

# Planteamiento para el desarrollo de un sistema que realice la migración de datos almacenados en Sybase a PostgreSQL

Ing. Géminis Lucía Mejía Layna<sup>1</sup>, MC Francisco Javier Gutiérrez Mata<sup>2</sup>,  
MTI Jorge Carranza Gómez<sup>3</sup>, Dr. Eduardo De La Cruz Gámez<sup>4</sup>

**Resumen** — Como respuesta a la problemática que enfrentan la mayoría de los Institutos del Tecnológico Nacional de México (TecNM), quienes utilizan la primera versión del Sistema Integral de Información (SII), el cual administra la base de datos a través del gestor Sybase, y la versión más reciente del SII está diseñada para trabajar con el gestor PostgreSQL, se plantea el desarrollo de un sistema que realice el proceso de migración, extrayendo los datos que originalmente se encuentran almacenados en Sybase para migrarlos a PostgreSQL. El artículo contiene los antecedentes conceptuales del planteamiento para el desarrollo de un sistema que realice la migración de la información de forma automática, la metodología propuesta para realizar la migración de los datos y un prototipo de interfaz gráfica para el sistema.

**Palabras clave** — Sybase, PostgreSQL, Base De Datos, Migración

## Introducción

La plataforma del Sistema Integral de Información, fue desarrollada para integrar y centralizar la información de las instituciones pertenecientes al TecNM. La primera versión fue creada para trabajar con el lenguaje PHP y el gestor de base de datos Sybase. Este administrador de base de datos de software libre, procesaba de manera óptima la cantidad de registros almacenados, además de que las consultas realizadas eran atendidas de forma inmediata. Con el potencial incremento en la cantidad de información que se generaba, la velocidad de respuesta a las consultas que se realizaban fue decreciendo, provocando que el SII se volviera lento, pues no era capaz de atender las peticiones efectuadas. Debido a estos problemas era necesario actualizar el motor de la base de datos. La empresa Sybase fue adquirida por la empresa SAP en el año 2010 por lo que su funcionamiento requería el pago de una licencia. En consecuencia, se desarrolló una nueva versión del SII ahora con el lenguaje PHP y el gestor de base de datos PostgreSQL, el cual, además de ser software libre, cuenta con las características necesarias para procesar grandes volúmenes de información.

Ahora surge otro problema, ¿Cómo migrar la base de datos desde Sybase a PostgreSQL? Solo algunos institutos han realizado la migración, pues cuentan con el personal calificado para llevar a cabo esta tarea. Pero en su mayoría, los institutos del TecNM continúan utilizando la primera versión de SII pues no tienen las herramientas necesarias para migrar sus datos.

## Marco teórico

### *Sistemas Gestores De Bases De Datos (SGBD)*

Se define como una colección de datos interrelacionados y un conjunto de programas para acceder a dichos datos. Permite almacenar y recuperar la información de una base de datos (BD) de manera práctica y eficiente (Universidad Politécnica de Puebla, s.f.).

Sus servicios principales son: Crear y definir una BD, Administrar los datos, Controlar el acceso a los datos, Mantener la integridad y consistencia, Controlar el acceso compartido, Recuperar y respaldar la información.

Los SGBD proporcionan lenguajes e interfaces específicas para cada usuario y sus lenguajes se clasifican en:

- Lenguaje de definición de datos
- Lenguaje de manipulación de datos
- Lenguajes de cuarta generación

<sup>1</sup> Ing. Géminis Lucía Mejía Layna, alumna de la maestría en sistemas computacionales del Instituto Tecnológico de Acapulco. [mejia.geminis86@gmail.com](mailto:mejia.geminis86@gmail.com)

<sup>2</sup> M.C. Francisco Javier Gutiérrez Mata, Profesor de la maestría en sistemas computacionales del Instituto Tecnológico de Acapulco. [fcomata84@hotmail.com](mailto:fcomata84@hotmail.com)

<sup>3</sup> M.T.I. Jorge Carranza Gómez, Profesor de la maestría en sistemas computacionales del Instituto Tecnológico de Acapulco. [jcarranzamx@hotmail.com](mailto:jcarranzamx@hotmail.com)

<sup>4</sup> Dr. Eduardo De La Cruz Gámez, Profesor de la maestría en sistemas computacionales del Instituto Tecnológico de Acapulco. [gameduardo@hotmail.com](mailto:gameduardo@hotmail.com)

### *Sybase*

Comenzó a mediados de los años ochenta como Sybase SQL Server. Durante varios años, Microsoft fue uno de sus distribuidores. En 1997, Sybase cambió el nombre de su producto a “ASE” (*Adaptive Server Enterprise*), el cual es un software de administración de bases de datos relacionales, versátil y de clase empresarial que es especialmente bueno para manejar cargas de trabajo OLTP (*OnLine Transaction Processing*). Hoy en día, Sybase es la empresa global de software más grande, exclusivamente enfocada en la gestión y movilización de información desde el centro de datos hasta el punto de acción. Sus soluciones abiertas y portables entregan información de manera segura, en cualquier momento y lugar, permitiendo a los clientes crear una ventaja de información. Entre sus principales características se encuentran: Optimizador de consultas completamente renovado y más inteligente, técnicas de particionamiento semántico de tablas que aumentan la velocidad de acceso a los datos, columnas cifradas para mayor seguridad de los datos, mejoras al lenguaje Transact-SQL, para mayor productividad, mejoras a los servicios de Java y XML en la base de datos, mejoras a los servicios para consumo y publicación de servicios Web y herramientas mejoradas para la administración y el monitoreo.

### *PostgreSQL*

Es un sistema de administración de bases de datos de propósito general y relacional de objetos, gratuito y de código abierto, desarrollado sobre POSTGRES 4.2 en el Departamento de Ciencias de la Computación Berkeley, de la Universidad de California. Fue diseñado para ejecutarse en plataformas similares a UNIX y a la vez ser portátil, de modo que pudiera ejecutarse en varias plataformas, como Mac OS X, Solaris y Windows. Fue el primer sistema de administración de bases de datos que implementó la función de control de concurrencia de múltiples versiones, incluso antes de Oracle. permite agregar funciones personalizadas desarrolladas usando diferentes lenguajes de programación como C / C ++, C# y Java. Está diseñado para ser extensible. Permite definir tipos de datos propios, tipos de índices, idiomas funcionales y desarrollar complementos personalizados. (POSTGRESQL TUTORIAL, 2012). Sus principales características se encuentra su compatibilidad con ACID, soporte completo para claves foráneas, uniones, vistas, disparadores y procedimientos almacenados (en varios lenguajes). Soporta almacenamiento de objetos binarios grandes, como imágenes, sonidos o vídeo. Cuenta con interfaces nativas de programación para C / C ++, Java, Net, Perl, Python, Ruby, Tcl, ODBC (Henríquez).

### *C#*

Ceballos lo describe como “un lenguaje orientado a objetos seguro y elegante que permite a los desarrolladores construir un amplio rango de aplicaciones seguras y robustas que se ejecutan sobre .NET Framework” (Ceballos, 2013). Cerezo lo define como un lenguaje orientado a objetos creado por Microsoft para su plataforma .NET, adecuando todas sus estructuras a las características y capacidades de dicha plataforma (Cerezo, 2007, pág. 2).

Algunas de las características no propiamente del lenguaje sino de la plataforma .NET pero que tienen una repercusión directa en el lenguaje son: sencillez, orientación a componentes, gestión automática de memoria, orientación a objetos (C# soporta todas las características propias del paradigma de programación orientada a objetos, encapsulación, herencia y polimorfismo), seguridad de tipos, instrucciones seguras, sistemas de tipos unificado, extensibilidad de tipos básicos, extensibilidad de operadores, extensibilidad de modificadores, versionable, eficiente y compatible (Gonzales, 2001).

### *Archivo XML*

Lenguaje de marcas extensible por si siglas en ingles *XML (Extensible Markup Language)*. Se trata de un metalenguaje (un lenguaje que se utiliza para decir algo acerca de otro) extensible de etiquetas. Boulanger lo menciona como “un archivo de texto especial compuesto por un determinado número de etiquetas en estructura de árbol, los elementos están vinculados entre ellos a través de las ramas que se crean” (Boulanger, 2015, pág. 21). permite jerarquizar, estructurar la información y describir los contenidos dentro del propio documento, así como la reutilización de partes del mismo.

### Metodología para la migración de Sybase a PostgreSQL

La migración de una base de datos, es un proceso que tiene un ciclo de vida, conformado por una serie de pasos que garantizan que la información será procesada de forma eficaz, conservando la integridad y calidad de los datos. Celis lo describe como una tarea compleja que no sólo supone transferir datos entre tipos de almacenaje y formatos de un servidor origen de base de datos a otro servidor destino, sino que también supone reescribir sentencias SQL, con la finalidad de adaptarlas al nuevo servidor de destino (Celis). De acuerdo con Fernández, para efectuar una migración de Bases de Datos, se debe realizar un estudio previo, para poder identificar cuáles son los mejores pasos a seguir para cada migración en particular (Fernández, 2011).

Considerando las definiciones descritas anteriormente, se propone la metodología mostrada en la figura 1, segmentado el proceso de migración de datos en cinco etapas.



Figura 1 Metodología para la migración de los datos de Sybase a PostgreSQL

En estas etapas o fases se describen en la tabla 1.

FASE	DESCRIPCIÓN
<b>ANÁLISIS</b>	Esta fase comprende el análisis de la estructura de la base de datos origen en Sybase, como están diseñadas las tablas y sus relaciones. Con base en este análisis, crear la base de datos destino (BDO), es decir la construcción del esqueleto en PostgreSQL, asegurando que los nombres de las tablas coincidan perfectamente con la base de datos origen (BDO).
<b>EXTRACCIÓN</b>	En esta fase se extraerán los datos almacenados en la BDO de Sybase generando un archivo XML que posteriormente será cargado a la memoria del sistema.
<b>PROCESAMIENTO</b>	En esta fase se aplicará un algoritmo de comparación, para determinar el tipo de datos al que pertenece cada campo. Al mismo tiempo el algoritmo hará una búsqueda en la BDD para ubicar la tabla y campo en el cual será insertada la información. Una vez determinadas las características anteriores, generará la consulta de inserción para realizar la migración correspondiente.
<b>MIGRACIÓN</b>	En esta fase se ejecutaran las consultas generadas en la etapa de procesamiento.
<b>VERIFICACIÓN</b>	En esta fase se generará un archivo que contendrá el estatus de cada migración, así, como el porcentaje de inserciones correctas y fallidas, para determinar si la migración se realizó con éxito.

Tabla 1 Descripción de las etapas de la metodología de migración de datos

### Prototipo de la interfaz

Una de las características que debe tener el sistema, es la de contar con una interfaz gráfica intuitiva, que permita al usuario realizar la migración de forma fácil. El sistema será desarrollado con el lenguaje de programación Visual C#, por las características descritas en el marco teórico. La propuesta de interfaz que tendrá el sistema se muestra en la figura 2. Este prototipo contiene el menú *Archivo* en el que se podrá seleccionar la base de datos que será migrada, así como la base de datos destino. También contará con el menú *Ayuda*, el cual contendrá información acerca del sistema. La opción *Compatibilidad* hará una comparación de las entidades en las bases de datos, tanto fuente como destino para verificar que los nombres de las tablas coincidan en ambas bases de datos, los resultados de este paso se verán reflejados en la parte superior derecha de la interfaz. La opción *Cargar Archivo XML* generará el archivo *XML* con la información de la base de datos origen, que posteriormente será volcado a memoria. En la opción *Transformar* se crearán las cláusulas con las cuales se insertarán los datos en PostgreSQL. Automáticamente al seleccionar el proceso de transformación, se ejecutarán las cláusulas generadas dando inicio al proceso de migración. En este paso, también se generará un archivo con la información de las inserciones realizadas, tanto exitosas como fallidas, para obtener un estadístico de estos factores y determinar si la migración se realizó con éxito.

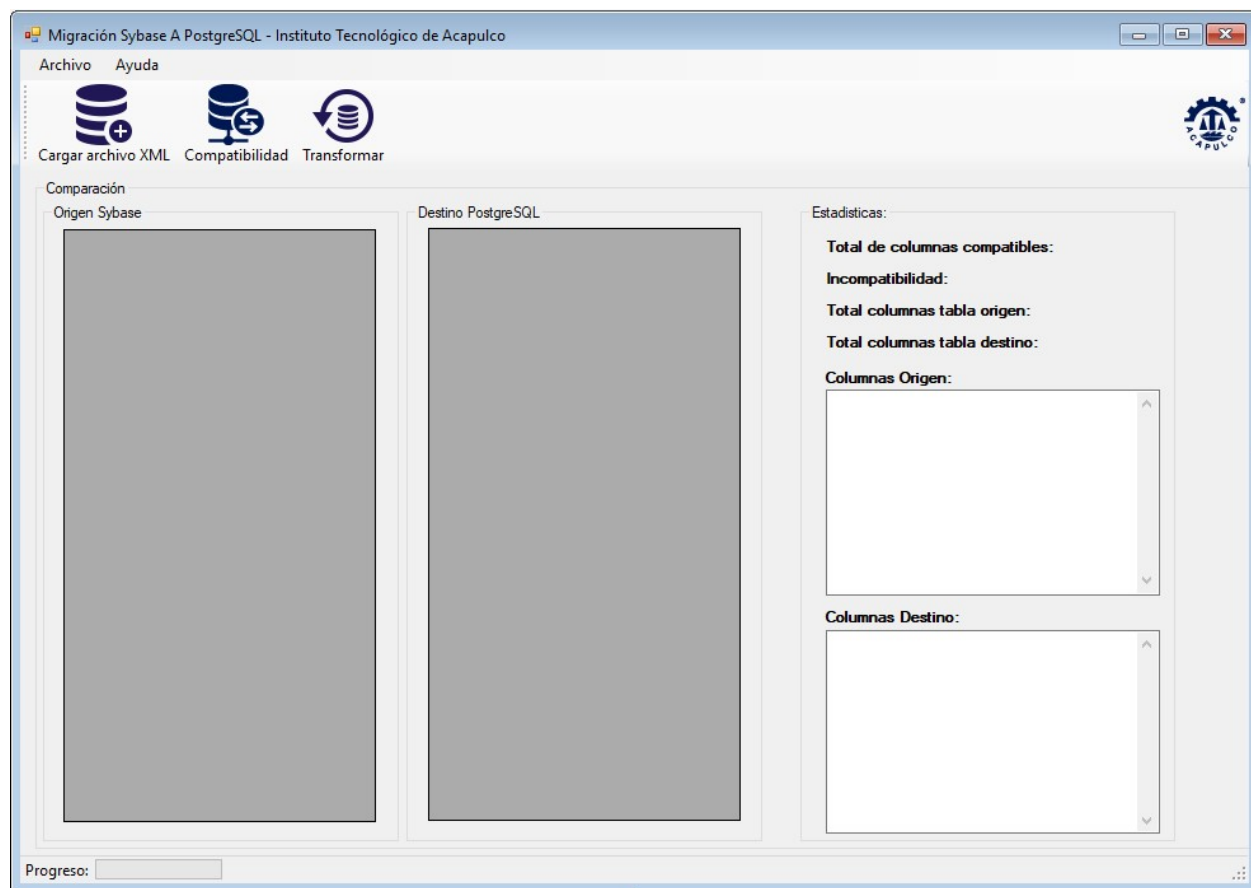


Figura 2 Prototipo de la interfaz del sistema para la migración de datos de Sybase a PostgreSQL

### Conclusión

Se optará por realizar la exportación de los datos a través de un archivo *XML*, debido a la estructura de este tipo de documentos, en las que se presentan los datos de forma ordenada y delimitada, gracias a que es un lenguaje basado en etiquetas, lo cual facilita su lectura y comprensión. Desarrollar el sistema para la migración de los datos de Sybase a PostgreSQL fomentará la evolución de los sistemas de información del TecNM, pues proveerá de una herramienta que permitirá llevar la información hacia un nuevo gestor de base de datos.

## Referencias

- Boulanger, T. (2015). *XML práctico: Bases esenciales, conceptos y casos prácticos (2ª edición)*. Ediciones ENI.
- Ceballos, F. (2013). *Enciclopedia de Microsoft® Visual C# Interfaces gráficas y aplicaciones para Internet con Windows Forms y ASP.NET 4.ª edición*. Madrid: RA-MA Editorial.
- Celis, A. (s.f.). Metodología para la migración de forma segura de Sistemas de Gestión de Bases de Datos Relacionales a software libre y estándares abiertos. *REVISTA PGI - INVESTIGACIÓN, CIENCIA Y TECNOLOGÍA*, 140-142.
- Cerezo, Y. (2007). *Iniciación a la programación en C#: un enfoque práctico*. Madrid: Delta Publicaciones.
- Fernández, J. (2011). METODOLÓGIA PARA LA CONSTRUCCIÓN DE UN MIGRADOR UNIVERSAL DE BASES DE DATOS. *DYNA*, 266-275.
- Gonzales, J. A. (2001). *el lenguaje de programacion c# jose antonio gonzalez seco*. "Documento en formato digital".
- Henríquez, N. (s.f.). Postgresql una alternativa efectiva en las empresas. 1-5.
- POSTGRESQL TUTORIAL*. (08 de 2012). Obtenido de <http://www.postgresqltutorial.com/what-is-postgresql/>
- Tecnológico Nacional de México. (s.f.). *TecNM*. Obtenido de <http://www.tecnm.mx/telecomunicaciones/sistema-integral-de-informacion-dgest>
- Universidad Politécnica de Puebla. (s.f.). *UPPuebla*. Obtenido de <http://repositorio.uppuebla.edu.mx:8080/xmlui/bitstream/handle/123456789/338/OA-UPPue-SGBD.pdf?sequence=1>