Gestión de Identidad en la Universidad de Sevilla

Identity Management at the University of Seville

Carmen López Herrera

Resumen

Este artículo pretende dar una visión general acerca de la gestión de identidad en una institución como la Universidad de Sevilla.

En él intentaremos centrarnos tanto en la problemática inicial en la que se encuentra una institución como la que nos ocupa, como en las soluciones adoptadas para intentar resolver los problemas iniciales.

Empezaremos planteando la problemática inicial para después, poco a poco, ir adentrándonos en las soluciones propuestas a todos los niveles que hemos visto convenientes, intentando así, por una parte, solventar los problemas propuestos y por otra, avanzar en una de las líneas más innovadoras en lo que a identidad se refiere, adaptándola a las necesidades actuales.

Palabras clave: Identidad, bases de datos, LDAP, SSO, autenticación.

Summary

The aim of this article is to provide an overview of identity management in an institution such as the University of Seville.

In it we address both the initial problems encountered by this kind of institution and the solutions adopted in an attempt to resolve these problems.

After setting out the initial problems, we then go on to examine the proposed solutions from every perspective deemed appropriate, not only with a view to resolving the problem but also to advancing in one of the most innovative lines in terms of identity, adapting it to current needs.

Keywords: Digital identity, databases, LDAP, SSO, authentication.

La mayoría de las instituciones tienen bases de datos corporativas con datos específicos de una persona

1. Problemática inicial

En la mayoría de instituciones nos encontramos con distintos repositorios de datos independientes como pueden ser:

- un conjunto de bases de datos corporativas aisladas (personal, nóminas, alumnos...) donde se encuentran la gran mayoría de datos específicos de una persona con una relación contractual con la universidad.
- Por otro lado, es habitual disponer de un directorio corporativo con una serie de datos mínimos que básicamente permiten autenticar un usuario o autorizar algún servicio en función de unos ciertos valores reflejados en los distintos atributos del directorio, y hasta es habitual disponer de un servicio Active Directory donde vuelve a haber más datos del usuario para aplicaciones o servicios concretos.
- Y no olvidemos además la gran cantidad de bases de datos asociadas a servicios independientes que se han ido creando y están en uso para investigadores, becas, enseñanzas propias, etc.

También se han ido creando gran cantidad de bases de datos asociados a servicios independientes

Suele ser habitual que cada uno de los repositorios de datos anteriormente citados dependa de grupos de personas de distintas áreas, incluso de servicios diferentes, sin la más mínima relación entre ellos.

Además, la provisión de datos se suele hacer de forma específica atendiendo a requerimientos concretos e incluso utilizando juegos de caracteres distintos si fuera necesario. Disponemos por tanto, de múltiples repositorios de datos distintos en los que hay datos repetidos, datos mal formados, equivocaciones que conllevan dobles identidades de cara a los usuarios, etc.

Bajo esta perspectiva, los usuarios se ven obligados a preocuparse de mantener sus datos al día en distintos repositorios de datos y a tener que usar claves distintas para acceder a diferentes servicios, con las consiguientes desventajas que ello provoca además del número de personal de apoyo para llevar a cabo toda esta casuística (administradores, personal de apoyo dedicado a la atención de usuarios, etc.)

No olvidemos por otra parte que de cara a los administradores también se plantean grandes problemas puesto que el aprovisionamiento de los datos se suele hacer de forma manual y por tanto, las modificaciones a los mismos se hace igualmente tediosa. Planteemos entonces algún mecanismo que nos permita gestionar toda la identidad de manera centralizada y automática.

El LDAP es el directorio corporativo de la universidad

2. Solución de identidad

Vamos a exponer los distintos mecanismos que hemos implementado en la Universidad de Sevilla para mejorar la situación de partida.

En el caso particular que nos ocupa contábamos con varios conjuntos de datos iniciales como eran las bases de datos corporativas, el servicio de directorio y un conjunto de bases de datos añadidas de distintos servicios universitarios.

2.1. Sincronización de datos y administración de los mismos

El primer paso que nos planteamos cubrir fue mantener un repositorio de datos actualizado, sincronizado en todo momento, con los datos maestros, i.e, los datos de bases de datos corporativas. Este repositorio de datos es el directorio corporativo de la universidad (LDAP).

La Universidad de Sevilla ha conseguido tener en el LDAP los datos al día de forma automática

Inicialmente desarrollamos un conjunto de herramientas que nos permitían, a través de ficheros compartidos, mantener la información del servicio directorio sincronizada con los datos de las bases de datos corporativas. Aunque inicialmente el procedimiento dio los resultados esperados, nos dimos cuenta que cada vez teníamos más requerimientos y que para cubrirlos, necesitábamos realizar más desarrollo específico y sobre todo, contar con un tiempo del que no disponíamos. Fue el momento entonces de evaluar las distintas opciones que había en el mercado y optamos por el producto Sun Identity Manager para gestionar la identidad.

Este producto nos ha permitido definir por un lado los repositorios de datos que hemos considerado necesarios mantener sincronizados y sincronizar así toda la información que se ha considerado oportuna, y por otro lado, nos ha suministrado herramientas que nos permiten, a través de perfiles, permitir a usuarios concretos (administradores) la posibilidad de realizar altas, bajas y modificaciones sobre el directorio corporativo.

Con esto lo que conseguimos es tener en LDAP los datos al día de manera automática a todos los niveles, tanto de altas como de bajas o modificaciones de usuarios. La piedra angular de un producto como el que tenemos entre manos es la lógica interna definida en los procesos de diseño del mismo.

Es crucial tener perfectamente definidos los tipos de usuarios que hay (en el caso de la universidad son PAS, PDI, ALUMNOS, EXPAS, EXPDI, EXALUMNOS, Alumnos de enseñanzas propias, PDI externo,

Profesores de secundaria, Alumnos de secundaria y Miscelánea) y la tabla de compatibilidad entre los mismos, así como las preferencias de perfil.

Esto nos permitirá implementar gran parte de la lógica de la aplicación para que funcione de manera automática sin errores.

Igualmente importantes son asuntos como el tipo de datos que tenemos de los usuarios en los distintos repositorios, cuáles de ellos queremos mantener sincronizados y dónde, qué política de claves queremos implantar, qué política de usuarios es la adecuada (tiempo de vida de usuarios), qué datos permitimos que puedan ser modificados por los propios usuarios para así mantenerlos al día sin otro tipo de intervención, etc.

Todo ello implica un trabajo minucioso y especialmente escrupuloso a la hora del diseño de la solución. El trabajo de coordinación entre las distintas personas y servicios responsables de los datos es una de las partes que puede llevar más tiempo a la hora de realizar un proyecto de este tipo. Un sistema de este tipo conlleva la automatización prácticamente completa de toda la identidad dentro de una institución y se hace realmente peligroso poder encontrarse con situaciones como son la anulación de un usuario por un mal diseño inicial.

Un sistema de este tipo genera la automatización completa de toda la identidad dentro de una institución

Recapitulando un poco, podemos decir que los objetivos conseguidos hasta este momento han sido los siguientes:

- Sincronización automática de los datos de los usuarios desde las distintos repositorios de datos "maestros".
- Altas, bajas y modificaciones de usuarios automáticas.
- Gestión de perfiles centralizado.
- Contraseña única. Sistema de autenticación única.
- Herramientas administrativas específicas de gestión de usuarios para casos concretos.
- Posibilidad de aportar datos por parte del propio usuario y mantenerlos actualizados.
- Posibilidad de activación de servicios como el correo, de manera autónoma por parte del usuario.
- Centralización del cambio de contraseña del usuario. Al haber una única contraseña, el lugar definido para el cambio de la misma es único igualmente.
- Posibilidad de acceso a las aplicaciones universitarias a través de certificados y tarjeta universitaria+PIN (además del acceso por usuario y clave).
- Implementación de mecanismos de control de cambios de forma automática.

2.2. Acceso a datos del usuario desde las distintas aplicaciones

Llegados a este punto se plantea una nueva problemática que consiste en definir los mecanismos para ofrecer los datos que necesitan las distintas aplicaciones de la universidad de manera coherente. Está claro que con lo que hemos conseguido en el punto anterior, estamos en condiciones de dar acceso a los datos actualizados que hay en LDAP, pero... ¿tiene el directorio corporativo de una institución TODOS los datos que pueden llegar a necesitar TODAS las aplicaciones de una institución tan grande como la nuestra?... La respuesta es no.

Se pueden ofrecer datos del directorio corporativo a través del protocolo LDAP

Ante esta situación siempre está la opción de ofrecer los datos que están en el directorio corporativo a través de protocolo LDAP y hacer que se consulte el resto de datos a través de otro tipo de protocolos o servicios como puede ser a través de web services ofertados desde el área de aplicaciones corporativas.

Entendemos que esto "complica la vida" al programador de aplicaciones puesto que tendrá que hacer uso de repositorios y protocolos distintos para conseguir lo que necesita.

Nuestra solución pasa por el siguiente mecanismo:

En la actualidad, el sistema LDAP de la Universidad de Sevilla permite autenticar usuarios a través de usuario y contraseña (únicas) además de autorizar acceso a través de atributos propios del usuario.

El número de atributos que tenemos en la actualidad en el directorio corporativo, aunque es grande, no llega a cubrir todas las necesidades de las distintas aplicaciones que necesitan de datos específicos de usuarios. De ahí que hayamos decidido resolver este problema a través de un directorio virtual que permita acceder a las distintas aplicaciones a todos los datos necesarios a través de protocolo LDAP manteniendo distintos repositorios de datos gestionados internamente por el propio directorio virtual que contemplen todas las necesidades requeridas hasta el momento.

Así, el directorio virtual o metadirectorio proporcionará acceso a todos los datos requeridos haciendo uso de varios tipos de repositorios como pueden ser el propio directorio corporativo y web services dedicados (por ejemplo para suministrar datos de matrícula de alumnos, plan docente, asignaturas, grupos…). Conseguimos así que los programadores de aplicaciones accedan mediante un único protocolo (LDAP en este caso) a un único lugar de donde pueden obtener todos los datos que deseen.

El sistema SSO da un acceso único y centralizado a todas las aplicaciones web de la universidad

2.3. Autenticación única

La siguiente necesidad que aparece es la de dar acceso único y centralizado a todas las aplicaciones web de la universidad a través de un sistema Single Sign on (SSO) que permita exactamente esto.

El sistema SSO por el que hemos apostado es OpenSSO Express Build8. Las opciones que hemos utilizado en la universidad para realizar la integración de las distintas aplicaciones con este sistema se han basado en:

- el uso de agentes específicos que "protegen" la ruta de la aplicación a integrar
- a través de un proxy de agente (para paliar la falta de agentes para distintas versiones de software o infraestructuras)
- o directamente a través de servicios web SOAP y REST que permiten una integración cómoda por parte de los programadores de aplicaciones.

Esta tarea no es fácil de implementar puesto que requiere un tiempo de adaptación y adecuación de las distintas aplicaciones a esta nueva forma de trabajo. Aunque entendemos que esto es un proceso que lleva un tiempo, estamos ya integrando aplicaciones con este nuevo sistema, tanto a través de agentes como a través de web services.

El usuario autenticado correctamente en la federación accede a cualquier servicio federado

2.4. Federación de identidades

El último y más avanzado aspecto en lo que a identidad se refiere es justamente el que tratamos en este punto, la federación de identidades. Nos planteamos la posibilidad de federar servicios de manera que un usuario que se ha autenticado correctamente en la federación, pueda acceder a cualquier servicio federado de la misma, sea o no un servicio suministrado por la institución de origen.

En la actualidad nos encontramos realizando un proyecto de federación de identidades a nivel andaluz (CONFIA) en el que se están integrando todos los servicios de docencia virtual de las distintas universidades (LMS).

Con este proyecto conseguiremos que cualquier usuario autenticado en una institución origen perteneciente a la federación pueda hacer uso de los distintos servicios de enseñanza virtual del resto de universidades federadas sin necesidad de duplicar datos de usuarios de unas universidades en otras.

El proyecto de federación está basado en SimpleSAMLphp tanto para los proveedores de identidad como

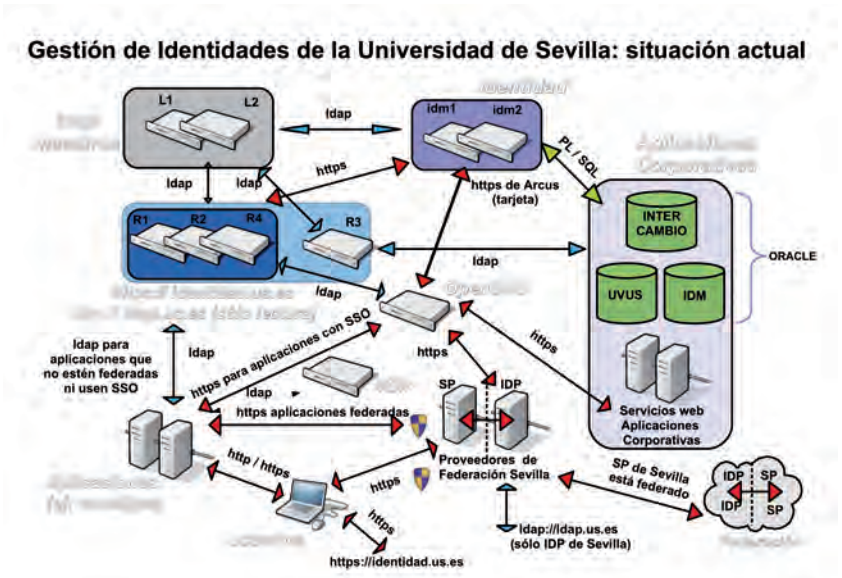
para los proveedores de servicios, aunque hay instituciones federadas utilizando otros sistemas de federación como puede ser Shibboleth. En cada universidad se despliega un proveedor de identidad encargado de validar al usuario en su propia universidad y todos los proveedores de servicios que se estimen oportunos.

Así, cuando un usuario accede a una URL perteneciente a la federación, habrá un mecanismo que le solicitará una institución de origen que pueda validarle, permitiendo posteriormente la redirección del usuario al proveedor de servicios al que se conectó inicialmente, si la autenticación se ha realizado correctamente.

2.5. ¿Y cómo conseguimos todo esto?

En la figura que se presenta a continuación se puede ver toda la infraestructura involucrada en la Universidad de Sevilla para gestionar la identidad de una manera eficiente y única.

FIGURA 1. GESTIÓN DE IDENTIDADES DE LA UNIVERSIDAD DE SEVILLA: SITUACIÓN ACTUAL



El proyecto de federación se basa en Simple SAMLphp

Cada universidad tiene un proveedor de identidad que valida al usuario

Carmen López Herrera
(carmen@us.es)
Universidad de Sevilla