软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2024,35(10):4681−4695 [doi: 10.13328/j.cnki.jos.006990]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

面向人机物融合的数联网标识解析技术*

张 宁[1], 柳 熠[2], 马新建[2], 姜海鸥[2], 王 璐[1], 景 翔[3], 黄 罡[2,4]

[1](西安电子科技大学 计算机科学与技术学院, 陕西 西安 710071)

[2](北京大数据先进技术研究院 大数据操作系统研究部, 北京 100195)

[3](北京大学 软件与微电子学院, 北京 100871)

[4](北京大学 计算机学院, 北京 100871)

通信作者: 黄罡, E-mail: hg@pku.edu.cn

摘 要: 以数据的深度挖掘与融合应用为主要特征的信息化 3.0 阶段正在开启, 传统静态环境下的软件正向人机物融合环境下开放动态的复杂软件演化. 如何在不可信不可控的互联网之上实现全网一体化可信可管可控的数据互联互通是当前亟待解决的难题, 以数字对象架构、标识解析技术为代表的数联网技术体系为上述挑战提供了一种可行思路. 针对互联网上数据资源共享交换过程中广泛存在的传输效率低、协调成本高、安全管控难等问题, 提出面向人机物融合的数联网标识解析技术规范, 并基于人机物融合环境中各实体之间数据资源可发现、可获取、可理解、可信任、可互操作等需求, 设计实现面向人机物融合的数联网数字对象标识解析协议以及数字对象标识系统. 最后, 对所实现的数字对象标识系统进行系统测试与评估, 并在实际应用场景中对其有效性进行验证.

关键词: 标识解析; 人机物融合; 数字对象; 数联网

中图法分类号: TP311

中文引用格式: 张宁, 柳熠, 马新建, 姜海鸥, 王璐, 景翔, 黄罡. 面向人机物融合的数联网标识解析技术. 软件学报, 2024, 35(10): 4681–4695. http://www.jos.org.cn/1000-9825/6990.htm

英文引用格式: Zhang N, Liu Y, Ma XJ, Jiang HO, Wang L, Jing X, Huang G. Identifier Resolution Technology for Human-cyber-physical Ternary Based on Internet of Data. Ruan Jian Xue Bao/Journal of Software, 2024, 35(10): 4681–4695 (in Chinese). http://www.jos.org.cn/1000-9825/6990.htm

Identifier Resolution Technology for Human-cyber-physical Ternary Based on Internet of Data

ZHANG Ning[1], LIU Yi[2], MA Xin-Jian[2], JIANG Hai-Ou[2], WANG Lu[1], JING Xiang[3], HUANG Gang[2,4]

[1](School of Computer Science and Technology, Xidian University, Xi'an 710071, China)

[2](Department of Big Data Operating System, Advanced Institute of Big Data, Beijing 100195, China)

[3](School of Software and Microelectronics, Peking University, Beijing 100871, China)

[4](School of Computer Science, Peking University, Beijing 100871, China)

Abstract: The informationization 3.0 represented by deep mining and fusion applications of big data is starting, and the software in the traditional static environment is evolving into complex software in the human-cyber-physical ternary environment which is open and dynamic. How to realize the trusted, manageable, and controllable data interconnection on the untrusted and uncontrollable Internet is an urgent problem to be solved. The Internet of Data technical system represented by digital object architecture and identifier resolution technology provides a feasible solution for these challenges. In order to solve the problems such as low transmission efficiency, high coordination cost, and security management issues in the process of data sharing on the Internet, this study proposes identifier resolution standard specifications for human-cyber-physical ternary environments. Moreover, to meet the demands that data resources owned by different entities need to be discoverable, accessible, understandable, trustworthy, and interoperable in the human-cyber-physical ternary

* 基金项目: 国家重点研发计划 (2021YFF0901100); 北京市科技新星计划 (Z211100002121159)
收稿时间: 2022-08-16; 修改时间: 2022-11-09, 2023-04-12; 采用时间: 2023-06-19; jos 在线出版时间: 2023-10-11
CNKI 网络首发时间: 2023-10-16

environment, this study designs the identifier resolution protocol and implements the identifier/resolution prototype system for human-cyber-physical ternary environments. At last, this study tests the performance of the prototype system, and the validity of the system is verified by applying it to application scenarios.

Key words: identifier resolution; human-cyber-physical ternary; digital object; Internet of Data

信息技术的飞速发展进一步推动了人类社会、信息空间和物理空间的融合, 形成了集人机物于一体的三元融合型泛在计算环境. 近年来云计算、移动终端、物联网的兴起, 使得作为数据来源的人机物规模进一步增长, 数据已经渗透到当今社会各个行业及其业务职能领域. 根据联合国《数字经济报告 2021》中统计[1], 2020 年 47 个国家的数字经济规模已经达到 32.6 万亿美元, 占 GDP 比重为 43.7%, 数据已成为创造私人价值和社会价值的基础性战略资源和关键性生产要素.

人机物融合的一个共性需求是数据互联互通和智能应用[2]. 然而, 现有以机器和设备为中心的互联网协议栈并不适用于人机物融合计算环境下全球数据治理的具体特点和应用需求, 亟需建立一个结合多边、多利益相关方和多学科参与的新型技术体系及协议栈, 以此推动全球范围内网络空间向数字空间的数字化转型. 其根本原因在于, 互联网的核心 TCP/IP 协议主要关注数据在“二进制位”层面的编解码和传输控制, 应用系统之间需要自行在数据的语法、语义、语用等层面协调一致, 并面临着协调成本高、责权利效难保障, 以及低效、易错、难溯源等挑战.

北京大学系统软件团队借鉴互联网的设计理念, 采用软件定义的思路, 通过以数据为中心的开放式软件体系结构和标准化互操作协议, 将各种异构数据平台和系统连接起来, 从而在“物理/机器”互联网之上形成一个“虚拟/数据”网络, 即“数据互联网” (Internet of Data), 简称“数联网”[3,4]. 数联网的核心理念是 Kahn 等人提出的数字对象架构 (digital object architecture, DOA)[5], 它是对现有互联网的一种逻辑扩展, 即在现有互联网基础之上, 将数据进行标记、聚合, 并将数据之间的关联关系进行清晰的表达, 有效降低互联网管理信息的障碍. 通过将现有互联网上分布在不同地域、不同架构、不同模式的信息系统协同工作、统一管理, 消除这些系统之间的数据烟囱和数据孤岛, 实现异构、异地、异主数据的互联互通和互操作.

综合上述分析不难发现, 数联网将会成为以人机物融合为标志的大数据时代新型信息基础设施, 以标识解析技术为代表的数联网相关核心技术问题亟待解决和突破. 基于此, 本文开展面向人机物融合的数联网数字对象标识解析技术研究, 通过借鉴通用领域标识解析技术, 建立数联网标识解析技术规范, 提出能够兼容已有标识编码规则的数联网标识编码体系, 支撑各类标识编码体系中的数据资源互访问和互操作. 设计实现数联网数字对象标识解析协议, 通过将网络层传输协议和标识解析路由策略解耦, 以适配人机物融合环境下不同应用场景. 并结合分布式账本、节点动态扩容、访问控制、异常检测等技术, 设计实现包括客户端 SDK 以及服务端组件在内的数字对象标识原型系统, 为数联网上层应用提供标识基础设施支撑能力.

本文第 1 节介绍数联网和标识解析技术的相关工作和研究现状. 第 2 节介绍本文提出的数联网标识解析技术规范. 第 3 节基于开源的标识解析协议设计实现面向人机物融合的数联网数字对象标识解析协议. 第 4 节开发实现数字对象标识系统所需的相关系统软件. 第 5 节对原型系统进行测试评估, 并在应用场景中对相关技术进行有效性验证. 第 6 节总结本文工作, 对未来工作进行展望和探讨.

1 相关工作

1.1 数联网及 DOA

国内外针对数联网及 DOA 的早期研究与探索主要是互联网发明人、图灵奖得主 Robert Kahn, 他在 2014 年与国际电信联盟 (ITU) 合作成立技术标准开放基金会 DONA, 开始对 DOA 的概念与标准进行推广[6,7]. 在此基础上, 国际电信联盟电信标准分局 (ITU-T) 发表了两个针对 DOA 的标准建议稿[8,9]. 基于 DOA 的最大规模应用是 DOI 系统, 它通过将互联网上数字资源构建为数字对象并分配唯一且持久的 DOI 标识, 使得在任意一个支持 DOI 的应用系统中都可以解析到该标识对应的文献实体[10]. 此外, 美国在 2019 年首次发布了《国防部数字现代化战

略白皮书》[11], 提出了在国防信息系统 (机器互联网络) 之上构造国防数字信息网络 (数据互联网络), 进而实现网络中心战向数据中心战转型. 软件定义网络发明人、斯坦福大学 Nick McKeown 教授联合普林斯顿大学和康奈尔大学于 2020 年承担了美国国防部先进计划研究局的项目 Pronto, 探索在新一代 5G 技术下建立一个覆盖云、网、边、端的端到端软件定义网络, 使得网络层资源全栈可编程、可验证[12]. 大数据核心系统软件 Spark 研制团队、加州大学伯克利分校的 Stoica 等人在 2021 年 HotOS 上提出了 Sky Computing, 旨在构建各个云计算数据中心之间的算力互联网, 让云计算应用更方便地使用不同云计算提供商的特色服务, 提供更好的安全性、可靠性和性能, 降低云计算应用的成本[13].

近年来, 我国在与数联网相关的云计算与大数据、互联网新型结构、软件定义技术、泛在系统软件等领域也有着众多的研究成果与技术实践. 北京大学、清华大学、复旦大学、上海交通大学、南京大学、香港理工大学、中国科学院等团队于 2018 年启动了数联网架构、技术与系统软件的研究, 同时在 2020 年牵头制定了 DOA 两大核心协议之一的 DOIP 新版标准[14], 并在 2021 年实现了 DOIP 开源软件栈[15]. 国家工业信息安全发展研究中心、北京中数创新科技股份有限公司、北京西恩多纳信息技术有限公司组成的“MPA 中国联合体”于 2014 年开始参与 DONA 的创建, 与多国平等共建 DOA 全球节点并开始运营服务[16]. 工信部下属的中国信息通信研究院于 2018 年开始建设国家工业互联网标识解析系统, 截至 2021 年 12 月, 已经建成北京、上海、广州、重庆和武汉等 5 个国家顶级节点以及 160 个二级节点, 覆盖 25 个省市自治区 31 个行业超过 4 万家企业, 标识注册量超过 600 亿, 日均解析量超过 9 000 万次[17].

1.2 标识解析技术

目前, 被广泛运用的标识解析技术方案主要包括 Handle、OID (object identifier)、Ecode (entity code for IoT)、EPC 和 GS1 等. 除此之外, 一些行业/企业会根据自身的实际需求制定许多私有化标识, 但这些标识编码规范大多停留在编码层面且为内部使用. Handle 是由美国国家研究创新机构 CNRI 提出, 用于分配、管理和解析互联网上数字对象和其他资源永久标识符的综合技术体系[18−21]. OID 是由 ISO/IEC、ITU-T 国际标准化组织于 20 世纪 80 年代联合提出的标识机制, 采用分层树形结构对任何类型的对象进行全球无歧义、唯一命名, 并通过分布式的解析系统实现对象信息查询[22,23]. Ecode 是我国为适应全球物联网发展自主创新出来的一整套物联网标识体系标准, 包括 Ecode 编码、数据标识、中间件、解析系统、信息查询和发现服务系统、安全保障系统等内容, 是目前唯一具有我国自主知识产权的标识体系[24]. GS1 编码标识体系主要面向商品流通, 其中编码体系是整个 GS1 系统的核心, 它是面向流通领域中所有产品与服务的标识代码及附加属性代码[25,26]. EPC 编码标识体系是建立在 GS1 基础上, 用以实现对单品进行标识, 目前已融入 GS1 体系并成为 GS1 标准体系的一部分[27]. 此外, 在国内面向工业互联网的标识解析体系也日趋完善, 它是实现工业全要素、各环节信息互通的关键枢纽.

近年来我国标识解析技术的发展虽然已经取得较大进步, 但在底层基础技术方面的瓶颈问题仍然突出, 关键核心技术受制于人的局面还没有得到根本性改变. 总体而言, 世界范围内还未形成统一的、成熟的、大规模部署的全球性标识解析体系, 现有的几种标识解析技术方案中的技术、标准、产业、应用、治理等方面还有待进一步完善, 亟需对标识解析核心技术和应用方案进行有组织、大规模的验证示范及推广.

2 数联网标识解析技术规范

标识一般是指不同物体或客观对象所具有的区分于其他事物的名称标记, 可以由数字、字母、符号、文字等内容以一定的规则组合而成. 标识编码是唯一识别数据、算法、文件等虚拟资源的身份符号, 而标识系统则通常根据标识编码查询目标数据所在的网络位置或其他相关状态信息. 数联网标识解析是以实现人机物融合环境中数据资源互联互通为目标的标识注册、解析、管理、安全的技术体系, 能够为其中的数据资源提供永久标识、动态链接和安全管理等基础服务, 其作用类似于 DNS[28], 但数据管理功能更加强大、安全机制更加完善.

针对目前多种标识编码体系之间兼容性不足, 难以实现不同标识编码规则间数据互联与共享等问题, 本文提出了一种面向人机物融合大数据场景且兼容多编码规则的数联网标识解析技术规范, 实现数据标识在数联网范围

内的唯一寻址以及与其他数据标识编码之间的数据映射. 标识编码规范作为数联网标识解析技术体系的基础规范以及标识系统的核心资源, 用于规范数字对象的标识分类、编码规则以及编码结构.

数联网标识编码规则采用分级编码结构, 分为前缀和后缀两部分, 标识前缀用于唯一标识单位主体或根域节点, 标识后缀用于唯一识别数字对象. 第 1 部分为标识前缀, 前缀定义标识编码的命名空间, 保障标识的全局唯一性, 下一级节点的前缀都需要向上级节点进行申请, 通常由 3 – n 个节点组成, 如图 1 所示. 第 1 个节点是根节点标识符, 第 2 个节点到 n – 1 个节点是域节点标识符, 分别对应一级域节点, 二级域节点,…, (n – 2) 级域节点, 域节点的级数可以按照需求自定义进行扩展, 例如 OS 是一级域节点标识, DOA 是二级域节点标识, 第 n 个节点是端节点, 即 dev 为端节点标识.

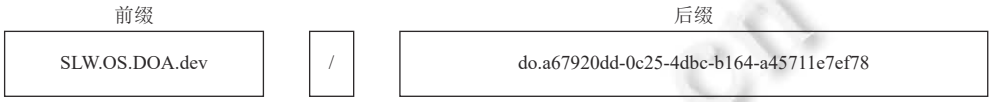


图 1 数联网标识编码结构示例

第 2 部分为标识后缀, 前缀一般由上级节点分配, 是相对“固定”的, 而后缀的编码规则由各个端节点的提供者自由决定, 在保证唯一的前提下, 可以使用 UTF-8 的编码进行自由定义, 既可以使用有意义的编码体现数据的规律性, 也可以使用无意义编码来保护数据的私密性, 由两部分标识符组成. 第 1 个是类型标识, 分为 do、dou、doip 等, 其中 do 代表数字对象, dou 代表参与者, doip 代表注册表和仓库等数联网组件. 第 2 个标识符是自定义代码, 代码位数不定长, 标识前缀部分和后缀部分用“/”隔开, 每一部分内部分级以“.”符号隔开, 面向人机物融合的数联网标识解析体系编码支持阿拉伯数字、英文字母组合, 默认情况下, 不区分大小写.

为了支持数联网标识编码体系对其他不同编码规则的适配, 本文扩展了标识属性和种类使其支持多种编码规则体系, 通过将前缀按照一定规则进行划分以及分配, 后缀部分保留使用现行的各类标识编码, 如 OID、Ecode、Handle 等. 同时通过对数联网数字对象标识解析协议参考实现进行改进, 对外提供可供用户使用并兼容多编码规则的 SDK, 用户在使用时只需通过统一的 SDK 即可调用或解析基于不同编码规则的数据标识, 从而在协议层面实现多编码标识体系下不同数据的访问以及互操作, 目前已支持对 Handle 标识的互操作.

3 数联网数字对象标识解析协议

本文基于 IETF (Internet engineering task force) 发布的 RFC 标识解析协议标准, 面向人机物融合环境中的应用需求, 对协议部分内容进行了扩展和修改, 使其能够兼容不同的标识编码体系, 支持服务节点间相互认证, 并增加了对解析策略、分布式账本等功能的配置和处理逻辑, 从而形成数联网数字对象标识解析协议参考实现. 该协议是用于数联网中创建、更新、删除和解析数字对象标识的客户端服务器应用协议, 描述了服务器如何向客户端解析给定数字对象标识的过程, 并定义了客户端和服务器之间进行标识操作时的消息交互格式以及应遵循的规范. 本文所实现的应用协议数据包的传输顺序遵循网络字节顺序, 即通常所说的大端模式 (big-endian), 消息可以通过 TCP/UDP 或其他传输层网络协议进行传输, 标识默认的编码方式是 UTF-8, 目前提供了可供用户 (开发者) 直接使用的 Java 版本 SDK, 并在国产 PK 体系上进行了适配验证与测试.

3.1 消息格式

每个消息报文都由 4 部分组成, 消息信封 (message envelope)、消息头 (message header)、消息体 (message body) 以及消息凭证 (message credential). 协议消息报文各字段结构示意图如图 2 所示.

(1) 消息信封

每条消息都以消息信封作为开头, 如果一条消息被截成多段进行传输, 那每段也必须以消息信封作为开头, 有 20 个字节的固定大小, 由 7 个字段组成. 消息信封中的字段不包含任何应用层相关的信息, 主要用于帮助在网络上进行消息传输, 不受消息凭证中数字签名的保护. MajorVersion 和 MinorVersion 两个字段用于指定标识解析协议的版本. MessageFlag 是消息标志位, 其中 CP 标志位用来表示消息是否进行了压缩, EC 标志位用来表示消息

是否进行了加密, TC 标志位用来表示该消息是否是一个截断消息, 在使用 UDP 进行大数据量的消息传递时通常会将其拆分成多个消息进行分段传输, 其余 13 位均置为 0 作为预留标志位. SessionId 字段用于指定客户端和服务器之间的通信会话, 通常由服务器进行分配, 当会话未建立时, 客户端需将请求消息中的该字段设置为 0, 服务器则会为每个新建会话分配一个唯一的非零 SessionId, 并自行终止一段时间内未使用的会话. 在已经建立的会话下进行消息交换时, 服务器和客户端需要维护相同的 SessionId. RequestId 字段是在客户端发起请求时由客户端指定并维护, 同一个客户端所有未完成的请求 RequestId 字段必须唯一, 当服务器处理完请求并返回响应时, 也必须包含对应的请求 RequestId 以方便客户端跟踪请求消息的处理结果. SequenceNumber 字段是在使用分段传输时标记每个截断消息的计数器, 当消息接收方进行处理时, 可以根据 SequenceNumber 的数值将原始消息进行重组. 对于分段传输的每条截断消息, 必须将其 TC 标志位置为 1 并从 0 开始计数. MessageLength 字段指定了每个消息的字节长度, 但不包括消息信封自身所占用的 20 个字节. 单个消息的长度不能超过 4 字节无符号整数的大小限制, 否则应截断分成多个消息进行传输, 并将 RequestId 置为相同的值.

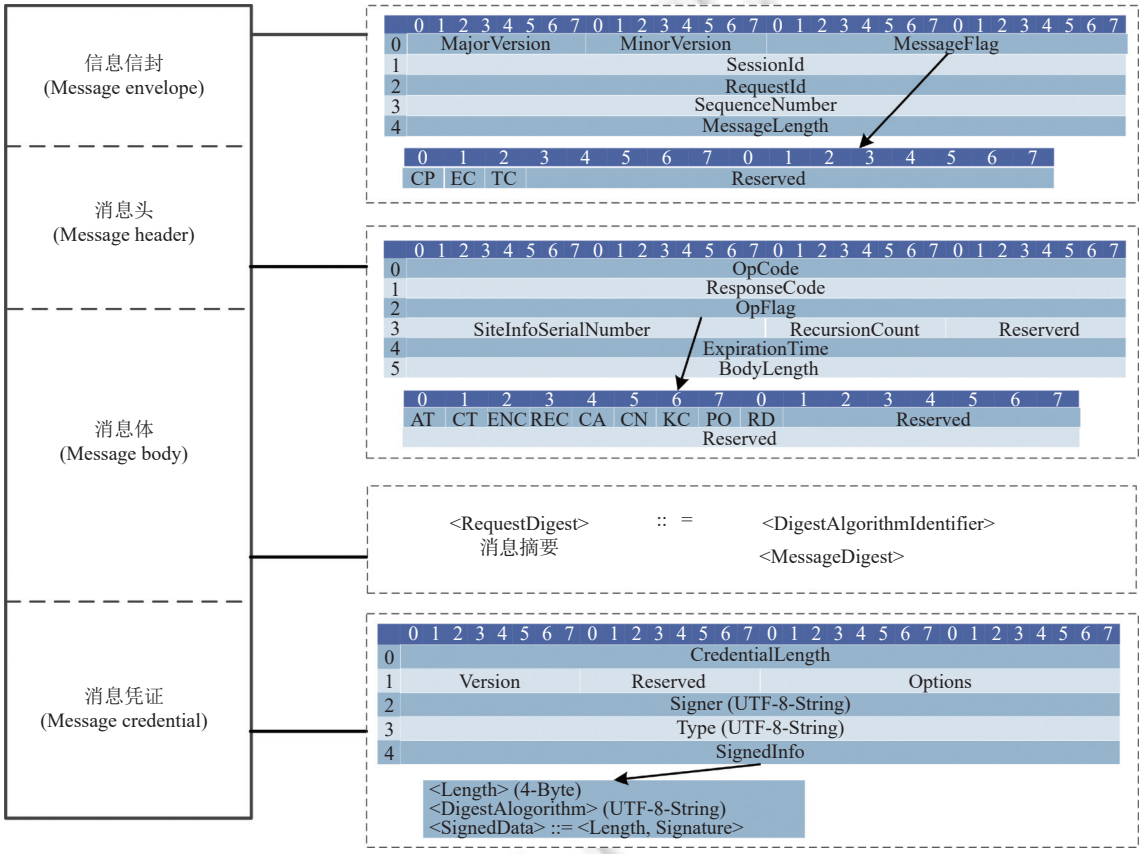


图 2 协议消息报文各字段结构示意图

(2) 消息头

消息头包含了客户端和服务器之间进行信息交互的公共数据字段, 有 24 个字节的固定大小, 包括操作码、响应码、控制标志位等部分. 与消息信封必须出现在每个消息开头不同, 当消息是单个进行传输时必须有消息头, 但如果进行了截断分开传输, 则只需在第一个消息中指定消息头, 其余分段消息可以不包含消息头. 其中 OpCode 字段用于指定消息的操作码, 客户端在发送消息时通常使用 OpCode 通知服务器需要完成的标识操作请求, 服务器处理完成后返回的响应消息中 OpCode 字段须与请求消息中的一致. ResponseCode 字段用于指定消息的响应码,

服务器对于任何请求的响应结果都通过该字段进行指定, 通常客户端在发送请求消息时将 ResponseCode 字段置为 0, 服务器通过非 0 的 ResponseCode 响应码来表示对请求处理的结果.

OpFlag 字段定义了协议操作的各种控制选项, 服务器在对请求进行处理时需要遵守 OpFlag 字段中所有指定的标志位, 如果无法满足请求的控制选项, 需要在响应码中给出错误信息. 其中 AT 标志位在请求消息中置为 1 表示请求应该定向到主服务器, 在响应消息中置为 1 表示该响应是从主服务器返回. CT 标志位在请求消息中置为 1 表示请求需要服务器使用数字签名对响应进行签名, 在响应消息中置为 1 表示响应已经进行了签名. ENC 标志位用来表示请求需要服务器使用预先建立的会话密钥对响应进行加密. REC 标志位置为 1 的请求表示服务器需要将该查询请求转发到其他合适的服务器进行处理, 并将处理结果返回给客户端, 服务器也可以通过设置特定响应码表示不支持此类请求. CA 标志位在请求消息中置为 1 表示需要缓存服务器代表客户端验证服务器的响应, 在响应消息中置为 1 表示缓存服务器已经对响应进行了验证. CN 标志位表示同一个请求或响应是否截断为多条消息进行发送. KC 标志位置为 1 时表示消息接受者需要保持 TCP 连接, 从而允许相同的 TCP 连接处理多个标识操作. PO 标志位主要用于查询请求. RD 标志位置为 1 的请求消息表示服务器需要在其响应消息中包含请求的消息摘要, 置为 1 的响应消息表示返回的响应消息正文第一个字段包含原始请求的消息摘要.

SiteInfoSerialNumber 字段在请求消息中指的是客户端使用的服务器版本信息. RecursionCount 字段用于指定服务递归的次数, 如果服务器将客户端的请求转发到另一台服务器时, 需将 RecursionCount 字段增加 1, 任何来自客户端的直接请求必须将 RecursionCount 字段置为 0, 当服务器完成处理返回响应消息时需要保持与请求消息中 RecursionCount 字段相同的值. 为了防止无限递归, 服务器可以配置为 RecursionCount 字段达到某个值时停止转发递归请求. ExpirationTime 指定了一个消息的过期时间, 通过 unix 时间戳以秒为单位进行表示, 如果置为 0 则表示该消息没有过期时间. BodyLength 用于指定消息体的字节长度, 不包括消息头以及消息凭证的字节长度.

(3) 消息体

消息体中存放的是每个协议操作所对应的数据, 其格式会根据消息头中操作码和响应码的不同而有所区别, 字节长度通过消息头中的 BodyLength 字段确定, 可以为空且受消息凭证中数字签名保护. 消息体在传输过程中可能被截断为多个部分进行传输. 对于任何响应消息, 如果消息头的 RD 标志位置为 1, 则消息体必须以原始请求的消息摘要开头. 消息摘要主要包括两个字段, 其中 DigestAlgorithmIdentifier 字段是单字节标志位, 指定了消息摘要的生成算法, MessageDigest 字段是消息摘要正文本身, 根据原始请求的消息头和消息体计算得出, 其长度根据不同的摘要生成算法而有所不同.

(4) 消息凭证

消息凭证为客户端和服务器之间进行信息交换提供了一种安全传输机制, 通常包含来自消息发送者的数字签名或者基于预先建立的会话密钥生成的单向消息验证码, 主要用于客户端和服务器间消息验证以及消息传输后数据完整性检查, 从而保护消息头和消息体中的内容在传输过程中不被篡改. CredentialLength 字段指定了消息凭证的字节长度, 如果消息不包含消息凭证, 需将该字段的值置为 0. Version 字段指定了消息凭证的版本号. Reserved 字段和 Options 字段均为预留字段. Signer 字段和 Type 字段均是一个 UTF-8 的字符串编码, 前者用来指定签名者的信息 (比如公钥或证书), 后者用来指定生成数字签名的方式. SignedInfo 字段是具体的数字签名信息. 数字签名的作用是证明服务器确实已响应客户端的请求并作为不可抵赖的凭证, 消息头和消息体都符合其消息凭证的消息表明该请求或响应确实来自发起者, 并确保在其传输过程中未被篡改.

3.2 操作响应码

客户端和服务器之间传递消息时通常需要在消息头中指定操作码 OpCode 以及响应码 ResponseCode, 这两个字段决定了每个消息中消息体的具体字段结构. 本文对开源协议规范中的 OpCode 以及 ResponseCode 进行了改进, 主要扩展的操作响应码如表 1 中加粗部分所示. 首先, 为了支持对其他标识编码规范的兼容, 保留了开源协议规范中的基本操作码和响应码, 如 RC_SUCCESS、RC_ERROR、RC_NA_DELEGATE 等, 增加了数字对象标识 DOID 增删改查的操作码和数联网中其他构件如 Repo (仓库) 的操作码以及对应的消息处理逻辑, 通过在标识解

析协议 SDK 实现过程中使用不同类型的 Response 对象进行区分处理, 使其可以兼容多种标识编码规范. 其次, 增加了 OC_VERIFY_IRS 等节点认证相关操作码使得标识服务节点 (根域) 间的认证也可以使用数联网数字对象标识解析协议统一进行消息交互. 最后, 改进了响应码中的错误类型以支持数联网数字对象标识系统不同解析策略的实现, 同时在消息发送和接收逻辑中增加了存证标志位及相关配置以支持在标识重要处理流程中接入分布式账本.

表 1 协议主要操作响应码

操作码	协议操作	响应码	处理结果
OC_RESOLUTION_DOID	标识解析	RC_RESERVED	请求预留码
OC_CREATE_DOID	标识创建	RC_SUCCESS	请求成功
OC_BATCH_CREATE_DOID	批量创建	RC_ERROR	请求失败
OC_UPDATE_DOID	标识更新	RC_NA_DELEGATE	转发请求
OC_DELETE_DOID	标识删除	RC_SERVER_BUSY	服务器繁忙
OC_GET_USERS	获取用户	**RC_DOID_NOT_FOUND**	标识不存在
OC_GET_REPOS	获取仓库	**RC_DOID_ALREADY_EXIST**	标识已存在
OC_VERIFY_IRS	验证域节点	**RC_INVALID_DOID**	无效标识

解析操作主要是客户端用来向服务器查询给定标识所对应的标识内容, 如果服务器在其数据库中可以找到标识则将查询结果放入消息体中返回给客户端, 如果待解析的标识在数据库中不存在, 则将错误信息返回给客户端. 如果服务器不负责所要查询的标识, 则应将响应码设为服务引用, 表示需要转发本次解析请求. 除此之外, 服务器还可以返回其他响应错误码, 这些情况均视为解析请求的处理结果为失败.

只有命名机构管理员才有权限在对应的域节点创建标识, 在服务器接收到创建标识的请求时, 应该首先对请求消息的签名进行验证, 确保发送请求的客户端具有当前节点的标识创建权限, 再进行后续创建操作, 否则视为失败. 服务器在创建标识时, 应将整个请求作为事务执行, 过程中有任何部分失败都应该回滚整个操作. 此外, 面向人机物融合的数字对象标识解析协议还提供了批量创建标识的操作码, 方便进行大批量的标识生成. 标识更新是指对已经创建的标识更新修改其对应的标识内容. 标识删除是指删除域标识节点中已经存在的标识, 标识更新和删除同样需要服务器对客户端的身份进行验证以确保其具有对应的操作权限, 否则的话视为失败.

域节点认证是指下级的标识域节点在启动时需要向自己的上级标识域 (根) 节点发送认证信息, 从而在上级标识节点中更新下级域节点的地址、用户签名等信息, 在发送认证请求前下级域节点需要首先向上级域节点管理机构申请并分配前缀并在本地域节点标识系统上进行配置与部署.

3.3 其他说明

面向人机物融合的数联网数字对象标识解析协议规范主要针对标识在传输过程中应遵循的编解码标准进行约定, 并未对标识解析服务器和客户端的具体实现进行约束, 在对标识解析服务器组件进行开发的过程中, 只需满足上述标识解析协议标准 (使用本文提供的 SDK) 就可以进行标识的互操作. 除此之外, 还应根据实际情况考虑如服务器最大并发进程/线程数、TCP 连接超时间隔、UDP 重连策略、递归解析策略、事务日志级别分类等内容, 并允许管理员对其进行配置和微调. 服务器系统软件还应该能够支持各种标识内容的存储, 并能容忍较高的客户端并发请求数量. 客户端在设计实现时也需要能够处理服务器端返回的各种数据标识, 并在进行递归解析请求处理时设置相关策略以避免陷入死循环.

4 数联网数字对象标识系统设计实现

传统的 DNS 服务日益暴露出在作为通用资源寻址技术方面的缺陷, 即脆弱的安全性、可靠性, 易受到 DNS 欺骗攻击、DDOS 攻击, 缺乏安全访问控制策略, 数据更新延迟大等. 未来的人机物融合环境中亟待一种具有灵活访问机制、管理机制以及安全机制的网络资源寻址定位技术, 数联网数字对象标识系统正是满足上述需求的一种解析技术体系, 它可以实现数据共享、数据确权、安全控制等传统 DNS 无法满足的功能. 本节主要基于数联网标

识解析技术规范和数字对象标识解析协议, 设计实现面向人机物融合的数联网数字对象标识原型系统, 为数联网上层应用提供标识寻址支撑能力.

4.1 标识系统设计

数联网数字对象标识系统能够为数联网提供数据标识管理、数据标识解析、数据标识业务管理、数据标识查询以及异构数据标识互操作等多个方面能力. 其中, 数据标识管理是指基于数联网标识编码规范, 数字对象标识系统能够对数据标识进行创建和存储. 数据标识解析是指标识系统能够提供高效稳定、安全可靠的标识解析服务, 同时具备相应的密钥安全管理、身份鉴别认证等功能. 数据标识业务管理是指标识系统具备用户管理、注册管理、日志管理、监控管理和统计分析等业务支撑和运维管理能力. 数据标识查询是指标识系统能够提供标识的数据信息查询能力, 向信息系统以及相关用户展示标识映射信息和关联信息. 异构数据标识互操作是指依托标识系统的强兼容特性, 对目前国内外存在的其他标识解析体系如 Handle、OID、Ecode 等, 在技术上具备实现异构标识融合解析的能力.

(1) 系统架构

数联网数字对象标识系统主要以树状的组织结构层次化地管理和分配数据标识, 并对外提供数字对象标识解析协议, 从上到下分别为根节点、域节点、端节点.

根节点在树状层次顶层, 根节点可以有多个, 所有根节点组成根节点网络, 不同根节点负责管理其下属域节点的增加、删除和更新, 同时维护其下属的域节点信息, 包括节点标识前缀、IP 地址、标识注册量和解析量等数据, 根节点之间互相同步其下属域节点的前缀和 IP 地址. 域节点是根节点的下属节点, 每个域节点只能归一个根节点管理, 域节点负责管理其下属端节点的增加、删除和更新, 同时维护其下属端节点信息. 域节点要把标识注册量、解析量等数据同步给其所属的根节点. 端节点是域节点的下属节点, 每个端节点只能归一个域节点管理, 端节点负责管理标识的注册、解析、删除和更新等操作, 以及标识访问权限管理. 端节点是直接负责数据标识存储的节点, 标识解析请求最终都会由端节点做出应答.

(2) 解析策略

数联网数字对象标识系统可以根据不同的网络拓扑结构以及应用场景配置合适的标识解析策略, 从而提高整个系统解析性能, 根据根节点是否同步全网所有的域节点前缀信息可以分为根节点标识解析以及递归标识解析. 图 3 为不同解析策略路由流程示意图.

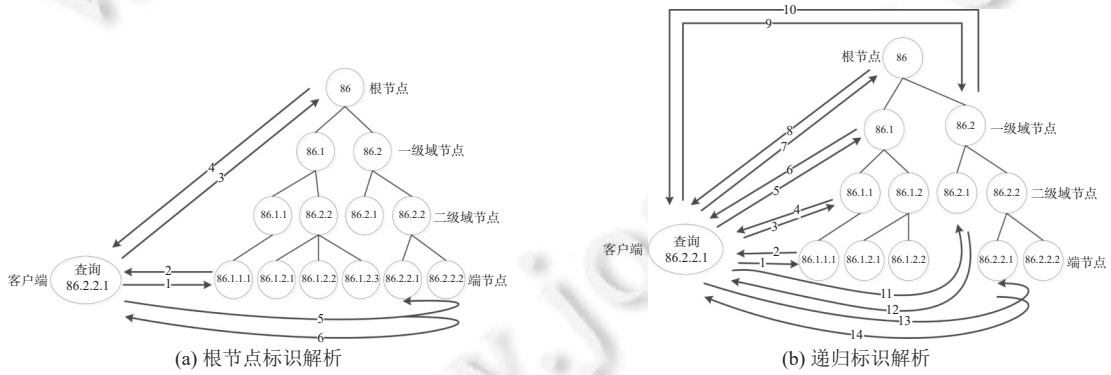


图 3 不同解析策略路由流程示意图

根节点标识解析策略的方法实现中根节点会存储全网所有域节点以及端节点的地址和前缀信息, 前缀的分配权限也需要通过根节点进行确认, 因此当端节点收到不是自己所属前缀的解析请求时, 都会先向根节点去请求目标标识前缀所在的域节点或者端节点信息, 进而完成整个解析流程. 递归标识解析策略的方法实现中根节点不会存储全网所有域节点以及端节点的地址和前缀信息, 每个域节点只会存储自己上一级节点 (根/域) 地址信息以及所管理的下一级所有节点 (域/端) 信息, 端节点只存储自身所管理的标识信息, 前缀的分配权限只需要通过上级节

点进行确认即可, 因此当某个节点收到不是自己所属前缀的解析请求时, 会先向自己的上一级节点或者下一级节点去进行递归搜索, 直到找到标识所在的前缀节点从而完成整个解析流程. 此外, 后续将通过引入分布式哈希表和动态缓存技术来提高系统在分布式网络环境中的解析性能.

(3) 支撑模块

数联网数字对象标识系统需要在根节点、域节点、端节点以及客户端等多主体之间进行数据交换以实现标识的解析注册, 面临着来源主体多、安全保障难、协作透明度低等问题. 为了降低不同主体间数据标识共享的潜在风险, 提升系统的公信力, 引入分布式账本技术, 为数联网数字对象标识系统提供内生的信任机制. 课题团队所研发的北大数瑞图式分布式账本具有高并发、高可用、线性扩展等特性, 十分契合人机物环境下的大数据标识解析应用场景. 通过对标识注册解析过程中涉及的相关重要操作进行可信存证, 保证整个过程不可篡改、全程溯源, 从而规范平台系统对标识以及数据的使用范畴、时机, 构建数据合规性评估机制, 防止数据标识的滥用, 形成可信的数据标识运行环境, 从而改变现有标识解析技术体系中行为缺乏监督、信息共享风险较高、出现问题后溯源难等痛点.

与此同时, 随着系统的节点增加、规模增长, 服务之间关联关系错综复杂, 会导致系统运行以及服务节点变更过程中易于发生故障, 因此需要智能化的运维手段快速发现、定位、解决各种类型的故障, 提高系统的可靠性. 异常检测模块以流式计算框架为基础, 基于日志数据利用实时计算、网络推断、因果推断等技术手段, 在系统日志数据中发现不符合预估行为的异常模式, 做到提前发现故障隐患以及自动化处理故障, 从而降低故障率并提升运维效率.

4.2 标识解析协议 SDK

标识解析协议 SDK 是本文针对数联网数字对象标识系统实现的客户端组件, 具有统一的操作调用接口, 接口主要包括标识解析、标识创建、标识更新、标识删除等, 提供了基本的标识管理以及用户认证等功能, 所有接口使用前均需先初始化客户端, 方式如下所示.

```
String url = “tcp://121.89.195.14:2641”;
JWK jwk = GlobalUtils.loadKeysFromJWKFile(“./keys/jwk/dou.TEST.jwk”);
IrpClientImpl client = new IrpClientImpl(jwk, “clientTest”, url);
client.connect(url);
```

上述代码中, url 参数是目标的域节点服务器地址, jwk 为本地存的用户私钥文件, 主要的功能接口如表 2 所示.

表 2 标识解析协议 SDK 主要功能接口

接口名称	功能	主要参数	备注
StateInfoBase resolve(String doid) throws IrpClientException	标识解析	待解析标识	StateInfoBase为数据标识的实体类. 标识解析接口默认不需要用户的认证信息
String register(StateInfoBase hr) throws IrpClientException List<String> batchRegister(StateInfoBase hr, int count) throws IrpClientException	标识创建/批量创建	待创建标识的数据对象、批量创建数量	
String reRegister(String doid, StateInfoBase hr) throws IrpClientException	标识更新	待更新标识及其数据对象	
String unRegister(String doid)	标识删除	待删除标识	

标识解析协议 SDK 底层网络传输是基于 Netty 实现, 它是 JBOSS 提供的一个异步、基于事件驱动的 Java 开源网络应用程序框架, 用以快速开发高性能、高可靠的网络 IO 程序. Netty 隐藏了 Java 原生 NIO 中客户端在处理断连重连、网络闪断、半包读写、失败缓存、网络拥塞和异常流控制等方面的复杂性, 并通过 IO 多路复用技术来同时处理多个客户端接入请求, 从而使得系统在单线程的情况下可以同时处理多个客户端请求. 图 4 给出了标识解析协议 SDK 中协议消息格式以及编解码等几个关键类的 UML 类图.

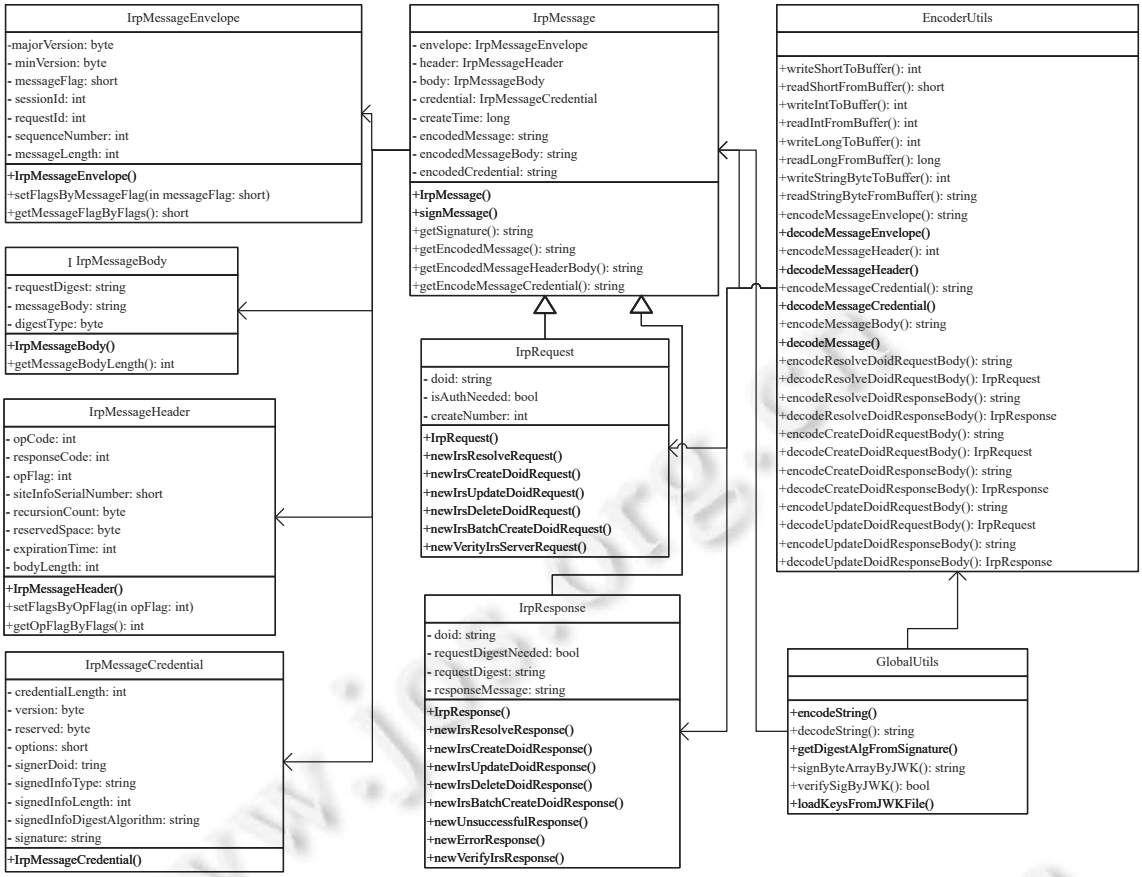


图 4　标识解析协议 SDK 组件部分 UML 类图

4.3　标识系统服务器组件

数联网数字对象标识系统服务器组件主要部署在标识系统的根域节点上, 负责对客户端以及其他域节点服务器的请求进行响应处理, 组件主要包括标识管理、标识解析、节点管理和用户管理 4 大功能模块, 如图 5 所示.

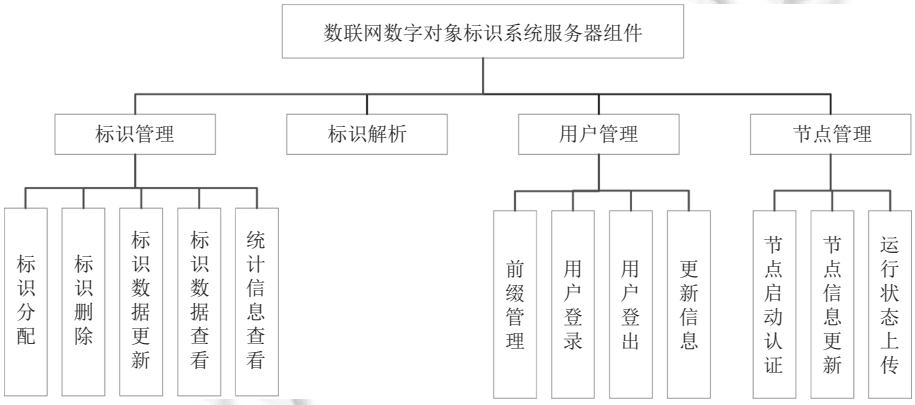


图 5　标识系统服务器组件功能架构图

在标识管理部分, 用户可以看到目前能够访问的数字对象以及可用的数联网基础设施列表, 包括数字对象仓库和数字对象注册表的信息, 同时也可以点击详情查看数字对象和基础设施的字段信息. 在标识解析部分,

用户可以通过标识 doid 解析已经存在的数据对象, 从而获得数字对象的具体内容. 利用管理员账号分配好的所有者标识和公钥, 用户可以以非管理员的身份登录标识系统服务器, 修改用户名、密码以及其他信息, 同时可以生成数字对象仓库标识. 管理员账号可以在前缀管理部分填写相应的组织名称、前缀、密钥和服务地址从而对不同机构进行前缀分配, 并支持预览以及密钥自动生成, 在审批通过后用户可以用此账号登录标识系统. 管理员也可以在此模块看到不同用户前缀之间的上下级关系树状图. 此外, 统计分析部分是将标识系统中存在的标识、前缀、用户 (机构) 等数据条目进行汇总统计并进行实时形式化展示, 方便管理员了解目前系统的运行状态.

5 系统测试与应用验证

本节主要对数联网数字对象标识系统的核心功能接口进行测试, 并对整个系统在实际应用场景中进行有效性验证.

系统测试包括两部分内容, 首先是将本文所实现的数联网数字对象标识原型系统与 Handle、国家工业互联网标识解析系统在相同环境下进行对比测试. 由于不具备其他标识解析系统域节点权限, 本文测试的均是不同标识解析系统客户端 (SDK) 与端节点间的调用性能. 其中 Handle 和数联网数字对象标识原型系统所选取的硬件配置环境均是阿里云部署的服务器单节点 (配置见表 3) 无负载的情况, 采用多个虚拟用户通过 CNRI hdllib 及本文所实现的 SDK 分别调用对应系统的服务接口进行负载测试, 国家工业互联网标识解析系统选取的是其标识解析二级节点配置选型参考规范中给出的推荐数据. 其次, 本文还测试了数联网数字对象标识原型系统对 Handle 的兼容性. 除此之外, 由于域节点 IRS 在启动时默认会向根节点 GRS 发送认证请求, 如果认证失败则无法处理标识服务请求, 标识服务器节点正常工作的前提是节点间认证流程已经完成, 因此上述系统测试已涵盖此部分的功能测试.

表 3 系统测试软硬件环境配置情况

名称	硬件配置	软件配置
服务器 (对比实验)	4核 8 GB内存 2 GHz CPU, 100 GB存储	CentOS 7.5
服务器 (低配置)	8核 16 GB内存 2 GHz CPU, 100 GB存储	CentOS 7.5
服务器 (高配置)	64核 256 GB内存 2.3 GHz CPU, 480 GB存储	CentOS 7.6
客户端	4核 8 GB内存 2 GHz CPU, 500 GB存储	CentOS 7.5
测试工具	—	JMeter 5.4.1

从测试结果 (表 4) 可以看出, 本文所实现的数联网数字对象标识原型系统不论在响应时间以及吞吐率上都要优于 Handle 以及国家工业互联网标识解析系统的参考性能指标. CNRI hdllib 由于其内部实现机制在连接 Handle 端节点时会先调用递归节点, 而递归节点部署在国外, 所以响应时间相对来说较长, 而基于本文所实现的数联网标识解析 SDK 兼容性测试过程中在直连 Handle 端节点时会直接跳过递归节点, 所以响应时间会大幅下降. 此外, Handle 系统在标识注册时需要身份认证, 因此只测试了数联网标识解析协议 SDK 对于 Handle 系统标识解析的兼容性.

表 4 不同标识解析系统端节点性能测试结果

标识解析系统名称	标识注册		标识解析	
	平均响应时间 (ms)	吞吐率 (TPS)	平均响应时间 (ms)	吞吐率 (TPS)
CNRI hdllib (Handle)	6 560	169	5 839	201
国家工业互联网标识解析系统	—	50–100	—	50–100
数联网数字对象标识系统	147	379	50	410
数联网SDK与Handle服务节点兼容性	—	—	105	190

原型系统自身性能测试分为两种情况, 第 1 种是标识解析服务器低配置无压力情况下的基础性能测试, 第 2 种是标识解析服务器高配置但系统中有存量数据 (100 万+) 的高性能测试, 这种情况也更接近系统的实际部署运行环境. 两项测试均在客户端使用 JMeter 虚拟多个用户通过标识解析协议 SDK 调用标识注册及解析接口进行, 以逐渐加压的方式稳定运行一段时间, 记录不同并发数下的服务接口响应时间、TPS、成功率等指标.

服务节点低配无压力的测试结果如表 5 和表 6 所示, 其中压力测试分别进行了 3 min 和 10 min, 客户端并发数从 20 逐渐增加到 5 000.

表 5 标识系统注册接口测试结果

并发数	3 min				10 min			
	平均响应时间 (ms)	95%请求响应时间 (ms)	请求失败率 (%)	吞吐率 (TPS)	平均响应时间 (ms)	95%请求响应时间 (ms)	请求失败率 (%)	吞吐率 (TPS)
20	16	23	0	388.1	16	23	0	391.6
50	26	39	0	449.4	26	39	0	453.2
100	36	55	0	453.6	36	55	0	441.8
200	47	71	0	363.3	47	72	0	405.6
500	77	118	0	370.0	78	125	0	364.1
1 000	153	228	0	318.7	146	249	0	339.5
2 000	349	459	0	325.7	282	395	0	327.6
3 000	595	5 918	0.94	319.2	291	418	0	318.8
4 000	661	6 166	0.08	313.5	212	495	0.06	312.8
5 000	1 282	10 801	0.03	280.2	962	6 317	0.30	266.6

表 6 标识系统解析接口测试结果

并发数	3 min				10 min			
	平均响应时间 (ms)	95%请求响应时间 (ms)	请求失败率 (%)	吞吐率 (TPS)	平均响应时间 (ms)	95%请求响应时间 (ms)	请求失败率 (%)	吞吐率 (TPS)
20	13	20	0	885.7	13	20	0	897.7
50	24	40	0	1 000.3	24	40	0	995.7
100	41	76	0	1 035.8	41	77	0	1 037.7
200	76	150	0	1 083.8	76	152	0	1 073.0
500	179	368	0	1 104.3	181	373	0	1 100.8
1 000	379	776	0	1 052.7	358	739	0	1 103.9
2 000	740	1 521	0	1 072.9	708	1 468	0	1 107.7
3 000	1 117	2 116	0	1 056.9	1 071	2 209	0	1 099.0
4 000	1 507	3 021	0.04	1 018.0	1 435	2 920	0	1 037.4
5 000	2 001	3 726	0.13	901.1	1 841	3 664	0.09	1 065.0

从测试结果可以看出, 原型系统当前在进行标识注册时, 当请求并发数达到 3 000 注册服务接口会出现响应错误, 响应时间明显增加, 系统最大吞吐量约为 450 TPS, 且负载测试时间长短对系统各项性能指标的影响不大. 标识注册服务接口单个服务节点当前配置下基本支持最大 2 000 并发数, 平均响应时间约为 400 ms, 吞吐量约为 330 TPS. 在进行标识解析时, 当请求并发数达到 2 000 时响应时间明显增加, 请求并发数达到 4 000 时解析服务接口会出现响应错误, 系统最大吞吐量约为 1 100 TPS, 且负载测试时间长短对系统各项性能指标的影响不大. 标识解析服务接口单个服务节点当前配置下支持最大 4 000 并发数, 平均响应时间约为 1 000 ms, 吞吐量约为 1 000 TPS.

服务节点高配、系统存有 100 万+真实数据的测试结果如表 7 和表 8 所示, 在前述基础性能测试过程中, 系统在并发数过高时偶尔会出现因链接释放不及时而导致请求失败, 为了解决实际应用中可能存在的高并发问题, 通过引入线程池来对系统进行改进优化, 并进行了测试. 其中压力测试进行了 10 min, 客户端并发数从 50 逐渐增加到 1 000.

表 7 未使用线程池解析注册接口测试结果

系统有无存量数据	并发数	标识注册		标识解析	
		平均响应时间 (ms)	吞吐率 (TPS)	平均响应时间 (ms)	吞吐率 (TPS)
无存量数据	50	2	17 794	2	18 461
	100	5	17 850	5	18 378
	200	11	17 629	10	18 474
	500	29	16 923	27	17 844
	1 000	60	16 518	57	17 207
100万+真实数据	50	28	1 775	3	13 343
	100	69	1 428	7	13 451
	200	98	2 032	15	12 806
	500	241	2 066	39	12 487
	1 000	492	2 026	80	12 452

表 8 使用线程池后解析注册接口测试结果

系统有无存量数据	并发数	标识注册		标识解析	
		平均响应时间 (ms)	吞吐率 (TPS)	平均响应时间 (ms)	吞吐率 (TPS)
无存量数据	50	2	20 073	1	46 914
	100	5	19 426	2	37 675
	200	10	18 833	6	31 787
	500	22	22 240	11	42 089
	1 000	46	21 670	25	38 478
100万+真实数据	50	27	1 810	2	22 364
	100	47	2 109	4	20 681
	200	94	2 106	9	20 336
	500	237	2 102	24	20 242
	1 000	476	2 097	49	20 107

从测试结果可以看出, 标识系统服务器所在节点的硬件配置高低以及系统中已有数据量的大小对整个系统的性能影响较大, 当系统中已有数据量过大时, 查找和操作数据库所用时间都会增加, 而这也会直接反映到整个系统的平均响应时间以及吞吐率上. 同时可以看到, 线程池的使用可以有效提升系统的整体性能, 降低系统出现链接资源耗尽的问题. 综合上述分析, 在使用线程池并且系统具有 100 万+真实存量数据的测试环境下, 系统的测试结果更能反映在实际场景部署运行过程中可能的性能状况.

此外, 本文还对数联网数字对象标识系统使用的分布式账本相关接口进行了测试和验证, 测试选取了阿里云、华为云、腾讯云在北京、成都、深圳、呼和浩特、贵阳等 10 个不同数据中心的 500 个部署节点 (8 核 16 GB), 进行了转账查询、区块查验、账户查验等分布式账本关键指标负载测试, 结果如表 9 所示. 在此基础上, 将上述分布式账本的存证接口与原型系统进行了对接, 基于数联网数字对象标识解析协议中预留的配置位以及处理逻辑, 使数联网数字对象标识系统具备内生的信任机制.

表 9 分布式账本核心功能测试结果

测试项	平均响应时间 (ms)	吞吐率 (TPS)
转账查验	26.09	1 679 343
区块查验	2800	34 030
账户查验	33.70	2 127 320

最后, 本文还将原型系统在山西省广域政务数联网应用示范、某情报数据管理平台、某数据治理信息管理平台等实际人机物环境中进行了部署与数据对接, 对研究的有效性进行了试验验证. 在山西省政务数联网应用场景中, 门户网站、协同办公、公文交换、电子签章、文件管理等不同信息系统隶属于省-市-区 (县) 多级部门和主体, 部署在信创云中不同架构 (如 x86、国产 PK 体系等) 的服务器上, 且具有不同的系统结构组成和运行模式 (如集中式或分布式), 针对上述应用场景特点, 本文基于信创云部署数联网数字对象标识系统各个组件, 通过建立省域、市域、区县端三级联动的政务数据数字对象标识注册管理和动态解析, 支持跨地域、跨部门、跨业务的政务数据可信协同与融合应用.

6 总 结

在大数据时代, 人机物融合的泛在环境中数据面临着开放互联和可信管控之间的矛盾, 基于 DOA 的数联网技术体系有效解决了数字对象多主体性与开放协同之间的信任难题. 本文结合标识解析、区块链、人工智能等新一代信息技术, 对面向人机物融合的数联网数字对象标识解析技术进行了研究, 提出了数联网标识解析技术规范, 设计实现了支持不同数据标识编码规范的数联网数字对象标识解析协议以及数联网数字对象标识系统, 以此达到

数字对象在数联网数字空间内的唯一寻址以及各类异构、异主、异地特征的关键基础数据资源之间的互联互通. 受限于时间和精力, 本文设计实现的数联网数字对象标识系统还处于原型系统阶段, 距离人机物融合环境中大规模应用及部署仍具有一定差距, 还需对系统不断进行功能改进和性能提升, 使其能够支撑超大规模的数联网基础设施建设和运营.

References:

[1] United Nations Conference on Trade and Development. Digital economy report 2021. 2021. https://unctad.org/webflyer/digital-economy-report-2021

[2] Mei H, Cao DG, Xie T. Ubiquitous operating system: Toward the blue ocean of human-cyber-physical ternary ubiquitous computing. Bulletin of Chinese Academy of Sciences, 2022, 37(1): 30–37 (in Chinese with English abstract). [doi: 10.16418/j.issn.1000-3045.20211117009]

[3] Mei H, Huang G, Cao DG. Understanding "software definition" from the perspective of software researchers. Communications of the CCF, 2015, 11(1): 68–72 (in Chinese with English abstract).

[4] Huang G. Internet of Data: An infrastructure of digital space. Communications of the CCF, 2021, 17(12): 60–62 (in Chinese with English abstract).

[5] Kahn R, Wilensky R. A framework for distributed digital object services. Int'l Journal on Digital Libraries, 2006, 6(2): 115–123. [doi: 10.1007/s00799-005-0128-x]

[6] DONA. Digital object interface protocol specification version 2.0. 2018. https://www.dona.net/specsandsoftware

[7] CNRI. Handle system overview. 2003. https://www.ietf.org/rfc/rfc3650.txt

[8] X.1255: Framework for discovery of identity management information. 2013. https://www.itu.int/rec/T-REC-X.1255

[9] Y.4459: Digital entity architecture framework for Internet of things interoperability. 2020. https://www.itu.int/rec/T-REC-Y.4459

[10] Paskin N. Digital object identifier (DOI®) system. In: Bates MJ, Maack MN, eds. Encyclopedia of Library and Information Sciences. 3rd ed., Boca Raton: CRC Press, 2009. 1586–1592.

[11] US Department of Defense. DOD digital modernization strategy. 2019. https://media.defense.gov/2019/Jul/12/2002156622/-1/-1/1/DOD-DIGITAL-MODERNIZATION-STRATEGY-2019.PDF

[12] Foster N, McKeown N, Rexford J, Parulkar G, Peterson L, Sunay O. Using deep programmability to put network owners in control. ACM SIGCOMM Computer Communication Review, 2020, 50(4): 82–88. [doi: 10.1145/3431832.3431842]

[13] Stoica I, Shenker S. From cloud computing to sky computing. In: Proc. of the 2021 Workshop on Hot Topics in Operating Systems. Ann Arbor: ACM, 2021. 26–32. [doi: 10.1145/3458336.3465301]

[14] ATSD. Digital object interface protocol specification version 2.1. 2021. http://doa-atsd.org/?page_id=599

[15] Open source software stack of DOIP protocol. 2021. https://gitee.com/bdoa/bdoa-core

[16] Coalition for Handle Services—China. 2019. https://www.dona.net/mpas

[17] Alliance of Industrial Internet. Identifier resolution of industrial Internet standardization white paper. 2021 (in Chinese). http://www.aii-alliance.org/upload/202102/0222_094332_553.pdf

[18] Sun S, Reilly S, Lannom L. Handle system namespace and service definition. 2003. https://www.rfc-editor.org/info/rfc3651

[19] Chandrakar R. Digital object identifier system: An overview. The Electronic Library, 2006, 24(4): 445–452. [doi: 10.1108/02640470610689151]

[20] Chen Y, Zheng ZB, Luo S, Li SQ. Research on application of handle system in the industrial Internet. In: Proc. of the 4th Int'l Conf. on Intelligent Robotics and Control Engineering (IRCE). Lanzhou: IEEE, 2021. 124–128. [doi: 10.1109/IRCE53649.2021.9570920]

[21] Guo XF, Li Y, Sun SX. Federated content rights management for research and academic publications using the handle system. D-Lib Magazine, 2010, 16(11–12). [doi: 10.1045/november2010-guo]

[22] Roussos G, Chartier P. Scalable ID/locator resolution for the IoT. In: Proc. of the 2011 Int'l Conf. on Internet of Things and the 4th Int'l Conf. on Cyber, Physical and Social Computing. Dalian: IEEE, 2011. 58–66. [doi: 10.1109/iThings/CPSCom.2011.66]

[23] Ren YZ, Xie RC, Zeng SQ, Zhao HR, Yu JY, Huo R, Huang T, Liu YJ. Survey of identity resolution system in industrial Internet of Things. Journal on Communications, 2019, 40(11): 138–155 (in Chinese with English abstract). [doi: 10.11959/j.issn.1000-436x.2019238]

[24] General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. GB/T 31866-2015, Identification system for Internet of Things—Entity code. Beijing: Standards Press of China, 2016 (in Chinese).

[25] Roussos G. Supply chain management standards in ubiquitous commerce. In: Roussos G, ed. Ubiquitous and Pervasive Commerce. London: Springer, 2006. 15–31. [doi: 10.1007/1-84628-321-3_2]

[26] Lee M, Kim Y. An extended identification system of cosmetics based on GS1-128. In: Proc. of the 2017 IEEE Int'l Conf. on Consumer Electronics (ICCE). Las Vegas: IEEE, 2017. 331–333. [doi: 10.1109/ICCE.2017.7889342]

[27] Thiesse F, Michahelles F. An overview of EPC technology. Sensor Review, 2006, 26(2): 101–105. [doi: 10.1108/02602280610652677]

[28] Mockapetris P, Dunlap KJ. Development of the domain name system. In: Proc. of the 1988 Symp. on Communications Architectures and Protocols. Stanford: ACM, 1988. 123–133. [doi: 10.1145/52324.52338]

附中文参考文献:

[2] 梅宏, 曹东刚, 谢涛. 泛在操作系统: 面向人机物融合泛在计算的新蓝海. 中国科学院院刊, 2022, 37(1): 30–37. [doi: 10.16418/j.issn.1000-3045.20211117009]

[3] 梅宏, 黄罡, 曹东刚. 从软件研究者的视角认识"软件定义". 中国计算机学会通讯, 2015, 11(1): 68–72.

[4] 黄罡. 数联网: 数字空间基础设施. 中国计算机学会通讯, 2021, 17(12): 60–62.

[17] 工业互联网产业联盟. 工业互联网标识解析标准化白皮书. 2021. http://www.aii-alliance.org/upload/202102/0222_094332_553.pdf

[23] 任语铮, 谢人超, 曾诗钦, 赵浩然, 喻嘉义, 霍如, 黄韬, 刘韵洁. 工业互联网标识解析体系综述. 通信学报, 2019, 40(11): 138–155. [doi: 10.11959/j.issn.1000-436x.2019238]

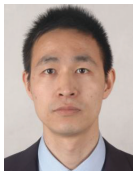
[24] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. GB/T 31866-2015 物联网标识体系 物品编码Ecode. 北京: 中国标准出版社, 2016.

张宁(1989－), 男, 博士, 讲师, 主要研究领域为系统软件, 工业互联网, 大数据互操作, 数联网.

王璐(1991－), 女, 博士, 副教授, CCF 高级会员, 主要研究领域为软件演化与自适应, 智能化软件运维.

柳熠(1993－), 男, 博士, 助理研究员, CCF 专业会员, 主要研究领域为服务计算, 移动计算, 数联网.

景翔(1979－), 男, 博士, 副教授, CCF 专业会员, 主要研究领域为操作系统, 分布式计算.



马新建(1987－), 男, 博士, 助理研究员, CCF 专业会员, 主要研究领域为信息安全, 云计算, 数联网.

黄罡(1975－), 男, 博士, 教授, 博士生导师, CCF 杰出会员, 主要研究领域为系统软件, 软件自适应, 数联网.

姜海鸥(1987－), 女, 博士, 助理研究员, CCF 专业会员, 主要研究领域为服务计算, 云计算, 数字对象架构, 数联网.