基于云计算的数字图书馆学术资源安全探讨*

邓胜利，刘宇
（武汉大学信息管理学院，武汉 430072）

摘要：云计算作为一种新兴的共享基础架构方法，应用范围逐渐扩大。首先概述了云计算环境下数字图书馆的概念、特点、原理；然后讨论了“云”图书馆在安全性上的利弊；最后结合云计算特点，研究了“云”数字图书馆安全架构，并制定了相应的安全策略。

关键词：云计算；数字图书馆；信息安全
中图分类号：G250.7
DOI：10.3772/j.issn.1673-2286.2015.10.002

1 引言

数字图书馆经历了“基于自动化的管理信息系统时期”、“基于局域网和城域网的电子图书馆时期”，正在走进“基于信息高速公路的数字图书馆”的“云计算”时代[1]。OCLC于2009年推出了基于WorldCat书目数据的在线图书馆服务项目，这项服务的目标是取代图书馆的集成管理服务，构建一种新型的云计算服务。这项服务被认为是图书馆云计算服务的开始。云计算在服务模式与技术上的突破，为高校图书馆数字信息资源管理与服务提供了一种全新的手段，因此，其产生之初就吸引了图情界的关注。当前，数字图书馆面临很多问题，巨量的信息资源迫切需要一种新的技术来支持它的组织、管理与传播，而云计算的出现，为图书馆的资源组织和利用模式的变革指明了方向。云计算用户最为关注的是安全问题，安全性位于所有影响因素的首位。Gartner报告显示70%以上的用户对云计算服务的安全性表示担忧[2]，如图1所示。

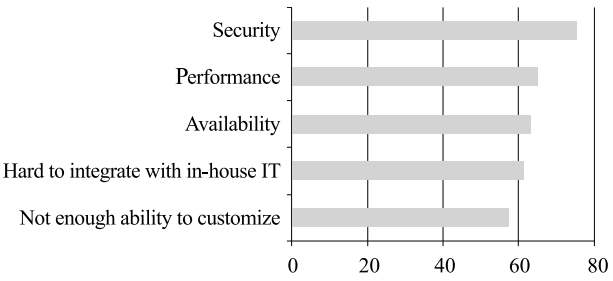

图1 云计算发展中面临的挑战

事实上，新服务模式也会给图书馆数字信息资源的安全带来一些隐患。图书馆将数据信息资源交由云服务提供商管理，窃取图书馆用户隐私、敏感数据成为不可避免的安全隐患；云服务提供商通过网络传输传递服务，图书馆数字资源同样不可避免面临着网络传输中的安全隐患；同时，云计算服务平台的安全问题也会影响图书馆数字资源的存储和安全利用。要想为图书馆提供可靠服务，云计算服务提供商必须提高基础设施的安全性。各大云服务提供商频繁爆发的数据中心物理设备安全事故，使得云服务平台物理设备和环境的安全问题成为高校图书馆云服务模式下不得不考虑的安全问题[3]。例如：2014年4月，由于内部网络云平台遭到攻击，中国快递1400万条信息泄露；2014年9月，震惊全球的iCloud云平台存储被黑客破解安全防护，导致许多好莱坞女星艳照泄露；2014年12月，12306云平台防火墙出现问题，被黑客盗取用户数据，泄露含身份证及密码信息等。鉴于此，本文从分析云环境下图书馆

* 本研究得到国家社科基金重大项目“云环境下国家数字学术资源信息安全保障体系研究”（编号：14ZDB168）资助。

学术资源面临的安全问题出发，重点研究相应的安全保障体系的构建以及实施保障，以期提高云图书馆信息服务的质量，发挥云图书馆在信息存储与传播中的作用。

2 国内外研究现状

2.1 国外研究与实践发展

数字图书馆国外研究比较早，也比较深入，西方国家在信息技术、计算机技术方面具有很大优势，他们解决了很多数字图书馆方面的基础性问题。数字图书馆作为一个新生事物，和传统图书馆有很大的不同，如访问时效问题。为了研究如何管理用户访问数字图书馆的时长，美国国会图书馆于2009年启动了一个研究项目，该项目的主要目的是研究是否能够通过云计算技术实现数字内容的永久访问[4]。数字图书馆的实现需要计算机网络技术的支持，云计算的最终目的是为用户提供数据的本地化服务，Dura Cloud项目研究了在不需要本地技术支持的情况下，用户在使用数字图书馆时的数据存储和访问问题。它对于数字图书馆的推广应用具有积极的推动作用。

2009年，云计算技术被列为图书馆界的重要发展方向，随着云计算在图书馆中应用的逐渐深入，各种基于云计算而开发的在线图书馆项目应运而生。目前世界上数字图书馆还没有一个统一的标准，基于不同规范的数字图书馆系统如雨后春笋般出现，如Koha系统、Web Feat Express跨库检索系统、SaaS系统、Google API系统等。Koha系统应用最广泛，世界上大多数数字图书馆系统都是基于Koha系统开发的。Koha系统部署在Lib Lime云计算平台上，功能非常全面，能够满足世界上大多数数字图书馆的功能要求。

2.2 国内研究与实践发展

我国图情界对图书馆与云计算的研究始于2008年，以李永先等的《云计算技术在图书馆中的应用探讨》作为开端。张晓林“数字化科技信息资源长期保存体系与政策机制”（国家科技图书文献中心项目，2004年立项）项目，构建了数字化科技信息资源长期保存的体系模型，提出了实现科技信息资源长期安全保存的政策制定建议。在云环境下的隐私保护方面，黄国彬“云计算环境下图书馆信息资源安全政策法律研究”（国家社科基金项目2011年立项）项目，认为政策法律是解决云计算环境下图书馆信息资源安全问题的有效途径，以此出发在国内相关政策法规梳理、国内外比较研究的基础上，提出了云计算环境下图书馆信息资源安全保障政策法律建设思路。

国内有些基于云计算的图书馆学术资源信息服务及安全保障的实践探索，如CALIS数字图书馆云服务平台、学科知识网络云平台、百链云图书馆等；依赖于第三方云平台进行学术信息资源的存储和服务的安全保障，如湖北省数字图书馆云平台采用了ISDM（IBM Service Delivery Manager）平台管理系统进行安全保障[5]。数字信息服务的安全保障，国内主要从图书馆服务组织角度进行研究。2010年，上海图书馆主持编写的《数字图书馆安全管理指南》获得全国数字图书馆建设与服务联席会议成员审议通过并正式发布，该指南定义了数字图书馆的概念，通过分析数字图书馆的特点，指出了数字图书馆安全管理的重点要素，安全管理指南分别从政策实施过程中的多个环节提出了数字图书馆安全管理的原则性意见。在信息系统的整个生命周期里，安全体系需要覆盖到它的各个方面，在软件的设计、开发、测试、运营以及维护阶段，都要考虑信息安全，其中包括安全需要的定义、安全机制的构建、安全技术的采用、风险的评估等等。

3 数字图书馆的云计算应用

经济社会的急速发展导致企业运营过程中产生了大量的数据，传统的数字技术难以处理如此多的数据，必须利用云计算来解决这一问题。作为一个新兴的应用，和传统图书馆相比，数字图书馆在很多地方都有自己独特的优势，然而要想吸引更多的用户，必须提高服务水平，引入云计算来改善服务质量。数字图书馆既是云计算的使用者，利用云计算商提供的服务，同时也是云计算的提供者，为用户提供数字图书、文献。对于数字图书馆，云计算主要提供信息的存储、编目和服务[6]。

3.1 信息资源的存储

信息技术极大地方便了人们的生活，数据的储存与使用变得非常简单，然而方便的同时也存在隐患，在信息资源的存储与使用过程中，云存储服务提供商扮演了

重要的角色。在数据的存储过程中，数字图书馆并不能够保护这些数据的安全，需要云计算提供商通过相关的技术手段保证数据的安全。数字图书馆的所有数据都在云计算提供商的数据库里，一旦提供商的数据存储设备出现问题或者被黑客入侵篡改数据，数字图书馆将承担极大的损失，这些损失很有可能是无法挽回的。作为公告服务的提供者，图书馆有义务保存信息资源。

数字图书馆的信息资源一般包括用户信息数据、书目信息数据、商业资源数据以及自建的特色资源等。各类数字图书馆的主要数据资源来源于商业数据库，商业数据库收录了世界上大多数图书、文献资料，并且这些数据每年都在增加，因此数字图书馆需要购买商业数据库。商业数据库的种类很多，包括文、法、理、工、农、医等各种数据库，数字图书馆通过购买不同的商业数据库组建自己的图书数据库。由于数据库的数据量非常大，并且时时刻刻都在增加，这些数据占据了很大的存储空间。数字图书馆可以将这些数据存储在云计算服务商的系统中，服务商的系统安全性较高，可以保证数据的安全性。

3.2 信息资源的编目

数据资源种类繁多，必须通过一定的标准对它们进行分类，以保证数据检索和数据管理的便捷性。数字图书数据量大，分类繁多，因此需要通过编目来简化数据的管理。编目是按照一定的规则对数据进行著录，通过组成目录使数据有序化，从而简化数字资源的检索。当前，数据呈现爆炸式增长的趋势，数据资源的增长速度已经难以通过简单的数据编目来有效管理，因此必须通过引入先进的技术来管理数据。云计算在数据处理方面具有很大的优势，可以通过其来实现一定区域内数据的编目。通过云计算技术，数字图书馆数据的编目将实现区域的联合与开放，改变过去数据编目时的封闭性，实现数据在更大范围内的传播，通过云计算技术将图书资源传递到更广阔的区域，减少资源的浪费。

美国OCLC图书馆在数字图书馆联合编目方面进行了深入的研究，其开发的图书馆联合编目系统具有云计算特征，世界范围内的图书馆都可以利用他们的系统进行图书数据的编目。通过这一系统，世界范围内的数据资源实现了扁平化，世界任何角落的人都可以通过WorldCat.org网站检索图书信息。

3.3 信息资源的云服务

数字图书馆引入云服务时，图书馆可以从各种渠道购买图书版权，如出版社或者图书、文献作者，也可以通过整理网络上的开放性资源来构建数字图书馆的云资源。作为一个专业的服务系统，数字图书馆需要通过对数据进行分类来实现数据的有序化，以提升用户的数据检索体验。通过引入云计算，数字图书馆的管理人员不必担心图书馆的数据受到外部的攻击，也不需要对数据的存储设备进行维护，工作的目标变得非常简单即为用户提供更好的服务。

数字技术的发展催生了很多世界级的信息技术公司，例如国内的阿里巴巴、腾讯、百度，国外的谷歌、亚马逊等，这些公司都拥有非常先进的云服务。2014年春运期间，阿里巴巴集团将自己的云计算平台提供给了铁道部的购票系统，极大地缓解了购票高峰时铁道部购票系统的压力。这一事例充分说明了互联网公司在云计算方面的巨大优势与能力，因此可以将数字图书馆的某些应用搭载到互联网公司的云计算平台上。通过利用互联网公司的云计算平台来实现数字资源的便捷访问与检索。

4 数字图书馆云应用中的安全问题

据IDC（Internet Data Center）的预测，2015年规模较大的云计算生态系统（公共云、私有云、授权IT和服务）开支将达到1180亿美元（2018年将接近2000亿美元），其中有700亿美元（2018年为1260亿美元）将投资到公共云上[7]。在使用云计算平台的服务时，广大用户最关注的是云服务的安全性，随着公众对个人隐私的关注程度逐渐提升，云服务的安全性成为其推广中的一个重要问题。随着数据资源在全球范围内的流动，数据资源变得越来越开放，与此同时对云计算平台的攻击可能来自于世界上的任何一个角落，因此需要云计算提供商开发更加安全可靠的服务平台来保证用户数据的安全性。当前，云计算平台还处在一个技术探索阶段，虽然市场上出现了很多云平台，但它们并没有一个统一的标准。因此，在云计算平台的发展过程中，需要构建一个统一的安全与技术规范来应对外部威胁[8]。

4.1 数据的云存储安全

数字图书馆使用云计算时会将大量的文献资源、

用户信息、OPAC系统等数据存储在云计算平台提供商的服务器中，这给数字图书馆的管理带来了极大方便。然而，当数字图书馆不再使用当前的云服务平台而将云服务平台中的数据删除时，系统中的数据可能不会被删除掉，由于这些数据具有极高的价值，当下一个用户使用时，极可能通过数据恢复功能还原数据，也极可能使用这些数据谋取非法利益，使用户遭受损失。

4.2 服务端和用户端之间的资源供求安全

在租赁用户使用数字图书馆时，其信息是可以被数字图书馆的云计算提供商所获得的，如果提供商有意窃取用户的信息，将对用户的隐私构成极大的威胁。因此，数字图书馆需要通过一定的技术措施构建一个安全系统，使用户和云计算提供商之间形成一道安全墙，保证用户的信息不会被不良服务提供商所窃取。另一方面，数字图书馆允许租赁用户上传数据到服务器端，可能为非法租赁用户通过所属虚拟机攻击服务器内其他的虚拟机。同时，租赁用户之间进行数据共享和传输过程中，可能出现非法攻击者和非法用户恶意篡改和泄露信息。

4.3 “云”平台可靠性与服务可持续性问题

数字图书馆的图书文献资源非常庞大，并且不断增加，基于“云”的数字图书馆系统是一个拓扑结构庞大的系统，其访问量很大，对数据安全性的要求也非常高，因此要求云计算提供商所提供的服务具有极高的安全性与可靠性。在云平台的运行中，软件出现故障将可能导致很大的事故，极有可能导致系统瘫痪。云计算平台是数字图书馆各项服务功能的基础，云计算平台系统软件与硬件的水平影响到数字图书馆的服务质量、安全性能。因此，需要保证云服务平台服务的连续性，通过软件与硬件构建安全防火墙，为数字图书馆提供一个安全的环境。

云计算平台中用户的数据对平台提供商是透明的，这就需要引入第三方安全审计平台来监督云服务提供商的行为，避免提供商恶意窃取用户信息，同时对提供商软硬件设备的安全性能进行审计，督促云计算提供商升级落后的设备与系统，避免遭受到黑客的攻击，保证数字图书馆数据的安全性。另一方面，由于数字图书馆的数据信息可能被存储在世界上很多的国家，而不同国家的法律法规是不同的，因此在数字图书馆的发展过程中必须考虑不同地区法律对自身业务的影响。

5 云环境下数字图书馆的安全策略

近年来发生了多起信息安全事件，导致用户信息大范围泄露。在云计算环境下，用户的信息对第三方即提供商是透明的，在用户信息的认证过程中也可能在中途遭受到攻击，数字图书馆对信息安全的要求更加高。只有认真分析“云”用户安全需求与“云”服务提供方式，密切结合“云”计算安全特点，分析“云”计算数字图书馆所面临的安全威胁，找到相应的对策，才能实现数字图书馆真正意义上的安全“云”[9]。

5.1 数字资源的安全存储与管理

数字图书馆的各种数据资源是存放在提供商服务器上的，数字图书馆构建方主要是通过数据接口为用户提供数据服务，数字图书馆不需要对数据存储单元进行过多的管理，数据的管理工作由云计算提供商进行，因此数据的安全性不高，极有可能发生数据丢失和被盗的事故，给图书馆的管理工作带来障碍，严重损害图书馆的利益。因此，数字图书馆在使用云服务商的服务时，必须考虑数据的安全性，不能仅仅听信服务商的一面之词，要详细查看第三方认证机构的评估报告以及它之前的业务记录。

数字图书馆在使用云服务商的云存储服务时，图书馆会将各种数据都存储在云端服务器，这样非常方便，用户在任何地方都可以通过终端访问云端的数据，对数据进行管理。云端的数据可以在不同终端间转移，也可以在不同云间转移。数据的转移是通过互联网进行的，在数据转移过程中，极有可能受到恶意攻击，导致数据被篡改或劫持。因此在数据的传输过程中需要对其进行加密，防止数据被窃取。数据加密方法有简单的也有复杂的，加密方法太简单的话，数据被窃取后易于破解，加密方法太复杂的话，用户下载数据后需要很长的数据解密时间，因此需要选择合适的数据加密方法。对数据进行加密时，可以针对不同密级的数据采用不同的加密方法，对于一般密级的数据可以采用简单的加密方法，对于密级高的数据可以采用复杂的方法，从而既能保证数据的安全，也能减小系统的开销。

5.2 数字图书馆终端用户机安全防范

用户在使用数字图书馆时，可以通过不同的用户终端访问系统，在登录数字图书馆认证通过后，用户可以使用数字图书馆的各种功能，如通过定制服务功能，构建独具特色的个人空间。部分用户在使用数字图书馆时，若同时访问其他网站，可能遭受到病毒的感染，因此用户可能成为攻击云服务平台的跳板。在使用云计算时，用户需要在终端上安装安全软件，及时的更新软件的病毒库，保证用户在访问数字图书馆时不会遭到病毒或者黑客的攻击。

有些用户在使用数字图书馆时会同时在虚拟机上进行工作，而虚拟机通常并没有安装补丁修复漏洞，是不安全的，在网络环境下极有可能受到攻击，从而成为黑客攻击数字图书馆的跳板。因此数字图书馆应该提出相关的规定，禁止用户在安装有虚拟机的终端上使用数字图书馆。

5.3 数字图书馆云平台服务安全

云服务需要使用不同的安全技术和防护策略来保障服务的安全。针对数字图书馆，云平台主要有三种服务模式：

软件即服务（software as a service，SaaS）模式的原理是用户通过网络访问“云”图书馆，在通过身份验证后获得软件的使用权，它是一种以云计算为基础的服务模式。“云”图书馆的管理员为用户提供用户名和密码，用户访问“云”图书馆时需要首先通过网站的权限认证。软件即服务的用户权限管理模块包括数字身份验证模块、访问控制模块，当用户通过网络访问“云”图书馆时，服务器首先验证用户的用户名和密码，通过验证后才会授予用户相应的权限。当前的网络安全问题非常严重，为了防止黑客获得网站的超级权限，图书馆的管理人员以及云计算供应商的所有者要经常的修改密码。由于供应商的客户不止“云”图书馆一个，因此在存储数据时，不同类型的客户之间可能会相互交叠的存储，在云计算提供商升级软件时，不同用户数据间的隔离可能会变得非常脆弱，甚至会发生数据混淆。因此，“云”图书馆的管理人员需要了解云计算供应商在存储数据时使用的虚拟数据存储架构以及预防机制，保证“云”图书馆的数据和其他用户数据有效隔离。

平台即服务（platform as a service, PaaS）模式的核心是为“云”图书馆提供专业的中间服务平台，它包括应用服务器和图书馆数据库。平台即服务模式能够大幅提高平台中资源的数量。平台即服务模式的运行安全性关系到“云”图书馆的正常运行，云计算供应商为用户提供安全保障。一般情况下，云系统中会安装有各种类型的插件或应用，这些应用的安全性由应用的提供商负责，但是第三方应用的漏洞对云系统的安全威胁依然很大，因此，“云”图书馆应该要求云计算供应商严格审查第三方应用的安全性，即时的更新软件，为应用的漏洞打补丁，保证用户数据的安全性。

基础设施即服务（infrastructure as a service，IaaS）模式中，云计算供应商为“云”图书馆提供服务器、存储单元以及管理工具等基础架构。基础设施即服务模式下“云”图书馆的安全性也是由云计算供应商负责。其主要负责“云”计算的网络安全、物理安全、系统安全、基础设施可靠性以及信息存储安全，为“云”图书馆提供一个安全的工作环境。“云”图书馆的数据维护以及具体的管理工作由“云”图书馆的所有者负责。

基于“云”计算的数字图书馆，是未来图书馆建设的方向与标准。相对于当前的数字技术，“云”计算具有非常明显的优势，它代表了信息技术的发展趋势。随着技术的进步，互联网应用变革的时间越来越短，技术的进步推动了网络的变革。数字图书馆的核心为数据资源，数据资源的安全性是构建“云”图书馆时首先要考虑到的。在建设“云”图书馆时必须将网络技术与网络安全有机结合在一起，为用户提供一个安全、可靠的“云”图书馆。

参考文献

[1] 吴娟.基于云服务的图书馆建设与服务策略[J].中国信息技术教育,2014(2):63.

[2] Schoder D, Fischbach K. Peer-to-peer prospects[J]. Communications of the ACM, 2003, 46(2): 27-29.

[3] Ahmad M,Abawajy J H. Digital Library Service Quality Assessment Model [J]. Procedi- Social and Behavioral Sciences, 2014, 129:571-580.

[4] 刘嗣辛,刘赛君,胡晓雯等.构建数字图书馆云服务平台的思考[J]. 图书馆工作与研究, 2014(4):30-32.

[5] 付佳,闫实.基于云计算环境下的数字图书馆云服务模式构建[J]. 图书馆学刊, 2014(11):85-87.

[6] 马晓亭,陈臣.数字图书馆云计算安全分析及管理策略研究[J].情报科学, 2011(08):1186-1191.

[7] 钛媒体.IDC预测2015十大科技趋势：中国和“亚马逊创新”将主导科技世界[EB/OL].[2015-08-10]. http://www.tmtpost.com/175001.html.

[8] 马苏华.云计算环境下数字图书馆安全管理研究[J]. 管理观察, 2014(29):37-38.

[9] 赵裕玲.基于云服务的图书馆建设与服务研究[J].科技视界, 2013(36):230,346.

作者简介

邓胜利，男，1979年生，博士，武汉大学信息管理学院教授，研究方向：数字信息资源管理与服务、网络用户信息行为，E-mail：victorydc@sina.com。

刘宇，男，1992年生，武汉大学硕士研究生，研究方向：信息资源管理与服务。

Discussion on Academic Resources Security of Digital Library Based on Cloud Computing

DENG ShengLi, LIU Yu

(School of Information Management, Wuhan University, Wuhan 430072, China)

Abstract: Cloud computing is a new method sharing infrastructure which provides the usage of service. Firstly, the concepts, characteristics and principles is described for the digital library based on cloud computing in this paper. And then, we provides a high-level discussion of the fundamental challenges and benefits of digital library under the cloud computing security. Finally, the basic concept and characteristic of cloud computing, combined with the characteristics of cloud computing, study of establish the appropriate framework for digital libraries cloud computing security, and to develop appropriate security policies.

Keywords: Cloud Computing; Digital Library; Information Security

（收稿日期：2015-09-18；编辑：雷雪）