文章编号:1000-5641(2015)S1-0471-05

基于 OAuth 的数据共享方案研究

欧阳荣彬, 杨 旭, 王倩宜, 刘云峰
(北京大学 计算中心,北京 100871)

摘要: 本文基于 OAuth 框架协议提出了一种数据共享方案,实现数据便捷共享,确保只能授权用户访问,并且有效防止授权后其访问方式被滥用.本文不仅介绍了方案的框架内容,还详细阐述了相关细节.在北京大学数据综合服务平台中的实践显示,本文提出的方案达到了预期目的.

关键词: 数据共享; OAuth 协议; 访问票据
中图分类号: TP315 **文献标识码**: A **DOI**:10.3969/j.issn.1000-5641.2015.z1.076

A research of data sharing based on OAuth protocol

OUYANG Rong-bin, YANG Xu, WANG Qian-yi, LIU Yun-feng
(*Computer Center, Peking University, Beijing* 100871, *China*)

Abstract: This paper presents a data sharing scheme based on OAuth protocol. With it, data can be shared conveniently, service can be accessed only by authorized user, and the access address will not be abused. This paper introduces the scheme, and also presents some implementation details. The scheme's application in PKU Data Service Platform shows that it achieves its design purpose.

Key words: data sharing; OAuth protocol; access token

0 引 言

随着高校信息化不断推进,各个业务部门陆续建设了不少的信息系统,这些信息系统在业务功能上围绕各自的核心业务功能建设,在技术架构上目前看来都具有其历史的局限性,因而形成了典型的"信息孤岛".解决信息孤岛问题,在业务层面上需要重构业务模型,打通各自的壁垒,实现业务协同;在数据层面上,则需要设计完整流畅的数据流,实现信息数据的集成和共享.

数据共享的方式一般可以分为两种:复制式和借阅式.复制式数据共享是指访问并占有共享数据(一般是另一份拷贝),拥有对其的全部权限;借阅式数据共享是指访问但不占有共享数据,只对其拥有受限的部分权限.不管哪种方式都需要权衡两个方面的问题:便捷性和安全性.复制式数据共享对于用户是透明的,用户并不清楚用的是哪一份共享数据,一旦原数据发生了变化,应当如何传播并共享其变化?而且即使是数据共享中心,其权限也应受到

收稿日期:2014-10
第一作者:欧阳荣彬,男,硕士,高级工程师.E-mail:ouyang@pku.edu.cn.

限制.借阅式共享一般能够提供便捷的访问方式,但是如何确保未经授权不得访问,以及一旦授权后访问途径不得被滥用?

本文基于 OAuth 的授权模型提出了一种方案,尝试应对解决上述数据共享面临的问题.

1 相关研究

近年来,为了应对"信息孤岛"问题,在数据集成与共享方面,研究人员发表了很多有益的成果.王倩宜等在[1]文中提出了通过基于代理的数据交换实现数据集成,并且发布数据服务实现数据共享,但是对于数据服务的进一步设计和安全性考量,文[1]中并未提及.Ming Li 等在文[2]中提出了一种通用的跨平台数据共享方案,采用 XML 描述数据元信息,指出数据使用方首先需要认证,然后才能访问数据提供方提供的数据服务,但是对于认证过程的具体要求和建议方案并未进一步阐述.

OAuth 是一种专门针对跨平台的应用之间授权而设计的框架协议,其整体思路是采用跨应用之间重定向的方式让用户显式而明确地参与授权过程,并且保护用户关键的凭证(口令)信息[3-4].在其 2.0 版本中,虽然为了更好地兼容各种应用系统而设计了多种模式,但其标准模式与 1.0 版本中的设计依然基本一致[5-7].

2 数据共享方案

本节将从三个方面详细阐述数据共享方案,首先介绍方案的整体框架,接下来阐述其交互流程,最后就有关具体方法进行详细介绍.

2.1 整体框架

如图 1 所示,整体框架主要由三大部分组成:数据服务中心、数据服务使用方和终端用户.终端用户是指最终访问数据的用户,数据服务使用方是指使用数据共享服务访问共享数据的应用,数据服务中心是指数据共享服务提供方.

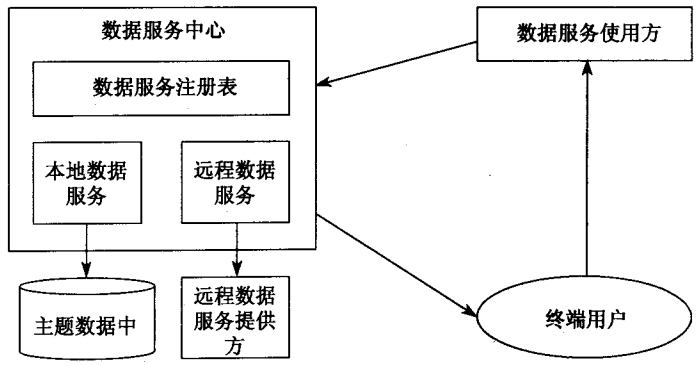

图 1 数据共享整体框架

数据服务中心的整体设计方案在文献[1]中已有探讨,它注册了所有可以发布的数据服务.这些数据服务大致可以分为两类,一类是由数据服务中心提供的本地数据服务,数据来源是通过数据交换建立的主题数据中心[1];另一类是其他的远程数据服务,其数据来源不是主题数据中心,只是通过数据服务中心转发.远程数据服务只是针对一些个案设计的特殊渠

道，本文不详加阐述.

2.2　交互流程

举例说明应用场景：数据服务中心存储了所有学生的照片信息，并对外提供照片访问服务；某学生访问学生系统打印学籍卡需要学生照片，而学生系统并没有学生照片，需要访问数据服务中心获得学生照片；因此学生系统首先会向数据服务中心请求验证并获得授权，授权通过之后返回访问票据，最后学生系统通过访问票据访问学生照片，学生从而完成了打印学籍卡.

总结起来，首先有如下两步作为前提必须完成：

(1) 数据服务中心发布数据服务；

(2) 应用向数据服务中心注册成为数据服务使用方.

在此基础上，通常的数据共享访问交互流程(如图 2)大致如下：

(1) 终端用户访问应用(数据使用方)，应用(数据使用方)需要使用共享数据；

(2) 应用(数据使用方)向数据服务中心请求身份验证和授权；

(3) 数据服务中心校验身份验证和授权请求，返回数据服务访问票据；若不通过，则流程终止；

(4) 应用(数据使用方)使用票据访问数据共享服务；

(5) 数据服务中心校验访问票据，若通过，返回数据内容；若不通过，则流程终止；

(6) 应用(数据使用方)使用返回的数据内容响应用户请求.

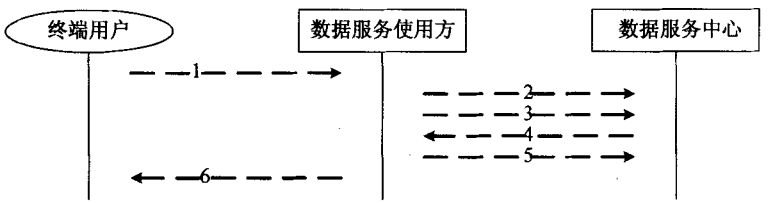

图 2　交互流程图

一般的，上述步骤(5)中可能直接使用应用的身份凭证信息(如应用 ID 和口令)访问数据共享服务，但是类似这种访问经常被直接集成在网页中，从而就很有可能暴露关键的凭证信息. 而且，数据共享服务的访问凭证应该是有过期时间的，不能长期有效，否则就又可能被滥用. 显然，如果使用应用的身份凭证信息(如应用 ID 和口令)访问数据共享服务，无法确保该访问途径不被滥用. 因此需要采用一种带有有效时长(即存在过期时间)的随机而唯一的访问票据，OAuth 在这方面有着独特的适用性.

OAuth 的标准模式中在身份验证和授权部分需要数据所有者显式的参与[4-5]，然而本文的交互流程方案中并没有数据所有者个体的参与. 因为从主题数据中心形成过程可以得出数据服务中心即是共享数据的所有者[1]，所以参考文[5]中的 OAuth 2.0 Client Credentials 模式形成了本文的交互流程方案，直接在应用身份验证过程中同时完成访问授权.

2.3　相关方法

2.3.1　应用注册

如 3.2 所述，应用需要首先向数据服务中心注册成为数据使用方，而且交互过程会在应用身份验证同时通过数据访问授权，因此应用注册信息成为关键的凭证信息. 应用注册同时

也是为了确保数据共享服务被信任的数据使用方访问,是避免数据泄露的必要手段.一般的,注册时会要求应用提供 ID、名称和口令等信息,但是实际应用中发现口令还是容易泄露,因此采用了同时登记应用 IP 的方式,身份验证时则验证应用的请求 IP 是否为注册 IP,以此确保数据使用方为登记注册的受信任方.

然而注册应用 IP 的方式也限制了应用的形式,其只能是 BS 形式的应用,对于那些 CS 形式的应用由于发布范围广,逐一登记 IP 不现实,因此针对 CS 形式应用依然登记口令,在身份验证时采用口令验证.

2.3.2 票据生成

文[4]和[5]中并没有明确票据的生成方法,本文的框架在具体实施过程中是采用 MD5 消息摘要的方式生成,与数据使用方、终端用户、数据服务 ID、请求时间等信息有关,以此得到一个随机而唯一的数据共享服务访问票据,详细过程本文不再赘述.

数据共享服务访问票据是有时效的.一般的,票据一旦使用应当马上失效,否则可能导致票据被滥用;如果在获得票据之后并没有马上访问数据共享服务,而是过了一段时间再访问,这个时间长度即有效时长,一般应限制为 5 分钟.但是,具体的也有数据共享服务并不要求访问之后立刻将票据失效,例如学院教务批量打印学籍卡,此时每次访问都将票据马上失效势必增加交互频度,加大系统负载.因此在具体实施过程中,这类数据共享服务的访问票据不马上失效,而且其有效时长会设置一个较长的时间片.然而为了确保这类数据共享服务不被滥用,要求每次访问都增加一个随机标识 NONCE,数据服务中心每次都校验其是否重复,如果重复则返回错误信息.

3 应用效果分析

北京大学数据综合服务平台在个人照片共享中应用了上述方案,数据共享平台发布了个人照片共享服务.北京大学校内信息门户中的个人基本信息功能、北京大学学生综合信息管理系统的学籍卡打印功能、北京大学人事综合信息管理系统职称考试准考证打印功能、北京大学组织工作综合信息管理系统干部简历功能等,均使用了其提供的照片共享服务.

在便捷性方面,照片共享服务对用户是透明的,应用系统使用照片共享服务无需本地存储照片,数据服务平台中照片的变化立刻可以反馈到各个应用系统中去.而且应用系统使用照片共享服务,相比较传统方式而言只是增加了一步身份验证.因此,对于用户和应用系统而言本文的方案都具有足够的便捷性.

在安全性方面,数据共享访问必须经过验证授权,而身份验证和授权只是在应用系统和数据综合服务平台之间进行,在网络信道安全(例如使用 SSL 协议)的情况下,可以确保身份凭证等关键信息不暴露.终端用户或者潜在的攻击者只能发现访问票据,而访问票据只是一个随机字符串,无法通过其获取身份凭证信息.同时,访问票据带有有效时长,会在一定时间内过期,并且还附带随机标识 NONCE,可以避免访问途径被滥用.

但是应用系统在使用共享服务提供的照片之后可能另外存储,或者可能被潜在的攻击者另外存储照片之后用于其他用途,本文方案针对这类安全问题并未涉及,需要进一步完善.

4 小　结

本文分析了高校信息化中数据共享的常见模式和面临的主要问题，基于 OAuth 协议框架提出了一种数据共享方案. 方案具有便捷性特点，数据共享访问必须经过验证授权，验证授权后采用带有效时长的随机性访问票据，避免了身份凭证等关键信息暴露，同时避免了访问途径被滥用，在一定程度上保证了共享资源的安全. 本文方案实际应用于北京大学数据综合服务平台和多个综合信息系统，实践结果显示，方案有效地达到了便捷共享和安全可靠的目的，应用效果良好.

[参　考　文　献]

[1] 王倩宜，李丽，欧阳荣彬，等. 数据综合服务管理平台主体方案探讨[J]. 实验技术与管理，2011，28(5)：4-6，16.

[2] LI M，LUO N L. Data sharing between web applications based on the request of user [C]. 2009 ISECS International Colloquium on Computing，Communication，Control，and Management，CCCM 2009. IEEE，2009：280-282.

[3] LEIBA B. OAuth Web Authorization Protocol [J]. IEEE Internet Computing，January/February 2012：74-77.

[4] HAMMER-LAHAV E. The OAuth 1. 0 Protocol，RFC5849 [S]. Internet Engineering Task Force (IETF). 2010.

[5] HARDT D. The OAuth 2. 0 Authorization Framework，RFC6749 [S]. Internet Engineering Task Force (IETF). 2012.

[6] LODDERSTEDT T，MCGLOIN M，HUNT P. OAuth 2. 0 Thread Model and Security Considerations，RFC6819 [S]. Internet Engineering Task Force (IETF). 2013.

[7] JONES M，HARDT D. The OAuth 2. 0 Authorization Framework：Bearer Token Usage，RFC6750 [S]. Internet Engineering Task Force (IETF). 2012.

(责任编辑　林　磊)