

RCCI Vol. 3, No. 3-4 JULIO-DICIEMBRE, 2009 p. 13-18

Recibido: 3/4/2010

Una experiencia en integración de aplicaciones empresariales

An experience in business application integration

Yasser Azán Basallo^{1*}, Anay Díaz Estrada² y Salvador González Gómez³

¹ Universidad de las Ciencias Informáticas. Departamento de Seguridad Informática, Carretera a San Antonio de los Baños. km 2 ½, Torrens, Boyeros. Ciudad de La Habana. Cuba.

² Universidad de las Ciencias Informáticas. Departamento de Gestión Universitaria, Centro de Informatización Universitaria, Carretera a San Antonio de los Baños. km 2 ½, Torrens, Boyeros. Ciudad de La Habana. Cuba.

³ Universidad de las Ciencias Informáticas. Departamento de Telecomunicaciones, Carretera a San Antonio de los Baños. km 2 ½, Torrens, Boyeros. Ciudad de La Habana. Cuba.

Autor para la correspondencia: yazan@uci.cu

Resumen

El artículo describe el proceso de informatización que se desarrolla en la Contraloría General de la República de Cuba (CGRC), donde existe la necesidad de integrar varias herramientas informáticas creadas con diferentes tecnologías y evitar lo que se conoce como "islas de información". Se detallan los casos en que se necesita la integración y se identifican las técnicas de integración adecuadas a la situación presentada; se mencionan otras soluciones similares existentes en el mundo.

Palabras clave: Contraloría General de la República de Cuba, integración de sistemas informáticos.

Abstract

This paper describes the process of computerization that it was developed in the Comptroller general of the Republic of Cuba (CGRC), where there is a need to integrate various software tools created by different technologies and in order to avoid what is known as "islands of information". It details where the integration is needed and identifies integration techniques appropriate to the situation presented and it mentions similar solutions in the world.

Keywords: General Comptroller of the Republic of Cuba, integration of information systems.

Introducción

En la actualidad muchas empresas no tienen un único sistema informático capaz de dar solución a todos los procesos que en estas se encuentran, sino más bien que en su infraestructura tecnológica, debido a muchos factores, existe una variedad de estas, respondiendo a diversas plataformas de desarrollo, distintos lenguajes y arquitectura. La heterogeneidad de tecnologías, crean las llamadas 'islas de información' donde la comunicación entre estos sistemas informáticos es insuficiente. Esta situación provoca la fragmentación de los procesos del negocio en las empresas y que la información estratégica de la entidad, esté dividida por áreas, departamentos o direcciones y no se tenga una mejor colaboración mutua y coordinación entre las actividades o procesos que se pueden llevar a cabo con más agilidad si estos sistemas informáticos se entendieran entre sí. Por eso, se dice que no hay una integración única del negocio, las entidades con esta característica en su negocio no son capaces de responder como un todo único o se requiere de un mayor esfuerzo para lograrlo.

Alejandro Lafarga, Director general de Software AG España y Portugal comenta sobre Serge Novaretti, un experto de la IDABC, la agencia de la Comisión Europea encargada de promover la sociedad de la información, afirmaba en Gijón durante un encuentro organizado por el consorcio W3C que el mayor desafío actual para la administración electrónica en Europa no eran ni el desarrollo de los servicios ni el mayor o menor grado de aceptación de estos servicios por parte de los ciudadanos, sino la necesidad de una mayor interoperabilidad entre las Administraciones europeas. Alejandro Lafarga explica además que la UE (Unión Europea) ha creado el eGov-Interop Observatory, un observatorio específico sobre interoperabilidad, para superar esas 'islas de información', una buena metáfora que ilustra la importancia de la comunicación y la compatibilidad entre los sistemas e infraestructuras informáticas. El motivo es que la mejora en la interoperabilidad permitirá desplegar servicios que aumenten la competitividad de las empresas y mejoren la vida de los ciudadanos en general.

Desarrollo

En la informatización que se está llevando en la CGRC, se ha identificado este tipo de caso donde existe esta versatilidad de tecnologías. Por eso se está realizando un proyecto informático donde además de automatizar los procesos, se ha requerido la inte-

gración entre los sistemas que conforman la solución. Se está hablando de un sistema realizado en lenguaje PHP por la Universidad de las Ciencias Informáticas¹ conocido como Sistema Informático de gestión de Auditoría y Control (SIGAC), para gestionar los procesos de la planificación y control de las auditorías y de otro sistema implementado en java para la gestión documental de todos los archivos que dentro de este órgano del estado se manejan, llamado Alfresco en su edición comunitaria. Pero estas dos aplicaciones con arquitectura cliente-servidor, necesitan comunicarse para complementarse la una de la otra por la misma necesidad de la lógica del negocio que se plantea dentro de esta entidad. Un requisito esencial es la posibilidad que en el momento que se apruebe un Plan Anual, se cree automáticamente el documento que contiene una orden de trabajo con datos iniciales en Alfresco y a partir de allí los jefes de grupos de auditores, empiezan su labor en la especialización que les corresponde. Ver figura 1.

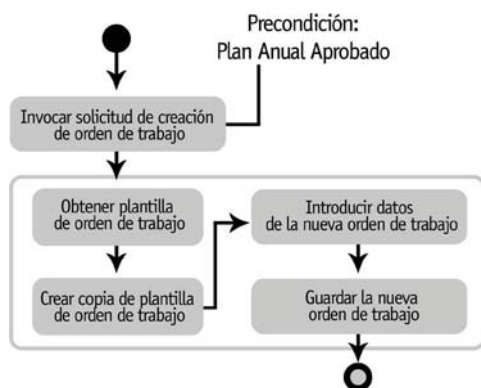


Figura 1. Proceso de crear orden de trabajo.

Luego el resultado final de la auditoría debe dar como reporte en el sistema SIGAC, pero debe ser sustraído del gestor documental, por lo que también fue otro punto de contacto entre ambas aplicaciones. Ver figura 2.

Ventajas de la integración de sistemas informáticos:

Creación de actividades de negocio coordinadas, manejadas por diferentes personas de modo que se pueda formar un proceso de negocio unificado. Permite derribar las barreras entre las unidades de la organización.

Reduce la duplicación de esfuerzos. (Pérez *et al*, 2009)

Y reduce los costos en desarrollo y mantenimiento de tecnologías.

¹ UCI: Universidad de Ciencias Informáticas. Habana, Cuba

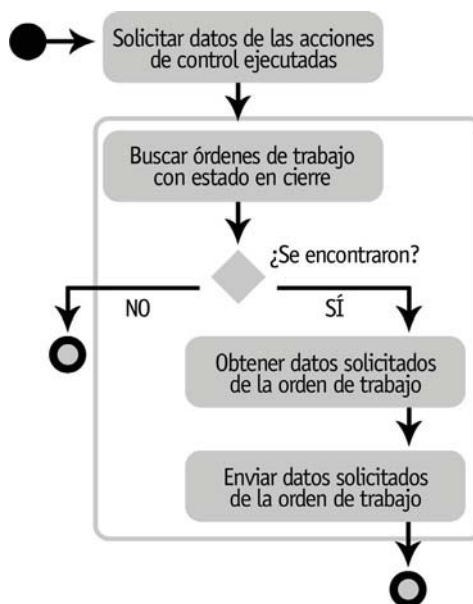


Figura 2. Proceso obtener controles de cumplimiento.

¿Cómo se logra integrar varias aplicaciones empresariales a partir de la problemática planteada anteriormente?

Materiales y Métodos

Técnicas de Integración de datos

El esquema tradicional cliente/servidor permite que diferentes módulos de aplicación se comuniquen directamente sin una capa intermedia. El problema se presenta en sistemas complejos con componentes de diversos proveedores donde resulta poco flexible e inoperante la comunicación (Mendoza, 2009).

La Visión de Integración es la entrega de información:

1. Precisa
2. Consistente
3. Oportuna
4. Coherente

La integración se puede enfocar de varias formas diferentes dependiendo de la idea de "Integración" que se tenga. Fundamentalmente existen cuatro:

- Replicación de Datos.
- Extracción, Transformación y Carga de Datos o ETL4.
- 1. Integración de Información Empresarial o EII5.
- 2. Integración de Aplicaciones Empresariales o EAI.

Replicación de datos

La Replicación de Bases de Datos es una técnica de integración que se basa en la creación y mantenimiento de múltiples copias de una misma base de datos. En la mayoría de las implementaciones de Replicación, un servidor mantiene la copia primaria de la base de datos y servidores adicionales mantienen las copias esclavas de la misma. (White)

- Lo escrito en la Base de Datos se envía al servidor primario y luego es replicado por las copias esclavas.
- Las lecturas de las Bases de Datos están divididas entre todos los servidores, lo cual trae una gran ventaja debido al uso compartido de carga.

La transferencia se realiza de Base de Datos a Base de Datos (este es el estilo más antiguo de integración), lo cual es una desventaja para la resolución del problema a resolver dado la heterogeneidad de las fuentes.

Extracción, Transformación y Carga de Datos (ETL)

La técnica Extracción Transformación y Carga de Datos (Extract, Transform and Load), como su nombre lo indica extrae información de un sistema fuente, transforma esos datos para satisfacer los requerimientos del negocio y carga el resultado en el sistema destino. Tanto la fuente como el destino son generalmente Bases de Datos y archivos.

La transformación puede implicar la reestructuración y reconciliación del registro de datos, limpieza del contenido de datos (es decir, revisados por si existen discrepancias y eliminación de datos obviamente falsos), y/o agregación del contenido de datos. Esto sucede mediante una serie de procedimientos especiales que permiten obtener un formato unificado común y mejorado. Sólo después de la revisión y unificación de los datos estos son cargados. Esta técnica se encarga de la integración de datos, no de aplicaciones, y obtiene los datos directamente de las Bases de Datos. De acuerdo a las características de esta técnica se concluyó que no es la más adecuada para la solución que se necesita. (Kimbal, 2009)

Integración de Información Empresarial (EII)

La Integración de Información Empresarial o EII es otra de las técnicas de integración. Es un

mecanismo de transformación y acceso a datos transparente y optimizado para suministrar una única interfaz a lo largo de los datos de las organizaciones. Dicha interfaz permite acceder a los datos y el resultado de este método es un Sistema de Información Heterogéneo Distribuido, Virtualmente Integrado. Este tipo de solución consiste en crear un intermediario que contenga los directorios de la Base de Datos y que a su vez sirva de canal de consulta y representación de la información recuperada. La información es capturada en tiempo real lo que implica que las fuentes de datos tengan una estructura tecnológica sólida y bien establecida. EII protege a las aplicaciones de la complejidad de recuperar datos de múltiples localizaciones, donde los datos pueden diferir en semántica y formato, y emplear diferentes interfaces de datos. Teniendo en cuenta estos aspectos, para la integración de Datos a Tiempo Real la técnica EII constituye una buena alternativa, sin embargo no es factible para la integración de aplicaciones. (Morgenthal, 2006)

Integración de Aplicaciones Empresariales (EAI)

EAI es el proceso de integrar múltiples aplicaciones desarrolladas independientemente, que utilizan tecnología incompatible y que son gestionadas de forma independiente, permitiendo que se comuniquen e intercambien transacciones de negocio, mensajes, y datos entre sí. Uno de los principales objetivos de EAI es proporcionar acceso transparente a la amplia gama de aplicaciones que existen en una organización. Las características más importantes de esta tecnología es que se utiliza para la integración de Aplicaciones – a – Aplicaciones y proporciona un enfoque de integración orientado a proceso basado en mensajes XML. Es generalmente utilizada para el procesamiento de transacciones de negocio operacional en tiempo real. (Denodo Technology, 2009)

La utilización de un motor de integración EAI reduce los costes de desarrollo y mantenimiento de las unidades de Tecnologías:

- Los analistas del mercado afirman que las organizaciones gastan entre el 40% y 60% de los presupuestos de Tecnologías en integración destinando la mayor parte de ese coste en mantener los interfaces actuales. EAI reduce el coste de creación de nuevas interfaces en un 50% y puede reducir el coste de mantenimiento de interfaces en un 80%.

- Un estudio del 2007 de Forrester Research sobre ahorro de coste comenta que mediante una inversión en una arquitectura de integración y planificación del trabajo, las empresas pueden reducir el coste de proyectos EAI en 710.000\$ en cinco años. Esto puede generar una idea sobre el retorno de inversión que puede aportar una estrategia de arquitectura de integración frente a desarrollo de interfaces punto a punto.

Resultados

Se utilizó la última técnica que es la más se ajusta a la situación del ministerio. Aunque se podría trabajar con la técnica ETL por ejemplo, no es recomendable por trabajar directo con la base de datos y hacer esto hace una solución frágil a futuros cambios de la base de datos, como ha ocurrido con Alfresco, aplicación que ha tenido variaciones en su base de datos en el paso de una versión a otra. Por lo que es mejor trabajar a un nivel más alto con transparencia del nivel de datos.

En la actualidad se emplea mucho para la integración el motor EAI. En la Figura 3 se muestra la diferencia existente en la integración cuando se usa un motor EAI y cuando no. "Un motor de integración (EAI) realiza una interacción entre dos o más aplicaciones, permitiendo la comunicación entre ellas normalmente mediante el paso de mensajes, ya sean HL7 v.2.x, HL7 v.3, ASTM, mensajería no estándar..."

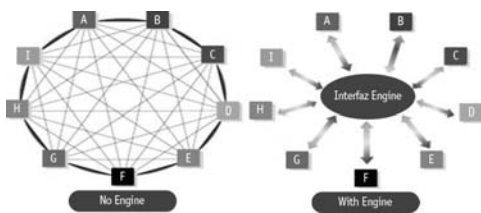


Figura 3. A la izquierda una integración sin motor EAI y la derecha con motor EAI.

Para llevar a cabo la integración, optamos por usar los servicios web que nos brindan la abstracción del lenguaje empleado y de la base de datos, manejando la información sin ningún impacto en los sistemas que intervienen. Alfresco en su arquitectura, ofrece una capa dedicada a ofertar los servicios web que se necesitan y poder interactuar con otras aplicaciones así basado en esta opción se creó la integración. Ver Figura 4.

No optamos por usar motores EAI como BizAgi o

Mirth que es de código abierto aunque este último está dedicado a la salud y sí crear uno propio para ganar en sencillez y porque además ya teníamos el grueso de la situación resuelta, que era contar con los servicios web, solo había que aprender a consumirlos.

Además de los servicios web, la empresa de mismo nombre al producto Alfresco, ofrece varios tipos de API para reducir la complejidad de la interacción con esta aplicación. Una en Javascript y la otra en PHP, ambas gratis para los desarrolladores. Como la necesidad de integración es a nivel de la lógica del negocio, el usuario nunca interviene en esta interacción entre los sistemas por tanto la API a nivel de cliente es innecesaria, por lo que se quedó para la solución la API de PHP.

En la Figura 4 se muestra cómo estaría distribuida la solución, contemplando la integración como mediador entre ambas aplicaciones.

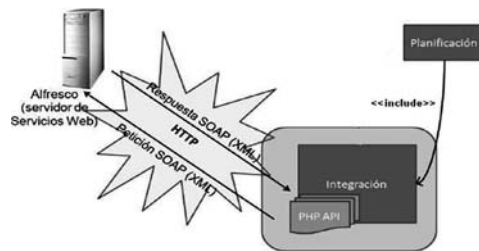


Figura 4. Arquitectura general de EAI en el sistema SIGAC.

Para crear la integración se pensó de varias formas. Una salida era hacerla independiente, como otra aplicación que hiciera función de mediadora de las demás o crearla como parte del módulo de la aplicación SIGAC. La segunda decisión fue la más viable ya que en todo caso la integración sería de otras aplicaciones con la aplicación SIGAC, así que más bien lo que hacía falta era que esta aplicación se entendiera con las demás y no todas con todas. Así de esa manera se aprovecha la arquitectura de esta, el framework que la soporta como programación base y la comunicación no tendría que ser por el protocolo HTTP y todo quedaría a un nivel interno.

En la Figura 5 se muestra como está distribuida la solución por componentes y la interacción existentes con el subsistema Alfresco a través de los servicios que esta implementa.

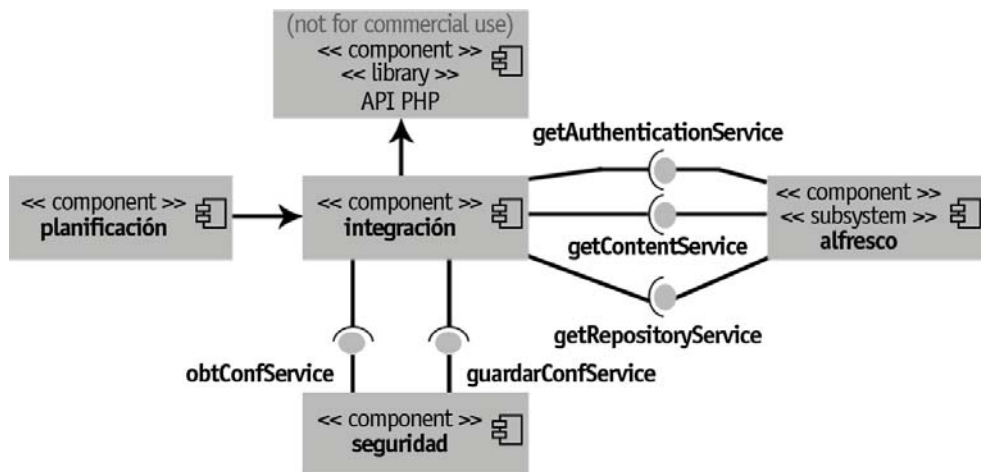


Figura 5. Vista de componentes de las partes involucradas en la integración y alguno de los servicios web de Alfresco.

Hay una tecnología que fue un factor importante para lograr la solución dada. Gracias a Lucene, que está soportada por Alfresco, se logró realizar consultas al sistema de gestión documental y que este devolviera no todo su contenido, sino aquellas partes necesarias, evitando todo un proceso engorroso de filtrado de la información superflua que y por consiguiente de tardanza excesiva análogo a ineficiencia absoluta de la aplicación.

Apache Lucene es una librería con todas las funciones de búsqueda de texto de alto rendimiento implementado en Java y apoyada por Apache Software Foundation. Además es un proyecto de código abierto por lo que no se incurre tampoco en un gasto monetario en su uso.

Conclusiones

Esta solución permitió integrar ambas aplicaciones y evitar la fragmentación del negocio automatizado y evitar dentro de la solución brindada la 'isla de información'. La Contraloría General de la República de Cuba puede tener una mejor visión de sí misma al terminar el proyecto de informatización. El costo en el desarrollo de esta tecnología es bastante bajo por la posibilidad de haber usado tecnologías gratuitas y la técnica EAI de integración de aplicaciones informáticas.

Referencias

Apache Software Foundation. Apache Lucene. [En línea] Apache. [Consultado el: 3 de agosto de 2009.] Disponible en: <http://lucene.apache.org/java/docs/>.

Denodo Technology. Denodo Technology. [En línea] [Consultado el: 3 de marzo de 2009] Disponible en: <http://www.denodo.com>.

Kimball, Ralph. The Data Warehouse ETL Toolkit.

Lafarga, Alejandro. 2007. Superar las islas de información en la empresa. Cincodias.com. [En línea] [Consultado el: 17 de septiembre de 2007]. Disponible en: http://www.cincodias.com/articulo/opinion/Superar-islas-informacionempresa/20070917cdscdiopi_5/cdsopi/.

Mendoza, Luis E. Estudio de Tecnologías Middleware para Sistemas Peer-to-Peer. Caracas: s.n.

Morgenthal, JP. Enterprise Information Integration: A Pragmatic Approach. s.l.: Bk&CD Rom edition. 2006.

Pérez, María, Mendoza, Luis E. y Carvajal, Yorka. 2009. Orientaciones para la selección de tecnologías DE. Caracas: Laboratorio de Investigación en Sistemas de Información (LISI), 2009.

Schmidt. 2000. Enabling Next-Generation Enterprises. 2000.

Shariff, Munwar. 2006. Alfresco Enterprise Content Management Implementation. Birmingham, Packt Publishing Ltd, 2006.

White, Colin. Data Integration: Using ETL, EAI, and EII Tools to Create an Integrated Enterprise.