外购数据库在读者门户系统中的集成认证研究

魏达贤 季士妍 范书云
(国家图书馆，北京 100081)

摘要：基于SAML和VPN认证技术，将外购数据库统一集成进读者门户系统内，旨在实现各个外购数据库之间的单点登录认证、权限管理、分类导航，使读者登录一次读者门户系统就可以访问其权限范围内的外购数据库，为读者提供“可见即可得”的资源服务。文章重点介绍了用户访问外购数据库的流程，分析了基于SAML和VPN认证的实现方式。

关键词：外购数据库；SAML认证；读者门户；VPN认证
中图分类号：G250
DOI：10.3772/j.issn.1673—2286.2014.04.005

随着互联网的普及，数字阅读以其便利性，逐渐成长为人们主要的阅读方式之一。作为公民终身学习课堂的图书馆纷纷投入数字图书馆建设，采购了大量的数据库资源。然而，由于数据库厂商的权限控制，大部分外购数据库只能在馆舍内访问，严重限制了互联网用户的使用。本文即以国家图书馆在读者门户系统中的集成认证实践为例，抛砖引玉，为业界同行提供参考。

近年来，国家图书馆一直致力于将更广泛的数字资源服务于读者，因此在采选、采购、租用外购数据库时，注重提高数据库内容的丰富性、数据库访问范围的广泛性、数据库使用的便利性。目前，在采购的200多个中外文数据库中，已经实现了100多个中外文数据库的互联网访问，打破了空间限制，为广大用户提供了访问资源的便利途径和方式。为了能让更多的互联网用户平等、自由地获取外购数据库资源，国家图书馆在“读者门户系统”内将外购数据库进行集成，使所有用户在任何时间、任何地点登录“读者门户系统”就可获取到所需的资源，极大提高了数据库的使用人群和利用率。

1 外购数据库在读者门户内的集成模式

1.1 基本思路

根据国家图书馆所购买/租用的外购数据库的访问授权范围和服务方式，对所有外购数据库进行集中管理、统一认证、授权控制，从而通过“国家图书馆读者门户”（http://mylib.nlc.gov.cn），用户可以一次性地获取到所有的在线资源。

（1）用户一次性获取资源

以往，用户访问国家图书馆的在线资源，需要从不同的入口进入，访问不同的应用系统，这往往给用户带来使用的不便，也会降低资源的使用效率。通过“读者门户系统”将所有外购数据库进行集中管理、统一认证、授权控制，可以为用户提供一个单一的资源访问入口。同时，根据授权控制，在用户统一认证后，实现了一次登录，可访问其权限许可的所有数据库的功能，避免了用户多次登录的麻烦，极大地方便了用户使用图书馆提供的在线数据库，也大大提升数据库的使用效率。

（2）统一认证无缝获取资源

“读者门户系统”对外购数据库进行统一的认证集成，实现用户的单点登录[1]，即一次登录就可以访问众多的数据库。系统依据数据库的系统软件能否修改，将其分为可修改资源和不可修改资源。“可修改资源”即可将单点登录组件嵌入到资源库的系统软件中，使之能接收和处理集成认证中心发送的各种消息，称之为信任域内资源，反之为信任域外资源[2]。对于信任域内的数据库，采用基于安全断言标记语言SAML规范实现分布式数据库的统一身份认证；对于信任域外的

资源，采用基于VPN的RADIUS认证协议实现统一身份认证。基于此种模式，打破了以往用户访问外购资源库，受到所处位置的限制，真正为用户提供了“无围墙”的数字图书馆服务。

（3）数据库的统一授权管理

为了满足不同层次用户的资源获取需求，同时根据外购数据的访问授权范围，将用户分为普通注册用户、实名认证用户、持有国家图书馆物理卡的用户等几种类型，每类的用户可使用的数据库种类和数量不同。系统对数据库进行统一的授权管理，根据读者的类型，授予每类用户不同的角色，每种角色被授予不同的数据库访问权限。这样，既严格遵守了各个外购数据库的授权要求，又为每种类型的用户提供了他们所能获取的资源的最大化。

1.2 访问流程

国家图书馆在设计“读者门户系统”时，本着“可见即可得”的原则，不同类型的用户登录后仅可以看到其权限范围内许可的数据库列表。在实现上，系统将众多的数据库以树状形式，按照资源类型进行分类导航，便于读者找到所需的数据库。用户具体的访问流程如图1所示。

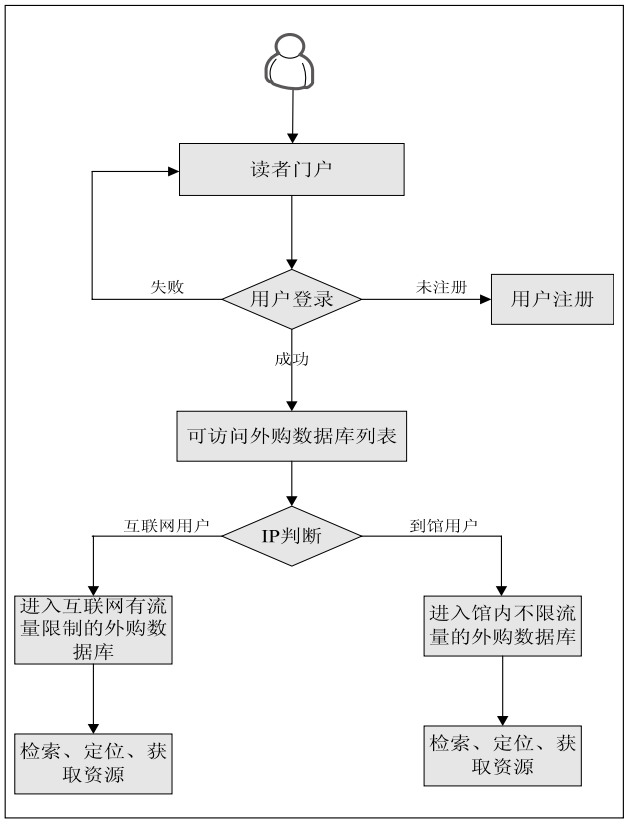

图1 用户的访问流程

当用户访问“读者门户系统”网站时，在未登录状态下，系统默认为用户呈现实名注册用户可访问的所有资源。若用户需要进一步访问资源库，则需进行用户登录，登录成功后就可以进入到其权限允许的数据库内了，同时“读者门户系统”资源导航树更新为相应权限允许的数据库列表。若用户未注册，则跳转到注册页面。在注册时，用户可以首先注册为普通用户，当用户想拥有更多的数据库使用权限，可进一步完成实名验证，从而成为“实名认证用户”。持有国家图书馆读者卡的用户不需要注册，使用读者卡号就可以登录。

另外，“读者门户系统”的外购数据库服务分馆舍内和互联网两种服务模式。系统可根据用户所处的位置，智能区分到馆用户服务和互联网用户服务。到馆用户服务，则是系统为用户提供不受流量限制的数据库服务，使得用户可以无限制地阅读、下载所需资源。同时，由于在馆域网内访问，速度较快；互联网用户服务，则是系统为互联网用户提供有一定限制的数据库服务，受国家图书馆购买的数据库互联网总流量的限制，一旦总流量使用完毕，则读者暂时无法使用，需要国家图书馆继续购买流量后方可使用。因此，为了更好地服务读者，系统自动根据访问用户的IP信息，智能地区分互联网用户和到馆用户，引导这两类用户进入各自的访问通道，最大限度地节省互联网访问流量。

1.3 访问权限控制

由于国家图书馆的用户和外购数据库较多，每类用户的资源需求不同。因此，采用RBAC访问控制模型[3]，即通过用户、角色、权限模型对用户访问控制进行管理。用户、角色、外购数据库的授权关系如图2所示。

每个用户属于不同的角色，系统对不同角色的访问权限进行授权，这样，属于不同角色的用户也就具有了相应的数据库访问权限。同时保留用户到数据库访问权限的直接分配，增加用户权限分配的灵活性。

基于此模型，“读者门户系统”针对不同用户群的资源需求，将用户分为普通读者、实名认证读者、物理卡读者（持有国家图书馆读者卡的读者），每类人群可使用的资源种类和数量各不相同。

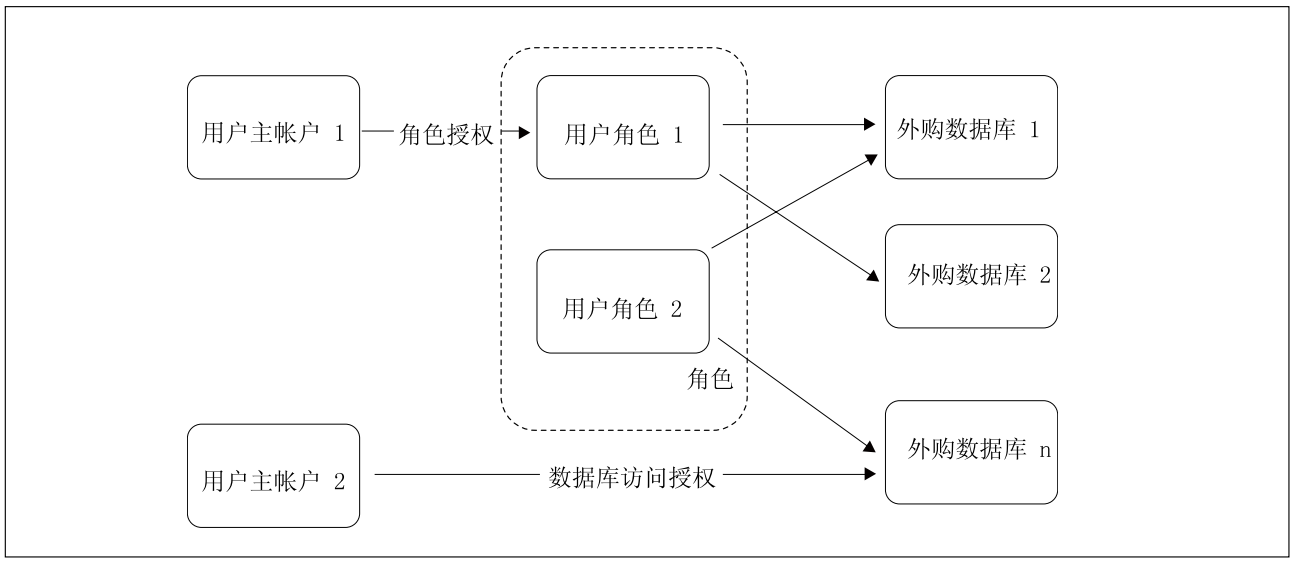

图2 访问权限授权

2 外购数据库集成平台的实现

为了能对外购数据库进行统一集成管理，让用户通过“读者门户系统”一次登录就可以使用众多外购数据库，本平台在架构设计上基于“统一用户管理系统”和“读者门户系统”两个子系统实现。两个子系统的关系如图3所示。

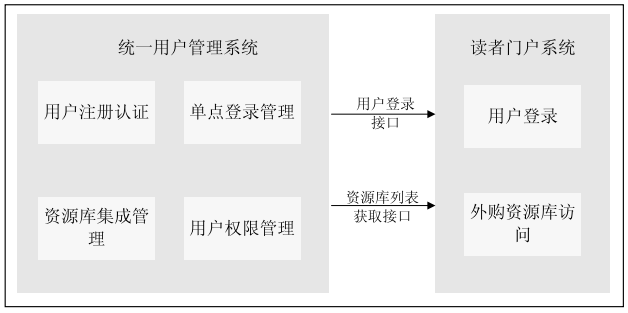

图3 平台总体架构

“统一用户管理系统”作为后台管理系统，主要实现用户注册登录认证、系统间的单点登录、外购数据库的集成及用户权限的管理等；“读者门户系统”是用户访问数字资源的前端入口，实现与统一用户管理系统进行界面、功能上的集成，通过用户登录接口为用户提供用户注册、登录、认证服务，通过数据库列表获取接口获取每位用户可访问的外购数据库列表，以分类导航树的方式提供给用户。

由此可见，“读者门户系统”负责外购数据库的分类揭示，提供统一的访问入口，而外购数据库的集成管理、用户管理等工作都是由“统一用户系统”负责，两个系统间通过接口进行数据的传输。这样降低了两个系统之间的耦合度，方便了对外购数据库增加、删除及权限变更等管理。

单点登录是本平台的重要功能模块，使用户进行一次身份认证便可以访问其被授权的所有网络资源。目前主流的单点登录实现方式是基于SAML框架协议的，该种方式需要对外购数据库进行接口改造，使其符合标准接口。因此，对于可进行接口改造的外购数据库，国家图书馆采用SAML框架协议进行单点登录认证；否则，采用基于VPN的RADIUS认证协议实现单点登录认证。

2.1 基于SAML的认证方式

SAML[4]是国际标准化组织OASIS安全服务协会制定的基于XML的安全标准，用于在Internet不同的安全域中交换身份验证和授权凭证。SAML建立了一种独立于协议和平台的验证和授权交换机制，且能够用于集中式、分散式以及联合式的部署场景。这样使得SAML具有以下特点：它提供单次登录身份验证的功能，可以大幅度地减少站点之间的复制安全性和身份验证信息的需求；SAML可令不同类型的安全服务系统之间实现交互；SAML不依赖于它所交互的任何系统，每个系统都可以为用户的身份验证和授权建立自己的策略。

2.1.1 基于SAML的单点登录认证模型

该模型是基于SAML框架，在用户浏览器、单点登录服务器与应用系统之间进行用户身份断言的分配、传输、验证，并采用SAML的安全机制（如SSL）保证信息传输的安全。该模型具体如图4所示。

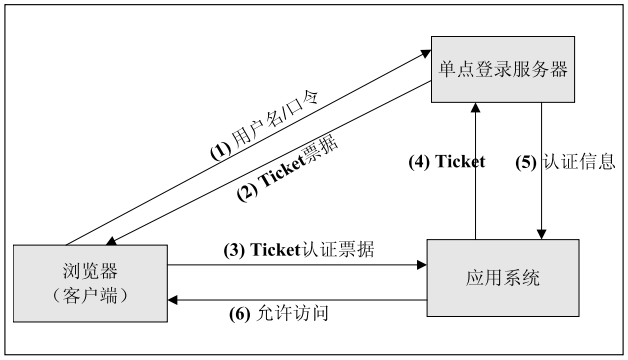

图4 SAML单点登录认证模型

第一步：用户从登录入口进行登录，单点登录服务器验证用户身份，并从读者用户库获取用户的应用系统访问权限列表。

第二步：单点登录服务器为用户生成SAML身份断言，完成断言的预签名并保存在数据库中，返回给用户应用系统权限列表，并发给用户身份Ticket票据。

用户Ticket票据作为用户身份凭证，由单点登录系统在统一登录后颁发，票据内提供用户的临时身份id、用户账号、有效期等信息，可作为用户访问应用系统时的身份验证凭证。票据内容经过加密保护，防止用户信息外泄。

第三步：用户从应用系统权限列表中选择所需访问的应用系统。

第四步：应用系统接收到用户的访问请求，获取用户的Ticket票据信息，并根据该票据信息生成SAML身份断言请求，询问单点登录系统用户的身份。

第五步：单点登录系统将该用户预生成的断言形成断言回复，发送到应用系统，应用系统验证用户身份断言的有效性后，准许用户的访问请求。

第六步：用户从应用系统中获取资源。

2.1.2 用户登录后访问外购数据库的具体流程

基于上述的单点登录模型，国家图书馆设计开发了外购数据库的SAML认证方式。用户登录后访问外购数据库的详细流程如图5所示。

在整个流程中，以浏览器cookie作为用户Ticket票据载体，在用户登录时由单点登录服务器写入到用户的客户端上。当用户访问外购数据库系统或访问单点登录服务器时，浏览器会携带该票据cookie进行访问请求，从而实现用户Ticket票据的传递。这样，在登录的有效期内，用户可以通过手动输入地址栏网址的方式访问外购数据库，用户体验较好。

2.1.3 外购数据库接入实例（以清华同方知网为例）

统一用户管理系统为外接系统提供了简单、快捷的接入方式，外接系统需要对自身进行部分代码改造。当外接系统接收到来自读者门户系统用户的访问时，外接系统运行单点访问验证流程，判断用户是否有效、是否有权限访问本资源，从而实现与统一用户管理系统间的单点登录与控制。下面以清华同方知网为例子，说明外接系统如何与统一用户管理系统实现对接。

（1）知网与统一用户管理系统的认证过程

下面主要描述用户登录读者门户后访问知网，知网与单点登录服务器之间的认证过程。

第一步：用户登录国家图书馆读者门户系统，访问同方知网资源链接，页面跳转到知网首页；

第二步：知网系统通过Ticket操作接口获取用户Ticket票据信息；

第三步：知网系统将Ticket票据信息通过验证接口发送给单点登录服务器作认证；

第四步：如果单点登录服务器返回认证失败标识，则认证结束，知网系统跳转到读者门户登录界面，提示用户进行登录；

第五步：如果单点登录服务器返回认证成功标识，知网系统则记录用户本地Session，认证完成，并打开同方知网页面。

（2）知网系统代码改造

第一步：配置文件

知网系统将统一用户管理系统提供的keystore.jks、log4j.properties、spconfig.properties三个文件放到系统项目的src目录下。keystore.jks为签名用的证书；log4j.properties为日志log4j的配置文件；spconfig.properties文件为单点登录所需的全局配置信息。

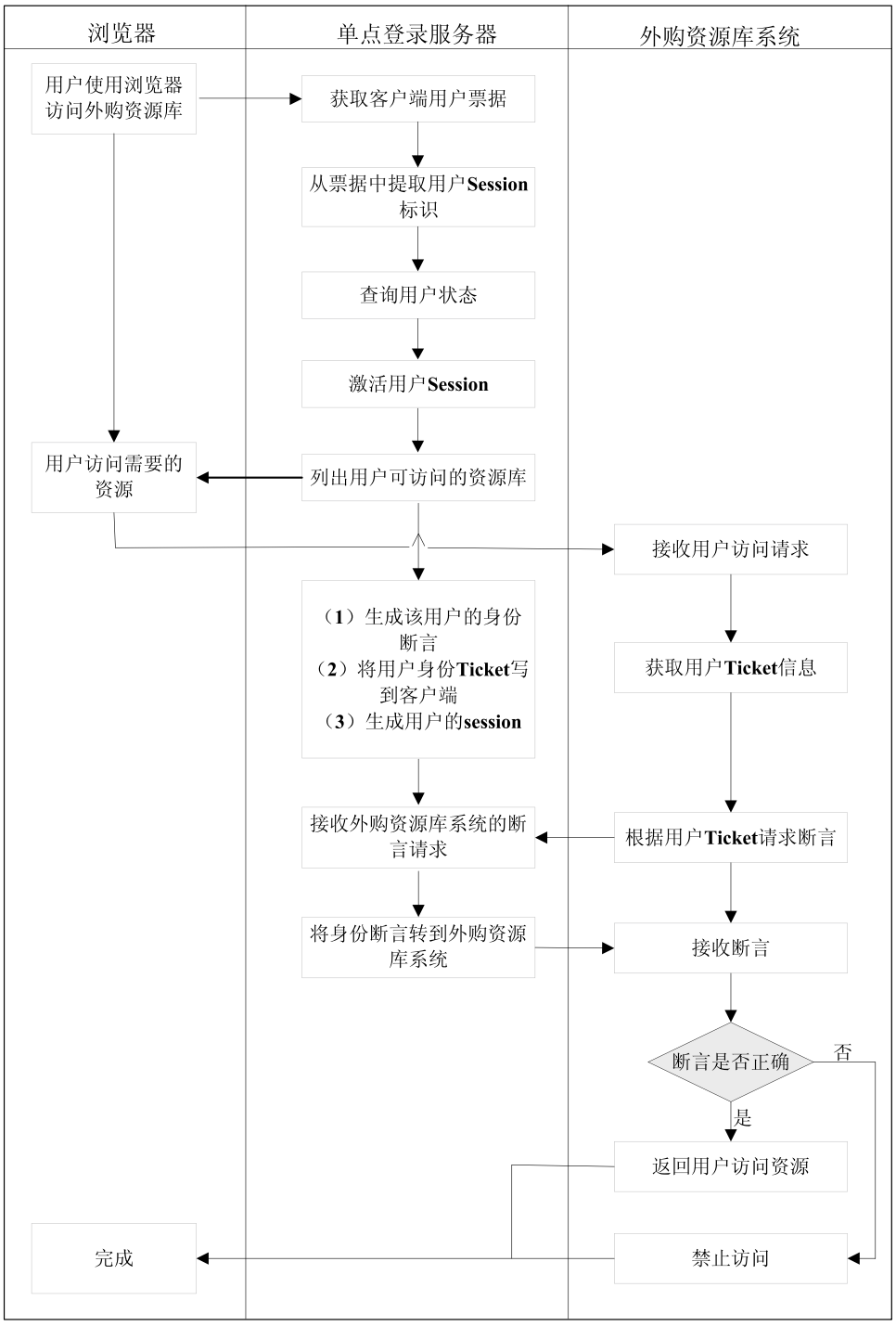

图5 用户登录后访问外购数据库的流程

第二步：所需JAVA类库

知网系统将统一用户管理系统提供所需的sso-cnlib.jar、opensaml-2.5.0.jar包复制到应用程序的\WEB-INF\lib目录下，重新启动WEB服务器后生效。其中sso-cnlib.jar为单点登录模块API所在jar包，opensaml-2.5.0.jar为SAML协议实现的jar包。

第三步：前端页面JS代码整合

a) 登录验证整合

首先，在知网系统的登录网页中的head部分加入代码片段1，以获得用户ticket票据信息：

```
<script language="javascript">
        document.write("<script language=\"javascript\" src=\" http:// sso1.nlc.gov.cn/sso-login/get-sso-ticket?rand="+Math.random()+"\">");
        document.write("<\/script>");
</script>
```

代码片段1：引入获取ticket的js

其中http://sso1.nlc.gov.cn/sso-login/为访问单点登录服务的根url。src属性中的rand参数是为了防止产生缓存。

然后编写ticket处理操作代码，如代码片段2所示：

```
varssoUtil = new SSOUtil();
varmessageArea = document.getElementById('loginMsg');
if (ssoUtil.ticketExist()) {//ticket存在
        ssoUtil.ssoProcess('sso-login',
{ 'sendType':1,'callback':function(returnObj)
{//验证成功，这里假设后台servlet输出的json结构为:
{'result':返回码,' errmsg' :' 错误信息' },
        //返回码为200表示成功
        if (returnObj.result == 200) {
        document.getElementById('loginOut').style.display = '';//显示登录成功
        messageArea.innerHTML ='欢迎您登录';
        } else {//验证失败
        messageArea.innerHTML ='非法登录';}
},'process':function() {
        messageArea.innerHTML = "正在登录...";},'type':'json'
        });} else
{messageArea.innerHTML = "ticket不存在";}
```

代码片段2：单点登录处理

b) 单点退出整合

在知网系统退出页面的head部分加入代码片段3，获取用户ticket信息，然后将ticket信息发送给单点登录服务器，单点登录服务器删除ticket信息，完成退出操作。

```
varssoUtil = new SSOUtil();
functionloginOut() {
        ssoUtil.ssoProcess('sso-logout');
        //sso-logout为应用系统处理单点退出的后台地址
        return false;
}
```

代码片段3：退出处理

第四步：后端业务逻辑JAVA代码整合

a) 登录验证整合

知网系统需依据代码片段2编写一个名称为sso-login的servlet，完成系统读者后台认证功能。系统通过代码片段1获取到ticket后，发送到sso-login，判断ticket是否有效，如果无效，则跳转到读者门户登录页面，如果有效，打开知网页面。

b) 退出整合

知网系统需依据代码片段3编写一个名称为sso-logout的servlet，完成系统读者后台退出功能。当读者执行退出系统操作时，系统将ticket信息发送到sso-logout，从而退出系统。

第五步：其他

完成以上四步以后，基本上完成了知网系统与统一用户管理系统的单点登录集成。由于知网系统使用国家图书馆读者数据库，不需要读者信息的同步，读者信息同步环节可以省略。如果外接系统有自己的读者信息库，则需要读者信息的同步工作，否则统一用户管理系统不能识别读者信息，不能完成系统间的单点登录。

2.2 基于VPN的认证方式

VPN属于远程访问技术，简单地说就是利用公网链路架设私有网络。在公用网络上建立专用网络的技术，也称为虚拟网，它涵盖了跨共享网络或公共网络的封装、加密和身份验证链接的专用网络的扩展[5]。

国家图书馆针对无法通过SAML集成认证的外购数据库，采用VPN认证的方式，使得通过认证的馆外用户通过VPN通道，访问国家图书馆的资源。

基于VPN的认证流程如图6所示。

整个流程中，统一用户管理系统负责用户信息的管理、用户身份的认证，将验证后的信息反馈给VPN设备；对通过验证的用户，VPN设备为其建立SSL访问通道，这样用户就可以访问到外购数据库了。具体的流程如下：

第一步：用户访问国家图书馆读者门户系统，通过系统登录页面填写已在统一用户管理系统注册的用户名和密码进行登录。

第二步：读者门户系统将用户填写的用户名和密码发送至统一用户管理系统进行认证，统一用户管理系统对用户下发单点登录Ticket票据。

第三步：统一用户管理系统将认证结果返回给读者门户系统，内容包括该用户可访问的外购数据库列表及权限等信息。

第四步：通过认证后的用户点击门户网站中受VPN保护的外购数据库链接，VPN设备获取用户Ticket票据，并发送到统一用户管理系统进行用户身份认证。

第五步：认证通过后，VPN设备建立用户与外购数据库之间的SSL访问通道，用户即可进入外购数据库系统。

3 结语

国家图书馆读者门户系统自2011年4月份上线运行以来，截至2014年1月份，通过统一用户管理系统集成管理的外购数据库已经有150多个，读者日访问量1万次，日注册用户平均1000人，日在线用户2000人左右。

通过将外购数据库集成进“读者门户系统”内，改变了用户访问资源的方式，使读者足不出户就可以访问到国家图书馆的数字资源，方便了资源的获取、扩大了资源的服务人群、拓宽了服务渠道、提升了服务水平。

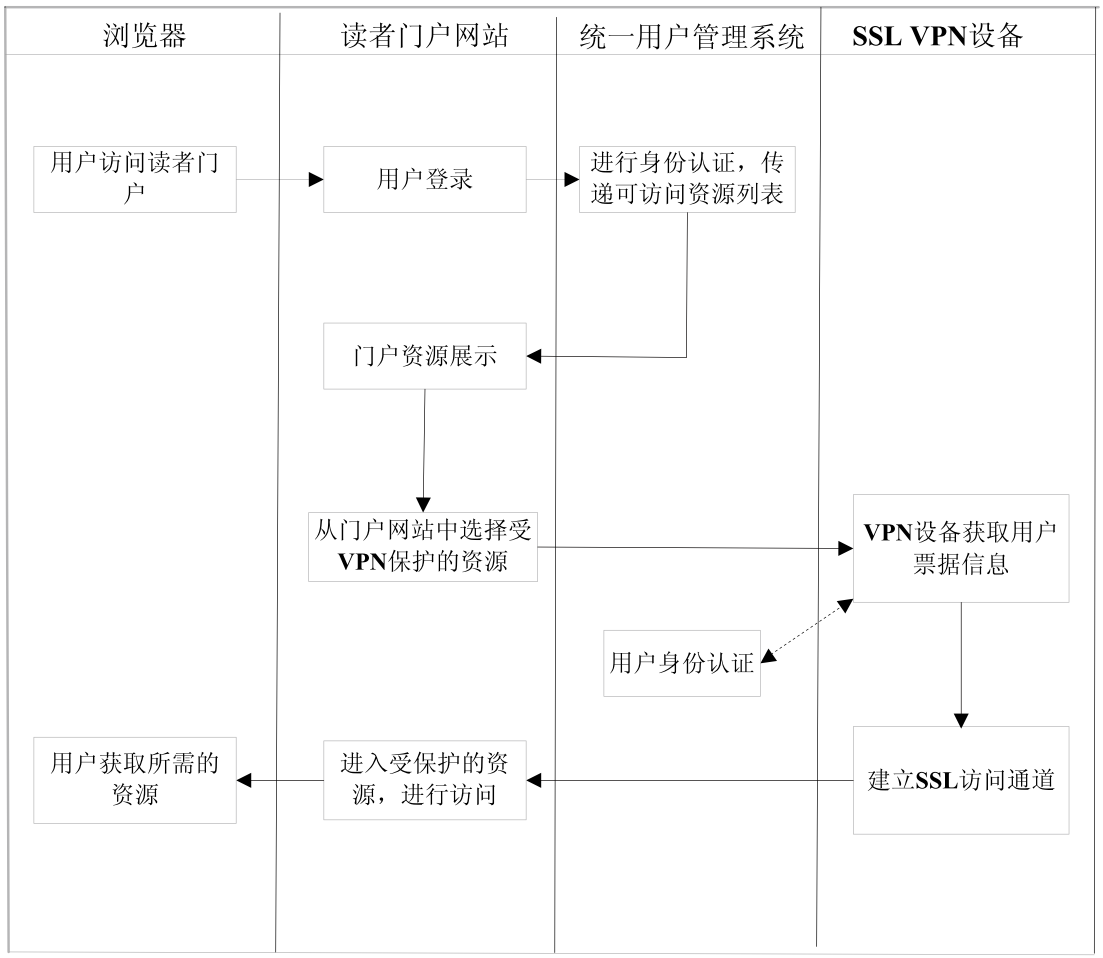

图6 基于VPN认证流程

参考文献

[1] 黄坤.基于SAML的单点登录技术研究[J].计算机与数字工程,2012,(8):90-93

[2] 王小梅.国家科学图书馆远程访问系统的设计与实现[J].现代图书情报技术,2008(8):79-83.

[3] 国家图书馆.国家数字图书馆统一用户管理系统项目需求书[R].2011-4.

[4] 罗兰.基于SAML 2.0的WebSphere Application Server SSO实现[EB/OL]. [2014-01-06]. http://www.ibm.com/developerworks/cn/websphere/library/techarticles/1111_luol_sso/1111_luol_sso.html.

[5] 王泽贤.VPN与ILAS III统一用户认证的设计与实现[J].现代图书情报技术,2011(12):79-82.

作者简介

魏达贤，男，1980年生，工程硕士，国家图书馆工程师。E-mail: weidx@nlc.gov.cn。
季士妍，女，1978年生，硕士生，国家图书馆高级工程师。
范书云，男，1982年生，硕士生，国家图书馆工程师。

Research on Integrated Authentication of Outsourcing Databases in the Reader Portal System

WEI DaXian JI ShiYan FAN ShuYun
（National Library of China, Beijing 100081, China）

Abstract: Outsourcing databases are integrated into the reader portal system based on the SAML and VPN authentication technology, aiming at the SSO, access control, classified navigation between different outsourcing databases. Readers log in the reader portal system once, then can access databases with its scope of authority, providing "visible can be available " for readers. The paper introduces the process of accessing outsourcing databases and analyzes the realization mode of SAML and VPN authentication.

Keywords: Outsourcing database; SAML authentication; Reader portal; VPN authentication

（收稿日期：2014-03-11）