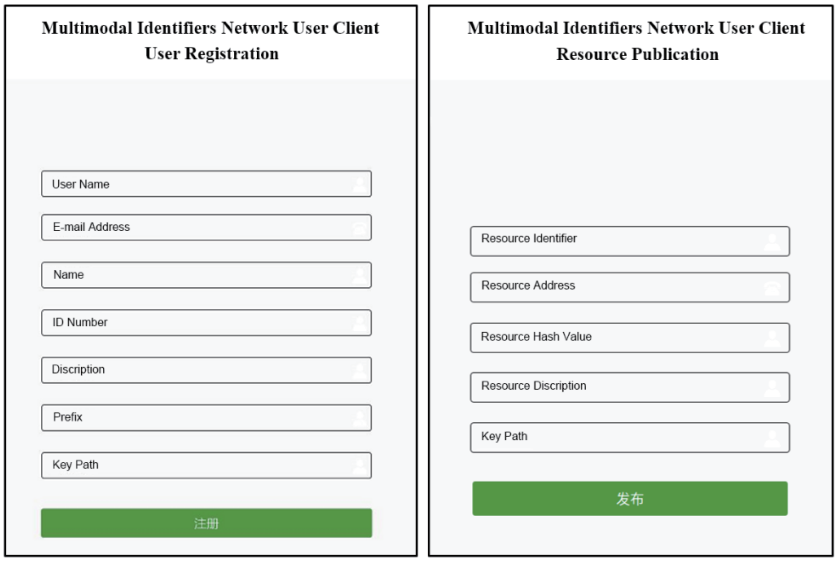

图 10 (网络版彩图) 用户注册与资源发布界面

Figure 10 (Color online) User registration and resource publishing interface

4.1.2 用户端功能描述

用户客户端的主要功能是为用户生成一对公私钥, 并使用将用户身份标识与公钥上传至区块链锁存的方式, 使得用户身份标识与公钥绑定. 同时, 发布资源的用户在注册时, 会申请一个内容名前缀, 从而也实现了内容标识与公钥的绑定. 注册成功的用户, 可以使用注册时申请的内容名前缀来发布一个内容标识, 资源发布操作需加上用户私钥签名信息, 如图 10 所示. 区块链收到发布资源的请求后, 会通过签名验证用户权限, 同时签名也将用户身份标识与发布的内容标识绑定, 以便于往路由寻址时的标识互译. 另外, 用户在发布资源时, 可以指定资源所在的实际位置. 这个实际位置可以是一个 CCN 网络内容标识, 也可以是一个 IP 网络标识. 所以, 当用户指定资源实际地址为 CCN 标识时, 就定义了一个身份标识向内容标识转换的映射; 而如果用户指定的源真实地址为 IP 标识, 则能得到一个身份标识向内容标识的转换映射. 这是各种标识转换表的来源. 用户网页客户端需要完整的展示一个用户从注册到发布资源, 再到查询资源和获取资源的过程, 现阶段的网络资源只有视频, 客户端提供视频在线点播的功能.

4.2 多模网络标识间相互隧道访问实验

本系统率先将 CCN 网络直接部署在 MAC 层, 独立于 IP 网络环境, 并且支持和 IP 网络互相承载. 原型系统完成了多边共同管理网络标识空间的生成管理解析功能, 以及不同包括 IP-CCN-IP, IP-CCN, CCN-IP, CCN-IP-CCN, CCN-CCN 多种网络情景下隧道传输和标识互访功能, 实现地址标识和内容标识转换和多种传输信道异构环境下的高清视频传输.

现网实验拓扑关系如图 11 所示, 原型系统在实际运营商级别网络上部署. 通过在北京大学深圳研究生院、中国电信、中国联通、广东省新一代通信与网络创新研究院、金山云、华南理工大学、广东工业大学、香港科技大学、香港中文大学、澳门科技大学组成的拓扑网络, 测试多边共同管理网络标识空间的生成管理解析功能.

图 11 中除了参与标识互访和互隧道的节点外, 其余节点作为区块链节点, 参与投票过程.

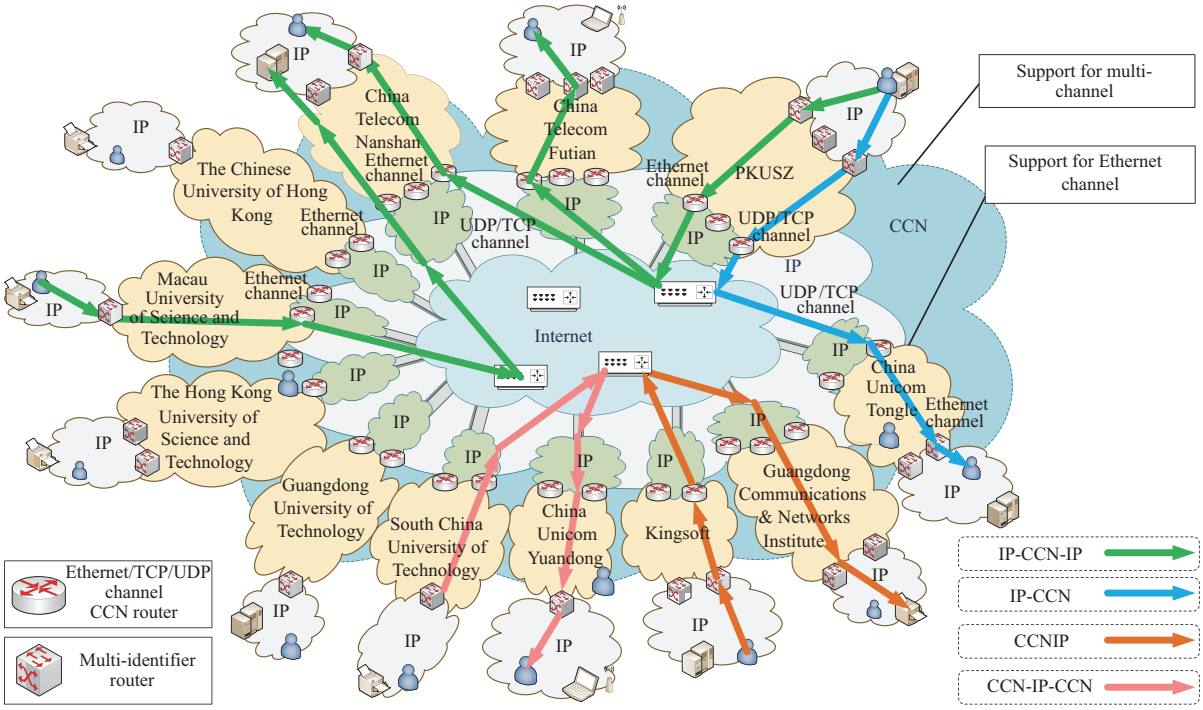

图 11　(网络版彩图) 原型系统测试概况

Figure 11　(Color online) Overview of prototype testing system

4.3　多模态网络互隧道测试及实验结果

实验中各个节点之间通过现有 Internet 实现通信, 各个 CCN 节点通过路由转发以此来实现 IP 和 CCN 间异构传输网络. CCN 节点使用 NFD (named data networking forwarding daemon) 模块通过 MAC 地址与本地其他节点通信, 本地机房与其他机房之间通过 IP 进行通信. 在 IP 网络与 CCN 网络的边界, 运行基于 libpcap, ndn-cxx 与 raw socket 的程序实现在 CCN 网络中运载 IP 包. 实验环境建立在真实的运营商网络上, 目前北京大学深圳研究生院本地服务器出口带宽为电信 100 M 专用宽带, 其余各个节点的带宽也都满足大于 50 M 的要求. 传输速率受实际网络环境影响, 例如不同运营商之间的限制, 内地与港澳台地区的连接限制. 实验数据表明, 我们的系统支持高清视频在线播放, 在大部分网络环境中, 系统支持超过 6 路 1080P 清晰度视频的在线播放. 我们的系统创新点是支持 IP Overlay CCN 的场景. 同时, 相比于以往的 CCN Overlay IP 的解决方案, 我们的系统将此场景下速率平均提升了 8 倍. 以下是我们的系统在 4 种场景下的实验结果.

4.3.1　IP-CCN-IP 互访场景验证及性能测试

把资源放在中国电信福田机房的 pkusz1 节点和南山机房的 pkusz3 节点上, 通过北京大学深圳研究生院的 node9 节点代理进行视频传输的测试实验.

我们通过代理从 node9 拉取福田机房 pkusz1 上的视频, 实验拓扑关系和传输结果如图 12 所示, 速率维持在 2.65 MB/s (21.2 Mb/s).

我们通过代理从 node9 拉取南山机房 pkusz3 上的视频, 实验拓扑关系和传输结果如图 13 所示, 速率为 2.44 MB/s (19.52Mb/s).

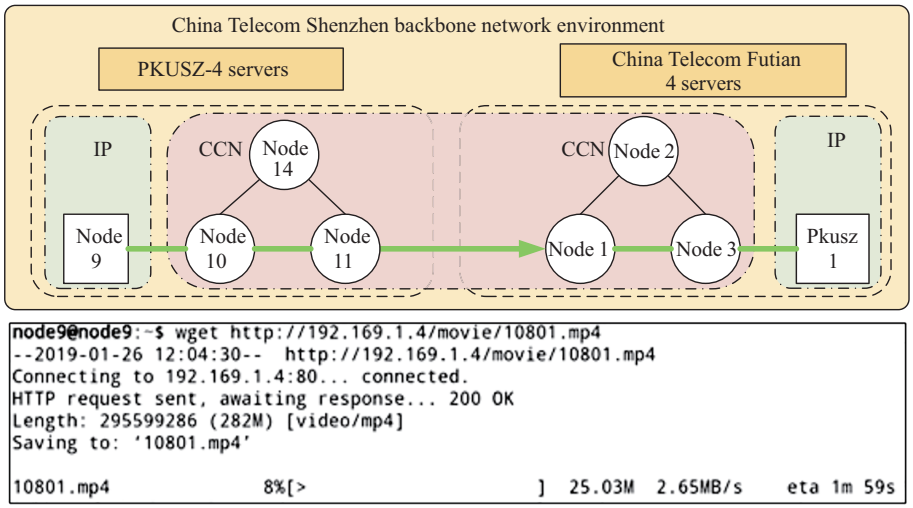

图 12　(网络版彩图) “北大深圳 – 电信福田” 连接拓扑关系和资源拉取速率

Figure 12　(Color online) Topology of “PKU-Shenzhen and Futian China Telecom” and its testing report

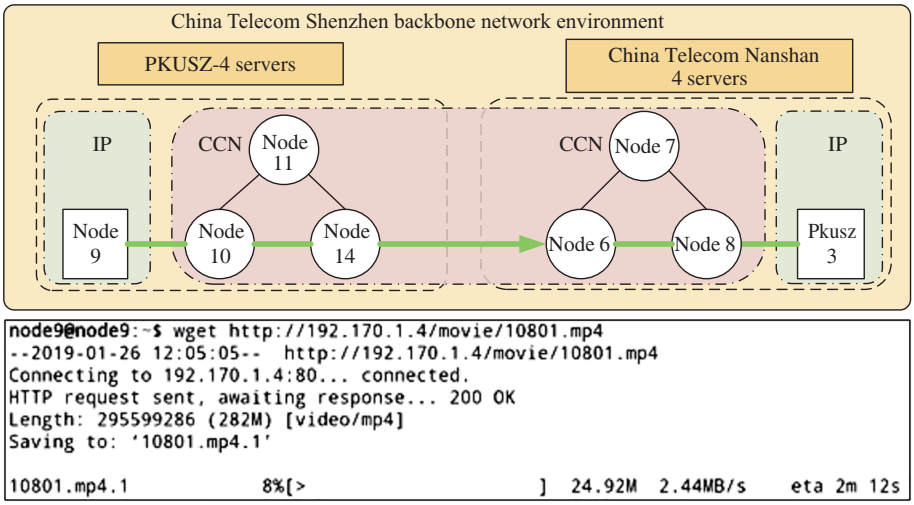

图 13　(网络版彩图) “北大深圳 – 电信南山” 连接拓扑关系和资源拉取速率

Figure 13　(Color online) Topology of “PKU-Shenzhen and Nanshan China Telecom” and its testing report

我们通过代理从 node9 同时拉取福田机房 pkusz1 和南山机房 pkusz3 上的视频, 实验传输结果如图 14 所示, 速率分别为 1.18 MB/s (9.44 Mb/s) 和 1.71 MB/s (13.68 Mb/s). 相比于单独拉取资源, 速率有所下降, 但总传输速率与之前保持一致水平. 原因主要是 node9 节点入口带宽受限, 成为性能瓶颈.

为了体现标识在不同网络环境下的传输可行性, 我们将数据源放在南山机房的 pkusz3 节点上, 通过澳门科技大学的 Server2 节点作为代理进行视频传输的实验. 受澳门科技大学的出口带宽影响. 平均速率为 1.2 MB/s (9.6 Mb/s), 但仍能传输高清视频. 实验拓扑关系和测试结果如图 15 所示.

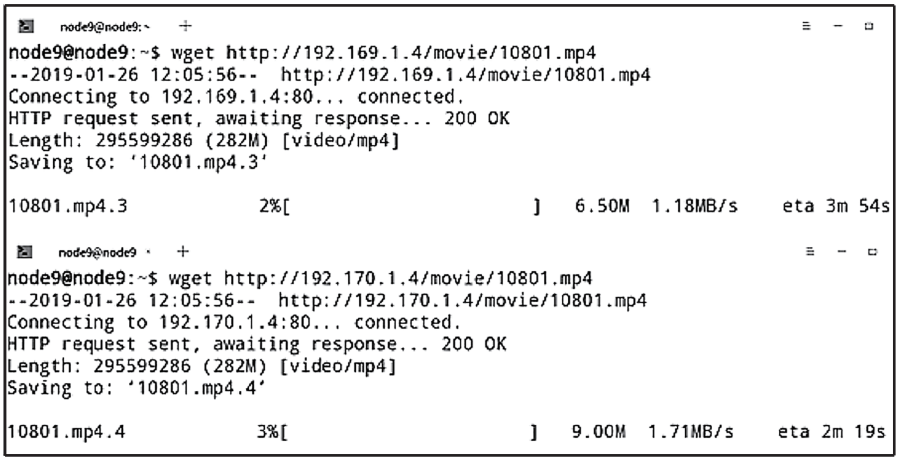

图 14 资源同时拉取速率

Figure 14 Testing report of 2-way resource request

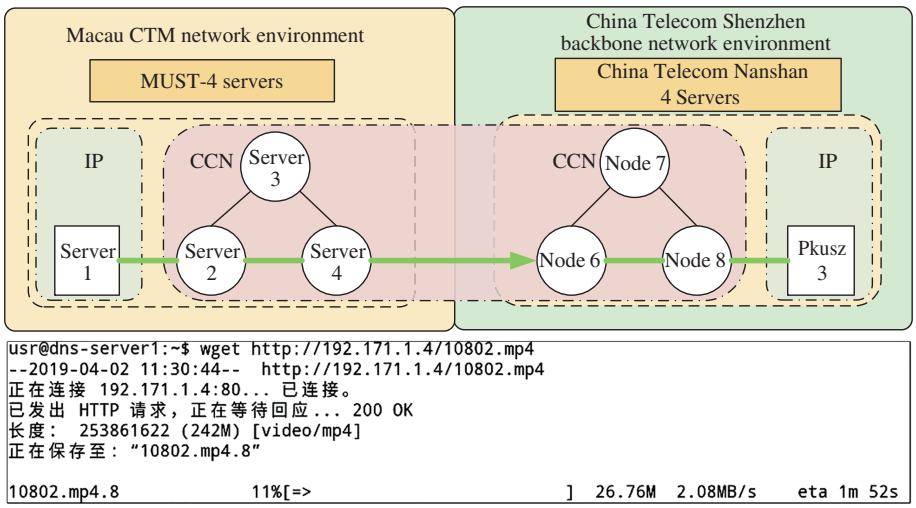

图 15 (网络版彩图) "澳门科大 – 电信南山" 连接拓扑关系和资源拉取速率

Figure 15 (Color online) Topology of "MUST and Nanshan China Telecom" and its testing report

4.3.2 IP-CCN 互访场景验证及性能测试

数据放在中国联通深圳同乐机房 host2 节点上, 我们用北京大学深圳研究生院 node10 作为多标识路由器, 本地主机为 IP 节点, 获取处于 CCN 网络环境上节点 host2 的资源. 此时数据流跨越两个不同的网络环境. 实验拓扑关系和测试结果如图 16 所示.

经测试, 接取视频数据的平均速率为 0.65 MBps (5.18 Mbps).

4.3.3 CCN-IP 互访场景验证及性能测试

数据放在 IP 环境下的北京大学深圳研究生院 node9 节点上, 我们以北京大学深圳研究生院 node10 作为多标识路由器, 处于 CCN 网络环境金山 host1 去拉取 node9 上的资源. 实验拓扑关系和测试结果如图 17 所示.

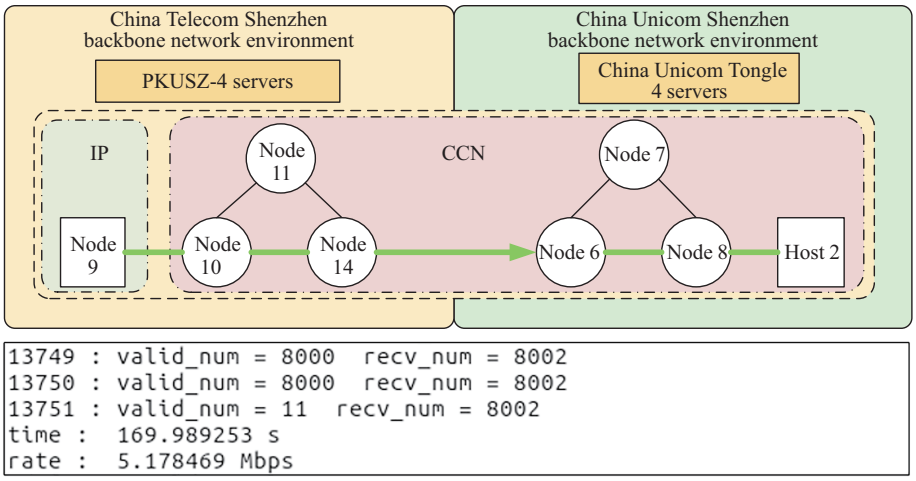

图 16　(网络版彩图) “北大深圳 – 深圳同乐” 连接拓扑关系和资源拉取速率

Figure 16　(Color online) Topology of “PKU-Shenzhen and Tongle China Unicom” and its testing report

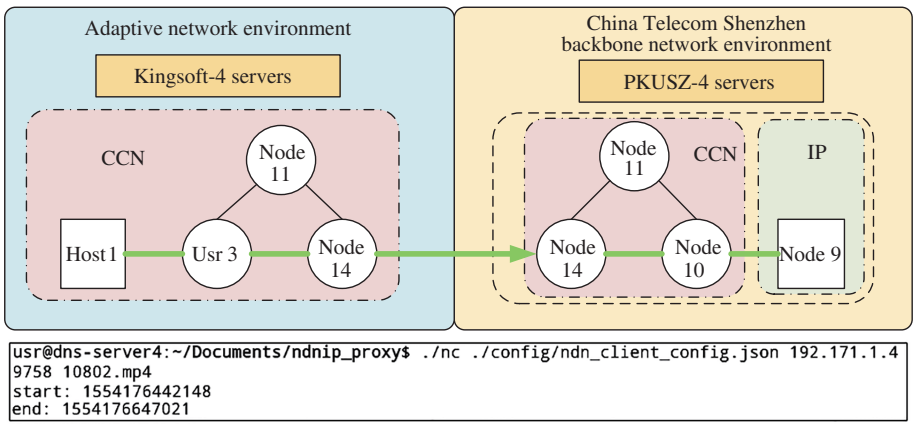

图 17　(网络版彩图) “金山云 – 北大深圳” 连接拓扑关系和资源拉取速率

Figure 17　(Color online) Topology of “Kingsoft and PKU-Shenzhen” and its testing report

经测试, 获取视频数据的平均速率为 1 MBps (8 Mbps).

4.3.4　CCN-IP-CCN 互访场景验证及性能测试

数据放在 CCN 环境下的金山云 host1 节点上, 我们用处于 CCN 网络环境的北京大学深圳研究生院 node10 去拉取金山云 host1 上的资源. 实验拓扑关系和测试结果如图 18 所示.

经测试, 资源平均传输速率为 2.65 MBps (21.2 Mbps).

5　总结

未来网络一定是平等开放、共享共治的, 逐步终结现有互联网体系下单一国家中心化管控的弊端, 支持国家网络空间主权正常有序发展. 没有网络安全就没有国家安全, 建立具有 “内生安全” 的新型域名解析系统的安全架构, 增强网络安全性. 在开放的网络架构下实施全维可定义, 网络的功能和业务

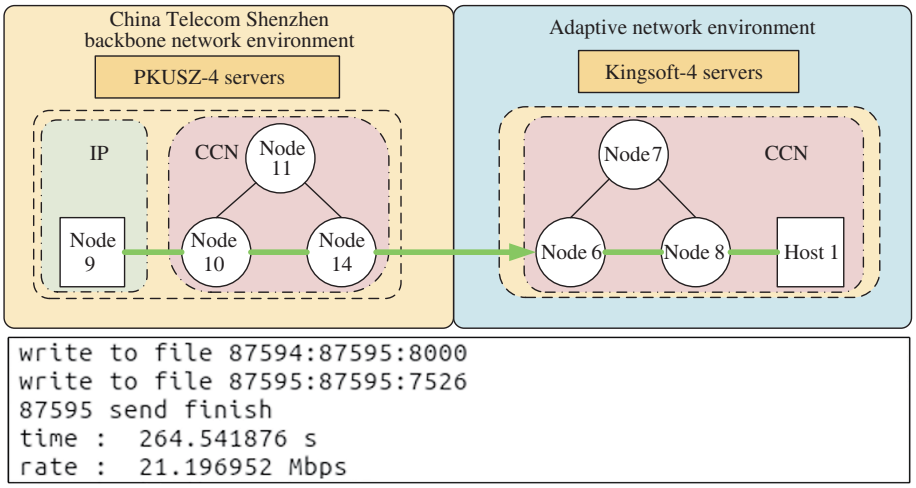

图 18 (网络版彩图) “北大深圳 – 金山云” 连接拓扑关系和资源拉取速率

Figure 18 (Color online) Topology of “PKU-Shenzhen and Kingsoft” and its testing report

机理都不再局限于现有架构、协议, 大幅提升网络服务能力, 充分满足新型网络不断演进的业务需求.

本文基于以上思考, 从共管共治共享的理念出发, 设计一套具有多边共管、多模寻址、内生安全、高效可用和隐私保护特点的新型域名管理系统. 该系统可以方便的应用于自主多模态标识空间下的主权网及高安全专网. 系统原型机已在北京大学深圳研究生院、中国电信、中国联通、广东省新一代通信与网络创新研究院、金山云、华南理工大学、广东工业大学、香港科技大学、香港中文大学、澳门科技大学组成的实际运营网络上演示了多边共同管理网络标识空间的生成管理解析功能及不同标识的互隧道和互访. 测试验证结果证明系统达到了设计要求, 并且性能指标良好.

参考文献

1 Lan J L, Cheng D N, Hu Y X. Research on reconfigurable information communication basal network architecture. J Commun, 2014, 35: 128–139 [兰巨龙, 程东年, 胡宇翔. 可重构信息通信基础网络体系研究. 通信学报, 2014, 35: 128–139]

2 Wu J X. Thoughts on the development of novel network technology. Sci Sin Inform, 2018, 48: 1102–1111 [邬江兴. 新型网络技术发展思考. 中国科学: 信息科学, 2018, 48: 1102–1111]

3 Wu J P, Lin S, Xu K, et al. Advances in evolvable new generation internet architecture. Chin J Comput, 2012, 35: 1094–1108 [吴建平, 林嵩, 徐恪, 等. 可演进的新一代互联网体系结构研究进展. 计算机学报, 2012, 35: 1094–1108]

4 Ali M, Nelson J, Shea R, et al. Blockstack: a global naming and storage system secured by block chains. In: Proceedings of USENIX Annual Technical Conference, Denver, 2016. 181–194

5 Wu H Q. Reflections on the reform of network architecture. ZTE Commun, 2019, 25: 1–5 [邬贺铨. 对网络体系变革的思考. 中兴通讯技术, 2019, 25: 1–5]

6 Lv P, Liu Q R, Wu J X, et al. New generation software-defined architecture. Sci Sin Inform, 2018, 48: 315–328[吕平, 刘勤让, 邬江兴, 等. 新一代软件定义体系结构. 中国科学: 信息科学, 2018, 48: 315–328]

7 Li H, Wang X G, Lin Z L, et al. US Patent, US10178069B2, 2019-01-08

8 Li H, Li K J, Chen Y L, et al. US Patent, US10178069B2, 2019-03-14

9 Li H, Wu J X, Zhang X C, et al. PCT No.CN2019/073507, 2019-01-28

10 Li K J, Li H, Hou H X, et al. Proof of vote: a high-performance consensus protocol based on vote mechanism & consortium blockchain. In: Proceedings of IEEE International Conference on High Performance Computing and Communications (HPCC), Bangkok, 2017. 18–20

11 Jacobson V, Smetters D K, Thornton J D, et al. Networking named content. Commun ACM, 2012, 55: 117–124

12 Moiseenko I, Oran D. TCP/ICN: carrying TCP over content centric and named data networks. In: Proceedings of ACM Conference on Information-Centric Networking (ICN), Kyoto, 2016. 112–121

13 Wu H, Shi J X, Wang Y X, et al. On incremental deployment of named data networking in local area networks. In: Proceedings of ACM/IEEE Symposium on Architectures for Networking & Communications Systems (ANCS), Beijing, 2017. 82–94
14 Kietzmann P, Gündogan C, Schmidt T C, et al. The need for a name to MAC address mapping in NDN: towards quantifying the resource gain. In: Proceedings of the 4th ACM Conference, Berlin, 2017. 36–42
15 Ou S W, Xie R C, Huang T, et al. A survey of communication between IP and IC. Inform Commun Technol, 2017, 11: 53–60 [欧思维, 谢人超, 黄韬, 等. IP 网络与 ICN 网络互通研究. 信息通信技术, 2017, 11: 53–60]

Prototype and testing report of a multi-identifier system for reconfigurable network architecture under co-governing

Hui LI[1*], Jiangxing WU[2], Kaixuan XING[1], Peng YI[2], Shisheng CHEN[3], Wei LIANG[3], Jinwu WEI[4], Wei LI[4], Fusheng ZHU[5], Kaiyan TIAN[6], Jiang ZHU[6], Yiqin LU[7], Ke XU[8], Jiaxing SONG[8], Yijun LIU[9], Yongji DONG[2], Yongxiang HAN[10], Hanxu HOU[10], Junfeng MA[11], Rui XU[1], Jianming QUE[1], Weihao YANG[12], Weihao MIU[13], Zefeng ZHENG[14], Tao SUN[15], Guohua WEI[1], Jiuhua QI[1], Ji LIU[1], Yongjie BAI[1], Chonghui NING[1], Han WANG[1], Xinchun ZHANG[1], Jiawei HU[1], Jiansen HUANG[1], Sai LV[1], Xinwei LIU[1] & Gengxin LI[1]

1. *Peking University Shenzhen Graduate School, Shenzhen* 518055, *China*;
2. *National Digital Switching System Engineering and Technological Research Center, Zhengzhou* 450002, *China*;
3. *China Telecom Corporation Limited, Beijing* 100033, *China*;
4. *China United Network Communications Limited, Beijing* 100032, *China*;
5. *Guangdong Communications & Networks Institute, Guangzhou* 510670, *China*;
6. *Kingsoft Cloud Network Technology Co., Ltd., Beijing* 100085, *China*;
7. *School of Electronic and Information Engineering, South China University of Technology, Guangzhou* 510006, *China*;
8. *School of Information Science and Technology, Tsinghua University, Beijing* 100084, *China*;
9. *School of Computers, Guangdong University of Technology, Guangzhou* 510006, *China*;
10. *School of Electrical Engineering and Intelligentization, Dongguan University of Technology, Dongguan* 523808, *China*;
11. *The China Academy of Information and Communications Technology, Beijing* 100191, *China*;
12. *Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, China*;
13. *Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, China*;
14. *Faculty of Information Technology, Macau University of Science and Technology, Macao, China*;
15. *Network & Information Center of Shenzhen University Town, Shenzhen* 518055, *China*

* Corresponding author. E-mail: lih64@pkusz.edu.cn

Abstract Even though the internet has become an essential infrastructure of modern society, the centralized domain name system (DNS) is unable to provide high-quality services and secure management and controlling. To face this challenge, we have proposed a reconfigurable multi-identifier network architecture and have developed a prototype of the multi-identifier system. We optimized network security using consortium blockchain, improved forwarding speed using HPT for FIB, and enhanced scalability through a tunnel algorithm. Experiments and tests of this prototype on the network of the two largest telecommunication service providers in China demonstrate that the system is considerably robust to support real-world traffic. Moreover, it can be applied to sovereign networks and other private networks with multiple identifiers, which may become a Chinese solution of network security to the world.

Keywords multilateral co-management, DNS, multimodal network, consortium blockchain consensus PoV, multi-identifier inter-translation, post IP high security private network

Hui LI was born in 1964. He received his B.Eng. and M.S. degrees in School of Information Engineering from Tsinghua University, China, in 1986 and 1989, followed by his Ph.D. degree in Department of Information Engineering from the Chinese University of Hong Kong in 2000. Currently, he is a Professor at Peking University, China. His research interests include future network architecture, cyberspace security, and blockchain technology.

Kaixuan XING was born in 1994. He is a postgraduate student at the Peking University Shenzhen Graduate School. His research interests include new architectures and new generations of information communication technology.

Jiangxing WU was born in 1953. He is an Academician of Chinese Academy of Engineering and a Ph.D. supervisor at National Digital Switching System Engineering & Technological Research Center (NDSC). His current research interests are information & communication technology and cyberspace security.