

Handle系统的发展及应用*

□ 郭晓峰 / 中国科学技术信息研究所 北京 100038

孙洵 / Corporation for National Research Initiatives, Reston, Virginia, USA, 20191

摘要: Handle系统作为一种通用的名称服务系统,能够为网络中的数字对象提供的永久标识、动态链接和安全管理等基础服务。文章重点介绍Handle系统的发展及应用情况,从数字图书馆、数字内容管理、数字出版、数字博物馆、远程教育到科学数据管理与网格计算、数字权益管理、信息安全管理与隐私保护等领域。近年来随着互联网的发展,以及物联网等新技术的兴起,Handle系统获得了更加广阔的应用前景。文章最后介绍了Handle系统作为联合国ITU推荐的互联网下一代链接技术,在技术上及管理方式上的最新发展情况,并对其应用前景进行展望。

关键词: 数字对象标识服务, Handle系统, 数字对象架构, DONA

DOI: 10.3772/j.issn.1673—2286.2013.08.004

前言

数字对象标识是数字对象架构(DOA, Digital Object Architecture)的关键组成部分,能够为数字对象提供命名与安全访问机制。数字对象标识的分配、管理与利用需要相应的系统提供服务。Handle系统^[1]于1995年由TCP/IP的联合发明人、互联网之父Robert Kahn博士领导的美国国家创新研究所(CNRI, Corporation for National Research Initiatives)开发,以Handle作为数字对象的唯一标识符,能为数字对象提供永久标识、动态解析和安全管理等服务。

Handle系统作为一种通用的名称服务系统,已在互联网上部署了超过17年,目前注册和管理了超过10亿个唯一标识符。该系统每天都在全球范围提供服务,证明了其具

有高度可靠性,是目前最先进、最成熟、应用最广泛的数字对象标识基础服务系统。Handle系统的早期应用主要集中于数字内容相关领域,利用其唯一标识的命名、注册功能,以及“简单”解析功能,为数字内容提供持久、唯一标识与动态链接、定位。诞生后10年左右,即2005年,CNRI公布了Handle系统的6.1版,并公布了公共许可证,开放了系统源代码,并取消了不允许商业机构使用的限制,只要求机构必须注册独立的前缀,同时简化了Handle前缀的申请流程,大大推动了Handle系统在更广泛的领域应用。不仅应用的范围更加广泛,而且应用的层次更加深入,逐渐利用了Handle系统对将标识符与元数据的绑定、对标识符相关信息进行安全管理等特性,在需要更复杂功能支持、安全性要求更高的网络计

算、数字内容追踪、隐私数据管理与共享等方面得以应用。本文将按照Handle系统的上述发展特征顺序介绍其应用。随着下一代互联网的研究与部署,Handle系统即将成为联合国ITU(国际电信联盟)监管下的下一代互联网链接技术设施进行全球化的管理和部署,为其应用打开更广阔的空间。本文在最后将对此进行重点介绍,并对Handle系统未来在物联网等更广泛的领域应用进行展望。

1 数字图书馆与内容仓储领域应用

诞生于第一代数字图书馆基础架构的Handle系统,先后为很多数字图书馆项目提供了支撑,最早的如美国DLI、NDLP等著名数字图书馆项目。于2000年之后先后诞

* 本文为中央级公益性科研院所基本科研业务费专项资金“基于数字对象唯一标识符的知识管理建设与研究”的成果之一,获得国家科技支撑计划“技术创新服务平台关键技术研究与应用示范”项目的资助(编号:2011BAH30B01)。

生的机构仓储系统Fedora(由美国由Virginia和Cornell大学联合开发)、DSpace(由美国麻省理工学院图书馆和美国惠普公司实验室联合开发)都内置了Handle系统,为数字内容提供永久定位功能。DSpace系统因其部署和应用的简便性得到了广泛的应用;Fedora系统则基于完整的DOA理论实现了灵活的数字对象模型,并提供了较强的扩展性,在近几年仍在数字仓储、长期保存等领域掀起了研究与应用的热潮。

2 数字出版领域应用

数字出版领域的DOI系统是Handle系统最著名的“品牌”。DOI系统基于Handle系统对数字内容的唯一标识符——DOI进行注册、管理和互联网解析,并已于2010年被批准为ISO国际标准。目前DOI的注册总量已超过8000万个,通过Handle系统从DOI解析链接到内容页面的次数每月接近1亿次。在美国、欧洲、中国、日本等都有授权的DOI注册机构(RA, Registration Agency)开展相关领域的DOI注册、解析及增值服务。目前,DOI系统服务的范围不仅从早期的期刊论文扩展到图书、会议论文、标准等各类出版物和非正式出版物,如学位论文、科技报告等,而且随着2010年为影视作品进行DOI注册的国际娱乐业RA——EIDR的成立,DOI已超越了文献领域,正式进入更广阔的内容产业领域^[2]。实际上,从2009年底开始,ISRC(国际录音制品编码)国际中心也一直在考虑使用DOI系统或者基于Handle系统建立与DOI类似的网络注册、服务体系。

3 非正式出版物管理领域应用

在非正式出版的内容领域,除使用DOI标识学位论文、科技报告外,美国国防部的DVIA(国防部虚拟信息架构)项目,基于CNRI的Handle系统和DOR(数字对象仓储)系统,建立了技术报告的注册系统,实现了技术报告的元数据及相关文件注册、存储、检索,并且结合OpenURL技术实现了场景化的链接与解析服务等(见图1)^[3]。DVIA利用DTIC(美国国防技术情报中心)报告数据库开展了示范应

使用Handle系统管理了超过6亿件内部资料,这些资料是不公开的,而且由于Handle系统分布式管理的特性,甚至连CNRI也不知道Los Alamos究竟注册与管理了多少个handle。

4 数字博物馆领域应用

在与数字图书馆相近的数字博物馆领域,也使用了Handle系统对馆藏进行标识与管理。2006年,中国教育部与惠普公司、北航等大学联合开展了中国数字博物馆建设项目,目标是创建一个大规模数字

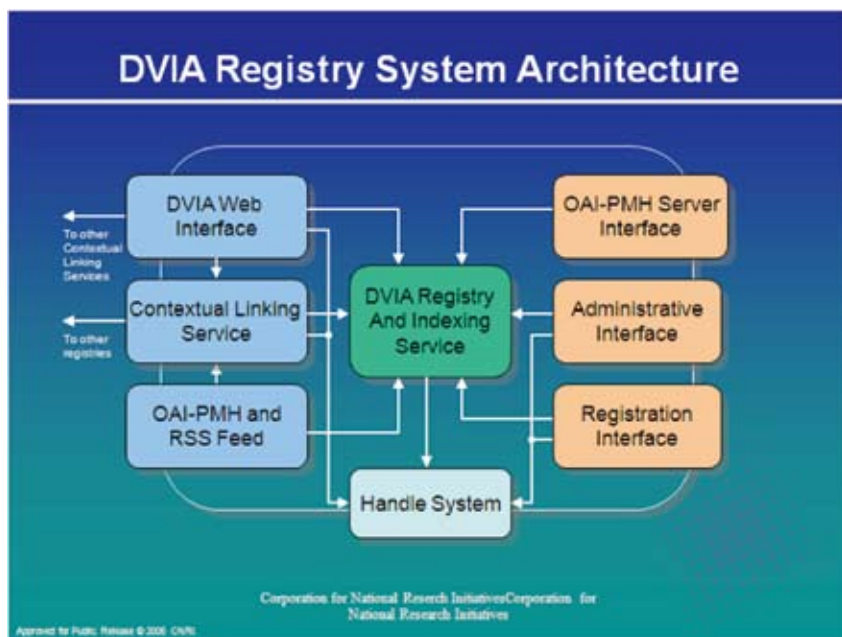


图1 DVIA注册系统架构

用,为DTIC存储的技术报告提供了唯一标识符,方法是在DTIC原有内部号码前加上Handle前缀。DTIC的实践为在某个行业或部门在现有数据库基础上应用全局化的唯一标识符提供了借鉴^[4]。截止到2007年,DVIA已注册、管理了60万件科技报告的元数据。

在档案与内部文件管理领域,美国的Los Alamos国家实验室早已

博物馆联盟,涵盖100所高校的博物馆藏品。项目解决方案使用惠普公司的DSpace系统存储各成员高校的博物馆数字化内容,并建立两个数据中心,从参加成员的博物馆中收割数字内容。该联盟系统使用Handle系统对博物馆资源(数字对象)进行唯一标识,同时用于定位可能存在于其他DSpace实例中的拷贝^[5]。

5 远程教育领域应用

1998年美国NSF正式启动了美国国家科学、数学、工程与技术教育数字图书馆计划(NSDL),发展为远程教育环境,向数百万学生和教师提供高质量的科学、数学、工程与技术教育资料。NSDL采用面向对象设计、底层由Handle系统支持的Fedora数字资源框架及仓储系统,截至2005年,NSDL的数字教育资源已经达到上百万件^[4]。

2005年,由美国国防部资助的先进分布式学习(ADL)计划,开展了内容对象仓储发现和注册/解析架构(CORDRA)的标准化工作,汇集了内容管理和网络计算方面现有的标准和规范,目标是极大地提高分布式学习内容的重用性、互操作性,以及与SCORM的兼容性。项目利用Handle系统为ADL CORDRA项目设计开发了ADL注册表(ADL-R)登记系统,并由国防技术信息中心(DTIC)进行部署,促进了学习内容在分布在美国国防部的多个仓储之间的发现和重用^[6]。

6 科研领域应用

2000年前后,网格技术在科研领域兴起。许多网格的部署都需要安全的途径,来发布和发现类似功能、政策、密钥及其他资源的元数据。这些资源本身都必须被“标识”。绑定到标识符的属性必须被安全地管理,这样,任何了解这些资源特性的人,就应该能够找到它的相关属性,并且可能需要对资源本身进行访问。2005年,通过Handle系统Globus Toolkit的集成项目,Handle系统被纳入建立网格系统和开发网

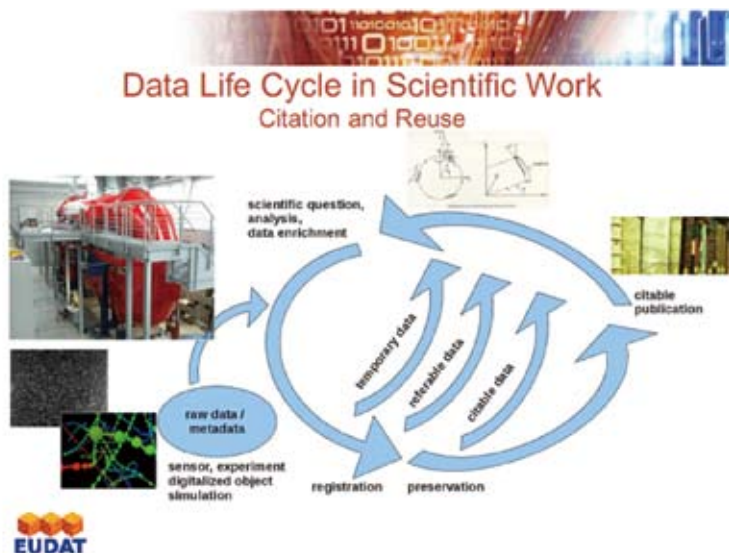


图2 科研工作中的数据生命周期——引用与重用

格软件事实的参考标准——Globus Toolkit标准工具包4.0版本中,为网格服务(Grid Service)提供命名和解析的功能。通过代理服务接口,可以实现包括SAML(安全声明标记语言)属性查询交换、WSRF(服务资源框架)服务组属性、XKMS(XML密钥管理规范),或可再生的端点引用等^[7]。

在科学数据管理方面,继DataCite(前身是德国科技图书馆TIB)于2005年开始使用DOI对科学数据进行标识、解析和引用之后,由EUDAT(欧洲数据基础设施)项目成立的EPIC(欧洲永久标识联盟)也即将基于Handle系统为欧洲科学研究社区提供科学数据的永久标识(PID)服务^[8]。EUDAT是欧盟第七框架计划(FP7/2007-2013)资金支持项目,目标是支持一个协作式的数据基础设施,能够使研究人员在社区内和社区之间共享数据,以开展有效的研究。EUDAT希望提供多种技术、工具和服务,支持数据的互操作性和信任,实现数据的广泛访问,并

以利用和重用数据为目的实现数据的长期保存。图2描述了在科研工作的数据生命周期中对于数据的引用和重用过程,通过对数据密集型科学研究和分析中产生的各类原始数据和元数据的注册、长期保存、出版,将以前的临时性数据变为可以引用和重用的数据。EUDAT认为在此过程中,关键的问题是对数据集进行永久性的标识与定位。因此于2009年成立了EPIC,并经过对多种标识、定位技术的对比之后,选择了Handle系统,完成关键性的解析功能和重要的对数据对象特征的描述与揭示功能^[9](图3)。EPIC的成员包括DKRZ(德国气候和地球系统研究中心)、STFC(英国科学与技术设施理事会)、INGV(意大利国家地理学暨火山学研究所)等多个欧洲著名研究机构,并与DataCite进行联合,已于2012年底发布了2个版本的API,用于规定PID的结构(主要是handle的后缀结构),PID的解析、创建和修改的方法,以及Web-service系统的使用方法等。

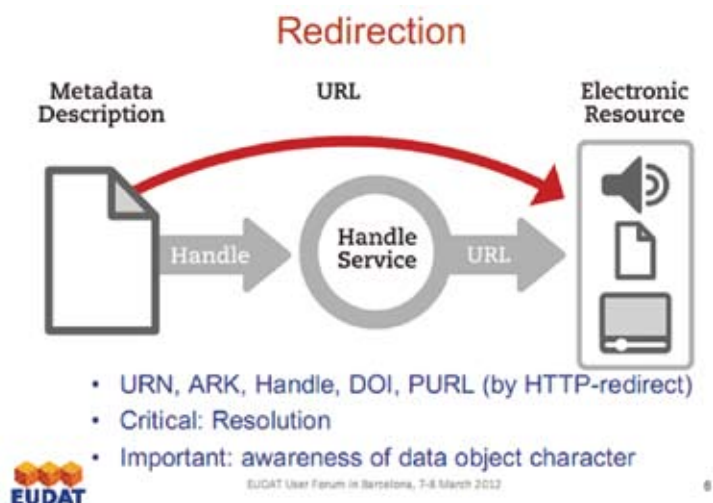


图3 EPIC中的Handle服务

随着以网格计算为基础的云计算在商业领域大行其道,以及数据资源突破科学研究领域的限制,发展为似乎将与社会生活息息相关的“大数据”,相信Handle系统也将在更广泛的数据资源、计算资源管理与服务领域大展身手。

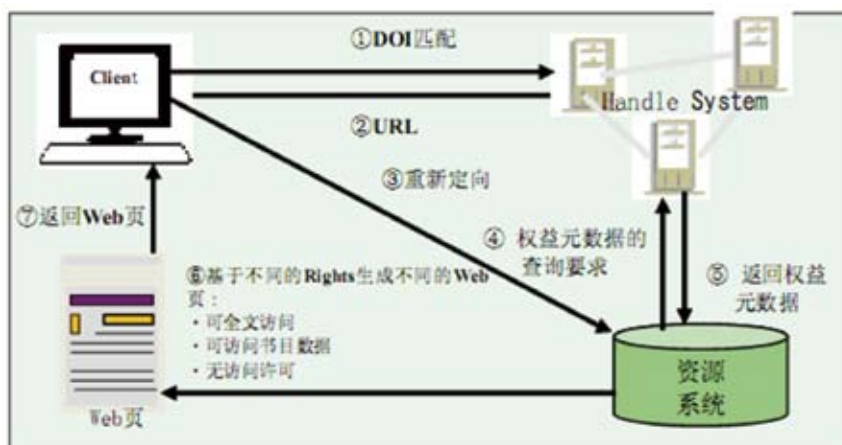
7 数字权益管理领域应用

无论是数字内容的版权保护还是更广义的数字权益管理,都需要在数字内容的全生命周期中对其进行有效(持久、一致)的识别和确认,因此,使用DOI或Handles技术对数字内容进行标识、追踪是非常合适的。

日本内容标识论坛(cIDf)于2007年发布的新技术规范中,就描述了利用Handle系统对分发的内容进行全球唯一标识并对内容元数据进行永久定位的具体方式。可以说cIDf只利用Handle系统最基本的标识与定位能力,并未对其强大的解析和安全管理能力进行充分的利用。而且cIDf在2007年之后就纳入日本宽带协会的管理之下,利用

Handle系统发行内容ID,即cID,促进内容的流通^[10]。

2007年至2010年,科技部资助的国际科技合作项目“建立中国数字对象唯一标识符体系的研究与应用”与CNRI合作开展了数字权益管理方面的研究与实验,建立了基于DOI/Handle的数字权益管理框架原型,主要思想是利用关键的Handle系统技术的安全性和分布式功能,以及标准的Web服务接口和权利元数据定义,支持内容权利的注册和发现^[11]。图4描述了实验的应用场景。

图4 基于DOI/Handle的数字权益管理框架应用场景^[12]

在数字内容追踪方面实际应用Handle系统的产品,当属由前微软员工创立的美国Infoflows公司于2006年发布的FedmarkSM平台,该平台为内容所有者提供了一个创新的方法来跟踪、监测、通过法律控制并恢复网络资产^[13]。Fedmark平台的核心技术主要是其特有的图像指纹识别技术以及CNRI专利的Handle系统,加上网络抓取技术、动态的视觉搜索能力等作为支持,力图成为涵盖所有形式的数字对象的知识产权保护标准——无论是图像、视频剪辑、音频文件,或软件应用程序。Fedmark平台利用Handle系统在互联网范围内对数字对象进行唯一和持久的标识、管理、追踪,并实现主动识别技术:通过在数字内容中注入很小的ID(Handle),可以解析到与数字内容相关的关键元数据,但不会给视觉效果带来任何影响。目前,Fedmark分别针对图像和视频内容提供两套解决方案。Fedmark既保护内容所有者,同时也帮助任何想使用数字内容的人能够找到它的主人。2008年初,Infoflows公司在美国版权局举办的“技术和孤儿作品”艺术活动中,重点向负责监

管版权法变化的国会工作人员演示了Fedmark技术如何帮助数字内容的用户努力找到孤儿作品的版权持有人,通常这些作品很难或者说不可能识别和定位版权持有人。在版权局的版权登记人员看来,Infoflows公司的演示反映了技术可以为版权法的相关变更提供支持。另外,Infoflows公司也由于在2010年和2012年两次与比尔·盖茨拥有的数字图像公司CORBIS的知识产权官司中胜诉,先后赢得2000万美元和1200万美元,而名噪一时。

8 信息安全管理与隐私保护领域应用

与公开出版物需要广泛传播利用相比,对于隐私数据的共享与利用同样存在着大量需求,甚至更为广泛。要满足这些需求,不仅会涉及数据访问权限的定义,还必然涉及用户的身份识别、角色定义、权利定义等一系列复杂问题,同时这些定义还必须被安全地存储和访问,避免被篡改、截获等。

2008年,CNRI与南加州大学合作,开展了医疗档案信息共享方面的研究项目并在加州的一所医院进行了实验。由于会诊、医疗研究等方面的需求,不同医疗机构之间需要共享病例信息。但病例的隐私性要求很高,只允许相关的人看到其需要的信息。此项目中利用Handle系统对病例数据以及需要访问数据的用户(机构、个人、用户组等)进行了唯一标识,并将数据的访问权限、用户的权限等作为handle的属性进行定义和安全管理,基于Handle系统内置的安全协议完成用户认证、授权等过程。

Handle系统对于标识及属性、值的安全管理与访问特性,使其能够在各种网络应用环境下对于各种对象的信息提供安全的管理、隐私的保护,实现对象信息可控制的共享和交换,也为其在基于网络的联合身份认证服务、数字权益管理、远程监控等对安全性要求较高的领域进行应用提供了很大潜力。从更广泛的意义上来说,为实现信息系统之间安全的信息传递和互操作、实现可信网络提供了基础。

9 最新发展——作为下一代网络基础设施的DONA (Digital Object Numbering Authority Foundation)

早在2005年,Handle系统就加入了下一代网络的研究。在GENI(全球网络创新环境)项目中,Handle系统成为数字对象注册表项目的一个组成部分^[6]。近年来联合国所属的国际电信标准制定机构——ITU正研究将Handle系统作为互联网的下一代基础设施进行统一管理,因此Handle系统在技术上和管理方式上都在进行一些重要的发展和变革。

Handle系统自创建以来,一直由CNRI运行和管理。Handle系统是开源软件,但要使用Handle系统首先要申请前缀(类似于域名),然后才能在属于自己的前缀下面注册完整的标识符。在目前的管理政策下,Handle系统的前缀都要向CNRI申请,CNRI为维持Handle系统的运转对前缀收取一定的费用。随着其在世界各国的应用不断扩展,CNRI希望将其作为服务于下一代互联网的公益性基础设施由全世界各个国家共同管理,并在世界多

个国家部署主根服务器,这样各个国家就拥有了对自己国家的Handle系统运营和服务的自治权。各国的相关管理机构可以为本国制定与执行相关的管理政策、开展相关服务、发放和管理前缀等。因此CNRI近年来不仅对Handle系统进行了升级,同时积极与ITU的相关机构磋商,获得了ITU的认可,从2011年初开始筹备DONA的组建工作,预计2013年可以正式成立。

另一方面,ITU也正在着手建立下一代互联网方案。目前互联网的名称服务系统——DNS(域名系统)在发展初期一直由美国公司NSI依据与美国政府的合同进行管理和运行,直至1998年成立非盈利组织ICANN,2009年10月才获准独立于美国政府之外,取得独立地位。因此,DNS无论在技术上还是在管理上都面临很大的危机。ITU接受了DNS的教训,在技术上选择更具优势的Handle系统作为下一代互联网基础设施的同时,在管理上也将其运行机构DONA置于ITU的监管之下,以避免将下一代互联网的基础设施置于某个国家或某个公司的控制之下。

DONA将管理Handles GHR(Global Handle Registry[®])的运行,最终目标是发展一组MPA(Multi-Primary Administrator,多主根服务器管理机构),这些MPA部分是政府机构,另一部分是非政府团体,并到达相对的平衡。DONA对各MPA提供技术需求和技术政策指导。ITU将对DONA进行公共政策框架方面的监管,以及赞助和保护,但DONA是独立的权力机构。DONA目前规划的组织结构见图5。

在技术方面,Handle系统于

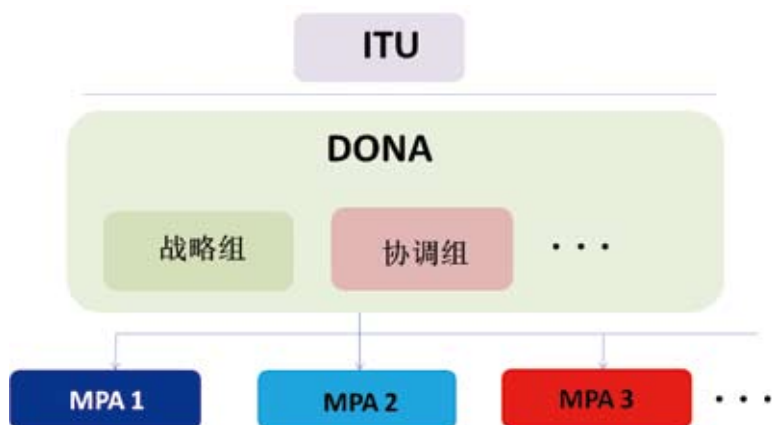


图5 DONA组织结构

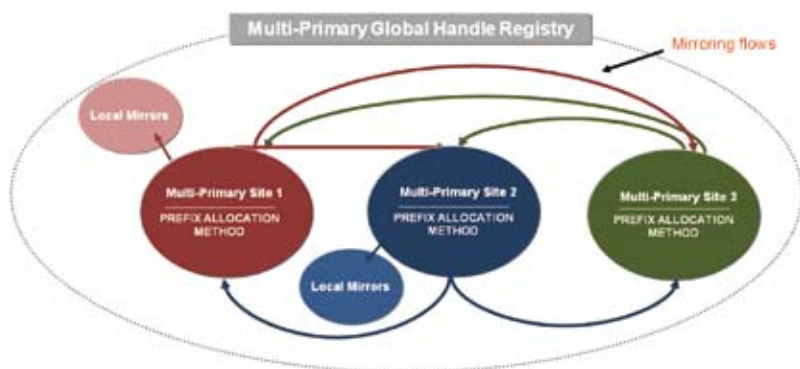


图6 多主根全球Handle注册系统结构

2010年12月发布了7.0版。除了增加了某些重要特性,包括基于模板一次性大批量创建handle、离线数字签名、支持按照国家、语言等自动选择多重解析定位、支持将DNS域名自动转换为handle、支持Python脚本与handle服务器的交互等外,最重要的是体系结构由单一的主根服务器发展为多主根服务器(图6),为下一步新的管理模式的发展提供了技术上的全面支持。

DONA已开始进行运营方面的准备,预计2013年上半年在瑞士联合成立非盈利组织,并组建理事会,继而与各MPA签署协议。目前在世界范围内已有多家机构确定作为MPA的候选机构,涉及美国、欧洲、中国等国家和地区。将来每

个MPA将分别运行一个全球的主根服务,可以创建Handle系统前缀,并有权对其进行管理。所有主根服务都将相互镜像。

10 未来应用展望

DOI并未满足于在内容领域发展,也在不断探索在其他行业应用的可能性,如作为金融服务行业的合法实体标识符,或者作为产品的唯一标识符等。但DOI在其他领域的发展也会带来一些困惑,比如可能会引起品牌定位不清晰、多个行业应用的社区基础不兼容等问题。

物联网的发展将互联网信息的范围大大扩展,Handle系统的应用范围自然也将随之扩展。标识技

术一直是物联网的重点研究领域,2009年发布的欧盟物联网研究战略路线图^[14]甚至将其作为最优先的研究重点,并将研究建立全球/全局统一的标识与解析体系和平台、研究标识的安全与保密技术、研究基于各种唯一标识体系建立全球/全局目录搜索与发现服务等问题,以及标识体系的互操作性问题等列为重点课题。对比Handle系统的特性以及已有的研究和应用可以看出,Handle系统完全可以满足以上要求。DOA理论的创立者、CNRI总裁Robert Kahn博士也认为:物联网实际上就是互联网,DOA提供的先进的基础设施对于管理“物”的信息仍然是非常理想和适合的^[15]。事实上,在物联网架构的研究中,越来越多的专家意识到IPV6无法解决物联网的统一标识问题,因为IPV6只是地址而非名称。因此,欧盟已开始与CNRI开展合作,将Handle系统技术与IPV6相结合,开展搭建物联网应用基础平台的研究与实验项目。

另一方面,DONA的正式成立将会进一步促进Handle系统在全球的大规模应用。以我国为例,早在几年前新闻出版、版权管理等行业就已对Handle技术青睐有加,但“Handle系统是美国技术”、“Handle的注册和解析必须依赖美国,我国没有自主管理权”这样的担心使可能的应用限于停顿。DONA正式成立后,将有中国的MPA在我国管理和运行Global Handle Service,拥有完全的管理权和完整的技术设施,上述对信息安全的担心就完全没有必要了。可以预见,不仅是新闻出版、科技信息管理等内容领域,包括物联网等更广阔的应用领域都可以开始放心

地利用Handle系统。目前在食品安全溯源、机构代码标识等方面已有利用Handle系统开展项目的计划。因此,可以说Handle系统即将打开

无限的应用空间,下一步的应用重点除了覆盖种类更加丰富的内容资源领域之外,在信息安全和隐私保护、数字权益管理、身份标识与识

别、物联网、大规模网络计算、大数据、信息系统集成和互操作等领域的深层次应用都将异彩纷呈,让我们拭目以待。

参考文献

- [1] CNRI. Handle System [OL]. [2013-03-20]. <http://www.handle.net/>.
- [2] IDF. Digital Object Identifier System [OL]. [2013-03-20]. <http://www.doi.org/>.
- [3] BLANCHI C. DVIA Contextual Linking Architecture. DVIA_HJ.ppt.
- [4] 毛军.数字科研、远程教育等领域的唯一标识符应用现状研究[R].2006.
- [5] TANSLEY R. Building a Distributed, Standards-based Repository Federation-The China Digital Museum Project [OL]. [2013-03-20]. <http://www.dlib.org/dlib/july06/tansley/07tansley.html>.
- [6] CNRI. Current Applications of the Handle System [OL]. [2013-03-20]. <http://www.handle.net/apps.html>.
- [7] Globus. Handle System - Globus Toolkit Integration Project [OL]. [2013-03-20]. http://www-unix.globus.org/toolkit/projects/handle_system.html.
- [8] EPIC [OL]. [2013-03-20]. <http://pidconsortium.eu/index.php?page=home>.
- [9] SCHWARDMANN U. PID-Systems for Digital Objects [OL]. [2013-03-20]. http://www.eudat.eu/system/files/PID_Systems_for_Digital_Objects.pdf.
- [10] cIDF [OL]. [2013-03-20]. <http://www.npo-ba.org/>.
- [11] GUO X, LI Y, SUN S X. Federated Content Rights Management for Research and Academic Publications Using the Handle System [J]. D-LibMagazine, 2010, 16 (11/12).
- [12] 李颖,赵蕴华,郭晓峰,等.中文DOI系统的应用研究与开发:从战略规划到国内外合作[J].数字图书馆论坛,2009(8):32-38.
- [13] InfoFlows [OL]. [2013-03-20]. <http://www.fedmark.com/Home.aspx>.
- [14] CERP-IoT. 欧盟物联网研究战略路线图[OL]. [2013-03-20]. http://www.internet-of-things-research.eu/pdf/IoT_Cluster_SRA_English2Chinese_Translation.pdf.
- [15] KAHN R. The Digital Object Architecture and its application to the Internet of things. Beijing, 2010.

作者简介

郭晓峰, 中国科技信息研究所DOI注册与服务中心高级工程师。主要从事中文DOI研究、开发、运行工作, 及唯一标识符研究开发工作。E-mail: doi_om@istic.ac.cn

孙洵, CNRI (Corporation for National Research Initiatives) 高级研究员。领导了Handle系统的架构设计和协议开发, 以及有关Handle系统协议规范的IETF RFCs编制。并作为项目经理, 领导、实施了Handle系统与DNS集成项目、中美科技合作项目, 以及将Handle系统应用于网格计算、医疗信息化等方面的合作项目。

The Development and Application of Handle System

Guo Xiaofeng / Institute of Scientific & Technical Information of China, Beijing, 100038
Sam X. Sun / CNRI, Reston, 20191

Abstract: As a generic name service system, Handle System has provided global services for digital objects over Internet, including persistent identifier, dynamic linking, security management, and other basic services. This article focuses on the developments and applications of the Handle System, from the areas as Digital Library, Digital Contents Management, Digital Publishing, Digital Museum, E-learning to Science Data Management and Grid Computing, Digital Rights Management, Information Security Management and Privacy Protection, etc. Recently, with the rise of new technologies of the Internet and the Internet of Things, more broad application prospects of Handle system are emerging. At the last section, the article introduces the state of the art in the technical and management of Handle System since it has been chosen by ITU as the next generation of the Internet linking technology, and finally prospects its future applications.

Keywords: Digital object identifier service, Handle system, DOA, DONA

(收稿日期: 2013-04-04)