

Comparativas Informáticas de SEA

Número 1: Bases de Datos

Versión Reducida
Por Ignacio Berenguer

Contenidos

1. Introducción	3
A quién va dirigido este documento	3
Cómo leer este documento	3
Para más información	5
2. Las Bases de Datos en la Actualidad	5
Definición de "sistema de gestión de base de datos"	5
Conceptos asociados a los sistemas de bases de datos	6
3. ¿Qué tipo de gestor de base de datos necesito?	9
Una primera distinción	9
Variables de decisión en sistemas de gestión de bases de datos personales	9
Variables de decisión en sistemas de gestión de bases de datos basados en servidor	11
4. Sistemas de Gestión de Bases de Datos Personales	14
Principales alternativas comerciales de bases de datos personales	14
Comparación de sistemas de gestión de bases de datos personales	15
Como elegir un sistema de gestión de base de datos personal	17
5. Sistemas de Gestión de Bases de Datos basados en Servidor	18
Características comunes de los 3 productos analizados	19
IBM DB2 Universal Database 8.1	19
SQL Server 2000	19
Oracle 9i Enterprise	20
Comparación de Precios de los Sistemas de Bases de Datos Analizados	20
6. Sistemas de Gestión de Bases de Datos para "Business Intelligence"	20
<i>Business Intelligence</i>	20
Infraestructura de Business Intelligence	21
Plataformas analizadas	21
Plataforma de Microsoft	22
Plataforma de IBM	22
Plataforma de Oracle	22
Comparación de las tres plataformas	23
Conclusión	24

Disclaimer

Este documento contiene referencias a empresas, productos y sus marcas registradas. Todas las marcas registradas son propiedad de sus respectivos dueños. La información utilizada para crear este documento ha sido obtenida de fuentes consideradas confiables. Ni el autor ni "Empresarios Alaveses SEA" tienen responsabilidad alguna por errores u omisiones en esta información. Todos los precios han sido obtenidos de la lista de precios pública disponible en el momento de escribir el documento, sin aplicar descuentos de ningún tipo.

Acerca del Autor

Ignacio Berenguer es Ingeniero Industrial, especialidad en Gestión y Organización por el ICAI (Universidad Pontificia de Comillas, Madrid). Posee una amplia experiencia en sistemas de informáticos corporativos. Actualmente trabaja como Director de Tecnología en opciona.com (Endesa Marketplace), empresa del Grupo Endesa que dedicada a la optimización de los procesos de sourcing y procurement mediante la utilización de tecnologías de comercio electrónico B2B (business to business). Opciona.com es el marketplace líder en el mercado español, y ofrece soluciones globales de e-sourcing y e-procurement. Es profesor del programa Temis de la Universidad de Navarra.

1. Introducción

A quién va dirigido este documento

El presente documento está dirigido a empresarios y gestores de empresas de cualquier tamaño... profesionales de todo tipo, usuarios más o menos avanzados de la informática, pero expuestos todos ellos a los avatares de la revolución digital y las tormentas de Internet.

Estos profesionales, preocupados lógicamente por la buena marcha de su compañía, no son necesariamente expertos en informática... detectan con indudable acierto que cada día que pasa, es mayor la cantidad de información acerca de su negocio que queda almacenada en almacenes de datos basados en diversas tecnologías.

En este contexto, las palabras "*base de datos*" han pasado a formar parte del vocabulario habitual de los profesionales de esta incipiente sociedad de la información... y sin embargo, a pesar de constituir el cimiento de los sistemas informáticos empresariales, es posible que la tecnología y posibilidades que hay detrás de esta expresión no hayan sido divulgadas correctamente.

Por este motivo, esta serie de artículos divulgativos se inicia con una comparación práctica y dirigida al usuario no técnico de los sistemas de base de datos más habituales en el mercado español y europeo.

Con la intención de conseguir una solución de base de datos para cualquiera que sea el tamaño, presupuesto y complejidad de la necesidad del lector, se tratan con igual profundidad los sistemas de base de datos personales, y los servidores de base de datos.

El ritmo de la sociedad actual, y la gran preparación que exige de los profesionales, obliga a aprovechar el tiempo al máximo, buscando formas rápidas de obtener conocimiento. Con el objetivo puesto en este público, impaciente y exigente, se ha tratado de preparar un texto sencillo y divulgativo. Abundan las llamadas a pie de página que intentan explicar los conceptos técnicos. Se ha sacrificado el rigor técnico por la facilidad de comprensión, y se han realizado ciertas simplificaciones de índole práctico a la hora de abordar problemas que, por su extraordinaria complejidad, serían muy extensos de tratar en rigor... el lector no técnico probablemente lo agradecerá, y para el experto en informática existe abundante bibliografía en la que se trata estos temas con más detalle y precisión.

Cómo leer este documento

El documento que tiene en sus manos está pensado para ser leído de principio a final. Si lo hace así, realizará en relativamente poco tiempo un intenso viaje por el mundo de la bases de datos, deteniéndose en prácticamente todos los aspectos relevantes.

Sin embargo, los capítulos en los que se compone pueden ser leídos de forma independiente.

En el **capítulo 2, "Las bases de datos en la actualidad"**, se introduce al lector a la importancia de este tipo de sistemas, dentro de la informática corporativa. También se presentan una serie de propiedades que distinguen a una base de datos de otros almacenes de datos. Se establece una distinción importante, necesaria ahora que este software es tan popular que puede ser encontrado en un *mainframe*¹ o en un *PDA*²: las bases de datos personales y los sistemas de base de datos basados en servidor. Por último se enuncia la diferencia entre OLAP y OLTP³, dos tipos de aplicaciones de bases de datos distintas surgidas de la popularidad de este tipo de sistemas.

El **capítulo 3, "¿Qué tipo de gestor de base de datos necesito?"** orienta al lector acerca de los elementos de decisión para elegir entre una base de datos personal y una base de datos basada en servidor. Cuestión siempre presente, con importantes consecuencias técnicas y económicas, y cuya respuesta en el caso de las empresas de tamaño mediano y pequeño no siempre es sencilla.

En caso de que el lector tenga prisa, es posible que con la información del capítulo 3 decida saltar o leer el **capítulo 4 "Los sistemas de base de datos personales"**. Para los que los lean, contiene la comparación de una serie de sistemas de bases de datos personales que se pueden encontrar fácilmente en el mercado.

Ya sea a través de capítulo 4, ya sea buscando apresuradamente respuesta a la información referente a los servidores de base de datos, el **capítulo 5 "Sistemas de gestión de bases de datos basados en servidor"**, proporciona al lector una introducción a las semejanzas y diferencias, a las posibilidades, ventajas e inconvenientes de los tres principales servidores de bases de datos comerciales. Cualquiera de ellos es un producto enormemente complejo, cuya documentación ocupa una librería completa... en este capítulo encontrará una comparación que puede leerse en 5 minutos.

Si el capítulo 5 se centraba principalmente en las características de base de datos general (es decir, en la faceta OLTP), el **capítulo 6 "Sistemas de bases de datos para *Business Inteligente*"** se adentra en las características de análisis (la faceta OLAP) de los servidores de bases de datos.

¹ *Mainframe* es un gran sistema informático corporativo, destinado a la gestión de procesos empresariales críticos.

² *PDA (Portable Digital Assistant)*, es una agenda electrónica, también conocido como asistente personal digital que cabe en el bolsillo de la chaqueta.

³ *OLAP* y *OLTP* serán explicadas con mayor detalle en el capítulo correspondiente.

Para más información

Tanto "Empresarios Alaveses SEA" como el autor agradecen las opiniones y colaboraciones de los lectores.

En caso de cualquier cuestión, animamos a enviarlas por correo electrónico a la dirección: basesdatos@sea.es.

2. Las Bases de Datos en la Actualidad

Posiblemente el común denominador de todos los sistemas informáticos de gestión es que están diseñados alrededor de una base de datos. Al igual que los procesadores de texto, las hojas electrónicas, el correo electrónico y los navegadores de Internet, las bases de datos son parte de los sistemas informáticos de todas las empresas, y están entre nosotros desde hace más de 20 años.

Sin embargo, las bases de datos tienen encomendada una de las misiones más críticas de las necesarias para el correcto funcionamiento de la empresa: la custodia de la información empresarial, su almacenamiento, y el acceso a la misma.

La elección de un sistema de gestión de base de datos es, por lo tanto, una decisión de gran trascendencia para las empresas, en la medida que condiciona de forma importante la capacidad del conjunto de sus sistemas informáticos.

El presente capítulo es una introducción destinada a profesionales no técnicos, que pretende presentar de una forma sencilla el complejo mundo de los distintos sistemas de gestión de bases de datos.

Definición de "sistema de gestión de base de datos"

Un *sistema de gestión de base de base de datos* (en inglés las siglas utilizadas para referirse a ellos son DBMS⁴) es una conjunto de programas de ordenador que permite almacenar, modificar y extraer información de una o más *bases de datos*. Hay muchos tipos de SGBD, desde aquellos destinados a ordenadores personales o incluso agendas (PDA⁵), a aquellos que necesitan grandes sistemas informáticos. En la mayor parte de las ocasiones, los sistemas de gestión de bases de datos que se encuentran en las empresas son *relacionales* (por lo tanto RDBMS⁶), aunque hay sistemas de gestión de bases de datos jerárquicos, de objetos... a lo largo de este artículo, se contemplarán únicamente los sistemas de gestión de bases de datos relacionales, que

⁴ DBMS, *Database Management System*

⁵ PDA, *Portable Digital Assistant*, también conocido como "agenda electrónica" u "organizador personal".

⁶ RDBMS, *Relational Database Management System*

representan la inmensa mayoría de los sistemas de bases de datos en funcionamiento en la actualidad.

Una *base de datos* es una colección de información acerca de un determinado tema, organizada de manera que un programa de ordenador pueda acceder rápidamente a los datos que necesita. Las bases de datos relacionales organizan la información en tablas, campos y registro.

Conceptos asociados a los sistemas de bases de datos

Al igual que en todas las disciplinas, en el ámbito de los sistemas de gestión de bases de datos existen una serie de conceptos que, aunque pueden resultar incomprensibles o irrelevantes para aquellas personas que abordan el tema desde una perspectiva no técnica, son indispensables para poder comprender correctamente las características de un sistema de gestión de base de datos.

Características que debe cumplir un sistema de base de datos

Hay una serie de conceptos que hacen referencia a la capacidad de un sistema de gestión de base de datos para poder cumplir su misión: *almacenar información de la empresa de forma segura, y permitir su acceso en cualquier momento y en el menor tiempo posible por parte de cualquier persona o aplicación que la necesite.*

De la propia lectura de la misión, se derivan una serie de propiedades o virtudes que debe cumplir un sistema de base de datos... a continuación se definen brevemente, aunque entraremos a verlas con más detalle a continuación:

- ? **Integridad:** la base de datos debe asegurar que la información que se almacena en la misma es exacta y que no es posible que se corrompa por el uso o a lo largo del tiempo.
- ? **Disponibilidad:** la base de datos debe estar disponible para almacenar, modificar o recuperar información en cualquier momento.
- ? **Concurrencia:** la base de datos debe permitir el acceso simultáneo a la información que contiene a muchos usuarios.
- ? **Escalabilidad:** la base de datos debe permitir almacenar grandes cantidades de información, y facilitar el acceso a los datos a gran velocidad aunque aumente la información almacenada, el número de usuarios o la actividad de los mismos.
- ? **Seguridad,** que incluye aspectos como asegurar la confidencialidad de la información, controlar los privilegios de acceso de información a los datos por parte de los usuarios, garantizar la recuperación de la información en casos de accidentes y desastres.

Tipos de sistemas de bases de datos

Como se verá a lo largo de este artículo, hay distintas soluciones de base de datos, diferenciándose unas de otras en lo bien que realizan estas 4 características. En una clasificación muy general es posible diferenciar dos tipos de sistemas de bases de datos:

- ? Las **bases de datos personales** (*desktop databases*): sistemas de gestión de bases de datos sencillos, baratos y fáciles de usar, destinados a usuarios particulares o a pequeños grupos de trabajo. Suelen utilizarse para aplicaciones de base de datos sencillas.
- ? Los **servidores de bases de datos** (*database servers*), productos de gama alta y multitud de opciones que permiten configurarlos para abordar hasta las aplicaciones de bases de datos más complejas. Lógicamente son más costosos y difíciles de utilizar correctamente.

Aplicaciones de las bases de datos

Como el lector puede fácilmente imaginar, hay innumerables aplicaciones de las bases de datos, puesto que cualquier aplicación que necesite almacenar una cantidad importante de información acaba, de una manera u otra, utilizando una base de datos.

Sin embargo, a partir de los 1980s comienzan a diferenciarse claramente 2 tipos de aplicaciones de bases de datos:

- ? OLTP (*OnLine⁷ Transaction Processing*): aplicación de base de datos en la cual el sistema debe responder inmediatamente a las peticiones del usuario. Cada petición se conoce como una *transacción*.
- ? OLAP (*OnLine Analytical Processing*): aplicación informática que permite el análisis de la información almacenada en una base de datos. Relacionado con esto están conceptos de *datawarehouse*, *data mining*...

Más información acerca de OLAP

Las aplicaciones de OLAP han experimentado un extraordinario crecimiento en los últimos años... de alguna manera, cuando las empresas disponen de grandes bases de datos fruto de la utilización masiva de aplicaciones OLTP, surge la necesidad de comenzar a analizar

⁷ *Online* (en línea, sinónimo de instantáneo, con el usuario delante) forma parte de cualquiera de los dos tipos de aplicaciones de base de datos. Hoy puede parecer sorprendente, pero debe tenerse en cuenta que en un pasado no muy lejano, los sistemas informáticos únicamente estaban preparados para responder a preguntas en base a trabajos planificados (*batch*). Imagínese el siguiente diálogo entre el Director Financiero y un operador del sistema: "Por favor, pregúntele al sistema cuanto hemos facturado el año pasado". "Sí, señor, su petición está en la cola, la respuesta dentro de tres días"... la prehistoria de la informática no es tan lejana.

una gran cantidad de datos mediante del desarrollo de una aplicación OLAP.

El objetivo⁸ de una aplicación OLAP se denomina *business intelligence (BI)*, es decir, el proceso de convertir los datos empresariales en información útil para el negocio. Para conseguir aplicaciones de business intelligence es frecuente recurrir a grandes almacenes de datos (bases de datos) separadas de las bases de datos utilizadas normalmente por la empresa (*bases de datos transaccionales*). Estos almacenes de datos de denominan *datawarehouses (DW)*.

El motivo de separar las bases de datos transaccionales de los datawarehouses es:

- ? La estructura de información de los datawarehouses es diferente. En general los DW almacenan únicamente información agregada.
- ? La tecnología o al menos la configuración del sistema de gestión de base de datos de un datawarehouse es diferente
- ? Evitar que las peticiones de información en el datawarehouse puedan repercutir en el rendimiento del sistema transaccional.
- ? Los requerimientos de acceso por parte de los usuarios
- ? En ocasiones los datawarehouse consolidan información de distintos sistemas de bases de datos

Cuando una empresa dispone de un Datawarehouse, generalmente puede considerar que es capaz de responder a preguntas complicadas acerca de su gestión, a partir de los datos almacenados en el mismo.

Dependiendo de la complejidad y el tamaño del Datawarehouse, es posible que la estrategia de Business Intelligence de la empresa no haya sido efectiva, ya que la empresa no es capaz de encontrar la información necesaria en el mismo.

En este caso, el Business Intelligence aporta una nueva herramienta: el *Data Mining*.

Se denomina *Data Mining* al conjunto de herramientas software que permiten descubrir hechos a través del análisis prácticamente automatizado de los datos transaccionales / datawarehouse de una empresa.

Aunque *Business Intelligence* y todo el conjunto de procesos y tecnologías asociados conforman un tema complejo y que excede la intención y extensión del presente artículo, en uno de los capítulos el lector podrá encontrar una introducción rápida y práctica a este tipo de sistemas basados en base de datos.

⁸ En este punto, se está tratando el concepto de *Business Intelligence* de forma bastante restringida... todo sea por simplificar los conceptos. En realidad, como se verá más adelante, *Business Intelligence* incluye otros aspectos además de OLTP.

3. *¿Qué tipo de gestor de base de datos necesito?*

Como el lector comprobará, existe una amplia gama de sistemas de bases de datos disponibles en el mercado. La pregunta acerca de la idoneidad de cada uno de ellos para atender las necesidades de una empresa concreta es, por lo tanto, más importante.

Una primera distinción

Frecuentemente la necesidad de un sistema de base de datos surge de la informatización de procesos empresariales de complejidad media o baja, en empresas o grupos de usuario de pequeño tamaño... a veces como la evolución de un sistema de trabajo basado en aplicaciones ofimáticas (documentos de texto, hojas electrónicas, correos electrónicos...).

En estas ocasiones es frecuente que los usuarios comiencen a desarrollar una pequeña base de datos por ellos mismos (frecuentemente liderados por algún usuario avanzado)

Si las necesidades de base de datos se encuentran dentro unos ciertos límites de número de usuarios, tamaño de la base de datos, y si la aplicación de base de datos no es muy crítica, entonces la elección está clara: se necesita un sistema de gestión de base de datos personal.

Por el contrario, si el sistema de base de datos va a ser utilizado por un número importante de usuarios, a almacenar grandes volúmenes de información o a dar soporte a aplicaciones críticas, entonces sólo debería emplearse un sistema de gestión de base de datos basado en servidor.

Variables de decisión en sistemas de gestión de bases de datos personales

Los sistemas de gestión de bases de datos personales deberían elegirse en atención a su sencillez de utilización y potencia, ya que de esta manera será más fácil que los usuarios de la empresa lleguen a utilizarlo y sean más eficientes.

Sencillez de uso

La sencillez de uso es un aspecto de singular importancia, toda vez que los usuarios finales del sistema no van a ser expertos en informática.

Debe tratarse de un sistema fácil de instalar, dotado de un interfaz gráfico de usuario sencillo y que no cause problemas al usuario final, no debiendo requerir ningún tipo de mantenimiento por parte de personal cualificado.

Las tareas más complicadas suelen simplificarse mediante el empleo de **asistentes**.

Mención especial en este apartado merece la compatibilidad con otras aplicaciones ofimáticas, es decir la fácil integración de información entre la base de datos y procesadores de texto, hojas de cálculo, sistemas de correo electrónico... también al revés, y por supuesto facilitar que los

usuarios habituales de un determinado programa ofimático se sientan “como en casa” al utilizar la base de datos.

En los últimos tiempos se ha tratado de conseguir el objetivo de integración de la base de datos con otras aplicaciones al máximo mediante la inclusión de la base de datos en las denominadas “suites ofimáticas”... el efecto de esto ha sido realmente efectivo en términos de integración y sencillez de uso, pero demoledor en lo que se refiere a la competencia en el mercado... hoy en día el sector de las bases de datos personales se encuentra en una situación de práctico monopolio.

Funcionalidades

Las funcionalidades de los sistemas de bases de datos personales incluyen:

- ? Diseño y creación de tablas⁹ de base de datos.
- ? Realización de consultas a la base de datos
- ? Creación de formularios para entrada de datos
- ? Creación de informes a partir de la información almacenada en la base de datos
- ? Definición de la seguridad de la base de datos
- ? Programación de aplicaciones (ver más detalles en el siguiente apartado).

Desarrollo de aplicaciones sobre bases de datos personales

Aunque el propósito principal de una base de datos personal es el almacenamiento de información, sucede con frecuencia que este primer objetivo, al tener éxito, evoluciona de manera que comienzan a crearse consultas de acceso a los datos predefinidas¹⁰, formularios de entrada de datos, informes a partir de la información... el usuario va consiguiendo más y más experiencia y acaba por descubrir que la base de datos incorpora un lenguaje de programación que le permite realizar operaciones cada vez más complicadas.

En efecto, los sistemas de gestión de base de datos personales suelen incluir un completo entorno de desarrollo que puede utilizarse para crear

⁹ En un apartado posterior el lector encontrará una introducción no técnica al diseño de las bases de datos relacionales. Por el momento, basta con mencionar que las bases de datos relacionales almacenan la información en una o más tablas compuestas de campos y registros. Una tabla es un almacén de información de hechos empresariales del mismo tipo (por ejemplo, una tabla de clientes). Cada uno de los hechos empresariales (cada cliente) es un registro. Las propiedades de cada registro son un campo (por ejemplo, la tabla de clientes tendría campos para el nombre, dirección, etc.).

¹⁰ Por ejemplo, se crea una consulta para acceder a los 10 clientes más importantes del mes.

verdaderas aplicaciones de gestión... estos lenguajes son sencillos y orientados al acceso a datos.

Es preciso tener en cuenta que, por poderoso que sea el entorno de desarrollo, las limitaciones de una base de datos personales siguen siendo las mismas¹¹.

Una vez presentados los principales sistemas comerciales de bases de datos personales (ver "Sistemas de Gestión de Bases de Datos Personales"), en el apartado "Como elegir un sistema de gestión de base de datos personal" se establecerán una serie de directrices para elegir correctamente el que más se adapta a las necesidades de cada usuario.

Variables de decisión en sistemas de gestión de bases de datos basados en servidor

Si los requerimientos de la aplicación de base de datos son medianamente serios, la primera decisión debe ser utilizar un sistema de gestión de base de datos basado en servidor.

Esta decisión implica una serie de requerimientos que, aunque se describirán con algún detalle en un apartado posterior, es posible resumir en:

- ? Compra de una licencia de software costosa
- ? Utilización de un equipo informático (un servidor) dedicado al efecto
- ? Necesidad de contar con personal especializado para la administración de la base de datos
- ? Necesidad de diseñar cuidadosamente la aplicación de base de datos.

En una primera aproximación, es preciso tener en cuenta que, en muchos casos, las aplicaciones informáticas (ERP, CRM, sistemas integrados) utilizadas por la empresa determinan el sistema de gestión de base de datos que se debe utilizar.

En otras ocasiones, es posible que la dirección de la empresa se encuentre en la tesitura de poder elegir entre varias posibilidades, debido a que las aplicaciones de negocio soportan dos o más sistemas de gestión de base de datos.

Algún que otro lector se preguntará acerca de las diferencias entre los diferentes sistemas de bases de datos que existen en el mercado... siendo una pregunta razonable e importante, debe tenerse en cuenta, no obstante, que la tecnología de las bases de datos está tremendamente avanzada, con lo que es posible concluir que:

¹¹ Nota para lectores avanzados: la fronteras son cada vez menos nítidas, es posible utilizar una sistema de gestión de base de datos personal como Microsoft Access para desarrollar una aplicación en la que se accede un sistema de base de datos basada en servidor SQL Server... lo que hace la informática tan maravillosa, tan potente y a la vez tan complicada es la posibilidad de utilizar los distintos componentes que ofrece en infinidad de combinaciones distintas.

- ? Los fabricantes llevan años copiándose unos a otros, de forma que prácticamente todas las características de un producto pueden encontrarse más o menos resueltas por los productos de la competencia.
- ? El exceso de potencia que tanto el hardware como el software actual ofrecen hacen que, excepto para las aplicaciones más exigentes (en términos de usuarios, tamaños o rendimientos o funcionalidades), la elección de una base de datos en base a las características de rendimiento no sea fundamental.

La respuesta es que, los servidores de bases de datos son un producto que se ha "comoditizado"¹², evolucionando de tal manera que el elemento diferenciador más importante entre 2 productos de similar segmento es el **precio**.

En relación con el **número de usuarios concurrentes**¹³ o simultáneos, aunque la mayor parte de los servidores de bases de datos comerciales están diseñados para gestionar un número muy alto de usuarios concurrentes es importante mencionar que puede afectar al precio, debido a los modelos de licenciamiento de los software de base de datos.

Otro elemento a considerar a la hora de elegir un sistema de gestión de base de datos es la **disponibilidad**¹⁴. En caso de requerirse una disponibilidad elevada, es frecuente que los fabricantes obliguen a utilizar módulos adicionales o versiones más costosas de su gestor de base de datos.

¹² "Comoditizado", horrible neologismo necesario. Procede del inglés *commodity*, es decir, producto o servicio que no se diferencia tan apenas de la competencia, por lo que el factor de decisión suele ser el precio. Ocurre en productos sencillos lo muy estandarizados, o en aquellos muy maduros, como en el caso que nos ocupa.

¹³ *Usuarios concurrentes* hace referencia a lo usuarios conectados simultáneamente a la base de datos que requiere una determinada aplicación empresarial. En general, a mayor número de usuarios concurrentes, mayores requerimientos de estabilidad y prestaciones al servidor de base de datos. En los últimos tiempos este parámetro ha perdido importancia debido a la aparición de tecnologías como las *aplicaciones de 3 capas*, el *middleware* y el *connection pooling*.

¹⁴ Se define *disponibilidad* como el porcentaje de tiempo que el sistema es capaz de atender las peticiones de los usuarios sobre el total de tiempo potencial de uso del sistema. La disponibilidad de una aplicación empresarial debería estar entre el 97% y el 99,99%. La primera cifra implica la posibilidad de 22 horas de no disponibilidad (*downtime*) al mes y es lo mínimo aceptable para una solución empresarial. La segunda cifra (también conocida como los 4 nueves) es representativa para una aplicación de misión crítica y representa menos de 5 minutos de *downtime* cada mes. Garantizar un 99,99% de disponibilidad implica normalmente un alto grado de tolerancia a fallos y mecanismos de recuperación ante desastres que repercuten en grandes inversiones y altísimos costes de operación... En resumen, sólo las aplicaciones más importantes de las grandes empresas logran, normalmente justificar el coste asociado a la *alta disponibilidad*.

El **rendimiento**, o la capacidad la base de datos de ejecutar más número de operaciones por segundo, no suele ser característica diferenciadora de los modernos sistemas de base de datos. Se trata más de un aspecto que se soluciona con un hardware potente¹⁵, con una correcta operación del sistema¹⁶ y muchas veces con un buen diseño de la estructura de la base de datos¹⁷.

La **escalabilidad**, es la capacidad de manejar bases de datos de mayor tamaño. Todos los sistemas de base de datos actuales están preparados para gestionar VLDB (*Very Large Databases*, Bases de Datos Muy Grandes), y generalmente es cuestión de utilizar el hardware adecuado, y operar correctamente el sistema.

El **tipo de aplicación de base de datos** –recuérdese la distinción entre OLAP y OLTP- frecuentemente condiciona la elección del gestor de base de datos a emplear... no tanto por la diferencia en funcionalidades, sino por el precio. Una vez más, no todos los sistemas de base de datos incluyen en el precio base los módulos necesarios para realizar OLAP.

En resumen

Precio, precio, precio... en los casos en los que se puede elegir, la elección del gestor de base de datos más adecuado para la empresa suele ser una cuestión de precio.

No conviene caer en el error de confundir el precio de la licencia con el coste total. En la industria se suele hablar de TCO (*Total Cost of Ownership*, Coste Total de Propiedad) para hacer referencia a todos los costes asociados a la posesión y operación de un sistema de base de datos... más adelante existirá ocasión de presentar algunos modelos y comparaciones más adelante, en este mismo artículo.

Modelos de licenciamiento de los servidores de bases de datos

Los servidores de bases de datos se licencian en base al número de usuarios concurrentes permitidos. De esta forma, el comprador debe estimar el número de usuarios concurrentes que van a utilizar la base de datos y adquirir un número de licencias suficiente.

Esta forma de licenciamiento no es adecuada para grandes sistemas empresariales o para aplicaciones de Internet donde el número de usuarios es desconocido y potencialmente muy grande, con lo que la

¹⁵ Nótese que no todos los sistemas de base de datos funcionan en plataformas hardware de gama alta. Para más información consúltese un apartado posterior en el que se habla de la importancia del hardware.

¹⁶ Ver apartado posterior en lo que se refiere a la importancia de la operación del sistema y los planes de mantenimiento.

¹⁷ Aunque elemental, consúltese el apartado de diseño de bases de datos.

forma de calcular¹⁸ el precio de los servidores de base de datos en la actualidad es, generalmente por CPU¹⁹.

4. Sistemas de Gestión de Bases de Datos Personales

En la introducción se posicionaron las bases de datos personales (*desktop databases*) como complementos de las aplicaciones ofimáticas. Si se encuentra leyendo esta sección, posiblemente se preguntará por la misión que tienen en el contexto de lo que hoy entendemos por ofimática, pero también por forma en la que este software ha evolucionado desde los comienzos, no tan lejanos, de la informática hasta la actualidad.

Principales alternativas comerciales de bases de datos personales

Como se puede observar en la historia anterior, en el mercado existe un buen número de sistemas de gestión de base de datos personales. En esta comparativa se han elegido las versiones más populares, más evolucionadas y conocidas.

Limitaciones

Las bases de datos personales, herramientas potentes para usuarios o pequeños grupos de trabajo, tienen importantes limitaciones que desaconsejan su uso para ciertas aplicaciones:

- ? No son indicadas más que para uso personal o un número de usuarios concurrentes bajo. Como regla, por encima de los 5 usuarios concurrentes, se debería utilizar un servidor de base de datos²⁰.
- ? No deben utilizarse para gestionar datos de procesos críticos del negocio.

¹⁸ Todavía es posible adquirir licencias por usuario, pero suele ser más caro. No se utiliza para entornos de Producción, pero sí para Preproducción o Desarrollo.

¹⁹ *CPU*, del inglés *Central Processing Unit*. Es la manera de referirse al microprocesador, al cerebro de un ordenador.

²⁰ Aunque la documentación del fabricante puede indicar que el sistema de gestión de base de datos soporta cifras mucho más altas (por ejemplo, Microsoft Access tiene un límite teórico de 255 usuarios concurrentes), los sistemas de gestión de bases de datos personales no pueden llegar a este nivel de usuarios concurrentes debido a que la base de datos es, esencialmente, un archivo del sistema operativo. Por lo tanto, el acceso concurrente de los usuarios se realiza a un mismo archivo, gestionado directamente por el sistema operativo. Si además se tiene en cuenta que el tamaño de este tipo de archivos es frecuentemente muy grande (varios MB), debemos ser mucho más conservadores a la hora de establecer la limitación del número máximo de usuarios concurrentes. Excederlo llevará asociado los riesgos habituales de corrupción de un fichero por fallos en el sistema operativo (pérdida de datos).

- ? No deben utilizarse para gestionar bases de datos de gran tamaño (cientos de miles de registros o más de 50 MB de tamaño total).
- ? No deben utilizarse para almacenar datos donde la confidencialidad sea importante²¹.
- ? No deben utilizarse para aplicaciones en red extensa²².
- ? Las aplicaciones basadas en bases de datos personales no son aptas para entornos que requieren una disponibilidad alta.

Comparación de sistemas de gestión de bases de datos personales

En el mercado existe un número considerable de sistemas de gestión de bases de datos personales. A continuación figura una lista de las más populares:

- ? Microsoft Access 2002
- ? FileMaker Pro 6
- ? Paradox 10
- ? Lotus Approach
- ? Microsoft Visual FoxPro 7

Microsoft Access 2002

Microsoft Access es un potente gestor de base de datos personal sobre Microsoft Windows que puede servir tanto para el usuario no avanzado como para desarrollo de aplicaciones. Dispone de un amplio soporte por parte de la industria, es fácil encontrar documentación, formación y usuarios entrenados.

El precio de Microsoft Access 2002 cuando se compra como producto independiente es de unos 300 €. El paquete ofimático integrado Microsoft Office XP incluye Access 2002 y tiene un precio de 725 €.

²¹ Aunque las bases de datos personales permiten definir una política de seguridad, el almacenamiento en un fichero que puede ser fácilmente borrado o robado, desaconseja su utilización en entornos donde la seguridad sea importante. Adicionalmente, debe tenerse en consideración que el nivel de seguridad es bajo: como dicen los expertos en seguridad, únicamente para "almacenar información que se quiera mantener a salvo de la curiosidad de la hermana pequeña".

²² Acceder a cualquier pieza de información requiere la transmisión del fichero completo de la base de datos. Suponiendo una aplicación de base de datos en la que un usuario en Madrid necesita acceder al dato de ventas trimestral (una consulta cuyo resultado ocupa muy poco espacio)... si utilizamos una base de datos personal que ocupa 50 MB, y la oficina de Madrid y Barcelona están conectadas a través de una línea de 256 Kbps (bastante rápida) la consulta necesitaría transmitir la base de datos completa... aproximadamente 30 minutos ocupando la capacidad de transmisión de la línea completamente, obviamente no es aconsejable.

FileMaker Pro 6

Se trata de una base de datos muy popular en el entorno Apple Macintosh, aunque también existen versiones para Windows. Tiene un interfaz de usuario intuitivo que facilita las tareas más complicadas.

FileMaker Pro 6 está disponible en varias versiones: **FileMaker Pro 6** para el usuario final, **FileMaker Pro 6 Unlimited** que permite hacer aplicaciones de base de datos accesibles desde Internet / Intranet²³, **FileMaker Pro 6 Developer** para el desarrollador profesional²⁴.

El precio de FileMaker Pro es muy ajustado, comenzando a partir de los 300 € para la versión base (la versión Unlimited ronda los 1000 €, Developer 500 €).

Paradox 10

Paradox 10 se vende como un componente de Corel Wordperfect Office 2002, edición profesional.

Algo más complicada de utilizar que sus competidores y sin una política de producto clara por parte de su fabricante (actualmente Corel), fue una buena alternativa a dBase y Access en el pasado, pero ahora prácticamente no se utiliza. Las bases de datos Paradox existentes deberían migrarse a Access con urgencia.

Este producto se vende dentro del paquete ofimático integrado Corel Wordperfect Office 2002, que tiene un precio aproximado de 475 €.

Lotus Approach (Millennium Edition)

Lotus Approach es la base de datos de Lotus (fabricante de software adquirido por IBM), integrada con del paquete integrado ofimático Lotus Smartsuite.

Este producto se vende dentro del paquete ofimático integrado Lotus Smartsuite, que tiene un precio aproximado de 475 €.

²³ La publicación de bases de datos en Internet / Intranet es una necesidad reciente de las empresas surgida a raíz de la popularización de este tipo de aplicaciones.

Aunque un producto como Filemaker permite simplificar esta tarea, los problemas asociados a publicar una base de datos en una red pública como Internet aconsejan encomendar esta tarea a desarrolladores profesionales que además deben utilizar una tecnología más robusta (servidores de bases de datos).

²⁴ La gama de productos FileMaker Pro 6 incluye además **FileMaker Pro 6 Server** que es un servidor de bases de datos. Este producto no ha sido incluido en el presente documento debido a que, aunque FileMaker Pro es una conocida y reputada aplicación de base de datos personal, su presencia en el mundo de los servidores de base de datos es prácticamente nula.

Microsoft Visual Foxpro 7

Microsoft Visual Foxpro es actualmente la última versión de Foxbase.

Hoy en día, MS Visual Foxpro se trata más de un entorno de desarrollo para aplicaciones de base de datos, y por lo tanto su público objetivo son más los programadores de aplicaciones que los usuarios finales. Para estos usuarios avanzados, Foxpro ofrece un motor de base de datos de alto rendimiento, compatibilidad con XBase y un entorno de desarrollo potente y productivo.

El precio de la Microsoft Visual Foxpro 7 está en el entorno de los 900 €.

Como elegir un sistema de gestión de base de datos personal

Para finalizar con este apartado acerca de las bases de datos personales, una serie de directrices acerca de cómo seleccionar la que más se adapta a las necesidades de cada empresa o usuario.

Para elegir un sistema de base de datos personal, se sugiere seguir el siguiente proceso:

- ? **¿Está seguro de que una base de datos personal cubre sus necesidades?** En caso contrario, deberá elegir un sistema de gestión de base de datos basado en servidor (consulte el apartado "¿Qué tipo de gestor de base de datos necesito?").
- ? **¿Va a utilizar el sistema en equipos Macintosh?** En caso afirmativo la recomendación es utilizar FileMaker Pro 6.
- ? **¿Dispone de algún paquete integrado de ofimática en su empresa?**
 - o Si dispone de Lotus Smartsuite, considere la posibilidad de utilizar Lotus Approach.
 - o Si dispone de Corel Wordperfect Office, considere la posibilidad de utilizar Corel Paradox.
 - o Si dispone de Microsoft Office, considere la posibilidad de utilizar Microsoft Access.
- ? **¿Es la sencillez de uso para los usuarios finales el criterio prioritario?** En este caso, la alternativa que ofrece la mayor sencillez de uso, a costa de sacrificar las posibilidades más avanzadas es FileMaker Pro.
- ? **¿Necesita desarrollar aplicaciones complejas de base de datos?**
 - o Si necesita un entorno potente y con mucha flexibilidad para el desarrollador, considere la posibilidad de utilizar Microsoft Access o Microsoft Visual Foxpro.
 - o Si necesita un entorno que permita a usuarios no expertos desarrollar aplicaciones sobre base de datos, las alternativas más adecuadas son Microsoft Access y FileMaker Pro.

- ? **¿Necesita un front-end para una aplicación de base de datos cliente-servidor?** En este caso, considere la posibilidad de utilizar Microsoft Access.
- ? **¿Está buscando la base de datos más extendida y popular del mercado?** Si lo que busca es compatibilidad, base instalada y facilidad de obtener formación, la decisión más acertada es Microsoft Access.
- ? **¿Necesita compatibilidad con DB2 o Lotus Domino / Notes?** En este caso, la alternativa más adecuada es Lotus Approach.
- ? **¿Desea disponer de una serie de aplicaciones de base de datos ya preparadas?** En este caso, las aplicaciones listas para utilizar que se distribuyen con FileMaker Pro pueden ser de utilidad. Microsoft Access también permite descargar algunas aplicaciones ya preparadas desde el sitio web de Microsoft (<http://office.microsoft.com>).
- ? **¿Desea una base de datos para una aplicación de Internet?** En este caso, aunque los productos Microsoft Access, FileMaker Pro, Lotus Approach y Paradox tienen características de publicación en Internet, debería considerar seriamente la posibilidad de utilizar un servidor de base de datos: no se arrepentirá.

Recomendación general

Para la gran mayoría de los lectores, dada la gran popularidad de la plataforma compuesta por PC / Microsoft Windows / Microsoft Office, la alternativa más acertada por sencillez de uso, potencia y funcionalidad es Microsoft Access.

5. Sistemas de Gestión de Bases de Datos basados en Servidor

Quizá una de las decisiones más complicadas que deben tomar los directores de sistemas es la elección del sistema de gestión de base de datos para una aplicación transaccional. Si la base de datos de clientes "se cae", se interrumpe el proceso de atención el cliente, ventas... lo mismo puede decirse de la práctica totalidad de los procesos empresariales que, hoy en día, están basados en mayor o menor medida en una base de datos.

En esta comparativa se analizarán con cierto detenimiento las características más importantes de los sistemas de gestión de bases de datos más importantes del mercado (datos de *Gartner 2001*):

- ? Oracle 9i, que cuenta con una cuota de mercado del 33,8%.
- ? IBM DB2 Universal Database 8.1, con una cuota de mercado del 30,1%

- ? Microsoft SQL Server 2002, con una cuota de mercado del 14,9%

Juntos, estos 3 productos representan el 90% del mercado de bases de datos (8.800 millones USD en 2001). Otros productos a destacar, pero que no serán incluidos en esta comparativa son:

- ? Informix Dynamic Server: este producto fue adquirido por IBM en Abril de 2001, y hoy continúa en el catálogo de productos de IBM.
- ? Sybase Adaptative Server Enterprise 12.5
- ? Productos open source²⁵ como MySQL o PostgreSQL.

Características comunes de los 3 productos analizados

Cualquiera de las tres bases de datos analizadas son productos maduros, potentes y ricos en funcionalidad, utilizados en aplicaciones críticas en muchas empresas del mundo. Todos ellos son capaces de gestionar grandes volúmenes de información y un elevado número de usuarios concurrentes. Todos garantizan la integridad y consistencia de la información que requiere una base de datos transaccional de misión crítica. Pueden ser utilizados tanto para aplicaciones transaccionales (OLTP) como para aplicaciones de análisis de datos (OLAP).

Los fabricantes de estos productos se encuentran entre las mayores empresas del mundo, y ofrecen un soporte excelente a los mismos, así como garantía de estabilidad y continuidad.

Cualquiera de ellos es una elección excelente como base de datos de una empresa.

Por lo tanto, como se mencionó en el apartado "Variables de decisión en sistemas de gestión de bases de datos basados en servidor", si no se requieren características extremadamente especializadas que puedan estar disponibles únicamente en uno de ellos, la gran diferencia de estos productos es su facilidad de administración y –de forma muy notable- el precio.

IBM DB2 Universal Database 8.1

DB2 Universal Database 7.2 es un excelente gestor de base de datos, que une a la potencia tradicional del producto nuevas características de facilidad de uso, incorpora las últimas tecnologías, constituye una sólida plataforma OLTP con una cierta potencia OLAP en su precio base y posibilidad de extenderse... todo esto a un precio muy competitivo.

SQL Server 2000

Microsoft SQL Server 2000 es una excelente base de datos sobre entorno Windows a precio muy competitivo, fácil de utilizar y con una gigantesca cantidad de funcionalidades incluidas en el precio. Si su empresa necesita todas las características de una base de datos de última generación, a buen precio y no requiere las prestaciones²⁶ típicas de los servidores UNIX de gama alta o mainframe, SQL Server 2000 es la mejor elección.

²⁵ Open Source: software con licencia abierta, gratuito.

Oracle 9i Enterprise

Oracle es la mejor solución de base de datos, capaz de resolver cualquier necesidad. El inconveniente es la elevada inversión que supone, y la dificultad de administración. Si dispone del presupuesto suficiente para hacer frente a la inversión y para contratar los servicios de un experimentado DBA, y necesita una solución de base de datos realmente potente, Oracle 9i Enterprise es su elección.

Comparación de Precios de los Sistemas de Bases de Datos Analizados

La comparación de precios entre los distintos productos de base de datos va a realizarse tomando como referencia los datos de SQL Server 2000, debido a que es el producto más competitivo de la comparación.

Se han utilizado en todo momento datos públicos, precios en dólares USA (USD) tomados del sitio web de Microsoft.

<http://www.microsoft.com/sql/evaluation/compare/default.asp>

Hay una gran diferencia de precios entre los productos de base de datos de Oracle y Microsoft. El aspecto más importante a resaltar es que Microsoft incluye las herramientas de OLAP en el precio base, mientras que en Oracle este módulo debe adquirirse –y pagarse- por separado. Como ejemplo, una empresa que necesita una licencia de base de datos para Business Intelligence, en un servidor de 8 CPU, debería invertir 640.000 USD en la solución de Oracle, en contraste con 159.992 USD en la solución Microsoft.

Entre ambos extremos del abanico de precios se sitúa IBM DB2. Utilizando la misma comparación anterior, una licencia para la instalación de IBM DB2 para BI sobre un servidor de 8 CPU tendría un coste de 323.000 USD.

6. Sistemas de Gestión de Bases de Datos para “Business Intelligence”

Business Intelligence

El término *Business Intelligence (BI)* se utiliza –quizá con demasiada facilidad- para designar a la capacidad, gracias a un sistema informático, de comprender y conseguir inteligencia de un negocio mediante el análisis de las operaciones del mismo, esto es, mediante el análisis de

²⁶ Hay que tener un gran cuidado al comparar las prestaciones de las bases de datos. En este documento existe un apartado exclusivamente dedicado a desmitificar las cifras de los *benchmark* de transacciones. La conclusión más útil para el lector es que los tres sistemas de bases de datos comparados (SQL Server 2000, Oracle 9i, DB2 UDB 7.2) son sistemas robustos y bien diseñados capaces de sacar el máximo partido al hardware sobre el que corren. No obstante, debe tenerse en cuenta que tanto DB2 como Oracle pueden correr en potentes plataformas UNIX de gama alta o mainframe (con precios de varios millones de euros)... si necesita esa escalabilidad, entonces debe elegir alguno de estos sistemas.

las operaciones.

Históricamente, el término *Business Intelligence* se ha aplicado también a sistemas de soporte de las decisiones (*Decision Support Systems*), análisis, *data warehousing*...

El análisis se inicia mediante la consolidación de información relativa a ventas, marketing, costes, servicio, comportamiento de los clientes ante las actividades de la empresa, y de los proveedores de la empresa.

Posteriormente esta información se almacena y organiza con una estructura que facilite su análisis.

Este almacenamiento y organización se realiza –obviamente– en un sistema de base de datos... la estructura, no obstante, poco tiene que ver con la de un sistema transaccional.

El análisis se realiza con una serie de técnicas como la elaboración de informes (*reporting*), consultas (*query's*), OLAP y *data mining*.

Infraestructura de Business Intelligence

La infraestructura de *Business Intelligence* provee a la empresa de la capacidad de implementar una serie de procesos orientados a *Business Intelligence*.

La infraestructura de *Business Intelligence* debe permitir las siguientes capacidades:

- ? **Data warehouse:** bases de datos relacionales y multidimensionales.
- ? **OLAP** (OnLine Analytical Processing), es la tecnología que permite a los usuarios de una empresa examinar y analizar los datos transaccionales para comprender la situación del negocio.
- ? **Data Mining**, análisis automatizado de los datos para descubrir información en los mismos.
- ? **Extract Transfer Load (ETL)**, es decir capacidad para alimentar periódicamente el data warehouse con los datos obtenidos de los almacenes de datos relacionales.

Plataformas analizadas

En esta comparación se van a analizar las plataformas de Business Intelligence de los mismos fabricantes de software de bases de datos que se compararon anteriormente:

- ? **Microsoft:** gracias a una política de precios muy agresiva, SQL Server 2000, un recién llegado al mundo del BI es ahora mismo la plataforma cuantitativamente más extendida y con mayor crecimiento.
- ? **Oracle:** el líder en sistemas de bases de datos transaccionales ha consolidado en su motor de base de datos Oracle 9i las capacidades de BI que anteriormente constituían módulos separados. Sigue siendo la solución segura, de gama alta... con alto coste.

- ? **IBM:** DB2 es una plataforma de BI potente, gracias a la integración de un producto específico: Hyperion.

Plataforma de Microsoft

Microsoft, con su producto de base de datos SQL Server 2000 Enterprise ha *acercado a todos los bolsillos* las posibilidades de los almacenes de datos para BI.

SQL Server 2000 aporta la plataforma de BI necesaria para datawarehousing, base de datos relacional y multidimensional, OLAP y ETL.

El producto es escalable, fácil de instalar y fácil de utilizar... y permite un rápido retorno de la inversión.

Es la elección adecuada para compañías de tamaño grande o mediano, que quieran un excelente motor de BI a un precio muy razonable (la plataforma de BI de Microsoft para un servidor con 4 CPU tiene un precio de lista de 19.999 USD).

Microsoft incluye en su versión SQL Server 2000 Enterprise Edition todos los componentes necesarios para *Business Intelligence*.

Plataforma de IBM

La plataforma de BI de IBM está basada en su base de datos DB2 Universal Database. DB2 aporta la capacidad de data warehousing, base de datos relacional, OLAP y ETL.

La capacidad de OLAP y ETL las proporciona un producto denominado *DB2 OLAP Server*, que es un producto de un tercero –Hyperion– adaptado para DB2 y comercializado por IBM.

IBM vende como opción adicional cada módulo de la plataforma de BI. De esta manera, la plataforma total puede resultar más cara:

- ? DB2 UDB 8.1 Enterprise Edition tiene un precio de 25.000 USD por cada CPU. Incluye la capacidad básica de ETL de Data Warehouse Center.
- ? DB2 OLAP Server tiene un precio de 28.000 USD por cada servidor y 1.500 USD por cada usuario adicional.
- ? DB2 Intelligent Miner, en su versión más avanzada tiene un precio de 75.000 USD por cada CPU.
- ? DB2 Warehouse Manager tiene un precio de 10.600 USD por cada CPU.

Plataforma de Oracle

La base de la plataforma de BI de Oracle es la base de datos Oracle 9i.

Las capacidades OLAP las aporta *Oracle 9i OLAP*, y data mining *Oracle 9i Data Mining*, mientras que ETL está soportado por el producto *Oracle 9i Warehouse Builder*.

Los precios de los componentes son los siguientes:

- ? Oracle 9i Enterprise, 40.000 USD por procesador.
- ? Oracle 9i OLAP, 20.000 USD por procesador.
- ? Oracle 9i Data Mining, 20.000 USD por procesador.
- ? Oracle 9i Warehouse Builder, 5.000 USD por usuario.

En conjunto, se trata de la plataforma más cara, con un precio 5 veces mayor que la plataforma de Microsoft.

Comparación de las tres plataformas

Cualquiera de las tres plataformas analizadas es una excelente base de datos relacional (ver más información en un capítulo anterior).

Comparación de características OLAP

La funcionalidad **OLAP** de las tres plataformas es bastante similar y, en general, satisfactoria. La implementación más madura es la de SQL Server 2000, incorporada en el producto en 1998. Oracle incorporó esta capacidad en 2001, y el producto de IBM (versión OEM de Hyperion) se comercializó por primera vez en 1999, aunque el producto de Hyperion es muy maduro y uno de los gestores OLAP más potentes del mercado.

Comparación de características Data Mining

En lo referente a **Data Mining**, las funcionalidades de SQL Server 2000 son las más fáciles de utilizar –basadas en asistentes- y es además la solución de menor coste, aunque el producto es todavía muy joven en esta tecnología. Prácticamente lo mismo se puede decir la implementación de Oracle 9i Data Mining, aunque dispone de más algoritmos de Data Mining que SQL Server 2000.

IBM DB2 Intelligent Data Miner es un producto mucho más maduro que los de Microsoft u Oracle. Compuesto de tres componentes (DB2 IM Visualization, DB2 IM Modeling y DB2 IM Scoring), se trata de un producto potente, con gran cantidad de algoritmos de data mining... su inconveniente es la complejidad.

ETL

En lo que se refiere a **ETL**, la plataforma de Microsoft es la más rica en funcionalidades y sencilla de utilizar, gracias a una serie de asistentes y herramientas gráficas. El proceso de carga de datos es fácilmente automatizable, y suficientemente flexible como para poder acometer cualquier escenario que se presente.

Oracle incorpora a la plataforma un potente gestor de ETL, aunque no dispone de alguna de las características y funcionalidades que se encuentran en SQL Server.

La plataforma de IBM, basada en DB2 es la más compleja, ya que hay funcionalidades incluidas en los módulos base, funcionalidades añadidas en módulos adquiridos a terceros... en resumen, una plataforma rica y potente en cuanto a ETL, pero excesivamente compleja de utilizar.

Conclusión

Cualquiera de las tres plataformas –Microsoft, IBM y Oracle- es capaz de satisfacer las necesidades de BI de la mayor parte de las empresas.

La plataforma de Microsoft es la elección perfecta sobre sistema operativo Windows. Con un coste muy ajustado, es apta para cualquier tarea, puede tener como limitaciones la gestión de grandísimos volúmenes de información, o unas capacidades de data mining todavía limitadas.

La plataforma de Oracle, con un alto coste, es una excelente base de datos relacional, pero las características OLAP y data mining no son todavía consistentes con el resto del producto.

Por último, la plataforma de IBM es la mejor solución para data mining, aunque en conjunto es una plataforma compleja y de alto coste, debido principalmente a la integración de productos de terceros.