

Por Ing. Guianella López Karell, Ing. Leosvani Núñez García e Ing. Yenlidie González Batista, Especialistas en Informática, Centro de Desarrollo de Software UCI en Villa Clara
gkarell@vcl.uci.cu, lnunez@vcl.uci.cu, ygbatista@vcl.uci.cu

Introducción

En el mundo empresarial moderno donde los negocios son cada vez más complejos y las necesidades de información de las empresas cada vez mayores, se ha incrementado notablemente el uso de aplicaciones informáticas de software libre (SL).

La denominación de software libre no se refiere al precio sino a la libertad: primero, a la libertad de copiar y redistribuir un programa para que otras personas lo puedan usar también; segundo, a la libertad de cambiar un programa a fin de que pueda ser controlado a través del código fuente [1]. Son estas características, precisamente, las que hacen que constituya una opción muy ventajosa para las empresas por lo que, a pesar de la incertidumbre y la desconfianza que se ha creado en torno a su implementación, es creciente el número de personas que lo emplean y dejan a un lado al software propietario.

Según un interesante estudio titulado “Valoración del software libre en la sociedad”, realizado por el sitio Web Portal Programas, arrojó que el 65% de los encuestados conocen y usan aplicaciones de software libre, mientras que el 6,5% declaró no usarlo no obstante haberlo intentado. Ello implica que, del total de personas que han probado alguna aplicación de SL, el 91% ha continuado y solo un 9% ha desistido [2].

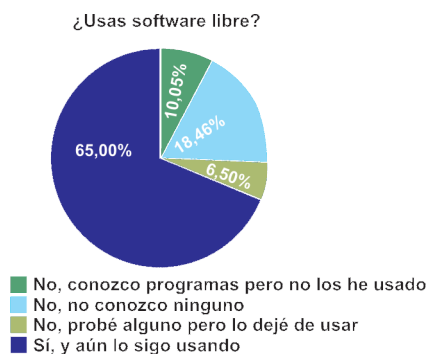


Figura 1 Grado de utilización que los usuarios tienen con el software libre. (Fuente: [2])

Todo esto demuestra que hoy en día el software libre tiene amplia aceptación y es el principal competidor de las empresas de software propietario.

En Cuba, desde hace algunos años, se trabaja en el desarrollo de la industria del software debido a la existencia de personal con el conocimiento y la calificación

requeridos y a la imposibilidad de comprar las licencias propietarias directamente en el mercado norteamericano, por pertenecer a la lista de países bloqueados. La mayoría de estas licencias se adquieren a través de terceros. En la búsqueda de una solución efectiva a este problema, se han considerado las potencialidades y ventajas del SL, que ha sido recomendado, además, como una alternativa práctica para los países en vías de desarrollo.

Herramientas de software libre para la migración de datos

Una vez que las empresas deciden usar tecnologías de bases de datos genuinamente libres, están obligadas a realizar procesos de migración que consisten en el traslado de la información o de los datos desde los sistemas existentes hacia nuevos sistemas.

La migración de datos constituye un procedimiento muy útil debido a que, en diversas ocasiones, las empresas necesitan dar respuesta a determinadas necesidades, que el gestor de bases de datos empleado no les satisface totalmente, como: mejorar el desempeño de la base de datos actual, cumplir con nuevos requerimientos de usuario, de la aplicación o de seguridad; así como estandarizar la tecnología de información en la organización, aumentar el volumen de datos o reducir costos en licencia.

No obstante, al ser un proceso tan complejo requiere que, de manera correcta, se siga cada uno de los pasos con el objetivo de asegurar el buen funcionamiento y la continuidad en la operación. Igualmente, durante la implementación de estas soluciones, es importante garantizar que las transferencias de datos a las nuevas plataformas de almacenamiento sean ejecutadas de manera precisa y oportuna, con una mínima interrupción de las operaciones [3].

En la actualidad existen alternativas informáticas que proveen herramientas de software libre para realizar las migraciones de datos.

En este sentido, PostgreSQL representa una solución con la calidad requerida para satisfacer las demandas de los usuarios y es equiparable a cualquier sistema gestor de base de datos (SGBD) propietario. Actualmente, es considerado como el sistema gestor de código abierto más avanzado a escala mundial por las características que posee, atribuible, por lo general, a los productos comerciales de alto calibre [4]. Dispone de una serie de funcionalidades inherentes a las bases de datos con altas prestaciones que lo hacen apto para la mayoría de las aplicaciones ya que fue diseñado y creado para tener un mantenimiento y ajuste mucho menor que otros productos. Además, ofrece ventajas significativas como elevada concurrencia, amplia variedad de tipos nativos, gran adaptabilidad, escalabilidad, entre otras, que justifican su preferencia.

Por otra parte, Kettle, de la Suite de Pentaho, y Apatar¹ constituyen las herramientas de código abierto líderes en extracción, transformación y carga. Sin embargo, ninguna de las herramientas mencionadas anteriormente provee suficiente información documentada sobre cómo realizar migraciones de datos específicas hacia PostgreSQL donde se expongan, por ejemplo, las razones concretas para migrar desde los gestores propietarios referidos o se describan cada una de las etapas a seguir, así como las herramientas libres utilizadas para llevar a cabo estos procesos, lo cual serviría como material de apoyo a las acciones de capacitación sobre estándares de código abierto y libre implementadas en nuestro país. La documentación es una parte esencial de cualquier software; sin embargo, en muchas ocasiones, las aplicaciones más importantes no vienen acompañadas de manuales completos. Cuando esto sucede, se crea una laguna de información [5]. En consecuencia, se puede afirmar que el problema fundamental que presentan las empresas involucradas en este tipo de proyectos radica, precisamente, en el hecho de no contar con la documentación que brinde un asesoramiento técnico integral para realizar los servicios de migración de datos hacia soluciones libres, lo cual dificulta el empleo de las herramientas definidas para este proceso. Surge así la necesidad de definir guías de migración de datos que orienten paso a paso la ejecución de este proceso desde los gestores propietarios hacia PostgreSQL, que contribuyen al soporte de los servicios de migración como parte de la transición hacia soluciones de software libre.

Propuesta de guía para la migración de datos

La guía propuesta [7] fue confeccionada por un grupo de especialistas del Centro de Desarrollo de Software (CDS) de la UCI en Villa Clara con el objetivo de brindar a los usuarios un documento que facilite la comprensión de los procesos de migración de datos a través del empleo de la herramienta Kettle. En la misma se describen los requisitos de instalación necesarios para trabajar con la herramienta, así como los pasos para migrar datos desde los gestores MySQL, FoxPro, Oracle, SQL Server y Access hacia PostgreSQL. La selección de la herramienta Kettle se debe a sus características: facilidad de uso, disponibilidad de transformaciones con opciones de configuración, factibilidad de implementación e integración con un sistema completo de inteligencia de negocios. Además, pertenece a la Suite de Pentaho [6], la cual cuenta con la Licencia MPL —*Mozilla Public License*— que ofrece garantías de uso al cumplir con todos los criterios del software libre. No obstante, es válido especificar que Kettle migra solo las tablas con los datos, de modo que si el alcance del proyecto incluye la migración de procedimientos almacenados, funciones, *triggers* u otros

¹ Para obtener más información sobre estas herramientas se pueden consultar los sitios Web <http://www.kettle.pentaho.org/> y <http://www.apatar.com/>, respectivamente.

objetos, dado el cambio de manejador de bases de datos, se debe considerar la posibilidad de programarlos nuevamente, ya que el código puede que no sea compatible.

Contenido de la Guía

La Guía para la migración de datos está compuesta por las siguientes secciones:

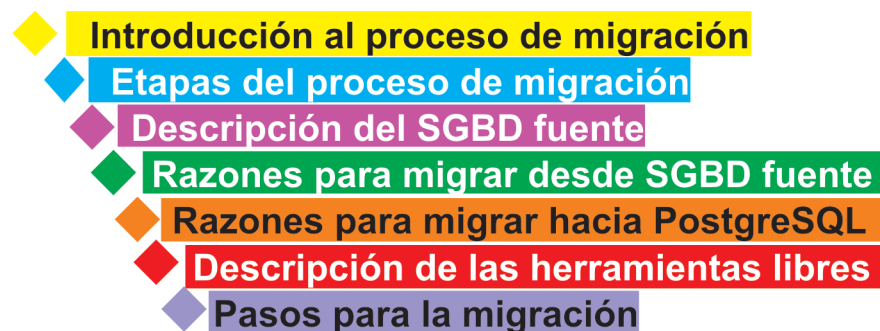


Figura 2 Secciones de la Guía de migración de datos. (Fuente: [7])

- 1. Introducción al proceso de migración:** ofrece una breve introducción sobre el proceso de migración de datos y sus objetivos, donde se señalan los beneficios de la migración de datos desde un sistema gestor propietario hacia PostgreSQL.
- 2. Etapas del proceso de migración:** abarca las fases para migrar debido a la importancia que tiene cada una de ellas a fin de garantizar el buen funcionamiento y la continuidad en la operación.
- 3. Descripción del SGBD fuente:** presenta las características fundamentales del gestor de bases de datos que constituye el origen de la migración, por ejemplo, su factibilidad de adquisición en el mercado, tipos de licencias que poseen, entre otras.
- 4. Razones para migrar desde el SGBD fuente:** expone los motivos por los cuales es necesario o factible realizar una migración desde ese sistema gestor fuente, se argumentan las limitaciones de este en relación con PostgreSQL.
- 5. Razones para migrar hacia PostgreSQL:** destaca las peculiaridades de PostgreSQL y los beneficios que implicaría su empleo, pues es considerado uno de los gestores más potentes dentro del mercado de SL por sus prestaciones.
- 6. Descripción de las herramientas libres:** describe las herramientas libres disponibles a fin de familiarizar al usuario con el empleo de las mismas. En esta sección también se fundamenta la selección de la herramienta Kettle para desarrollar el proceso.
- 7. Pasos para la migración:** detalla cada uno de los pasos para la migración con la ayuda de imágenes para lograr mejor comprensión.

Resultados

La presente Guía ha sido aplicada, hasta el momento, en el Banco de Crédito y Comercio (BANDEC) del municipio Santa Clara debido a la necesidad real de migrar su base de datos hacia un gestor libre, como parte de la estrategia de cambio a software libre trazada a nivel nacional.

El proceso fue satisfactorio. Se logró migrar completamente la base de datos del Banco en FoxPro con 623 294 registros hacia el servidor con PostgreSQL de manera exitosa. La herramienta Kettle de Pentaho arrojó los resultados que se muestran en la figura 3.

p/a	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo	Finalizado	Tiempo	Velocidad (r/s)	Pri/E/S
0	623294	623294	623294	623294	0	0	0	0	Finalizado	200.2	3113.5	-
0	623294	623294	623294	623294	0	0	0	0	Finalizado	202.8	3072.7	-

Figura 3 Visualización de los resultados de ejecución del proceso de migración de datos. (Fuente: elaboración propia)

Los inconvenientes encontrados fueron durante el traslado de los campos de tipo MEMO en FoxPro hacia el gestor de base de datos de PostgreSQL pues fue necesario realizar conversiones a tipo CHARACTER.

Esta Guía también ha sido utilizada como material para apoyar la capacitación de algunos desarrolladores del CDS de Villa Clara en la realización de migraciones de datos.

Conclusiones

La Guía para la migración de datos constituye un material didáctico de gran utilidad. La organización de su contenido facilita el trabajo para efectuar migraciones exitosas desde los gestores de datos propietarios MySQL, FoxPro, Oracle, SQL Server y Access hacia PostgreSQL; sirve, además, como apoyo al proceso de capacitación implementado en nuestro país en aras de impulsar la migración hacia estándares de código abierto y libre.

En ella se incluyen todos los recursos para que un usuario básico pueda dar sus primeros pasos en la ejecución de los procesos de migración de datos con soluciones libres de manera satisfactoria, por lo que constituye una fuente documental de gran valor que hace aún más atractiva la adopción software libre en Cuba. ■

Referencias bibliográficas

- [1] Stallman, Richard M. "What is the Free Software Foundation?". *GNU's Bulletin*, vol. 1, no. 2 (febrero, 1986). http://web.cecs.pdx.edu/~trent/gnu/bull/02/nb_toc.html#SEC. (acceso diciembre 5, 2009).
- [2] PortalProgramas.com. *Valoración del Software Libre en la sociedad 2010*. <http://www.portalprogramas.com/software-libre/informe>. (acceso febrero 25, 2010).
- [3] IBM S.A. Servicios de migración de datos. <http://www.05.ibm.com/services/es/sds/0905000000.html>. (acceso diciembre 6, 2009).
- [4] Sabater, Jaume. "Open-Xchange: una suite de trabajo en equipo". http://linuxsilo.net/articles/openxchange.html#configuracion_postgresql. (acceso diciembre 5, 2009).
- [5] Stallman, Richard M. *Software libre para una sociedad libre*. Madrid: Traficantes de Sueños, 2004.
- [6] Álvarez Serra, Ángel. *Migración de datos con OpenERP-Kettle*. <http://www.slideshare.net/raimonesteve/migracin-de-datos-con-openerpkettle>. (acceso diciembre 6, 2009).
- [7] López Karell, Guianella, et al. "Guías para la migración de datos". Documento inédito. CDS Villa Clara-UCI, 2009, p. 18.