Arquitectura para la Integrabilidad de Servicios sobre el protocolo HTTP basado en PKI[1]

Claudio Vaucheret, Jorge Sznek
Facultad de Informática, Universidad Nacional del Comahue
Buenos Aires 1400, Neuquén
{vaucheret,jsznek}@gmail.com

Resumen: Los modelos de e-government, soportados por la tecnología de Internet, tienen la necesidad de permitir el diálogo fluido entre los sistemas de los organismos gubernamentales, quienes proveen servicios a los ciudadanos. Tantos los sistemas como los organismos son de naturaleza heterogénea, por lo que se hace necesario definir un ambiente donde esa heterogeneidad pueda ser transparente. El ambiente propuesto es una Arquitectura para permitir la Integrabilidad, la que incluye en su infraestructura los aspectos de autenticación, autorización y auditoria para garantizar la confiabilidad de las transacciones y los servicios provistos.

1. Introducción

Este trabajo ha sido desarrollado como parte del proyecto FONSOFT-ANR 2008 - Proyecto N° NA 153/08 de THINKNET S.A., Titulado: "**Integrabilidad: Componentes de software basados en un Modelo Conceptual para lograr la macro integración informática de las organizaciones que operan interrelacionadamente**".

La Facultad de Informática (en ese momento, Departamento de Ciencias de la Computación) de la Universidad Nacional del Comahue fue contratada para participar en el desarrollo de los protocolos de seguridad entre los actores del Modelo de Integrabilidad que utilice como esquema de autorizaciones un modelo ABAC [15] de múltiples niveles denominado MABAC.

A partir de allí, se plantearon dos objetivos principales. En primer lugar la definición del protocolo de seguridad apoyándose principalmente en la infraestructura de clave pública (PKI - Public Key Infrastructure). Para ello se definieron desde cero los requerimientos de seguridad necesarios en la integración de los actores del modelo de Integrabilidad, buscando lograr el protocolo más adecuado sin tener en cuenta soluciones preexistentes desarrolladas para otros escenarios.

En segundo lugar se buscó el planteo de líneas de investigación fuera del alcance de este proyecto, que surgieron de las particularidades de este modelo frente a otros existentes.

En este documento se describen la arquitectura y los componentes participantes en una organización de

sistemas integrables que, aún estando basados en distintas tecnologías y plataformas, pueden colaborar, interactuar y proveer un servicio amplio contando con los recursos de todos los componentes.

Su aplicación es general para todo sistema basado en la tecnología de Internet y web-services, pero donde más impacto tiene es en los modelos de e-government [5] donde los actos administrativos y trámites deben tener su representación informática. En estos sistemas, por ejemplo, la identificación de un usuario, la firma de un documento, la autorización para obtener una información o el registro de operaciones debe tener su contrapartida informática con la misma solidez legal y operativa como lo tiene la manera tradicional.

Los modelos de e-government utilizan como plataforma la tecnología de Internet para intercambiar información, proveer servicios y transacciones entre los ciudadanos, las empresas y las reparticiones gubernamentales. Sin embargo el grado de heterogeneidad de los sistemas de los organismos públicos, desarrollados sobre distintas tecnologías y plataformas, requiere de una métrica especial del grado de integrabilidad [4] y de una arquitectura y organización que permita la interoperabilidad de todos los sistemas.

Esta propuesta de arquitectura descansa en la utilización de los estándares ya consolidados para la integración de los sistemas (HTTPS, XML, SOAP [13], UDDI [11], WSDL [12], LDAP [14]), y asume la disponibilidad de una infraestructura de claves públicas (PKI).

En la sección siguiente se presentarán los componentes del sistema o actores, en la sección 3 las interacciones entre los mismos y en la sección 4 se describe el protocolo de comunicación entre los actores. Luego en la sección 5 se describen las propiedades fundamentales de este enfoque y las diferencias con otros modelos. Por último en la sección 6 se plantean lineas de investigación futura para mejorar y complementar este diseño.

Es de resaltar que el modelo de Integrabilidad aquí descripto se proyecta aplicar en el ámbito de gobierno de la Provincia del Neuquén.

2. Actores

Los actores que intervienen son Sistemas Cliente, Fuentes Auténticas y los Coordinadores. Los Sistemas Clientes solicitan servicios que son provistos por las Fuentes Autenticas, interactuando a través de un Coordinador que realiza las funciones de autorización y registro de las operaciones para facilitar la posterior auditoria. Toda transacción comienza con un pedido de un servicio por un Sistema Cliente al Coordinador; el Coordinador deriva el pedido a la Fuente Autentica que provee el servicio y, tras la respuesta de esta última, el Coordinador la devuelve al Sistema Cliente.

Los componentes del sistema se pueden apreciar en la figura 1. Toda comunicación entre ellos se realiza sobre el protocolo HTTP.

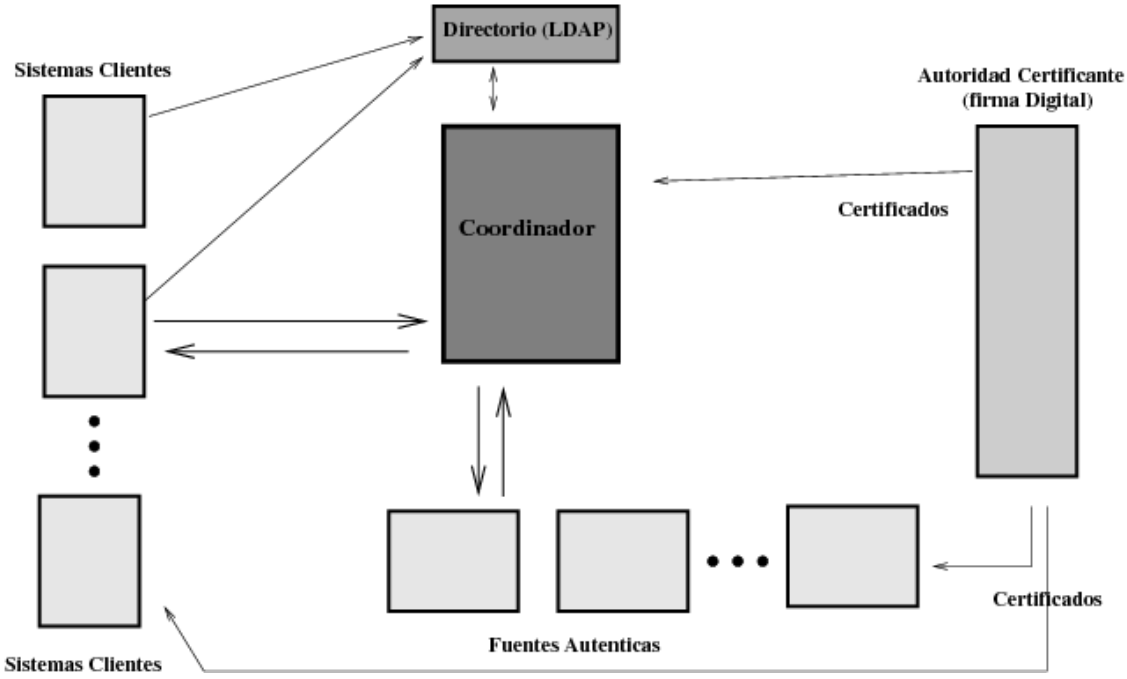

Figura 1: Esquema General

Sistema Cliente: es toda entidad que solicita un servicio o una información a alguna Fuente Auténtica. Un sistema cliente puede ser un sistema accedido por ciudadanos que realizan un trámite o puede ser un sistema de un organismo del estado que necesita un servicio de otro sistema de otro organismo. De esta manera un sistema puede ser fuente autentica y a su vez sistema cliente en otra operación. Un tramite puede involucrar la participación de varias fuentes autenticas.

Fuente Auténtica: es una entidad que provee servicios. Una Fuente Auténtica recibe a través del Coordinador pedidos firmados de los Sistemas Clientes; atiende estos requerimientos y los devuelve al Coordinador quien se encargará de entregar la respuesta a los Sistemas Clientes solicitantes. Una Fuente Auténtica entrega la información firmada y, eventualmente, cifrada.

Coordinador: es el mediador entre los Sistemas Cliente y las Fuentes Auténticas. Tanto los Sistemas Cliente como las Fuentes Auténticas son usuarios del sistema Coordinador. El Coordinador se encarga de autorizar las operaciones que los clientes solicitan a los proveedores mediante un sistema de control de acceso que contiene las reglas dinámicas que dictaminan a cuál información de cuál fuente auténtica puede acceder cuál sistema cliente [9,15,1,7,8]. Además, mantiene un registro de las operaciones realizadas a los efectos de facilitar auditorias posteriores. Vale la pena destacar que el Coordinador debe estar enmarcado en la propia legislación de la jurisdicción donde pertenece.

Además de los actores mencionados más arriba, existen otros que son necesarios para garantizar la correcta operación del modelo, que se explican a continuación:

Autenticador: Permite el control de acceso al sistema. Se implementa con LDAP y su función es permitir el ingreso con password al sistema. Es requisito para un cliente estar dado de alta en el Directorio. El autenticador además entrega tickets para que los clientes puedan solicitar servicios a fuentes auténticas a través del coordinador.

Autoridad de Certificación: Provee de certificado digital a todos los integrantes del sistema, lo que permite la posibilidad de firmar pedidos y respuestas, dando autenticidad a las operaciones. También provee el medio de cifrar (encrypt) elementos de las transacciones asegurando la confidencialidad de las operaciones que la legislación vigente requiera. La solicitud de un cliente o la respuesta de una fuente autentica puede contener datos que sólo pueden ser leídos por el destinatario y no deben ser accesibles para el coordinador.

3. Interacciones

En los trámites y operaciones en un contexto tradicional, no informático, se distinguen las instancias de autenticación, autorización, certificación y confidencialidad. En el contexto informático estas instancias muchas veces se confunden o se superponen. Por ejemplo, es necesario distinguir entre los trámites que requieren que un ciudadano se identifique presentando un DNI y llene una solicitud, de aquellos trámites que requieren que a su vez firme dicha solicitud y también de aquellos donde la solicitud deba ser enviada firmada por correo sin la presencia del ciudadano.

En combinación con esos casos un trámite puede requerir que la información sea presentada en sobre cerrado, como lo pueden ser propuestas en una licitación. Es decir que los componentes de autenticación, autorización, certificación y confidencialidad son dimensiones independientes y deben tener esa propiedad también con el soporte informático.

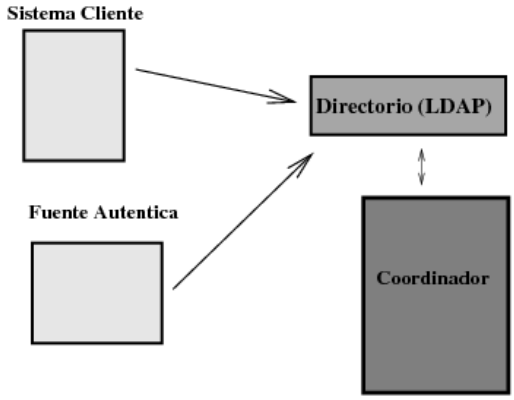

Figura 2: Autenticación

Autenticación: La autenticación es el proceso donde se verifica la identidad a un usuario del sistema (figura 2). Para que un sistema cliente pueda solicitar un servicio a una fuente auténtica, tanto él como la fuente auténtica deben estar dados de alta en el directorio LDAP.

Para darse de alta en el directorio LDAP una entidad debe presentar, junto a toda la documentación y requisitos que requiere el autenticador, el correspondiente certificado digital emitido por la autoridad de certificación

El directorio LDAP [14] funciona como un agente que certifica la identidad digital de un usuario de la misma manera que el registro nacional de personas con el DNI certifica la identidad de un ciudadano.

Autorización: El coordinador recibe los requerimientos de los clientes y transmite estos requerimientos a las fuentes auténticas realizando las funciones de autorización. Es decir el coordinador vincula los pedidos de los clientes a las fuentes auténticas sólo para aquellos clientes que están autorizados a operar con esas fuentes auténticas (figura 3).

La autorización es la principal función del coordinador. La otra es la de registrar las operaciones para funciones de auditoría. En este caso el mismo coordinador puede funcionar como fuente auténtica para el servicio de proveer información de auditoría.

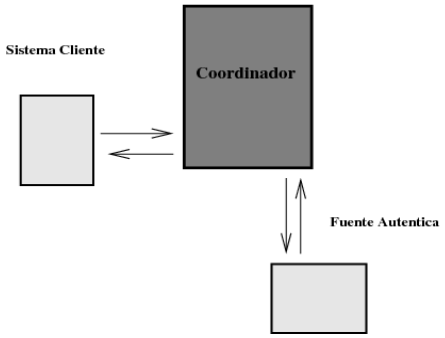

Figura 3: Autorización

Firma y Confidencialidad: Independientemente de la autenticación, algunos pedidos pueden requerir por motivos legales la firma digital de los mensajes. A su vez, partes de los mensajes, tanto de los pedidos como de las respuestas, pueden ser confidenciales y no visibles para la auditoría, por lo que deben cifrarse.

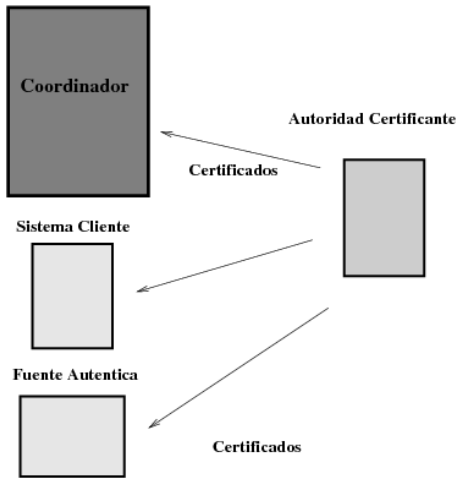

Figura 4: Confidencialidad

Todos los actores intervinientes en el sistema deben obtener de la Autoridad de Certificación su clave pública y generar su clave privada para las funciones de firma digital y cifrado. De la misma manera que un cliente de un banco registra su firma para operar en una cuenta, un usuario del sistema debe obtener de la Autoridad de Certificación su Certificado Digital que asocia su identidad con su clave pública (figura 4).

Un sistema cliente que solicita una información a una fuente auténtica debe conocer la clave pública de la misma, tanto para enviar en forma cifrada datos confidenciales del pedido como para verificar la firma digital que certifica la información entregada por la fuente auténtica. Para ello recibe del coordinador, cuando solicita un servicio, el correspondiente certificado con la información de la clave pública de la fuente auténtica.

Auditoría: El coordinador debe registrar con el fin de la auditoría sólo cuatro datos: fecha, solicitante, fuente auténtica y objeto de la solicitud. El resto de la información de la solicitud y la respuesta de la fuente auténtica puede estar cifrado y sólo visible para la fuente auténtica y el solicitante respectivamente.

4. Protocolo de Comunicación entre los Actores

La satisfacción de un servicio solicitado por un sistema cliente puede involucrar la participación de varias fuentes auténticas, las cuales a su vez pueden tomar el rol de sistemas clientes de otras fuentes auténticas. Por ejemplo, una municipalidad puede necesitar para completar un trámite el libre deuda del cliente otorgado por la compañia de agua. Por lo tanto para poder satisfacer el pedido del cliente debe ingresar como cliente al sistema y solicitar a la fuente auténtica compañía de agua el servicio de otorgar el libre deuda de un determinado cliente.

A continuación se describen los pasos sucesivos para resolver una petición atómica de un cliente a una fuente auténtica.

El CLIENTE accede al sistema y se identifica con su usuario y password. El AUTENTICADOR verifica la validez del usuario y lo autentica consultando al directorio LDAP. Luego el AUTENTICADOR le devuelve al CLIENTE un ticket que posee un tiempo de validez (figura 5).

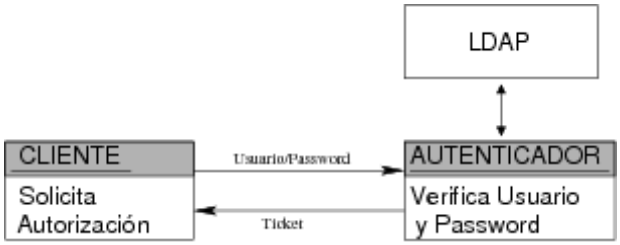

Figura 5: Autenticación

El CLIENTE se conecta con el COORDINADOR con el ticket para ver que servicios tiene disponibles. El COORDINADOR verifica con el AUTENTICADOR la validez y vigencia del ticket y luego le muestra al CLIENTE los servicios disponibles a los que puede acceder (figura 6).

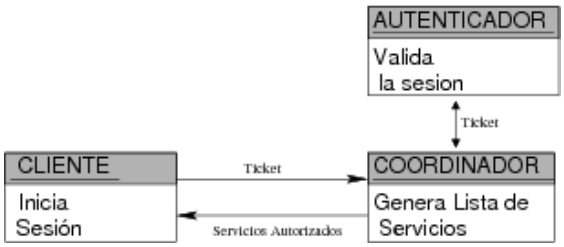

Figura 6: Autorización

Entonces el CLIENTE solicita un servicio. Para ello recibe del COORDINADOR el certificado digital de la FUENTE AUTENTICA que va a proveer el servicio. El CLIENTE firma con su clave privada la solicitud y cifra el pedido con la clave pública de la FUENTE AUTENTICA obtenida del Certificado. Luego, se lo envía al COORDINADOR(figura 7), quien chequea que el CLIENTE esté autorizado a acceder a la FUENTE AUTENTICA para obtener lo que requiere (Control de Acceso).

El COORDINADOR a su vez solicita un ticket al AUTENTICADOR para acceder a la FUENTE AUTENTICA a la que solicitó acceso el CLIENTE. El AUTENTICADOR devuelve el ticket al COORDINADOR, y éste solicita servicio del CLIENTE a la FA (figura 8). La FA entonces, chequea ante el AUTENTICADOR la validez y vigencia del ticket presentado por el COORDINADOR.

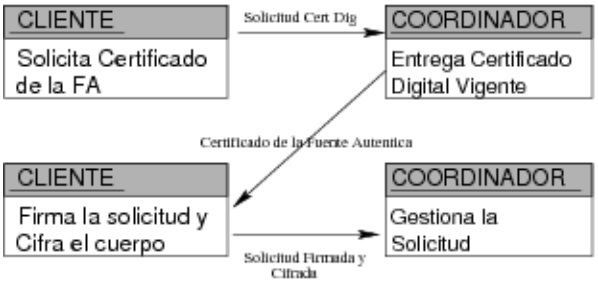

Figura 7: Solicitud de Servicio

La FUENTE AUTENTICA toma la solicitud, la descifra con su clave pública y chequea la validez de la firma del CLIENTE. Luego resuelve el servicio solicitado, lo firma con su clave privada y lo cifra con la clave pública del CLIENTE (figura 9).

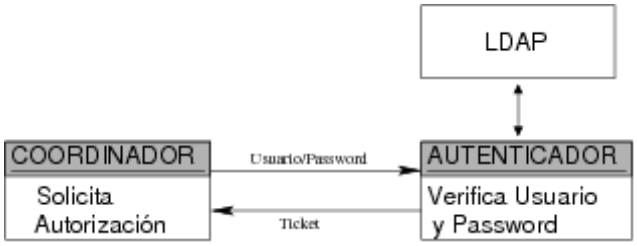

Figura 8: Autenticación del Coordinador

Despues, La FA le entrega al COORDINADOR el resultado del servicio solicitado, éste hace registro de la transacción (solo fecha, cliente que solicita, FA requerida, objeto solicitado).

El COORDINADOR le entrega al CLIENTE el resultado, éste lo descifra con su clave privada y chequea la firma de la FA. Por último usa los datos. En la figura 10 se puede observar el flujo completo de la secuencia de operaciones.

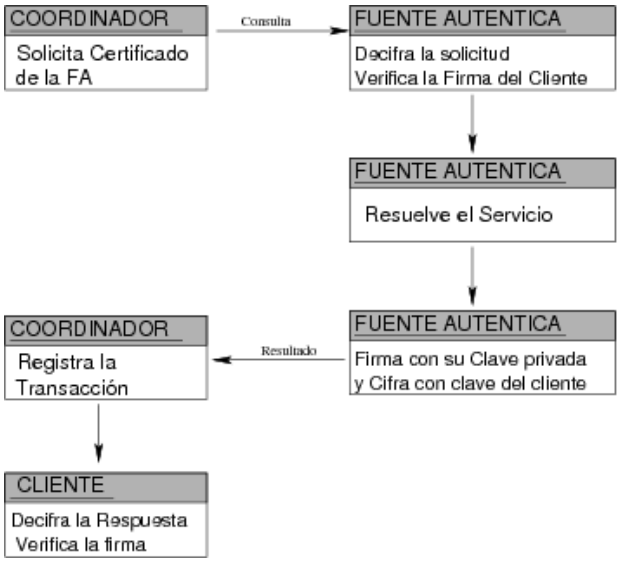

Figura 9: Resolución de Servicio

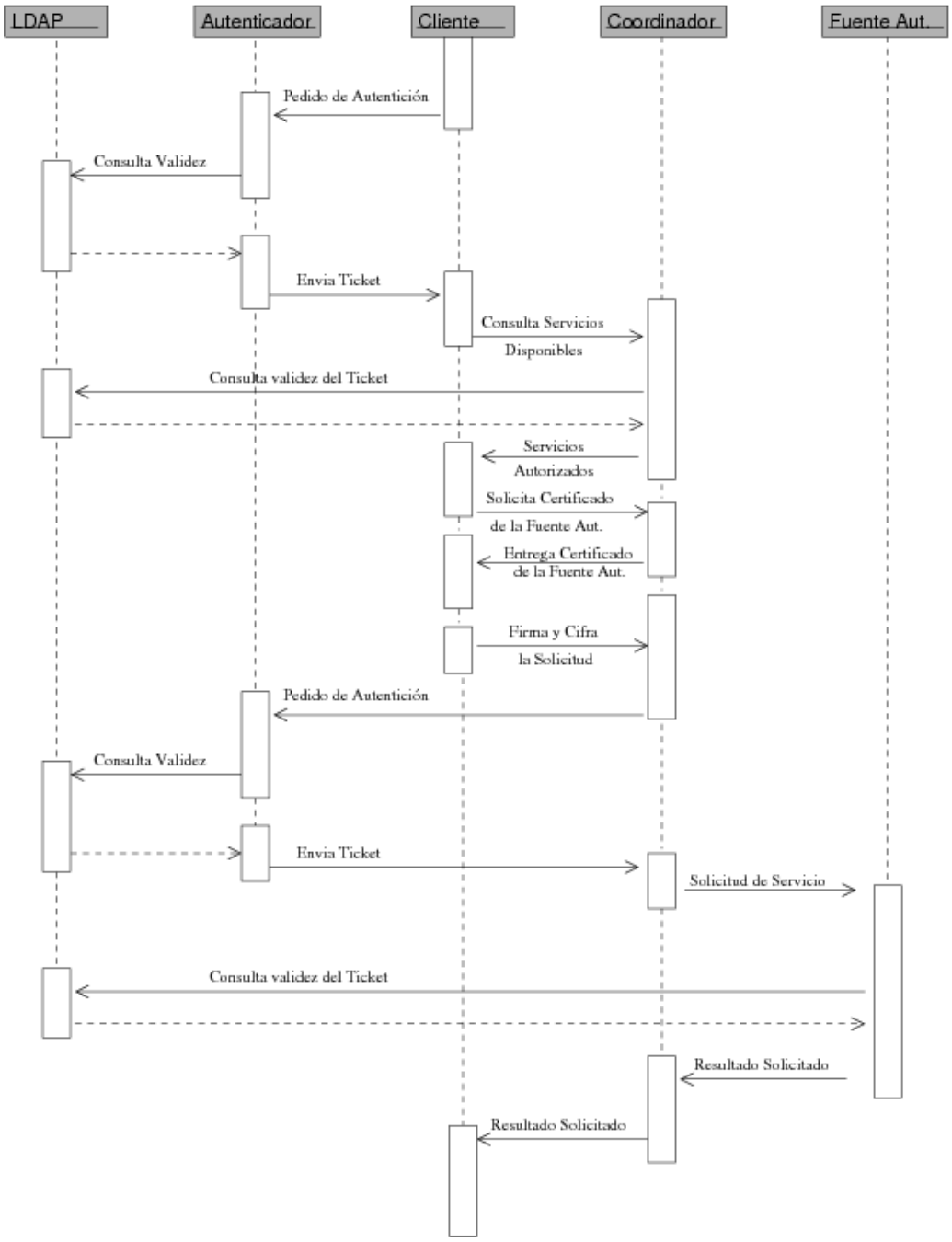

Figura 10: Secuencia de Operaciones Completa

5. Propiedades

Hemos presentado el diseño de una arquitectura para la integrabilidad de una red de sistemas que colaboran para producir servicios. Las líneas directrices de la propuesta buscan maximizar al mismo tiempo tanto la seguridad como la flexibilidad.

Después de considerar varias alternativas, incluídos protocolos de seguridad existentes, se ha considerado un diseño específico para los actores intervinientes que cumpliera con los requerimientos de seguridad impuestos y fuera el mas adecuado por su flexibilidad para el desenvolvimiento de los actores participantes en el sistema de

integrabilidad.

Algunas de las propiedades que lo distinguen de otros modelos son las siguientes:

Distinción entre Autenticación y Autorización: Los procedimientos de Autenticación y Autorización están conceptual y modularmente separados. La implementación de la Autenticación puede incluir, sin afectar la flexibilidad, la integración de mayor seguridad por medio de múltiples factores, por ejemplo, password + celular, password + tarjeta ID, etc.

El proceso de Autorización a su vez puede desarrollarse descansando en complejos modelos de control de acceso. El carácter distribuido del modelo y la posibilidad de asignación dinámica de autorizaciones descansa en la implementación de un modelo de control de acceso basado en atributos de múltiples niveles (MABAC) [2] . Al contrario de por ejemplo los modelos RBAC [7] o ABAC [15], que están basados en un modelo centralizado y de asignación estática.

Protección de Datos: Con respecto a la confidencialidad de los datos se ha tomado la decisión de proteger el dato mismo más que proteger el canal de comunicación. Este enfoque es compartido por [10]. Siguiendo la analogía de qué conviene proteger: si el camino donde va a circular el rey o el carruaje mismo del rey, hemos elegido proteger al rey, es decir al dato mismo, y poder disponer como plataforma de comunicación la internet pública mediante el protocolo HTTP. Además, dentro del modelo se implementan directamente las operaciones de cifrado y firma digital en lugar de basarse en un protocolo SSL. Esto posibilita la necesaria flexibilidad para reproducir los distintos niveles de confidencialidad que se pueden encontrar en la administración pública. De esta manera, un mismo mensaje puede tener distintos sectores cifrados para distintos actores, implementando la idea de sobres lacrados dentro de otros sobres.

Por definición, puede decirse que Internet es un medio inseguro. La información que viaja por ese medio tan heterogéneo se encuentra expuesta al husmeo por parte de quien disponga de los medios técnicos y motivaciones como para hacerlo. Una manera de contrarrestar esta debilidad es la de aplicar técnicas esencialmente criptográficas para implementar canales seguros. El uso de SSL, IPSec, VPN y otras tecnologías son conducentes a lograr la securización de los vínculos lógicos por donde viaja la información. Sin embargo, si bien estas técnicas producen resultados satisfactorios, restringen su usabilidad a solo aquellos actores que posean los medios tecnológicos para implementarlas, con lo que se desdibuja el sentido de universabilidad que pretende imprimir el modelo de Integrabilidad discutido en este documento. Un modelo de e-government debería ser capaz de integrar técnicas de protección de datos sobre datos individuales que viajen por medios inseguros, junto con técnicas de implementación de canales seguros.

Autenticidad de los Datos: El modelo descripto exige la responsabilidad de los datos emitidos por parte de la Fuente Auténtica. Cada registro es firmado digitalmente por la Fuente Auténtica que lo emite. Cada Fuente Auténtica es dueña de la información de la que es responsable y la información que circula por el sistema global está ``certificada'' por las Fuentes Auténticas. Esto valoriza el dato, evita la duplicación de datos en bases concentradoras y lo mantiene actualizado dado que se trata de registros atómicos altamente distribuidos.

Transparencia: El registro de las operaciones con los datos no confidenciales, permite la total transparencia en el uso de los web services. Todo intercambio es auditable por todos los actores involucrados a partir de la existencia de los registros de las transacciones(logs).

De este modo, un ciudadano podrá saber como ha sido tratada su petición, cuáles fuentes auténticas intervinieron y cuándo. El grado de auditabilidad es grande, lo que aumenta el nivel de confianza del modelo más allá de plantear problemas adicionales relacionados con la naturaleza de los logs, por ejemplo, almacenamiento y conservación en el tiempo.

6. Futuras Líneas de Investigación

Como trabajo futuro se plantean considerar los siguientes items:

1. Utilizar la implementación de Kerberos o una extensión de la misma, al modelo de Integrabilidad en el otorgamiento de tickets.

   Kerberos es un protocolo que facilita el procedimiento de autenticación en servidores de parte de los clientes que requieren servicios a los mismos, basado en la existencia de un servidor central cuya misión es la de entregar tickets a los clientes para que éstos puedan presentarlos ante los servidores y obtener de parte de éstos los servicios solicitados. Tanto los clientes como los servidores deben confiar en Kerberos y toda la infraestructura debería estar ``kerberizada''. Existen diversas implementaciones y extensiones de Kerberos, las cuales podrán ser estudiadas para ver la adaptabilidad al trabajo expuesto a los efectos de mejorar el procedimiento de otorgamiento de tickets para acceso a fuentes auténticas.
2. Analizar modelos de log distribuidos en los servidores de FA o la generación de un actor de auditoría independiente

   Los logs se caracterizan por necesitar un tratamiento particular, debido a su tamaño y a la información sensible que registran. En un modelo de interacción ciudadano-gobierno, deben quedar registros de todas las transacciones a los efectos de su auditabilidad. Este requerimiento hace que deba estudiarse con detenimiento la conveniencia o no de distribuir los logs entre los actores, o bien de centralizar los mismos en algún servidor específico y debidamente protegido o bien adoptar una solución intermedia,

donde parte de los logs estén en un servidor dedicado y también que las Fuentes Autenticas posean su propio registro de las operaciones hechas sobre él.

En todos los casos deberá estudiarse la implementación de una nueva entidad en el modelo, que se ocupe exclusivamente de las tareas de auditoria, con lo cual funcionará como una FA con particulares requisitos de seguridad (sólo determinados actores podrán acceder a él).

3. Extender los servicios de autenticación con los estándares emergentes de Open ID.

   Las soluciones de autenticación son muy variadas en modelos donde se presentan interacciones complejas entre múltiples clientes y múltiples servidores; en todos los casos son fuertemente dependientes de lo que se pretende obtener. Para arquitecturas donde sea necesario establecer la identidad de un cliente que pretenda el acceso a un recurso de un servidor deben implementarse esquemas basados en el concepto de Single Sign-On.

   Una propuesta interesante es el estándar Open Id [3], que es un sistema de identificación digital descentralizado, con el que un usuario puede identificarse en una página web a través de una URL y puede ser verificado por cualquier servidor que soporte el protocolo. Para su implementación debe considerarse la existencia de un proveedor de identidad, que es un servidor que se ocupa de verificar OpenId; cualquier cliente con su identidad OpenId la confirmará en este servidor de identidad.

   A diferencia de arquitecturas Single Sign-On, OpenId no especifica el mecanismo de autenticación. Por lo tanto, la seguridad de una conexión OpenId depende de la confianza que tenga el cliente OpenID en el proveedor de identidad.

4. Analizar escenarios de interacción complejos con 4 niveles, relaciones Internacionales, gobierno Nacional, gobiernos provinciales y municipios.

   Es esperable que la evolución del modelo de Integrabilidad, en la medida que vaya siendo adoptado por otros niveles gubernamentales, se plantee la necesidad de interactuar entre esos niveles. Un primer paso será la firma de convenios apropiados, los que deben complementarse con decisiones técnicas y tecnológicas que faciliten esa interacción.

   La utilización de certificados digitales con autoridades de certificación nacionales, la distribución de los logs, la adopción de protocolos de identificación y autenticación, son algunos de los desafíos que se deben encarar. Para una interacción con alcance extranacional, los aspectos a contemplar son similares, con el agregado que deberá plantearse un modelo de PKI que contemple a los países involucrados. Debe tenerse en cuenta que una cuestión de fondo es la referida a la legislación, la que

debería tender a ser uniforme para permitir las interacciones.

El que implementa la legislación de una jurisdicción particular en el modelo es el coordinador por medio de su política de control de acceso. Por tal razón puede haber coordinadores de ámbito municipal, provincial, nacional e internacional.

5. Estudiar mecanismos de seguridad proactiva y/o basados en Inteligencia Artificial ante ataques que eviten caer en denegación de servicios entre otros.

   En este tópico, habrá que estudiar como integrar herramientas de detección de intrusos (IDS) que sean capaces de ``aprender'' sobre las técnicas de ataques posibles a la infraestructura planteada. Podrían estudiarse por ejemplo las técnicas de ataque por medio de Honeynets[6].

Bibliografía

[1] D. Ferraiolo and R Kuhn.
Role-based access control.
In *NIST-NSA National (USA) Computer Security Conference*, pages 554-563, 1992.

[2] G.Giorgetti and C. Vaucheret.
Mbac, multi level attribute based access control.
White Paper, THINKNET, march 2008.

[3] Openid. http://openid.net/.

[4] K. Orr and G. Giorgetti.
Integrabilidad.
White Paper, Thinknet SA, December 2007.

[5] Shailendra C. Jain Palvia and Sushil S. Sharma.
E-government and e-governance: Definitions/domain framework and status around the world.
http://www.iceg.net/2007/books/1/1_369.pdf, 2007.
ICEG.

[6] The Honeynet Project.
Know Your Enemy.
Addison Wesley, 2002.

[7] R Sandhu.
Role activation hierachies.
In *Third ACM Workshop on Role-Based Access Control*, pages 33-40, Fairfax, Oct 1998.

[8] R. Sandhu, D. Ferraiolo, and R. Kuhn.
The nist model for role-based access control: Towards a unified standard.
In *Fifth ACM Workshop on Role-Based Access Control*, pages 47-63, Berlin, July 2000.

[9] Ravi S. Sandhu.
Lattice-based access control models.
IEEE Computer, 26(11):9-19, 1993.

[10] Uuno Vallner.
The estonian it interoperability framework.
Baltic IT & T Review, 41(2):42-47, 2006.

[11] W3schools.
Universal description, discovery and integration (uddi), tutorial.
http://www.w3schools.com/WSDL/wsdl_uddi.asp

[12] W3schools.
Web services description language (wsdl), tutorial.
http://www.w3schools.com/WSDL/default.asp.

[13] W3shools.
Simple object access protocol (soap), tutorial.
http://www.w3schools.com/soap/default.asp.

[14] M. Wahl, T. Howes, and S. Kille.
Lightweight directory access protocol (v3),.
www.ietf.org/rfcrfc2251, Dec. 1997.
[15] Eric Yuan and Jin Tong.
Attributed based access control (abac) for web services.
In *ICWS*, pages 561-569. IEEE Computer Society, 2005.

Notas al pie

... PKI[1]

Este artículo ha sido soportado por el Proyecto Nº NA 153/08 - FONSOFT-ANR 2008