



estudios de postgrado
en computación



Bases de datos avanzadas

Universidad de Los Andes

Postgrado en Computación

Prof. Isabel M. Besembel Carrera

*Unidad II. Sesiones 31 y 32. Integración de modelos y
sistemas multibases de datos.*

Introducción

- Durante los últimos 30 años las empresas han desarrollado múltiples aplicaciones sobre formatos diversos (archivos o BD variadas)
- Datos están diseminados en aplicaciones y sistemas heterogéneos
- Integración y federación de datos o de aplicaciones heterogéneas es la respuesta a las necesidades de consulta de los usuarios
 - ✓ SGBD que integre las fuentes de datos heterogéneas asegurando la transparencia de distribución y heterogeneidad
 - Heterogeneidad sintáctica: modelos de datos diferentes para representar los mismos datos
 - Heterogeneidad semántica: conceptos diferentes utilizados para codificar informaciones similares



Federación

- Tecnologías:
 - ✓ BD virtuales compuestas a partir de fuentes heterogéneas o BD federadas, HDI (Heterogeneous Database Integration)
 - ✓ Plataforma de integración abierta según un datawarehouse, ODS (Open Data Source)
- Ambas tecnologías migran a XML como lenguaje de intercambio
- BD federada:
 - ✓ Varias fuentes de datos heterogéneas (archivos, BD, documentos, etc.) accedidos como una sola, una vista y un lenguaje común
 - ✓ Motivado a las limitaciones de las BD distribuidas para la integración de datos: necesidad de gestionar cada conjunto de datos por un SGBD, fuerte dependencia entre las BD, dificultad para integrar documentos y archivos, etc.



Modelo global de intercambio XML

- Integración punto a punto: técnica de integración de datos que realiza traductores BD a BD implementada con pasarelas entre sistemas
 - ✓ M fuentes de datos accediendo a N fuentes de datos, MxN traductores
- Integración por modelo global: técnica de integración de datos que pasa a un modelo pivote y se implementan los traductores para cada fuente al modelo global permitiendo la unificación de la representación de los datos
 - ✓ M fuentes de datos accediendo a N fuentes de datos, M+N traductores
 - ✓ La modificación de una fuente solo impacta al traductor asociado
 - ✓ XML es el modelo global de intercambio
 - ✓ XML cubre los conceptos de BD jerárquicas, redes, relacional y objeto



Virtualización vs. Materialización

- Vistas integradas virtuales (Virtualización)
- Materialización de datos federados en datawarehouse
- Ventajas:
 - ✓ Posibilidad de tomar fuentes en el ambiente integrado y una interfaz de consulta
 - ✓ Posibilidad de tomar en cuenta datos geográficamente dispersos y semánticamente diversos
- Desventajas:
 - ✓ Complejidad para hacer colaborar los modelos de datos y los lenguajes de consulta diferentes
 - ✓ Falta de experiencia en la integración de fuentes heterogéneas en Web
 - ✓ Distribución del control de datos en varios lugares

➤ Mediadores XML

- ✓ XML Mediator de e-XMLMedia: Middleware de federación de datos con XMLizer (traductor XML<->BDR) y XForms Engine, con consultas en XQuery
- ✓ Liquid Data de BEA: modelo integrado y homogéneo de acceso a los datos permitiendo consultas de vistas XML reutilizables en XQuery
- ✓ Nimble de Nimble.com: integra BDR, anuarios, archivos y páginas web.
- ✓ Enosys Server de Enosys Software: transforma las fuentes en XML y consulta con XQuery.
- ✓ OLE-DB.NET de Microsoft: ADO.NET integra OLE-DB y los adaptadores de diversas fuentes de datos tipo SQL o conjunto de registros

Data warehouse

- Conjunto de datos históricos constituidos por medio de la extracción a partir de BD de aplicaciones o archivos, organizados por temas específicos, consolidados en una BD única gestionada por un SGBD y que soporta la toma de decisiones de una empresa
- Data mining: conjunto de técnicas que permiten extraer modelos de conocimiento de una BD histórica mediante el razonamiento estadístico (enfoques inductivo o deductivo) a fin de describir el comportamiento actual y/o de predecir el comportamiento futuro
- Lenguaje de intercambio de modelos XMI (XML Metadata Interchange)
- Middlewares para datawarehouses: ETL (Extract Transform and Load) herramienta capaz de extraer datos de una fuente, transformarlos e integrarlos en un datawarehouse

Data warehouse

- Ejemplos de ETL:
 - ✓ Acta (Acta Works), DataStage (AscentialSoftware), DecisionStream (Cognos), ETI Extract (ETI), Genio (Hummingbird), iWay Software (iWay), PowerCenter (Informatica), XMLizer (e-XMLMedia)
- Datawarehouse: datawarehouse de datos decisionales y de documentos obtenidos de web ligados entre ellos
- Text mining: técnicas de análisis lingüístico y de data mining para analizar los contenidos textuales
- Con la integración de documentos y la búsqueda de elementos textuales gracias a XML, las BD alcanzan la capacidad de Penta Bytes (PB) (10^{15})



Integración de aplicaciones heterogéneas

- EAI (Enterprise Application Integrator): plataforma de enrutamiento , transformación e integración de mensajes entre aplicaciones en una empresa
 - ✓ Orquestador: componente que encadena las operaciones y las transferencias en el corazón del EAI apoyándose en una red de tareas a ejecutar (workflow)
 - ✓ Ejemplos: Biztalk MS, Web Methods, iPlanetIS (Sun), MQSeries (IBM), etc.
- Portales de aplicación (Web Application Portal): portal que permite presentar las aplicaciones heterogéneas de la empresa, por medio de páginas web con ejecución de operaciones y despliegue de resultados, navegando entre las aplicaciones en forma integrada
- Infraestructuras B2B: estandarización del comercio electrónico a nivel mundial



*Nuestro
enfoque*

Integración de aplicaciones heterogéneas

➤ Objetivo:

- ✓ Integrar las aplicaciones heterogéneas de los procesos de negocios, implícitas en la Arquitectura de Aplicaciones, a través de agentes, objetos/reglas de negocios

➤ Modelado de Negocios

Para cada tarea:

- ✓ La entrada específica para poder tomar la decisión
- ✓ El espacio de decisión descrito con el conjunto de posibles alternativas
- ✓ La regla de decisión (cómo tomar la decisión)
- ✓ La condición que activa el disparo de la decisión

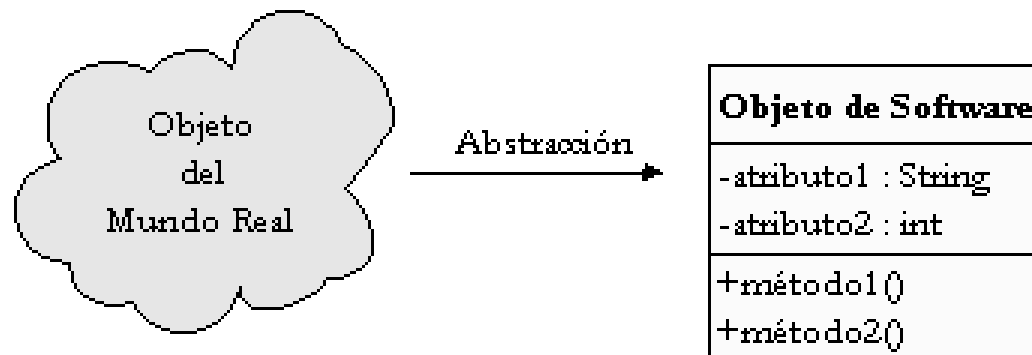
Modelado en términos de:

- los objetos que lo caracterizan y los reflejan (p.e. inventarios y facturas, clientes y vendedores)
- los objetos que representan eventos (p.e. comprar, vender)

Integración de aplicaciones heterogéneas

➤ Objetos de Negocios

- ✓ Representación mental de elementos comúnmente utilizados en un dominio particular del negocio (p.e. orden de compra, inventario, capacidad de producción, un empleado,...)
- ✓ Colección de objetos de software.
- ✓ Abstracción de los objetos del mundo real en objetos de software





Integración de aplicaciones heterogéneas

➤ Reglas de Negocios

- ✓ Reflejan parte del funcionamiento del mundo real
 - ✓ Restricciones que existen en un negocio dado
 - ✓ Cada aplicación, en principio, debe reflejar las restricciones que existen en el negocio dado, de modo que nunca sea posible llevar a cabo acciones no válidas.
 - ✓ Las reglas que debe seguir la aplicación para garantizar lo anterior, se denominan: reglas de negocio o *business rules*.
 - ✓ Ejemplos: no permitir crear facturas pertenecientes a clientes inexistentes, controlar que el saldo negativo de un cliente nunca sobrepase cierta cantidad, etc.
- El hecho de que la información sea manipulada por diversos programas hace más difícil garantizar que todos respetan las reglas, especialmente si las aplicaciones corren en diversas máquinas, bajo distintos sistemas operativos y están desarrolladas con distintos lenguajes y herramientas



Integración de aplicaciones heterogéneas

➤ Clasificación:

- 1.- Del modelo de datos (dominio del atributo)
 - 2.- De relación
 - 3.- De derivación (datos calculados en base a otros)
 - 4.- De restricción (de rango)
 - 5.- De flujo (cómo fluye la información)
 - 6.- De decisión (para la toma de decisiones)
- Soportadas por el SMBD: 1 y 4 reglas de validación de datos, 3 reglas de cálculo de datos y 2 reglas de integridad referencial.

Ejemplo:

Todo pedido debe ser realizado por un cliente y atendido de inmediato

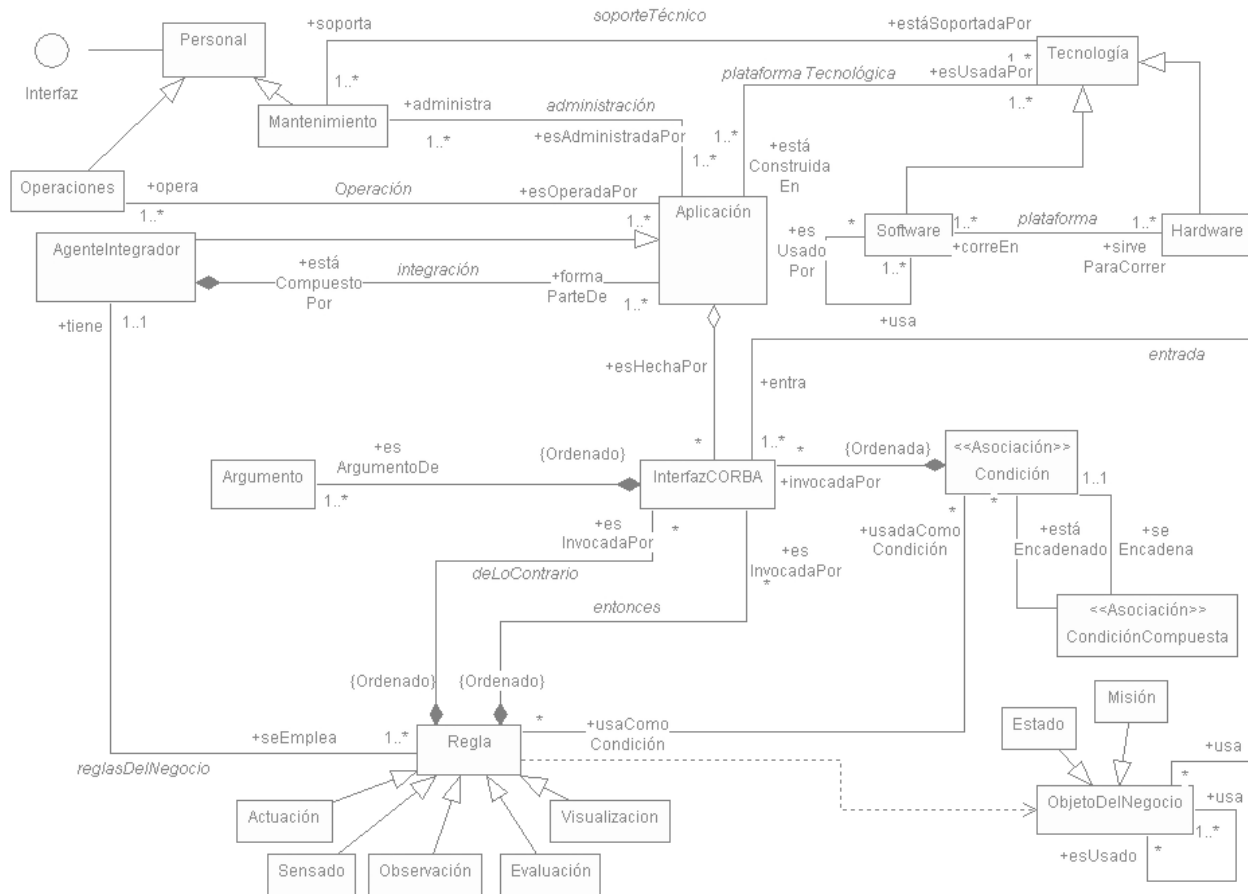
No se puede eliminar un cliente que tiene pedidos pendientes

Si el cliente tiene una línea de crédito aprobada

entonces el representante del cliente toma su pedido



Integración de aplicaciones heterogéneas





Agentes de software

- **Agente:** Programa autocontenido que es capaz de controlar sus acciones y decisiones para alcanzar unos objetivos, basándose en su percepción del entorno [Jennings, 1996]
 - ✓ Autonomía: control sobre sus acciones y estado interno.
 - ✓ Sociabilidad: interacción con otros agentes según un lenguaje
 - ✓ Reactividad: perciben su entorno y responden a cambios.
 - ✓ Pro-actividad: toman iniciativa según sus objetivos.
- **Estructura:** Agente = programa + arquitectura hardware
 - Programa: Implementa la relación que existe entre las percepciones y las acciones de deben realizarse
 - Arquitectura: Hace que las percepciones de los sensores estén disponibles para el programa



Modelo del Agente Integrador

- **Núcleo:** Las aplicaciones que conforman una UP
- **Objetivos:**
 - Sincronizar, mantener y controlarlas aplicaciones integradas
 - Manipular los modelos de negocios, representados por los objetos/reglas de negocios del dominio
- **Características:** Semi-autónomo, reactivo, pro-activo y social
- **Arquitectura**
 - Los sensores perciben los eventos generados por las aplicaciones
 - Estas medidas actualizan los modelos
 - El interpretador, basado en el modelo y sus capacidades, instancia un plan que se ejecuta en un hilo independiente
 - Un plan instanciado puede implicar la colaboración con otros agentes o acciones sobre objetos



Modelo del Agente Integrador

