

Selección de Bases de Datos

Administración de recursos
Ing. En sistemas de Información
FRBA - UTN - ARGENTINA

Bases de datos

Que se busca con los SGBD

- Abstracción de la información
- Independencia
- Redundancia minima
- Consistencia
- Seguridad
- Integridad
- Respaldo y Recuperación
- Control de concurrencia
- Tiempo de Respuesta

Bases de datos

Los motores seleccionados para la comparación son:



Bases de datos

Aspectos funcionales

1. Distribución y replicas
2. Conurrencia y niveles de aislamiento
3. Tipos de índices
4. Backups
5. Vistas materializada
6. Optimización de consultas
7. Particionamiento

Bases de datos

Aspectos funcionales

1. Distribución y replicas

La replicación es un conjunto de herramientas para copiar y distribuir datos y objetos de bases de datos de una base de datos a otra y, a continuación, sincronizar las diferentes bases de datos para mantener la coherencia.

Los tipos básicos de replicación son:

- **replicación de instantáneas**
- **replicación transaccional**
- **replicación de heterogénea**

Bases de datos

Aspectos funcionales

1. Distribución y replicas

	Replicación
	
	
	

Bases de datos

Aspectos funcionales

2. Conurrencia y niveles de aislamiento

El mejor modelo de concurrencia utilizado hoy en día por los sistemas de base de datos relacionales es el sistema de control de concurrencia por múltiples versiones.

MVCC provee de cada usuario conectado con la base de datos una “foto” de la base de datos para esa persona. Ningun cambio realizado será visualizado por otros usuarios de la base de datos hasta que la transacción haya sido confirmada.

Bases de datos

Aspectos funcionales

2. Conurrencia y niveles de aislamiento

	Bloquea a	Lectura	Escritura
Conurrencia con multiversionado	Lectura	NO	NO
Locks tradicionales	Escritura	NO	SI
	Lectura	NO	¿?
	Escritura	SI	SI

Bases de datos

Aspectos funcionales

2. Conurrencia y niveles de aislamiento

	MVCC
	
	 (InnoDB o Falcon)
	

Bases de datos

Aspectos funcionales

2. Tipos de índices

Índices Hash

Arbol B

Índices bitmap : Son índices que consumen menor almacenamiento respecto del Btree. En un índice bitmap, en vez de una lista de rowids, se crea un mapa de bits para cada valor clave del índice.

Árbol R-/R+: Es un tipo de árbol similar al Árbol B, que es el tipo de árbol utilizado en la mayoría de los índices de las bases de datos, pero optimizado para acceder a información multidimensional, generalmente utilizado para coordenadas geográficas.

Índices parcial : es aquel que se aplica a no toda las filas de una tabla sino a aquellas que cumplen una determinada condición, permitiendo generar un índice para un propósito más específico que un índice común, y evitando indexar filas que no son de interés.

Bases de datos

Aspectos funcionales

3. Tipos de índices

Índices de expresión : Los índices de expresión, o índices basados en funciones, se basan en el indexado de una función o expresión aplicada sobre las columnas del registro, y es eso lo que se indexa. Por ej. índice por **upper(col1)**.

Fulltext: Índice para búsqueda de texto completo. Permite la búsqueda de n palabras combinadas.

Bases de datos

Aspectos funcionales

3. Tipos de índices

	Árbol B	Árbol R-/R+	Hash	Expresión	Parcial	Full text	Bitmap
							
		MyISAM 				MyISAM 	
							

Bases de datos

Aspectos funcionales

4. Backups



Tipos de backups

- Online
- Offline
- Completo
- Tablespace
- Datafile
- Archivo de Control
- Redo Log Archivado

Herramientas de backups

- Recovery Manager (RMAN)
- Oracle Enterprise Manager
- Oracle Data Pump
- Export/Import

Bases de datos

Aspectos funcionales

4. Backups



Tipos de backups

- Online
- Offline
- Completo
- Diferencial
- Log de Transacciones
- Archivo / Filegroup
- Archivo diferencial

Herramientas de backups

Desde el SQL Server Management Studio (herramienta provista por el fabricante con todas las versiones) se pueden realizar todos los tipos de backups antes mencionado

Bases de datos

Aspectos funcionales

4. Backups



Tipos de backups

- Lógico
- Físico
- Online
- Offline
- Snapshot
- Full
- Incremental

Herramientas de backups

Herramientas visuales de terceros

Bases de datos

Aspectos funcionales

5. Vistas materializadas

Una vista es una consulta guardada en la base de datos, que permite reutilizarse como si reutilizamos una consulta ya hecha. En este caso lo que se guarda no es el resultado de la consulta, sino las relaciones utilizadas para poder volver a ejecutarla.

Una vista materializada es aquella donde la consulta se guarda en una tabla temporal y se actualiza para mantenerse consistente.

	Vistas Materializadas
	
	
	

Bases de datos

Aspectos funcionales

6. Optimización de consultas

La optimización de consultas es el proceso de selección del plan de evaluación de las consultas más eficiente de entre las muchas estrategias generalmente disponibles para el procesamiento de una consulta dada, especialmente si la consulta es compleja.

Por el contrario, se espera que el sistema cree un plan de evaluación de las consultas que minimice el coste de la evaluación de las consultas. Aquí es donde entra en acción la optimización de consultas.

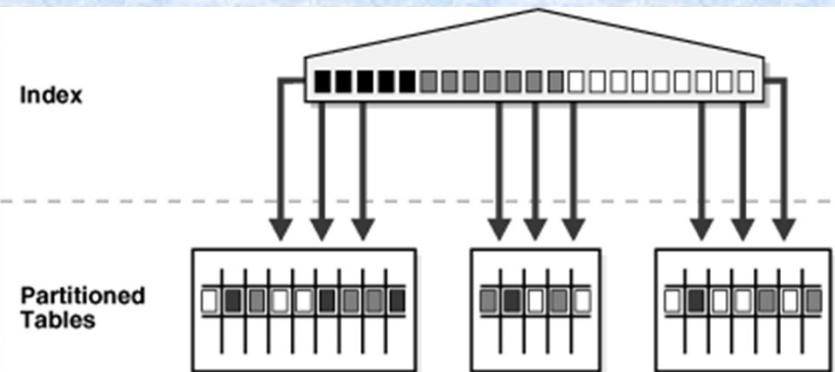
- Optimización de consultas

	Costos	Hints
	 (InnoDB)	
		
		

Bases de datos

Aspectos Funcionales

7. Particionamiento



El particionamiento es una técnica para dividir tablas grandes en tablas pequeñas más manejables, y permite que se manipulen desde el punto de vista de la aplicación de una manera uniforme.

En general, las técnicas de particionado mejoran la performance ya que una consulta solo necesita acceder a un subconjunto de las particiones de una tabla. Es conveniente para tablas grandes, y para tablas que tienen un conjunto de datos históricos y otros actuales, en las cuales se pueden colocar los datos históricos en una partición y los actuales en otra.

Bases de datos

Aspectos Funcionales

7. Particionamiento

	Por Rango	Por Hash
		
		
		

Bases de datos

Comparación de rendimientos

TPC-H es un *punto de referencia* apoyo a las decisiones que consiste en una cadena de consultas ad-hoc (Cuando el usuario real genera consultas de acuerdo a sus necesidades de cotejar la información de una forma nunca vista, y con métodos que conducen al descubrimiento de lo que la demanda), dirigida a la empresa y para la modificación de datos simultáneamente.

Comparación de rendimientos

System	Scale Factor	QphH	Price/QphH	Database
HP ProLiant DL380 G6	100	51.422	1.07 USD	Microsoft SQL Server 2008 Enterprise x64 Edt SP1
Kickfire Appliance 2300	100	49.228	0.70 USD	MySQL 5.1 w/ Kickfire database
HP ProLiant DL585 G7	300	107.561	1.08 USD	Microsoft SQL Server 2008 R2 Enterprise Edition
Kickfire Appliance 2400	300	54.895	0.89 USD	MySQL 5.1 w/ Kickfire database
HP BladeSystem c-Class 128P RAC	1000	1,166,976	5.42 USD	Oracle Database 11g Release 2 Enterprise Edt.
HP ProLiant DL785 G6	1000	81.514	2.90 USD	Microsoft SQL Server 2008 Enterprise Edition
Sun SPARC Enterprise M9000	3000	188.229	20.19 USD	Oracle Database 11g Release 2 Enterprise Edt.
HP ProLiant DL980 G7	3000	162.601	2.68 USD	Microsoft SQL Server 2008 R2 Enterprise Edition
Unisys ES7000 Model 7600R Enterprise Server(16s)	3000	102.778	21.05 USD	Microsoft SQL Server 2008 R2 Datacenter Edition
HP Integrity Superdome/Dual-Core Itanium/1.6 GHz	10000	208.457	27.97 USD	Oracle Database 11g Enterprise Edition
Unisys ES7000 Model 7600R Enterprise Server(16s)	10000	80.172	18.95 USD	Microsoft SQL Server 2008 Enterprise x64 Edition

Bases de datos

Aspectos técnicos

			
Tamaño maximo de la base	Máximo S.O.	Máximo S.O.	Máximo S.O.
Memoria máxima	Máximo del S.O.	Windows 2 GB	Máximo del S.O.
Cantidad máxima de procesadores	Ilimitado		Ilimitado
SO sobre los que funciona	Windows	Windows, Mac OS X, Linux, BSD, UNIX, AmigaOS, Symbian, z/OS 1	Windows, Mac OS X, Linux, UNIX

Bases de datos

Aspectos diferenciales

Tecnología Oracle Flashback

Flashback query es una nueva característica desde oracle 9i para recuperar de errores humanos. Permite a los usuarios ver el estado del dato, trasladado a un punto o instante anterior en el tiempo sin requerir de cambios estructurales a la base de datos. Usando la opción de flashback, usted puede ejecutar consultas contra un snapshot de datos consistentes a un punto en el tiempo especificando la hora del sistema o un SCN Numero de Cambio del Sistema.

Limitaciones

Aunque las ventajas de Flashback son muchas, tiene algunas limitaciones:

1. Si la corrupción de la base de datos se da a nivel físico, no se puede usar Flashback como medio de recuperación.
2. Es un producto pensado para volver atrás cambios no deseados realizados en un pasado cercano. **Cuanto tiempoatrás uno puede ir, es configurable, pero limitado según la cantidad de transacciones que se realicen en la base y el espacio disponible para guardar las transacciones realizadas.**

Bases de datos

MySQL

Característica	InnoDB	MyISAM
ACID - MVCC - Restricciones Integridad		
Fulltext Search		
Bloqueo a nivel tabla		
Máximo tamaño por tabla de 64 TB		

Bases de datos

Aspectos diferenciales

Tecnología Oracle Flashback

- Flashback Query
- Flashback Table
- Flashback Data Archive
- Flashback Transaction
- Flashback base de datos mediante RESETLOGS
- Flashback Drop

Bases de datos

Costos y licencias



Database Products	Named User Plus	Software Update License & Support	Processor License	Software Update License & Support
Oracle Database				
Standard Edition One	180	39.60	5,800	1,276.00
Standard Edition	350	77.00	17,500	3,850.00
Enterprise Edition	950	209.00	47,500	10,450.00
Personal Edition	460	101.20	-	-
Lite Mobile Server	-	-	23,000	5,060.00
Lite Client	60	13.20	-	-
 <i>Enterprise Edition Options:</i>				
Real Application Clusters	460	101.20	23,000	5,060.00
Real Application Clusters One Node	200	44.00	10,000	2,200.00
Active Data Guard	200	44.00	10,000	2,200.00
Partitioning	230	50.60	11,500	2,530.00
Real Application Testing	230	50.60	11,500	2,530.00
Advanced Compression	230	50.60	11,500	2,530.00
Total Recall	120	26.40	5,800	1,276.00
Advanced Security	230	50.60	11,500	2,530.00
Label Security	230	50.60	11,500	2,530.00
Database Vault	460	101.20	23,000	5,060.00
OLAP	460	101.20	23,000	5,060.00
Data Mining	460	101.20	23,000	5,060.00

Bases de datos

Costos y licencias



	Per Processor	Server/CAL	
Editions	License	License	User/Device CAL
SQL Server Datacenter*	\$ 54,990.00	NA	NA
SQL Server Enterprise	\$ 27,495.00	\$ 8,592.00	NA
SQL Server Standard	\$ 7,171.00	\$ 898.00	NA
SQL CAL	NA	NA	\$ 164.00
SQL Server Workgroup		\$ 3,743.00 \$730.00 (includes 5 Workgroup CALs)	NA
SQL Workgroup CAL	NA	NA	\$ 148.00
SQL Server Web*	\$3,500.00 (or \$15 per month in SPLA)	NA	NA

Precios en U\$S

Bases de datos

Características a tener en cuenta al momento de evaluar.

Aspectos Funcionales

- Rendimiento en sistemas similares, transacciones por minuto, máxima concurrencia, benchmarks
- Distribución y Replica
- OLAP y Datamining
- Niveles de aislamiento, concurrencia, bloqueos
- Administración de usuarios y permisos
- Manejo del almacenamiento
- Tipo de Índices. Indización de documentos , full text search
- Backups, recuperación , mantenimiento, tipos, herramientas
- Auditorias

Aspectos Técnicos

- S.O donde funcionan
- Arquitecturas de HW sobre las que funciona
- Herramientas de administración de memoria
- Tamaño máximo de una base de datos

Bases de datos

Características a tener en cuenta al momento de evaluar.

Características Diferenciales

- Facilidad para el desarrollador, herramientas
- Herramientas de diseño
- Acceso a datos
- Manejo de tipos de datos no estándar.
- Creación de tipos de datos complejos
- Manejo de herencia

Características de Servicio

- Capacitación
- Soporte en línea
- Servicios que ofrece el proveedor

Aspectos Económicos

- Costos de licencias
- Tipos de Licencias
- Representación en el mercado

Bases de datos – NO SQL

Ventajas

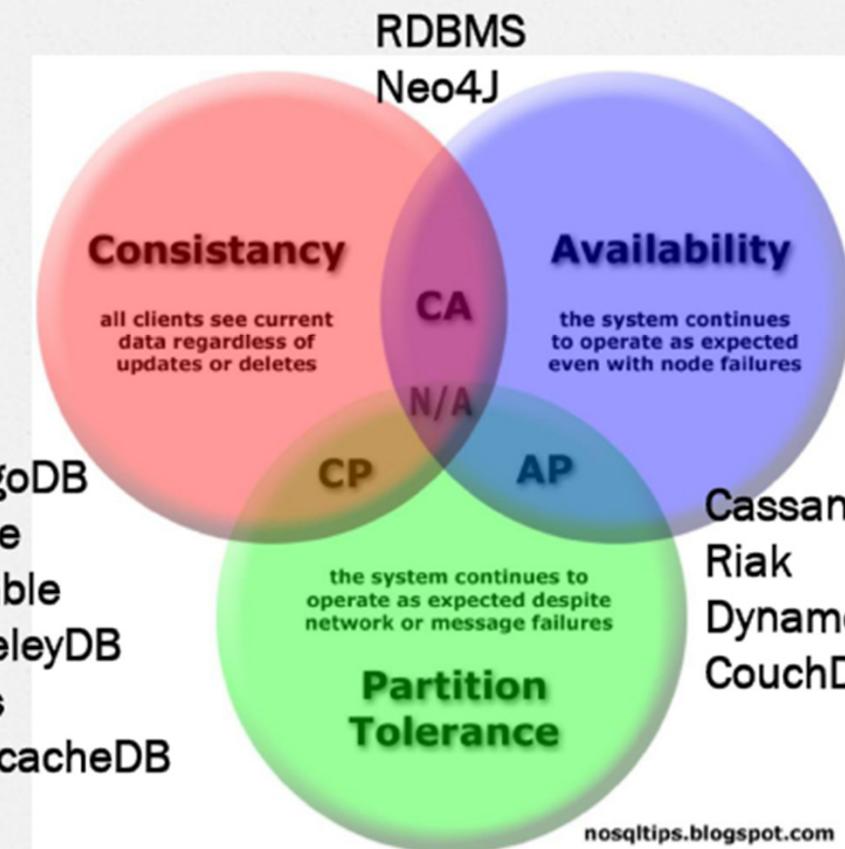
- Resuelven problemas de escalabilidad horizontal
- Manejo de enorme cantidades de datos
- Fácil de escalar
- Excelentes tiempos de respuesta en almacenamiento y recuperación

Desventajas

- No cumplen con ACID
- Utilizan mayormente memoria en vez de disco como la principal ubicación de escritura
- Falta de madurez
- Falta de estándares

Bases de datos – NO SQL

Teorema de CAP



Bases de datos – NO SQL

Clasificación

Clave-Valor: los datos se almacenan en pares del tipo clave-valor. El valor es un dato de tipo blob. Ej. Riak, Dynamo, Azure, Redis

Column Family: permiten almacenar claves mapeadas a valores y esos valores agrupados en múltiples familias de columnas siendo cada columna un mapa de datos. Ej. Cassandra, HBase, Amazon SimpleDB

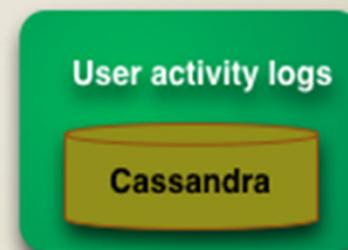
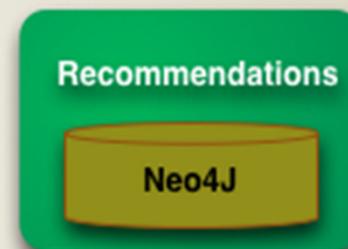
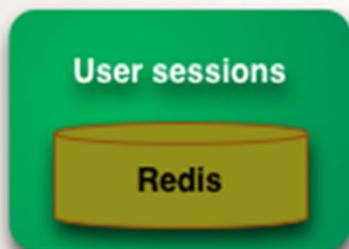
Basadas en documentos: La base de datos almacena y recupera documentos que pueden estar en XML, JSON o BSON. Ej. MongoDB, Couchbase, CouchDB, Lotus Notes, Oracle NoSQL Database

Basadas en grafos: permiten almacenar entidades y relaciones entre esas entidades. Tanto los nodos como las relaciones tienen sus propiedades asociadas. Ej Neo4J, InfiniteGraph, OrientDB, FlockDB

Bases de datos – NO SQL

Persistencia Políglota

Speculative Retailers Web Application



Bases de datos – Business Intelligence

“Business Intelligence (BI) es un término paraguas que abarca **los procesos, las herramientas, y las tecnologías** para **convertir datos en información, información en conocimiento** y planes para conducir de forma eficaz las actividades de los negocios. BI abarca las tecnologías de datawarehousing los procesos en el ‘back end’, consultas, informes, análisis y las herramientas para mostrar información (estas son las herramientas de BI) y los procesos en el ‘front end’.”

The datawarehouse Institute

Bases de datos – Business Intelligence

Objetivo: asistir en la toma de decisiones y, posteriormente descubrir conocimiento desconocido para la organización hasta el momento.

Los beneficios que se pueden obtener a través del uso de BI son:

- **Beneficios tangibles:** reducción de costes, generación de ingresos, reducción de tiempos para las distintas actividades del negocio.

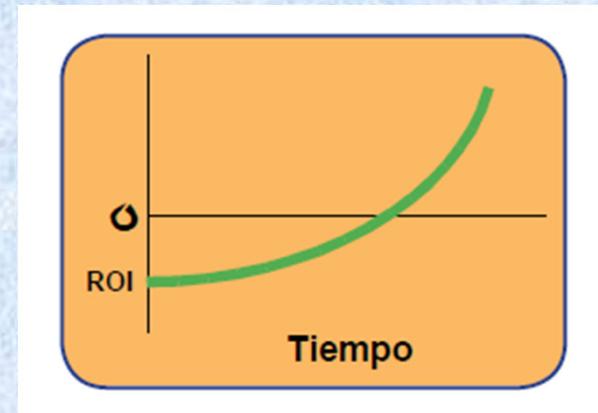
Bases de datos – Business Intelligence

- ***Beneficios intangibles:*** el hecho de que tengamos disponible la información para la toma de decisiones hará que más usuarios utilicen dicha información para tomar decisiones y mejorar nuestra posición competitiva.
- ***Beneficios estratégicos:*** Todos aquellos que nos facilitan la formulación de la estrategia, es decir, a qué clientes, mercados o con qué productos dirigirnos.

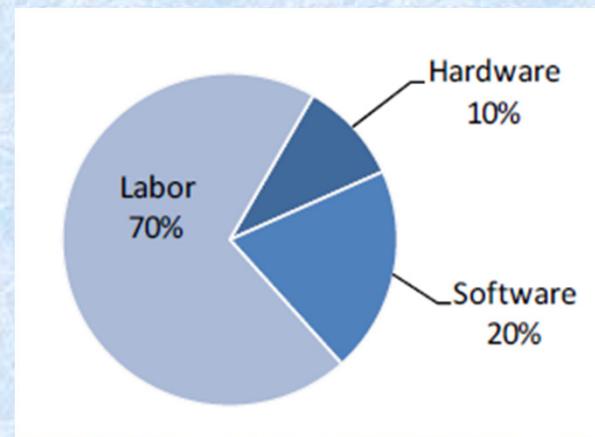
Bases de datos – Business Intelligence

- *Calculo del ROI en proyectos de BI:*

$$ROI = \frac{\text{Valor para el negocio}}{\text{Coste del proyecto}}$$

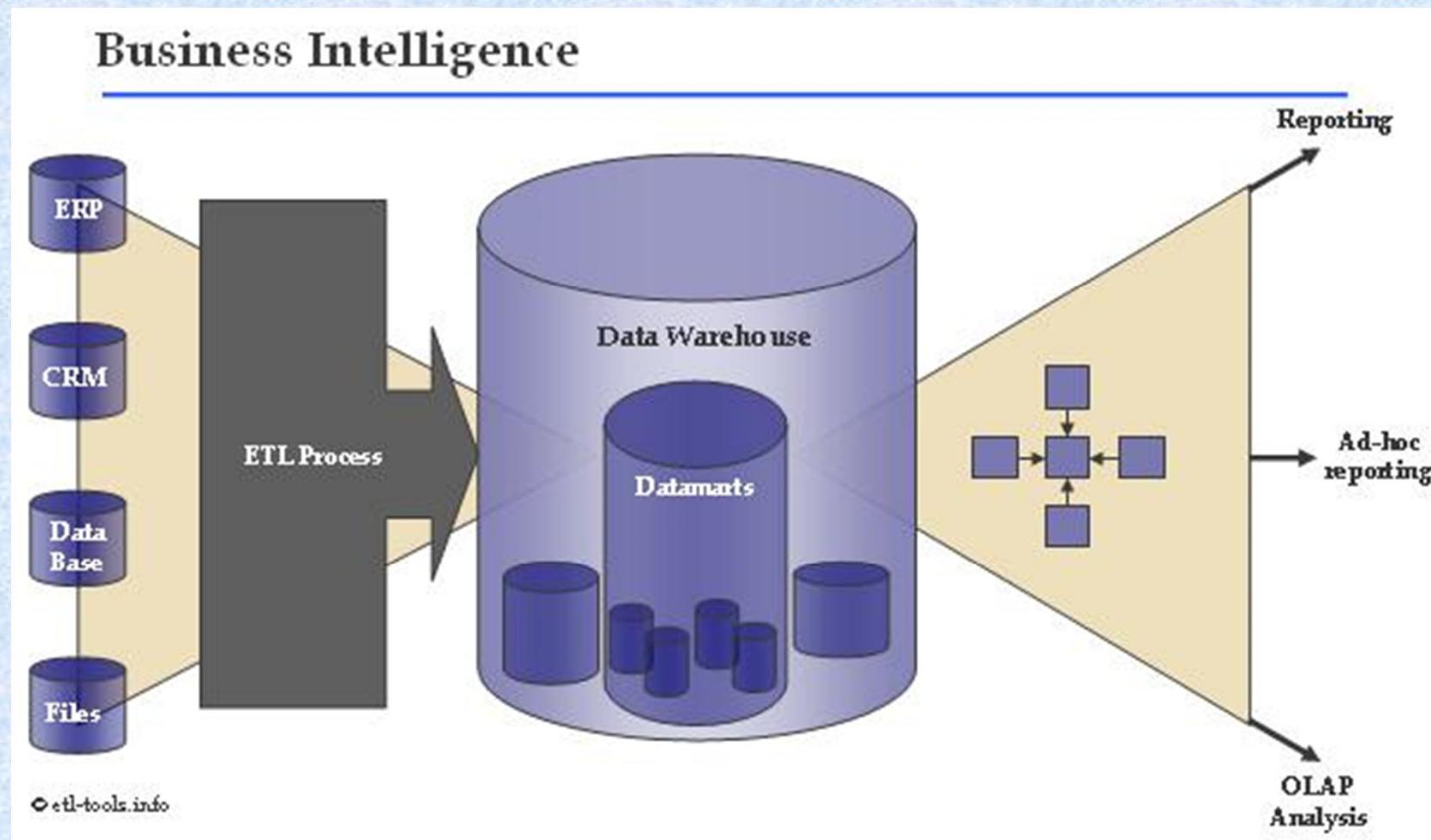


- *Distribución de costos en proyectos de BI:*



Bases de datos – Business Intelligence

- **Componentes**



Bases de datos – Business Intelligence

Cuadrante Mágico de Gartner

Representación grafica de la situación de mercado de los productos de BI realizada por el grupo empresarial Gartner.

El Cuadrante Mágico está dividido en cuatro partes en donde se distribuyen las principales compañías en función de su tipología y la de sus productos.

1) Líderes (leaders): aquellos que tienen la mayor puntuación combinan su habilidad para ejecutar (lo bien que un vendedor vende y ofrece soporte a sus productos y servicios a nivel global) y el alcance de visión, que se refiere a su potencial.

2) Aspirantes (challengers): caracterizados por ofrecer buenas funcionalidades y un número considerable de instalaciones del producto, pero sin la visión de los líderes.

Bases de datos – Business Intelligence

Cuadrante Mágico de Gartner

- 3) Visionarios (visionaries):** estos pueden tener todas las capacidades que ha de ofrecer un ECM de forma nativa, o mediante alianzas con otros socios, lo cual significa un fuerte impulso a la integración de programas y plataformas así como una habilidad para anticiparse a las necesidades del mercado que ellos no puedan cubrir.

- 4) Nichos específicos (niche players):** enfocados a determinadas áreas de las tecnologías ECM, pero sin disponer de una suite completa.

“Enterprise Content Management (ECM) son las estrategias, métodos y herramientas usadas para capturar, manejar, salvaguardar, preservar y entregar contenido y documentos relacionados con procesos organizacionales. ECM cubre la gestión de la información dentro del ámbito completo de una empresa, ya sea que esa información esté en forma de documento de papel, un archivo electrónico, una base de datos impresa e incluso un email.”

Bases de datos – Business Intelligence



Situación Actual de mercado.

El Cuadrante Mágico de Gartner

Bases de datos – Business Intelligence

Evolución del Cuadrante Mágico de Gartner



SAP – Business Object



Informes y análisis

Cuadros de mando

Exploración de datos

Móvil

Plataforma BI

- [SAP Crystal Reports](#): Conéctese a informes interactivos de prácticamente cualquier fuente de datos y con cualquier diseño o formato y compártalos de forma interna o externa.
- [SAP BusinessObjects Analysis, edición para OLAP](#): Realice análisis de conjuntos de datos multidimensionales con esta herramienta OLAP diseñada para realizar análisis financieros y empresariales.
- [SAP BusinessObjects Analysis, edición para Microsoft](#): Realice análisis avanzados de datos residentes en su almacén empresarial (BW) con una interfaz de usuario intuitiva en su entorno de Microsoft Office.
- [SAP BusinessObjects Web Intelligence](#): Realice consultas y análisis intuitivos en fuentes de datos heterogéneas online u offline.
- [SAP BusinessObjects Predictive Workbench](#): Extraiga sus activos de información para predecir tendencias futuras.

Microstrategy

MicroStrategy®

Agile Analytics

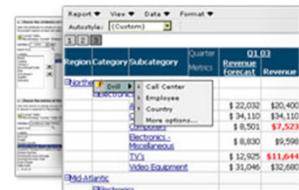
Intuitive visual data discovery and exploration. Build rapid-fire dashboards in minutes and deploy them with enterprise scale and governance.



[Learn more »](#)

Advanced and Predictive Analytics

Gives power business users and information analysts full investigative reach into any corner of the data warehouse



Scorecards and Dashboards

Conveys information at-a-glance with personalized views for executives and through every level of the organization



[Learn more »](#)

High Performance BI

MicroStrategy's Business Intelligence platform is specifically engineered to deploy scalable, secure BI applications



Enterprise Reporting

Provides Business Intelligence to the masses by delivering detailed information to decision-makers throughout the enterprise



[Learn more »](#)

INTELLIGENCE

Fully Integrated Platform
for Business Intelligence



MOBILE

Mobile App Platform
for All Essential Apps



Big Data

Lets companies manage large volumes of complex data efficiently



CLOUD

More Powerful, Faster,
More Cost Effective



EXPRESS

Agile Business Analytics
to the Masses



ALERT

Mobile Commerce Platform
For Retailers



WISDOM

The World's Most Extensive
Market Intelligence Network



IBM - Cognos



Business Intelligence

Freely explore information, analyze key facts and collaborate with key stakeholders to make the decisions that drive your business forward.

Reports

Use comprehensive query and reporting to make smarter decisions.

Dashboards

Monitor, measure and manage corporate performance at a glance.

Scorecards

Define quantifiable goals and targets, track performance and quickly identify performance areas that need attention.

Analysis

Drive better, smarter business decisions with wide-ranging analysis.

Advanced visualization

Your data can come to life for better perspective and analysis.

Collaboration

Align objectives, build relationships and gain consensus for important decision-making.

Mobile business intelligence

Interact with information like never before in a rich, visual and interactive experience.

Real-time monitoring

View operational data in motion to identify fleeting opportunities and take quick action.

Statistics

Add statistics to your analysis and reports to drive more informed decisions.

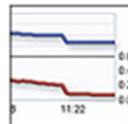
Engineered Systems



Get the answers to your toughest business questions in the blink of an eye. Oracle Exalytics is a powerful package of optimized hardware and software that delivers a high-performance and scalable platform for BI.

› [Learn more about Oracle Exalytics](#)

Enterprise Reporting



Implement a single solution for authoring, managing, and distributing highly formatted reports and forms to get more results with less effort.

› [Learn more about Oracle BI Publisher](#)

OLAP Analytics



Gain insights not only into what happened and what's happening in your business but also what the future may hold - with the most widely deployed OLAP server in the World.

› [Learn more about Essbase](#)

Mobile BI



Business doesn't stop just because you're on the go. Make sure critical information is reaching you wherever you are.

› [Learn more about Oracle BI Mobile](#)

BI Platform

Transform IT from a cost center to a business asset by standardizing on a single, scalable BI platform that empowers business users to easily create their own reports with information relevant to them.



› [Learn more about Oracle BI Foundation](#)

Information Discovery

New business opportunities are sometimes hidden in data sources that extend beyond your data warehouse. Search and explore structured and unstructured data - within the enterprise and beyond.



› [Learn more about Oracle Endeca Information Discovery](#)

Scorecards and Strategy Management

Oversee operations but keep an eye out for tomorrow's opportunities. Execution and strategy are inseparable for success. Implement a solution that does both.



› [Learn more about Oracle Scorecard and Strategy Management](#)

Real-time Decision Management

Optimize and manage recurring, high-volume operational decisions with a self-adapting decision services framework that adapts in real time to change and uncertainty.



- ✓ Consolidating relevant data from multiple sources into a single application
- ✓ Exploring the associations in your data
- ✓ Enabling social decision making through secure, real-time collaboration
- ✓ Visualizing data with engaging, state-of-the-art graphics
- ✓ Searching across all data—directly and indirectly
- ✓ Interacting with dynamic apps, dashboards and analytics
- ✓ Accessing, analyzing and capturing data from mobile devices

Bases de datos – Business Intelligence

Referencias:

•**BUSINESS INTELLIGENCE: COMPETIR CON
INFORMACIÓN**

•<http://www.gartner.com>

•<http://tdwi.org>