doi:10.3969/j.issn.1000-436x.2013.z2.014

基于联盟身份认证和 CALIS 联合认证的图书馆资源访问方案

吕洁[1,2]，陈萍[1,2]，王文清[3]，张扬[1,2]，张蓓[1,2]
(1. 北京大学 计算中心，北京 100871；2. 北京大学 网络与软件安全保障教育部重点实验室，北京 100871；
3 中国高等教育文献保障系统(CALIS)管理中心，北京 100871)

摘 要：针对图书馆电子资源的访问控制问题，对国际上广泛采用的联盟身份认证技术和在国内图书馆大范围部署的 CALIS 联合认证进行了分析，提出了将联盟身份认证与 CALIS 联合认证相结合的方案，并在 CARSI 联盟的平台上进行了开发、部署和验证，实验结果表明，联盟身份认证与 CALIS 联合认证相结合的方案可以有效、灵活地对电子资源进行访问控制。

关键词：联盟身份认证；访问控制；CALIS 联合认证
中图分类号：TP393 **文献标识码**：A **文章编号**：1000-436X(2013)Z2-0069-05

Library resource sharing solution based on federated authentication and CALIS unified authentication

LV Jie[1,2], CHEN Ping[1,2], WANG Wen-qing[3], ZHANG Yang[1,2], ZHANG Bei[1,2]
(1. Computer Center, Peking University, Beijing 100871, China;
2. Key Laboratory of Network and Software Security Assurance Ministry of Education, Peking University, Beijing 100871, China;
3. Management Center of China Academic Library and Information System (CALIS), Beijing 100871, China)

Abstract: To address the problem of enforcing access control for library digital resources, federated authentication technologies were analyze which are widely used in the world, and CALIS unified authentication technologies that are widely deployed in China's libraries were also analyzed. A solution integrating federated authentication and CALIS unified authentication technologies was proposed, which was developed, deployed and tested on the test-bed of CARSI federation. Test results show the proposed solution enforces effective and flexible access control for library resources.

Key words: federated authentication; access control; CALIS unified authentication

1 引言

目前，我国高校图书馆普遍采用 IP 认证的方式对已购买的数据库资源进行访问控制。各个高校图书馆把自己学校的 IP 地址段通报给数据库商。针对师生的校外访问，高校图书馆普遍使用代理服务器或者 VPN，将用户 IP 转换成校内 IP。

针对图书馆资源保护问题，联盟身份认证技术提供了新的解决方案。联盟身份认证技术以安全断言标记语言[1](SAML，security assertion markup language)标准为基础，允许用户使用由自己所在机构管理的用户名和密码进行登录，登录成功后访问其他机构提供的电子资源。联盟身份认证使用最广泛的软件实现为 Shibboleth。

为了把联盟身份认证技术与代理服务器结合，TAKAAKI 等[2]部署实现了基于联盟身份认证的代理服务器。图书馆中使用较多的商用代理软件 EZproxy[3]也提供接口，支持联盟身份认证。RIEGER 等[4]将 EZproxy 与 Squid 代理服务器相结合，可以把通过联盟身份认证的不同用户映射为不同 IP。由于进行了 IP 转换，代理技术在确定用户身份方面始终存在一定缺陷，因此，把联盟身份认证技术与身份认证系

收稿日期：2013-09-13
基金项目：国家发展改革委 2011 年国家信息安全专项“基于可信身份联盟和云计算的数字资源安全防护服务”基金资助项目
Foundation Item: 2011 National Information Security Special Project “A Digital Resource Security Protection Service Based on Reliable Identity Federation and Cloud Computing” Progect Funded by National Development and Reform Commission

统直接结合是较好的选择。

在我国的高校图书馆中，中国高等教育文献保障系统管理中心(CALIS)已经建立了 CALIS 联合认证系统，接入了全国上百所高校的用户库。如果能将联盟身份认证技术与 CALIS 联合认证相结合，无需分别接入各个高校认证系统就可以让高校的用户进行联盟认证。

本文的主要工作为：首先对联盟身份认证技术和 CALIS 联合认证分别进行分析，提出了将联盟身份认证技术与 CALIS 联合认证相结合的解决方案，并对该解决方案进行了部署和验证。

2 相关技术

2.1 联盟身份认证技术

联盟身份认证技术以身份联盟的形式确认用户所属机构和资源提供机构之间的信任关系，在用户所属机构建立身份提供者(IDP，identity provider)，通过身份提供者接入认证系统和用户库，由身份提供者对外提供用户的身份信息；在资源提供机构内建立服务提供者(SP，service provider)，通过服务提供者对应用资源实施访问控制。

身份联盟中，用户的访问过程如图 1 所示：1) 用户试图访问电子资源；2) 服务提供者将用户重定向到用户所属机构的身份提供者；3) 身份提供者对用户进行认证；4) 用户通过认证后，身份提供者把用户重定向到服务提供者；5) 服务提供者取得身份提供者传来的用户信息，把电子资源返回给用户。

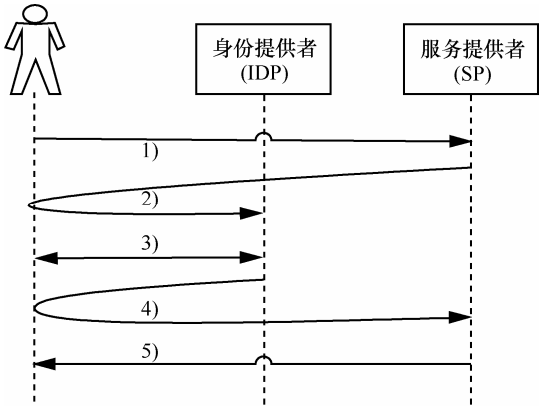

图 1 身份联盟中用户的访问过程

当服务提供者同时为来自多个身份提供者的用户提供服务时，需要在进行第 2)步之前，要求用户选择其所属的身份提供者。另外，服务提供者也可以把用户重定向到联盟的发现服务(discovery service)，根据发现服务提供的信息把用户重定向到相应的身份提供者。

为了提高联盟身份认证的灵活性，HARDING 等[5]提出了动态安全断言标记语言，以应对联盟中的服务提供者或身份提供者动态增加的情况；KOMURA 等[6]提出允许一个用户将自己通过联盟认证后的访问权限交给另一个用户。SATO 等[7]提出通过把身份提供者划分层次，允许一个身份提供者接受其他身份提供者的认证结果。针对多个联盟共存的情况，YAMAJI 等[8]提出方案使得一个身份提供者接受另一联盟的某个身份提供者的认证结果。这些工作都是基于安全断言标记语言标准，没有解决如何与遵守不同规范的身份提供者和服务提供者交互的问题。

2.2 CALIS 联合认证

CALIS 联合认证是由中国高等教育文献保障系统管理中心独立研发的统一身份认证技术。与联盟身份认证技术相似，CALIS 联合认证在用户所属机构内建立 CALIS 园区认证服务，并要求接入 CALIS 联合认证的应用接收 CALIS 园区认证服务发来的用户身份信息。同时，CALIS 还建立了 CALIS 中心认证服务，该服务会要求用户选择自己所属的机构，同时该服务也能够接收 CALIS 园区认证服务发来的用户身份信息。

CALIS 联合认证中，用户的访问过程如图 2 所示：1) 用户试图访问电子资源；2) 用户被重定向到 CALIS 中心认证服务；3) CALIS 中心认证服务要求用户选择自己所属的机构；4) CALIS 中心认证服务把用户重定向到其所属机构的 CALIS 园区认证服务；5) 用户进行认证；6) CALIS 园区认证服务将用户重定向到 CALIS 中心认证服务；7) CALIS 中心认证服务将用户重定向到 CALIS 应用；8) CALIS 应用取得用户的身份信息，允许用户访问相关资源。

联盟身份认证和 CALIS 联合认证最大的不同之处在于：两者遵守不同的标准。联盟身份认证中，第 2)步和第 4)步中身份提供者和服务提供者之间的信息交互遵守安全声明标记语言标准，而 CALIS 联合认证中，第 2)、4)、6)、7)步都使用 CALIS 联合认证特定的规范。因此，经过 CALIS 园区认证服务认证的用户无法直接访问联盟身份认证的服务提供者，2 个认证过程并不兼容。

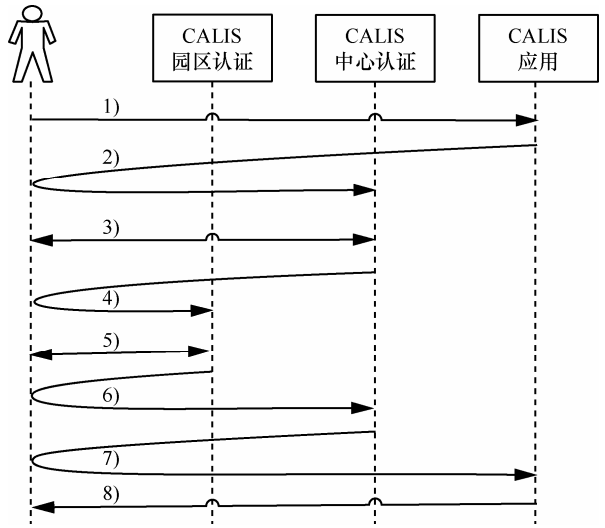

图 2　CALIS 联合认证中的用户访问过程

3　方案设计

如何将遵守不同协议的认证过程结合起来，始终是一个技术难点。REIGER 等[9]提出了由身份提供者提供用户代理，通过用户代理取得遵守其他标准的身份提供者的认证结果和用户信息。但是这一方案主要是解决基于安全断言标记语言标准与基于 OpenID 标准的联盟之间的互操作问题。

针对联盟身份认证和 CALIS 联合认证不兼容问题的解决方案的核心思想为：在用户所属机构内既建立联盟身份认证的身份提供者，也建立 CALIS 园区认证服务。服务提供者按照安全声明标记语言标准将用户重定向到身份提供者，身份提供者按照 CALIS 联合认证的规范将用户重定向到 CALIS 园区认证服务，CALIS 园区认证服务对用户进行认证后，遵守 CALIS 联合认证的规范将用户重定向到身份提供者，身份提供者解析出用户信息并按照安全声明标记断言标准将用户信息提供给服务提供者。

该方案有 2 个关键之处：一是认证过程的结合；二是用户属性信息的转换。

3.1　认证过程的结合

联盟身份认证和 CALIS 联合认证结合后的用户访问流程如图 3 所示：1) 用户试图访问服务提供者所保护的电子资源；2) 服务提供者将用户重定向到身份提供者；3) 身份提供者将用户重定向到同一机构的 CALIS 园区认证服务；4) CALIS 园区认证服务对用户进行认证；5) 用户认证通过后，CALIS 园区认证服务把用户重定向到身份提供者；6) 身份提供者把用户信息解析出来，把用户重定向到服务提供者；7) 服务提供者把电子资源返回给用户。

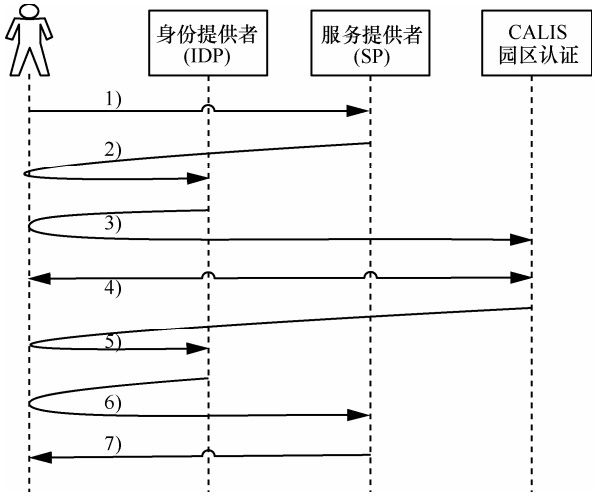

图 3　联盟身份认证与 CALIS 联合认证兼容解决方案

3.2　用户属性信息的转换

联盟身份认证中，身份联盟确定了身份提供者和服务提供者之间的信任关系，同时也对身份提供者和服务提供者之间的属性传递进行了规范，包括属性名称、可能的取值、各个取值的语义等。目前国际上最为通行的属性集是 Internet2 制订的 EduPerson 属性集[10]。

CALIS 联合认证对于 CALIS 园区认证服务和 CALIS 应用之间的用户信息传递也进行了规范。

以下用户类别这一属性为例进行说明。根据 EduPerson 属性集的规定，用户类别这一属性的属性名为 eduPersonAffiliation, 可能的取值包括 faculty、student、staff、alum、member、affiliate、employee 等。而按照 CALIS 联合认证的规范，该属性的属性名为 userType，取值为整数，不同的数值代表不同的用户类别，如取值为“1”表示本科生，“2”表示研究生，“3”表示博士等。

当把联盟身份认证与 CALIS 联合认证进行结合时，需要把 CALIS 联合认证结果的用户属性值转换为符合身份提供者所使用的属性集的用户属性值。

在图 3 的第 5)步中，身份提供者可以取得用户的身份信息，并记录在日志文件中。在第 6)步中，身份提供者根据安全断言标记语言标准，把用户的属性信息提供给了服务提供者，这个属性交换过程在身份提供者和服务提供者上都会有对应的日志。所以，如果服务提供者发现某个用户有恶意下载行为，可以与身份提供者核对相应日志记录，从而确定用户的身份。

4 实现和验证

对联盟身份认证和 CALIS 联合认证的解决方案的实现和验证在 CARSI 联盟的平台上进行。

CARSI 联盟是由北京大学发起的，致力于在 CERNET 高校内推动统一认证和资源共享的项目。CARSI 联盟采用 Shibboleth 软件作为安全断言标记语言的软件实现。目前，CARSI 联盟已得到多家知名数据库商的认可和加盟，包括 IEEE、汤姆森路透、Nature 出版集团等。

在实践中，数据库商的应用系统所接收的属性集主要为 EduPerson 属性集，对用户进行访问控制主要依据用户类别 eduPersonAffiliation 属性值，图书馆可以要求数据库商把应用系统配置允许 eduPersonAffiliation 属性值为哪些值的用户可以访问，否则无法访问。

实验环境如图 4 所示。以南京大学的用户和用户库进行实验，相应的 CALIS 联合认证园区认证服务为 uas.js.calis.edu.cn。在南京大学部署 Shibboleth 软件建立身份提供者 idp.nju.edu.cn。实验环境中的电子资源为 IEEE 的 Xplore 数据库和 Nature 出版集团的期刊资源。

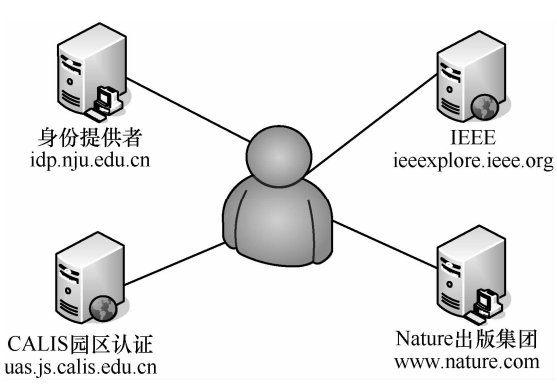

图 4 联盟身份认证和 CALIS 联合认证解决方案实验环境

根据 idp.nju.edu.cn 的认证需求和 CALIS 联合认证的技术规范，开发了实现认证过程和进行用户属性转换的程序，在 idp.nju.edu.cn 上进行了部署。

为了保证足够的灵活性，在程序中允许指定哪些 CALIS 联合认证的用户类型取值对应于哪个 EduPerson 用户类别值。实验中，配置了 userType 取值为 1、2、3、5 的用户类别为 student，userType 取值为 4 的用户类别为 faculty，userType 取值为其他值的用户类别为 other。跟数据库商的约定是只有 eduPersonAffiliation 属性值为 student 或者 faculty 的用户才可以访问。

相关的测试表明，用户使用南京大学的用户名和密码在 uas.js.calis.edu.cn 登录后，如果用户的 userType 为 1、2、3、4、5，则可以访问电子资源，否则无法访问电子资源。完全达到访问控制的目标。

5 结束语

本文对联盟身份认证和 CALIS 联合认证技术进行了分析，提出了解决这 2 种技术不兼容的技术方案，并在 CARSI 联盟环境中，基于 Shibboleth 软件进行了相应的软件实现，部署并测试了相应的功能。测试结果表明：通过联盟身份认证技术和 CALIS 联合认证技术的结合，可以为高校图书馆提供有效灵活的电子资源访问控制。当用户存在恶意下载时，通过数据库商的应用和身份提供者的日志记录，可以确定用户身份，从而有效地保护了电子资源。

参考文献：

[1] RAGOUZIS N, HUGHES J, PHILPOTT R, *et al*. Security assertion markup language(SAML) V2.0 technical overview[EB/OL]. http://www.oasis-open.org/committees/download.php/27819/sstc-saml-tech - overview- 2.0-cd-02.pdf, 2008.

[2] TAKAAKI K, HIROAKI S, NORITOSHI D, *et al*. Design and implementation of web forward proxy with shibboleth authentication[A]. 2011 IEEE/IPSJ 11th International Symposium on Applications and the Internet (SAINT)[C]. Munich, Germany, 2011. 321-326.

[3] EZProxy authentication and access software[EB/OL]. http://www.oclc.org/ezproxy/.

[4] RIEGER S. Using federated identities to access IP-protected web resources in multi-customer environments[A]. 2010 Fifth International Conference on Internet and Web Applications and Services (ICIW)[C]. Ohio, USA, 2010. 478-483.

[5] HARDING P, JOHANSSON L, KLINGENSTEIN N. Dynamic security assertion markup language: simplifying single sign-on[J]. IEEE Security & Privacy, 2008, 6(2): 83-85.

[6] KOMURA T, NAGAI Y, HASHIMOTO S, *et al*. Proposal of delegation using electronic certificates on single sign-on system with SAML-protocol[A]. 2009 Ninth Annual International Symposium on Applications and the Internet (SAINT)[C]. Seattle, USA, 2009. 235-238.

[7] SATO H, NISHIMURA T. Federated authentication in a hierarchy of IdPs by using shibboleth[A]. 2011 IEEE/IPSJ 11th International Symposium on Applications and the Internet(SAINT)[C]. Munich, Germany, 2011. 327-332.

[8] YAMAJI K, KATAOKA T, NAKAMURA M, *et al*. Attribute aggregating system for shibboleth based access management federation[A]. 2010 10th IEEE/IPSJ International Symposium on Applications and the Internet (SAINT)[C]. Seoul, Korea, 2010. 281-284.

[9] RIEGER S. User-centric identity management in heterogeneous federations[A]. 2009 Fourth International Conference on Internet and

Web Applications and Services(ICIW)[C]. Cape Town, South Africa, 2009. 527-532.

[10] Internet2 middleware architecture committee for education directory working group, eduPerson specification[EB/OL]. http://middleware. Interet2.edu/eduperson/docs/internet2-mace-dir-eduperson-200312.pdf, 2013.

作者简介：

吕洁（1982-），女，河南郑州人，硕士，北京大学计算中心工程师，主要研究方向为联盟身份认证、网络安全等。

陈萍（1971-），女，辽宁辽阳人，博士，北京大学计算中心教授级高级工程师，主要研究方向为校园网运行管理、联盟身份认证、网络安全等。

王文清（1965-），男，山西太原人，博士，中国高等教育文献保障系统（CALIS）管理中心总工程师、教授级高级工程师，主要研究方向为数字图书馆技术。

张扬（1976-），男，重庆人，硕士，北京大学计算中心工程师，主要研究方向为联盟身份认证、网络安全、网络应用等。

张蓓（1963-），女，辽宁沈阳人，北京大学教授，主要研究方向为计算机网络及应用。