

“Desarrollo de un Administrador de Base de Datos Relacional TecnoDB”

Autores:

Iris Gastañaga

Ing. en Sistemas de Información y Especialista en Docencia Universitaria, Investigadora Categoría III. Teléfono: 0351-4564245 igastañaga@cpcipc.org . Directora del Proyecto

Calixto Maldonado

Ing. en Sistemas de Información. Teléfono: 0351-155141710 calixtomaldonado@hotmail.com

Cesar Spessot

Ing. en Sistemas de Información. Teléfono: 0351-155099690 cspessot@sistemas.frc.utn.edu.ar

Anabel Ruiz

Estudiante de la carrera de Sistemas De Información. Tel:03548-421426

anabelru@bbs.frc.utn.edu.ar

Horacio de Oro

Estudiante de la carrera de Sistemas De Información. Tel 0351-155098073 hdeoro@uol.com.ar

Las Actividades son desarrolladas en el Laboratorio de Investigación de Software del Departamento de Sistemas de la Facultad Regional Córdoba de la Universidad Tecnológica Nacional.

Córdoba, 08/04/2004.

Resumen Técnico

Se construirá un motor que mostrará el funcionamiento de administración de una base de datos que permita conocer en forma práctica la resolución de consultas y transacciones (select, insert, update, delete, etc.) con filas de tablas almacenadas en el mismo.

El motor de Base de datos a construir contará con un monitor que permita dar visibilidad sobre el proceso que se ejecuta internamente ante cada consulta, reflejando el álgebra relacional y permitiendo evaluar si los resultados son los pretendidos, a la vez que se pueda comparar la eficiencia lograda con diferentes construcciones lógicas. Esta funcionalidad no se encuentra presente en los motores comerciales de difusión actual.

Para lograr estas funcionalidades deberá contar con un modelo de datos que almacene los Metadatos (información sobre las tablas, columnas, constraints o restricciones, índices existentes en la Base de datos) llamado generalmente Diccionario. Asimismo, el motor deberá proporcionar información sobre tiempos de búsqueda, inserción, borrado y actualización de filas en las tablas de la Base de Datos, conjuntamente con comparaciones de los mismos utilizando diferentes índices y sin valerse de éstos a los efectos de que pueda realizarse comparaciones.

El motor prestará servicios a través de redes TCP/IP y para la transmisión de consultas y sus respectivos resultados se utilizará XML.

Para posibilitar el acceso al motor desde aplicaciones se desarrollará un driver JDBC.

Se desarrollará complementariamente, una interfaz gráfica para facilitar la creación de consultas en lenguaje SQL. Además, se prevé que podrá establecer filtros, subconsultas, criterios y restricciones de agrupamientos y reglas de ordenamiento. Se pretende que la interfaz del motor desarrollado satisfaga el FASMI TEST, a la vez que se convierte en una herramienta útil para soportar la generación de procesos cognitivos en los estudiantes de cátedras que traten temas de bases de datos.

Grado de Avance:

A la fecha se ha logrado desarrollar una arquitectura que representa los datos de cada tabla e índices en sendos archivos del sistema operativo. Se construyeron objetos que permitieron construir en lenguaje C++, un intérprete de consultas SQL capaz de ejecutar las sentencias básicas, Create table, Select, Insert, Update y Delete, excluyendo las de seguridad y manejo de transacciones, soportando objetos como índices y tablas.

El trabajo ha sido publicado en *The Journal of Computer Science and Technology (JCS&T) Vol. 3 - No. 2 - October 2003 - ISSN: 1666-6038*, y aceptado para Congresos Internacionales tales como CACIC 2002 (UBA, 2002) y CIMAC II (Peru, 2003).

OBJETIVO PRINCIPAL

- Estudiar, Investigar y Desarrollar un Administrador de Base de Datos que provea herramientas para facilitar el estudio de procesos internos y los aspectos asociados a la performance de los algoritmos aplicados dentro del motor. La flexibilidad del Motor de Base de Datos lo convertirá en una buena plataforma de experimentación que permita su adaptación para diferentes patrones de aplicación. El origen del proyecto esta en el desarrollo de un compilador SQL para una materia de la carrera, lo que llevo a crear desde cero el proyecto, sin tener referencias o usar otras soluciones tecnológicas ya desarrolladas y disponibles en licencias abiertas sobre los fuentes.

OBJETIVOS PARTICULARES

- Estudio de los algoritmos y procesos de los Administradores de Base de Datos, lo que nos posibilitará tener una visión sobre cuales son los aspectos a mejorar e implementar.
- Formalización de un Lenguaje de Consulta y Manipulación de Datos.
- Diseño de un lenguaje nativo del Motor.
- Implementación de un intérprete de Lenguaje.
- Desarrollo de una Arquitectura flexible con el objetivo de proveer herramientas para el estudio e investigación de los procesos internos del Motor de Base de Datos.
- Desarrollo e incorporación de un monitor que permita la visualización de los procesos y lógica que se ejecutan durante cada requerimiento brindando índices de desempeño.
- Diseño de patrones de comparación que faciliten el análisis de la performance del motor de Base de Datos.
- Implementación de una estructura y mecanismos de comunicación TCP/IP para brindar servicios a través de la red.
- Implementación de Drivers de Datos para interactuar con otros lenguajes de programación.

Metodología

La metodología de Investigación y Desarrollo del proyecto se ha definido de acuerdo a los pasos lógicos y probados de análisis, diseño y construcción de un software de manera iterativa e incremental. Teniendo en cuenta los siguientes etapas funcionales:

1. Se estudiarán los algoritmos y procesos de los Administradores de Base de Datos que existen actualmente tanto comerciales como de investigación, y como abordan éstos el problema de la interoperabilidad y la flexibilidad, prestando especial atención a los mecanismos de indexación, análisis de código fuente y comunicación. En este estudio se conseguirá un conocimiento profundo de los modos de abordar dichas características. Asimismo, la visión de conjunto permitirá realizar comparativas que lleven a un análisis crítico respecto a los sistemas existentes y sus carencias.
2. Se definirán los requisitos deseables para el diseño óptimo de un Administrador de Base de Datos Relacional. Los aspectos generales de un sistema de gestión de base de datos y los aspectos relacionados con la arquitectura del sistema integral tendrán mucho que ver con la elección de los requisitos de partida. Cabe destacar que se pone como principal requerimiento la implementación de monitores que permitan el estudio de bajo nivel del Motor de Base de Datos.
3. Como consecuencia de los requisitos impuestos y los conocimientos adquiridos en la primera fase se especifica el núcleo del Administrador de Base de Datos, centrado en el mecanismo de indexación y ejecución de código fuente. Esta fase se concibe como resultado del proceso de crítica, análisis y selección de posibilidades iniciado tras la primera fase. A ello hay que añadir un proceso de síntesis y especificación de un modelo formal que responda a todos los aspectos definidos.
4. Una vez formalizado el sistema, se procede a la implementación en lenguaje de computación del Administrador de Base de Datos. Se programará los siguientes módulos:
 - Intérprete de Lenguaje de Consulta y Manipulación de Datos
 - Motor de Base de Datos Relacional
 - Procesos y Algoritmos de Indexación

- Módulo de comunicación del motor a través de una red TCP/IP.
- Driver JDBC que permita utilizar el motor desde cualquier aplicación Java.
- Generador de Consultas SQL con el objetivo de facilitar el aprendizaje del desarrollo de las mismas.
- Monitor de desempeño.

Esta etapa se llevará a cabo mediante un proceso iterativo e incremental de desarrollo de software, a través de los modelos propuestos por el Lenguaje Unificado de Desarrollo presentado por Jacobson, Booch y Rumbaugh [JacBR00].

5. Se desarrollará un proceso de implementación conjunta con las cátedras de Gestión de Datos y Sistemas de Soporte de Decisión de la UTN – FRC, que actuarán como beta tester del producto final y trabajarán con el grupo de investigación generando proyectos académicos sustentados en la herramienta e iniciando una retroalimentación y posible incorporación de nuevos subgrupos de integrantes para la ampliación de los alcances del producto y nuevos objetivos a implementar.

Estimación de los aportes esperables del proyecto

Tecnológicos

Debido al propósito principal de nuestro proyecto, soporte al proceso de enseñanza-aprendizaje, se desprenden los siguientes argumentos que contribuyen al avance científico:

- Desarrollar una herramienta que contribuya al proceso de enseñanza de toda la arquitectura de un motor de base de datos, la sintaxis de SQL y posibilitar el aprendizaje de la confección de consultas, disminuyendo así el índice de errores en la selección de datos en una base de datos con grandes cantidades de registros.
- Comprender la información implícita en la base de datos por medio de diversos mecanismos para la recuperación de información substancial para una organización.
- Facilitar estrategias educativas para favorecer el proceso de aprendizaje en alumnos poco avanzados en el conocimiento de las bases de datos y sentencias SQL.
- Desarrollar habilidades para aproximarse a los procesos de confección de consultas SQL y visualizar el proceso interno que realiza el motor para poder combinar las tablas, coordinar los datos y de ésta manera obtener el resultado de la misma.

Este proyecto recrea una solución técnica al modelo relacional, brindará conocimientos a las cátedras de nuestra Facultad Regional orientadas las disciplinas que estén relacionadas con Bases de Datos y Compiladores, ya que aporta un ejemplo práctico a las mismas.

Además podrá ser usado para otras investigaciones posteriores relacionadas con las disciplinas antes mencionadas.

Es de recalcar que éste proyecto podría tener importantes repercusiones tecnológicas que darían lugar a futuros proyectos de investigación, por ejemplo una herramienta que permita realizar la elaboración de un modelo multidimensional aplicado a la toma de decisiones, ampliar el gestor realizado para que soporte el manejo de datos en el ámbito temporal, etc.

Formativos

Esencialmente nuestro proyecto tiene una finalidad académica destinada a contribuir al proceso de enseñanza-aprendizaje.

La realización de este proyecto servirá a las Cátedras de Diseño de Lenguaje de Consultas, Sintaxis y Semántica del Lenguaje, Tecnología de Software de Base, Paradigmas de Programación y Gestión de Datos. Cabe destacar que se realizarán una serie de cursos y seminarios sobre los temas tratados en este proyecto de investigación. El producto se instalará en las máquinas de un laboratorio de la

facultad a los fines de ser utilizado por alumnos de la cátedra de Gestión de Datos para mostrarles su funcionamiento, enseñarles a construir sentencias de SQL y que realicen los prácticos asignados por la cátedra.

También se formarán a dos estudiantes avanzados en la Carrera de Ingeniería en sistemas de Información, con la finalidad de que inicien su formación en investigación científica y tecnológica, profundicen sus conocimientos en temas significativos en la actualidad y desarrollen su perspectiva con respecto a los sistemas gestores de bases de datos.

Inicialmente la cobertura del proyecto serán las Cátedras de Gestión de Datos y Sistemas de Soporte de Decisión de la Universidad Tecnológica Nacional Facultad Regional Córdoba y extensión aúlica Punilla. Posteriormente se contactarán otras Universidades, ya que Córdoba cuenta con 6 Universidades que dictan materias del tema, a los efectos de realizar transferencia tecnológica a las cátedras.

Bibliografía

- [Ullman-1997] ULLMAN, J. D. - Widom, J. (1997) "*A first course in database System*", Prentice Hall
- [Leonidovich-1996] Leonidovich, K. M. (1996) "*Research and Development of a Lenguaje Subsystem SQL*", Russian Academy of Sciences, Institute of System Programming, Moscow
- [Date-1990] DATE, C. J. (1990)
"*An Introduction to Database Systems*", Addison-Wesley Publishing Company, Vol I.
- [Silberchatz-1998] Silberschatz, A. – Korth, H.F. – Sudarshan, S. (1998)
"*Database System Concepts*", Mc Graw-Hill, 3 Edition.
- [Aho-1990] Aho, A. V. – Sethi, R. – Ullman, J. (1990)
"*Compiladores Principios, Técnicas y Herramientas*", Addison-Wesley Iberoamericana
- [Langsam-1997] Langsam, Y. – Augentein, M. J. - Tennenbaum, A. M. (1997)
"*Estructura de Datos con C y C++*", Prentice Hall, 2 Edición.
- [Chorafas-1983] Chorafas, D. N. (1983)
"*DBMS for Distributed Computers and Networks*", Petrocelli Books
- [Elmasri-2002] Elmasri, R.- Navathe, S. (2002)
"*Fundamentals of Database Systems*", Addison-Wesley, 3rd edition.
- [Ramakrishnan-2000] Ramakrishnan, R. – Gehrke, J. (2000)
"*Database Management System*", McGraw-Hill, , 2nd edition
- Novática Número 140 Bases de Datos. Julio/Agosto 1999
- Algunas Páginas web visitadas período agosto-diciembre de 2003.
- <http://www.ucse.edu.ar/fma/compiladores/>
- <http://www.me.gov.ar/index1.html>
- <http://www.ispras.ru/~kml/gss/>
- <http://sal.kachinatech.com/H/1/>
- <http://leap.sourceforge.net/>
- <http://www.jepstone.net/tinySQL>
- <http://amsoft.ru/easysql>
- <http://macine.epublish.cl/tesis/index.html>
- <http://www.lsi.us.es/colabora/fieware/programas/sbase.html#basica>
- <http://www.cindoc.csic.es/isis/indice.htm>
- <http://members.tripod.com/~MoisesRBB/sql.html>
- <http://www3.uji.es/~mmarques/f47/apun/>
- http://www.cicese.mx/~orodrigu/database/resumen_01.htm
- <http://espanol.geocities.com/michelsoftpe/bdatos.htm>
- <http://www.dsic.upv.es/~jorallo/docent/BDA/DisiplinaBD.pdf>
- <http://www.monografias.com/trabajos13/trsqlinf/trsqlinf.shtml#BREVE>
- <http://es.tldp.org/Postgresql-es/web/navegable/tutorial/sql.html>