

Capítulo 2. Sistemas de multibase de datos existentes

La necesidad de llevar a cabo la integración de BDC's, con problemas de heterogeneidad, ha llevado a diversos organismos de investigación y empresas comerciales a lanzar una serie de productos con características propias. A pesar de la gran variedad de Sistema de Multibases de Datos (MDBMS), no es posible concebir uno solo, capaz de resolver la mayoría de problemas que implica la integración de diversos componentes de datos. Cada sistema propone funciones, metodología de integración o alguna otra característica para lograr una integración consistente, sin embargo, por restricciones tecnológicas y características propias de la información, no es posible aún contar con un sistema que cumpla con las expectativas que demanda la industria actual.

La revisión de MDBMS, tales como: Multibase, Ingres, Pegasus, Sybase y CORDS se presenta en este capítulo. Los componentes arquitectónicos y las características de cada uno de estos sistemas, se describen de manera general con el fin de dar un panorama operacional de cada uno de ellos. Finalmente se presenta una tabla comparativa de dichos sistemas.

2.1 MULTIBASE

MULTIBASE usa un lenguaje de manipulación y definición de datos llamado DATAPLEX, el cual esta basado en el modelo de datos funcional. Debido a que el modelo de datos funcional es lo suficientemente rico en significado, se pueden manipular modelos de red, jerárquicos y el relacional.

Información mas detallada sobre este sistema se encuentra en [Ceri y Pelagatti 1985].

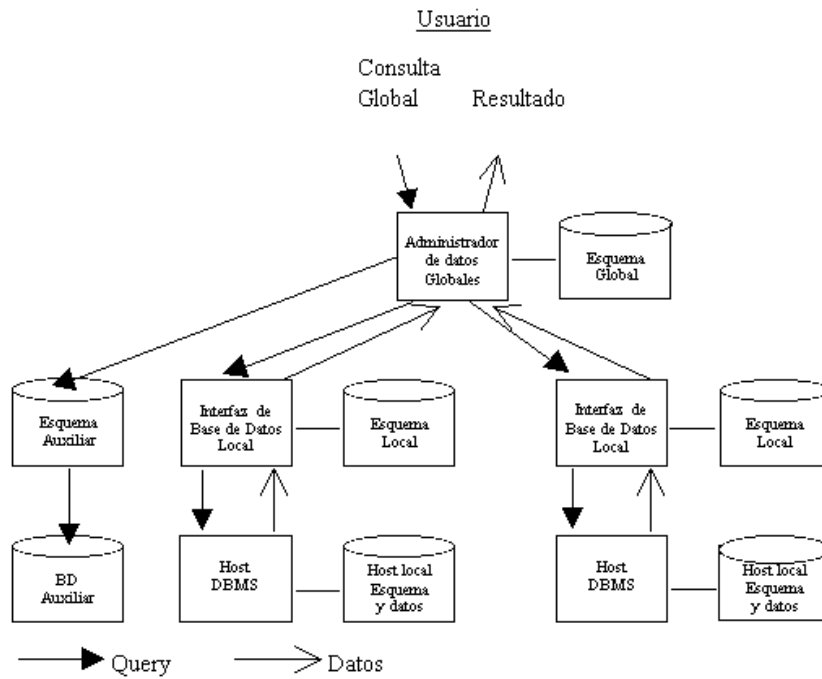
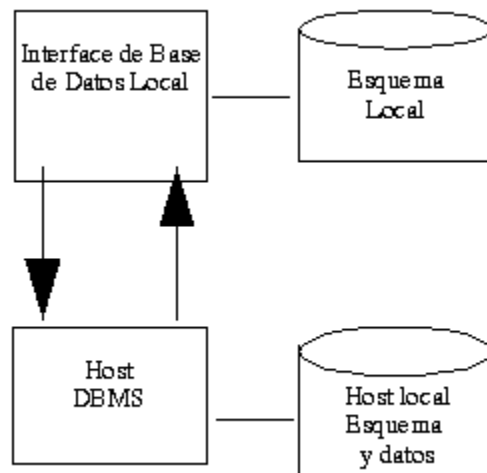


Figura 2.1 Arquitectura de MULTIBASE



2.2 INGRES

Las funciones de Ingres/STAR se proveen por siete módulos principales descritos a continuación:

1. **Servicio de Comunicación General** (GCF, General Communication Facility) provee la intercomunicación sobre instancias de Ingres/STAR, DBMS's de Ingres e Ingres/Gateways.
2. **Servicio de proceso de transacciones** (TPF, Transaction Processing

Facility), monitorea el estado de las transacciones de los diversos DBMS's de Ingres y mantiene un estado de las transacciones distribuidas, a través del TPC (Two Phase Commit) protocolo de dos fases.

3. **Servicio de evaluación de consultas** (QEF, Query Evaluation Facility), administra la ejecución actual de las consultas, envía subconsultas a otros participantes en una sesión, manipula los resultados retornados y regresa el resultado final al cliente Ingres/STAR.
4. **Servicio de consultas remotas** (RQF, Remote Query Facility), recibe instrucciones del QEF o del TPF, formatea las instrucciones, las envía a otros participantes en la sesión Ingres/STAR y retorna respuesta a quien la requirió.
5. **Servicio de descripción de relación** (RDF, Relation Description Facility), provee acceso eficiente al catálogo de información.
6. **Servicio de parser** (PSF, Parser Facility), verifica la sintaxis de la consulta.
7. **Servicio de optimización** (OPF, Optimizer Facility). toma la verificación de la consulta y planea el método para ejecutarla.

2.2.2 Características de Ingres/STAR

El DBMS Ingres permite a los usuarios acceder una base de datos distribuida, la cual se define como una colección de tablas de una o más bases de datos Ingres. Un conjunto de tablas de un grupo de bases de datos Ingres pueden combinarse para formar una nueva base de datos Ingres/STAR. Un servidor Ingres/STAR puede servir múltiples bases de datos distribuidas y múltiples servidores Ingres/STAR pueden existir en la misma red. De esta forma el sistema Ingres/STAR se define por [Sheth y Larson 1990] como un sistema federado débilmente acoplado el cual soporta múltiples sistemas federados.

El acceso a una Base de Datos Distribuida (BDD) Ingres/STAR es transparente en el sentido que los usuarios del sistema no necesitan conocer sobre la existencia de las bases de datos individuales Ingres que conforman la BDD.

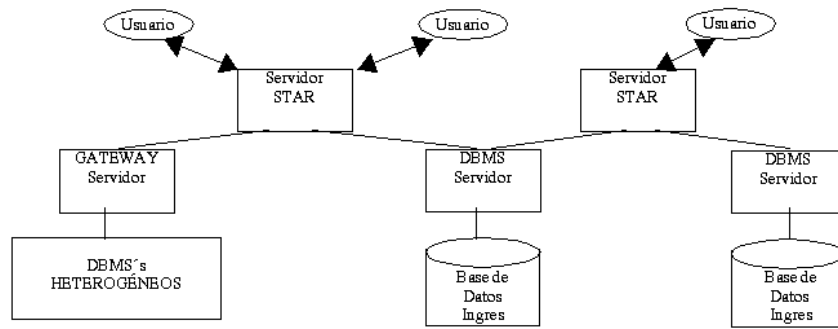


Figura 2.2 Configuración de las Bases de Datos y componentes de Igres/STAR

2.3 PEGASUS

Información detallada de cada uno de los componentes del sistema Pegasus se encuentra en [Ahmed et. al. 1991].

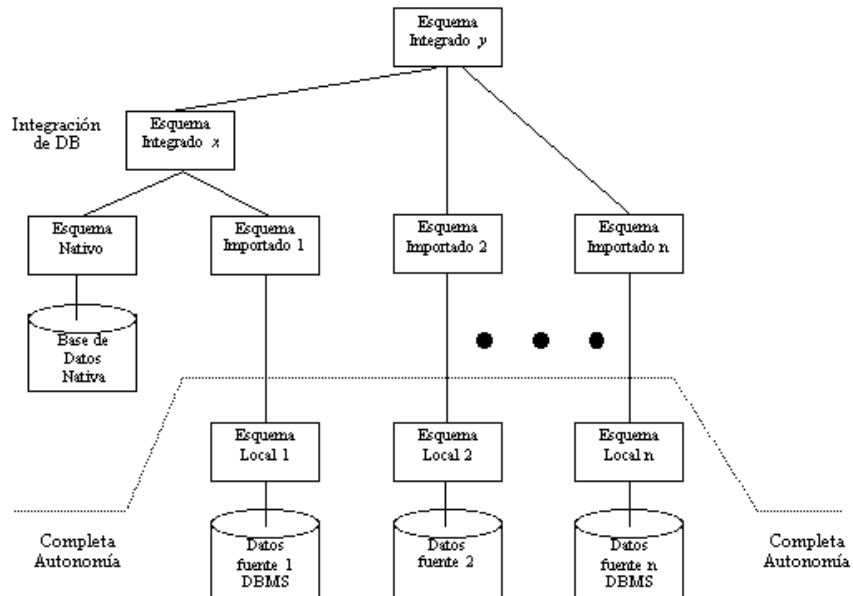


Figura 2.3 Configuración del sistema de base de datos de Pegasus

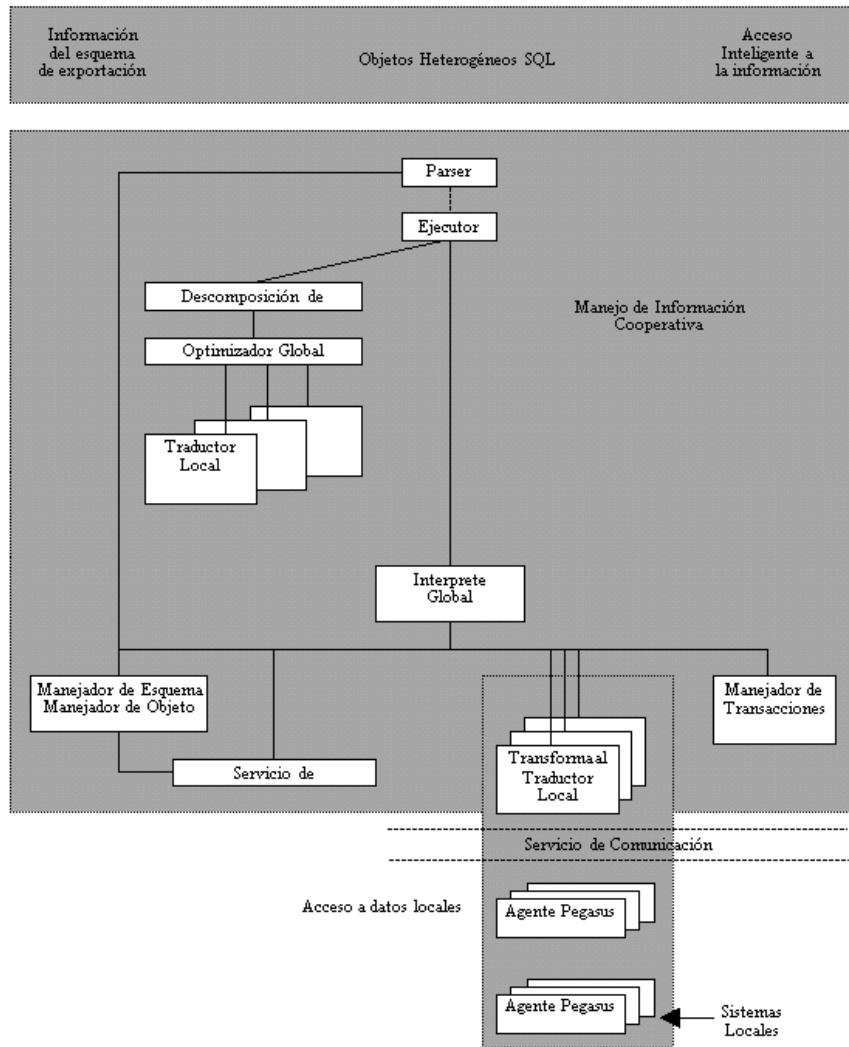


Figura 2.4 Niveles funcionales de Pegasus

2.4 SYBASE

Sybase, Inc. fundada en 1984 con el objetivo de brindar DBMS's distribuidos de alto desempeño al mercado. La necesidad de integrar una gran variedad de aplicaciones con múltiples fuentes de datos es en la actualidad un claro requerimiento comercial. Desde septiembre de 1989 Sybase introduce el Open Server, un producto que extiende las capacidades distribuidas de Sybase a fuentes de datos heterogéneas. Este producto complementa el Open Client, un API que se usa para enviar SQL o RPC's a un servidor SQL. Juntos forman la interfaz Cliente/Servidor, la base para BDD heterogéneas.

2.4.1 Características del sistema Sybase

Sybase se basa en el modelo relacional y soporta acceso programado e interactivo al servidor de SQL o alguna aplicación de Open Server. El lenguaje de consultas básicas es SQL. Múltiples sentencias SQL pueden aumentarse con la programación de constructores, tales como lógica condicional, llamadas a procedimientos y variables locales, estos pueden combinarse en un objeto de base de datos llamado un procedimiento de almacenamiento. Los procedimientos pueden regresar hileras de datos y mensajes de error, además de regresar valores en variables de programación en el programa de aplicación.

El servidor SQL también soporta disparadores como objetos independientes en la BD [Date 1983], estos tienen las capacidades de los procedimientos con tres extensiones importantes.

1. Ellos no pueden ejecutarse directamente, sólo responden al cumplimiento de una condición.
2. Un disparador puede restaurar o modificar los resultados de una transacción del usuario.
3. El disparador puede ver los cambios hechos a los datos

Además el servidor abierto de Sybase provee un método consistente para recibir requerimientos SQL o RPC's desde una aplicación basada en el conjunto de herramientas de SQL Sybase o desde una aplicación que usa la interfaz de cliente abierto de Sybase.

Sybase soporta actualizaciones distribuidas que se replican en localizaciones múltiples. TPC (Two-Phase Commit Protocol) se usa para mantener la consistencia de las transacciones. Información complementaria sobre este sistema puede encontrarse en [Sheth y Larson 1990].

2.4.2 Arquitectura del sistema Sybase

El sistema Abierto de Sybase presenta sus componentes básicos de la arquitectura cliente servidor en la Figura 2.5.

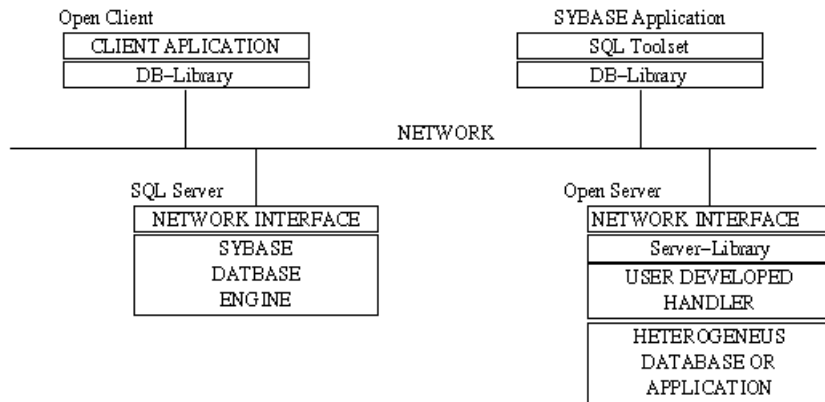


Figura 2.5 Arquitectura del sistema abierto cliente/servidor de SYBASE

2.5 CORDS

Exportar.- Exporta un esquema local (o porción de un esquema local), hace al esquema y sus datos correspondientes, disponibles para el MDBMS. La tarea de exportación involucra dos actividades: transformar el esquema local en su representación correspondiente en el modelo de datos común, y proveer el contexto a través del cual los datos serán interpretados.

Resolver Conflictos a Nivel de Atributos.- La correspondencia de atributos de los esquemas exportados se identifica y transforman a un atributo del MDBMS, usando la definición de vistas del MDBMS. Los conflictos de nombrado se resuelven al momento de la transformación. Otro tipo de conflictos que involucran facetas de contexto, se resuelven por la introducción de funciones de transformación para traducir el contexto del atributo exportado al contexto del atributo del MDBMS.

Resolver Conflictos a Nivel de Relación.- Las relaciones de los diversos esquemas exportados que representan objetos semánticamente equivalentes se identifican y los conflictos que ocurren se resuelven usando funciones de transformación.

Resolver Conflictos a Nivel de Esquema.- Los conflictos a nivel de esquema involucran mas de una relación. Los conflictos de este tipo se resuelven usualmente combinando las relaciones en una relación lógica usando un join o unión dentro del MDBMS.

Unir.- Un MDBMS se crea combinando las vistas correspondientes de los esquemas exportados en los pasos previos.

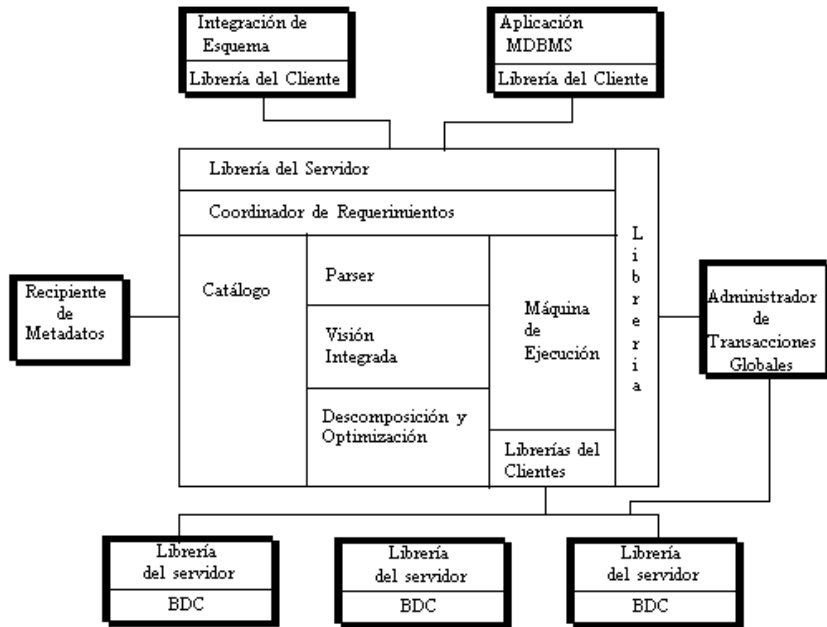


Figura 2.6 Arquitectura del MDBMS CORDS

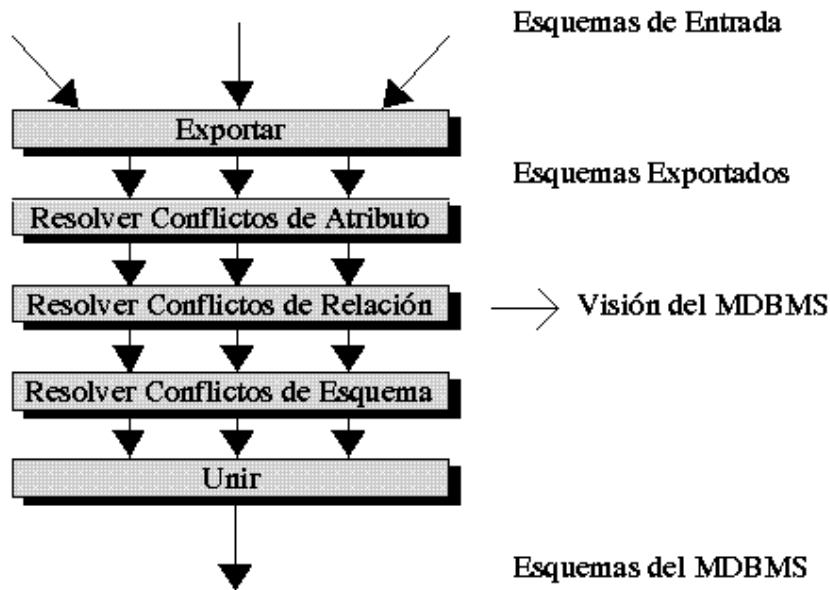


Figura 2.7 Proceso de Integración de Esquemas

2.6 ANÁLISIS COMPARATIVO DE LOS SISTEMAS DE MULTIBASE DE DATOS

La descripción y análisis comparativo de los sistemas Sybase, Ingres, Pegasus, Multibase y CORDS permiten establecer un patrón de referencia, para la propuesta de integración de esquemas planteada en los capítulos

4 y 5. Además este capítulo describe un elemento común en los sistemas, que es la resolución de conflictos para lograr un esquema global o federado.

La descripción detallada de los conflictos semánticos y de datos se describen en el siguiente capítulo.

Tabla 2.1 Tabla comparativa de Multibase, Ingres, Pegasus y Sybase

CARACTERÍSTICAS	MULTIBASE	INGRES	PEGASUS	SYBASE	COR
Federación fuertemente acoplada	SI	SI	SI	NO	SI
Manejo de Transacciones	NO	SI	SI	NO	SI
Uso de esquema auxiliar	SI	SI	SI	---	SI
Manejo de fragmentos	NO	SI	SI	---	---
Modelo de datos usado	Red, Jerárquico y Relacional	Relacional	OO, Relacional y de Red	Relacional	Red, Jerár Relat y OO
Interfaz proporcionada Consultas Lenguaje programación empotrado	SI	SI	SI SI	SI	SI
Optimización de consultas	SI	SI	SI	---	SI
TPC (Two Phase Commit)	NO	SI	SI	NO	---

Federación múltiple	NO	SI	---	NO	NO
Transparencia de acceso					
Acceso transparente	SI	SI	SI	SI	SI
Acceso no transparente	NO	NO	SI	NO	
Complejidad en el proceso de integración	Media	Alta	Alta	Media	Baja

Alvarez Carrión, G. 1999. **Integración de esquemas en bases de datos heterogéneas fuertemente acopladas**. Tesis Maestría. Ciencias con Especialidad en Ingeniería en Sistemas Computacionales. Departamento de Ingeniería en Sistemas Computacionales, Escuela de Ingeniería, Universidad de las Américas Puebla. Mayo. Derechos Reservados © 1999.