



Conceptos de Business Intelligence

José Mayorga Vindas
Consultor Business Intelligence

IN: llevando la TI a niveles estratégicos de negocios

- Visión transaccional de la TI:
 - Ejecuta procesos de negocios de las empresas
 - Facilita las actividades normales
- El papel de la TI en “la aldea global” es fundamental: muchas actividades no se pueden ejecutar sin infraestructura de TI.
- La TI tradicionalmente se ubica en ámbitos transaccionales:
 - Aplicaciones de usuario final
 - Bases de datos
 - Telecomunicaciones
 - ...
- La TI se diseña y optimiza con objetivos transaccionales

BI: llevando la TI a niveles estratégicos de negocios

- En el otro extremo, existen necesidades de información para la alta gerencia:
 - Información estratégica, resumida e histórica
- En **Inteligencia de Negocios**, TI se aleja del mundo transaccional, se acerca a la alta gerencia
- Se desarrollan nuevos paradigmas de:
 - Bases de datos
 - Técnicas para diseño y modelación de sistemas
 - Métodos de visualización
 - Metodologías de administración de proyectos
 - ...
- IN es una especialización que involucra varias áreas

Definiciones generales de Inteligencia de Negocios

- Inteligencia de negocios:
 - Hace referencia a un conjunto de tecnologías de Bases de Datos, aplicaciones, plataformas tecnológicas, modelos de razonamiento, ..., a una tendencia que integra a tales elementos
 - La IN genera información estratégica y la pone a disposición de los tomadores de decisiones.
 - Con la IN se logra que la TI de valor agregado a la a la organización, mas allá de los aspectos operacionales

Definiciones generales de Inteligencia de Negocios

- La IN es un concepto multifacético, ayuda a diversos objetivos:
 - Ayuda a tomar mejores decisiones y más rápidamente
 - Convierte los datos en información
 - Ayuda a utilizar un enfoque racional en la administración

Definiciones generales de Inteligencia de Negocios

- Sistemas de IN son aquellos que responden a la visión de la “Tecnología que genera Información Estratégica para la toma de decisiones”:
 - Sistemas sofisticados: bases de datos, herramientas de consultas, interfaces de usuario final, algoritmos para generación de información, ...
 - Un simple *Balanced Scorecard* en una hoja electrónica.
- IN requiere un componente interdisciplinario: el conocimiento de los negocios desde una perspectiva gerencial estratégica

Definiciones generales de Inteligencia de Negocios

- Deberíamos hablar de la Infraestructura de la IN:
 - Data Warehouse Corporativo
 - Sistemas de Minería de Datos
 - Balanced Scorecard
 - Portales de Información Corporativa
 - Bases de Datos Multidimensionales, sistemas OLAP, etc.
- En IN se generaliza la integración de datos y sistemas

Data Warehouse, Data Marts y Tecnología OLAP

- OLTP = On Line Transactional Processing
 - Son los sistemas informáticos transaccionales
 - Generan los datos básicos
- OLAP= On Line Analytical Processing
 - Bases de datos para generar reportes Ah-Hoc
 - Se analiza y almacena información analítica

Data Warehouse, Data Marts y Tecnología OLAP

- Cubos Multidimensionales, Dimensiones y Medidas
 - **Dimensiones**
 - Marca del producto
 - Fecha de la venta
 - Clasificación Geográfica
 - Estado Civil
 - **Medidas**
 - Costo de producción de la venta
 - Venta Bruta
 - Margen Neto de Ventas
 - Comisión de vendedor

Data Warehouse, Data Marts y Tecnología OLAP

- Cubos Multidimensionales:
 - Un Cubo es la unidad fundamental de almacenamiento y análisis de las Bases de Datos Multidimensionales.
 - Están compuestos por Dimensiones y Medidas

Ventas de vehículos usados, según Fecha de Venta, Categoría de Vehículo y Marca

Estado del Vehículo Usado

Venta Bruta Colones			Marca Vehiculo		
Año	Trimestre	Mes	Categoría Vehiculo	Honda	Nissan Gran Total
2002			4 x 4		87.816.374
			Sedan	610.311.012	696.698.509
2002 Total				610.311.012	784.514.883
					1.394.825.895
2003					
	Trim 1-2003		4 x 4		22.300.000
			Sedan	103.100.000	169.900.000
	Trim 1-2003 Total			103.100.000	192.200.000
					295.300.000
	Trim 2-2003				
		Abril-2003	4 x 4		10.500.000
			Sedan	60.100.000	62.900.000
		Abril-2003 Total		60.100.000	73.400.000
					133.500.000
		Mayo-2003	4 x 4		15.500.000
			Sedan	43.500.000	65.000.000
		Mayo-2003 Total		43.500.000	80.500.000
					124.000.000
		Junio-2003	4 x 4		12.000.000
			Sedan	53.000.000	62.400.000
		Junio-2003 Total		53.000.000	74.400.000
					127.400.000
	Trim 2-2003 Total			156.600.000	228.300.000
					384.900.000
	Trim 3-2003		4 x 4		49.500.000
			Sedan	188.900.000	201.700.000
	Trim 3-2003 Total			188.900.000	251.200.000
					440.100.000
	Trim 4-2003		4 x 4		46.000.000
			Sedan	269.100.000	228.700.000
	Trim 4-2003 Total			269.100.000	274.700.000
					543.800.000
2003 Total				717.700.000	946.400.000
					1.664.100.000
Gran Total				1.328.011.012	1.730.914.883
					3.058.925.895

Data Warehouse, Data Marts y Tecnología OLAP

- William Inmon: Data Warehouse es un repositorio de información, tal que:
 - Es orientado a temas específicos. Se enfocan en categorías de los datos.
 - Integrados: la información del DW involucra a toda la institución.
 - No volátiles: La información del DW es solo de consulta, solo recibe dos transacciones: carga y lectura
 - Históricos: almacenan resultados resumidos, período tras período. La historia normalmente no más de 3 años.

Data Warehouse, Data Marts y Tecnología OLAP

- Data Mart: subconjunto del DW, hereda los conceptos básicos de la definición y los objetivos de diseño y requerimientos:
 - Deben ser históricos
 - No volátiles
 - Orientados a temas específicos
 - Integrados (en sus alcances y con los demás DM)
- Ralph Kimball: “un Data Mart es un trozo completo del pastel tomado del pastel completo que es el Data Warehouse”
- Un Data Mart es un sistema de IN para una unidad de negocios específica (Data Mart de Ventas, Data Mart del Área de Crédito, ...)

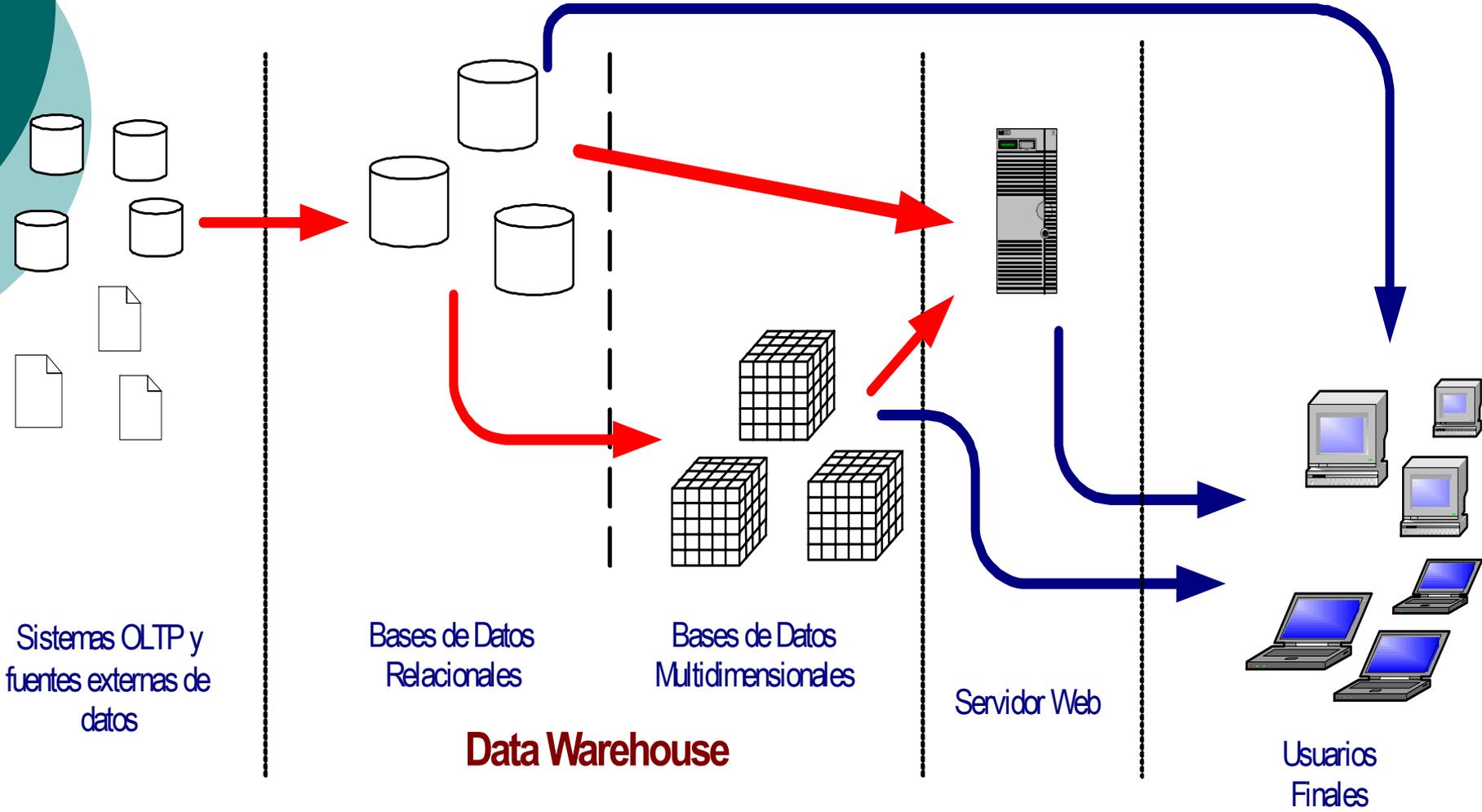
Data Warehouse, Data Marts y Tecnología OLAP

Característica	Sistemas transaccionales	Data Warehouse
Orientación del sistema	Ejecución y procesamiento de transacciones del 'día a día'	Generación de información estratégica e histórica.
Usuarios	Oficinistas, contadores, personal informático, jefes de departamentos operativos, clientes, etc.	Gerentes, ejecutivos, Juntas Directivas, analistas de información.
Tipo de diseño de base de datos	Modelos Entidad Relación, y/o sistemas de bases de datos orientados a las aplicaciones OLTP	Bases de datos Multidimensionales, Esquemas Relacionales del tipo Estrellas, con objetivos estratégicos en la información.
Nivel de detalle de los datos	Se almacenen con el mayor detalle, se trata de las transacciones específicas	Datos agregados en distintos niveles, no interesa el detalle sino los resúmenes de los datos.

Data Warehouse, Data Marts y Tecnología OLAP

Características del hardware y su configuración	Servidores de “pequeños a medianos”, sistemas de alta redundancia, configurados para tener recuperaciones ante fallas y optimizados para realizar transacciones puntuales en línea y con “muchos usuarios”.	Servidores de “grandes a gigantes”, optimizados para almacenar grandes volúmenes de datos y responder a consultas complejas que involucran mucha información, y con “pocos usuarios”.
Operaciones normales	Mucha lectura y escritura: actualizaciones, inserciones, sistemas de seguridad con alta redundancia, consultas.	Básicamente lectura de los datos: consultas complejas de los usuarios.
Volúmenes de datos	Dado que la información es siempre la actual, el volumen de datos no responde a la cantidad de transacciones que se almacenen. De 100 MB a 1 o 2 GB.	Dado que se acumula información histórica, los DW crecen constantemente. Los volúmenes se miden en Gigabytes a TeraBytes

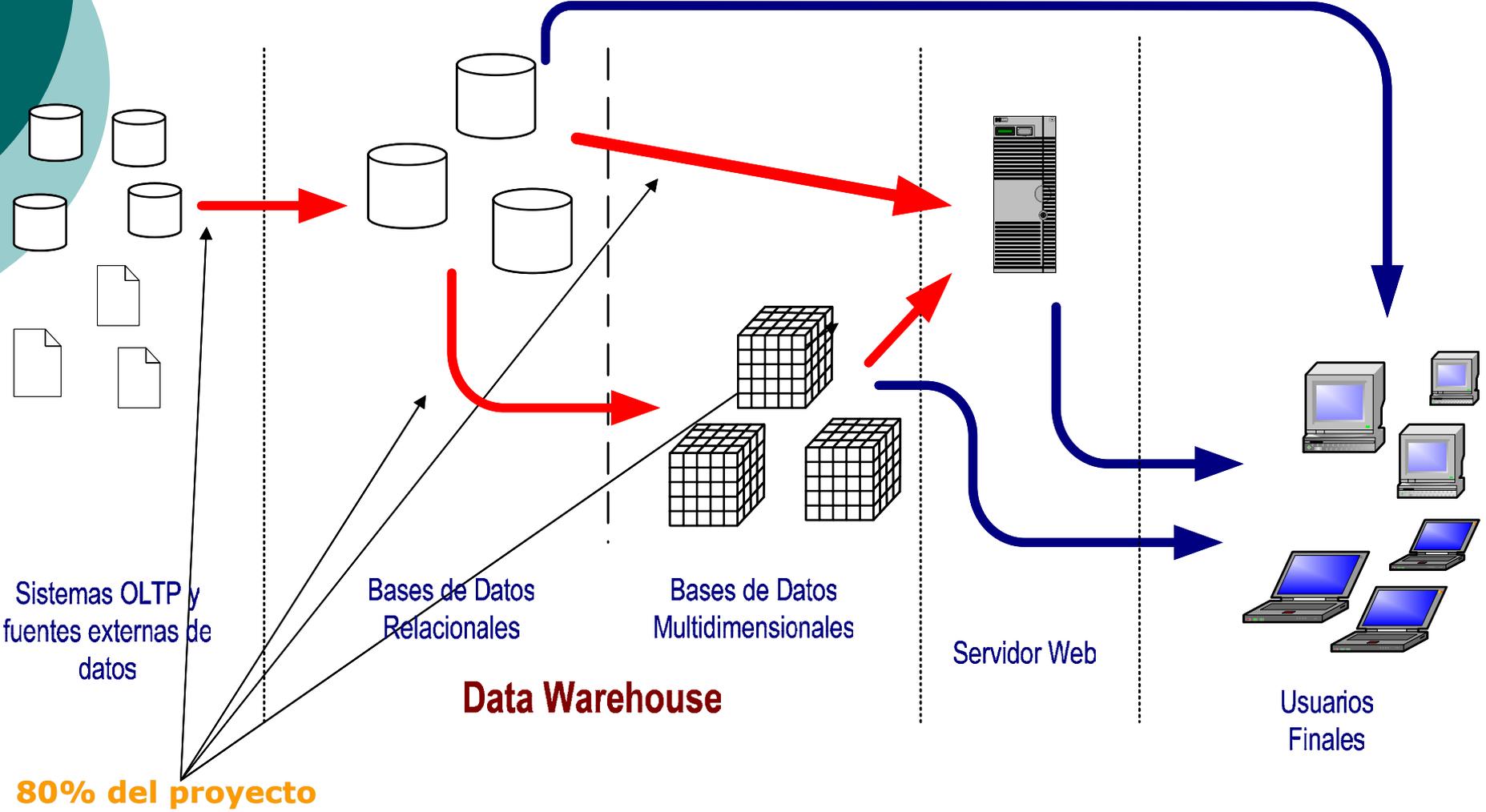
Estructura de un Data Warehouse con un servidor de aplicaciones Web



Data Warehouse: Metodología de desarrollo e implementación

- Utilización de los recursos en los proyectos de DW:
 - 80%: extracción, migración y depuración de datos
 - 20%: análisis y diseño, especificaciones, pruebas, ajustes, capacitación, modelos lógicos y físicos
 - En un sistema pre-construido se cuenta con un modelo que se debe adaptar a la realidad, o sea, se cuenta con un 20% que se debe adaptar
 - A la medida, el tiempo de desarrollo es igual o menor que al implementar una solución preconstruida
 - **Factor Crítico de Éxito fundamental:**
 - A la medida o pre-construido, el sistema debe estar 100% adaptado a la realidad de la empresa

Estructura de un Data Warehouse con un servidor de aplicaciones Web



Data Warehouse: caso real 1

- Implementación de enlatado (pre-construido) banco en Sur América:
 - El proyecto tarda más de un año
 - Consiste en poner en producción unos 30 cubos de datos
 - Dado que el sistema “está listo”, las adaptaciones son mínimas
 - No cumple con expectativas y no atiende requerimientos, la utilización es casi nula
 - La cultura de IN es nula
 - Luego de 1 año en “producción” se hace un backup y se formatea el servidor, ...
 - El banco pierde más de \$350.000

Data Warehouse: caso real 2

- Desarrollo a la medida otro banco sur americano:
 - El proyecto tarda un año
 - Consiste en desarrollar e implementar el Data Warehouse: 5 Data Marts completos
 - Dado que es a la medida el sistema “queda 100% adaptado”
 - La infraestructura es insuficiente
 - La Cultura de IN es intermedia
 - Luego de un año en producción, la utilización es baja
 - El banco invierte poco y aprovecha poco

Data Warehouse: caso real 3

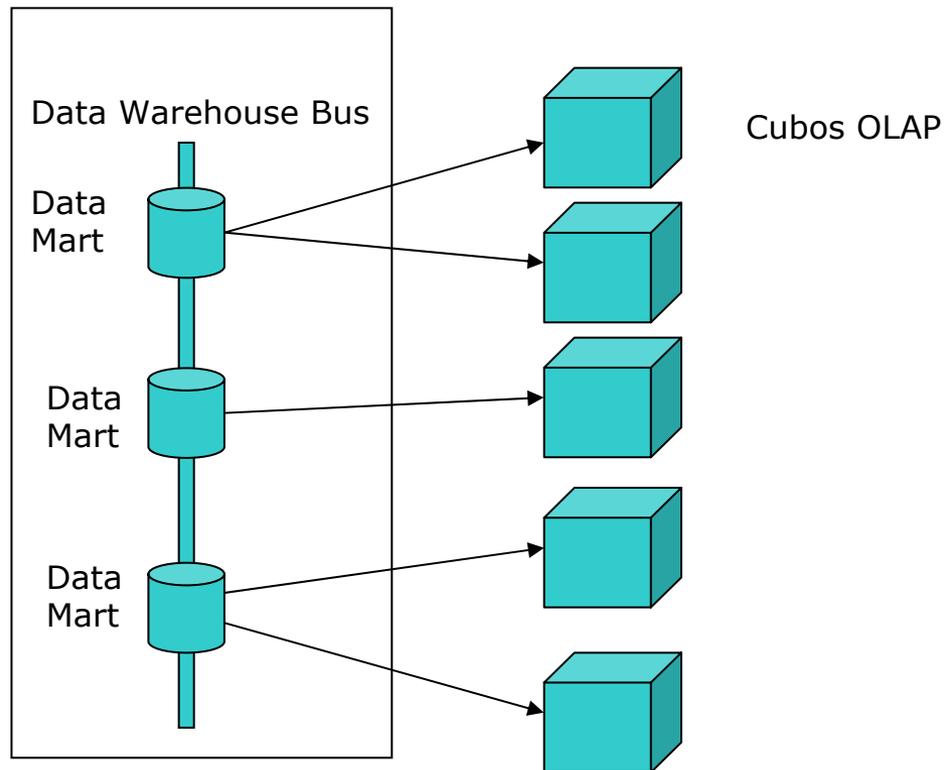
- Desarrollo a la medida, empresa manufactura CR
 - El proyecto tarda 3 meses
 - Consiste en desarrollar e implementar Data Mart para el área comercial
 - Dado que es a la medida el sistema “la adaptación es del 100%”
 - La infraestructura es adecuada y la Cultura IN alta
 - Luego de un año en producción, la utilización es intensiva
 - Recientemente la empresa decide enfrentar un proyecto de DW completo siguiendo la misma metodología

Data Warehouse: Metodologías de desarrollo e implementación

- William Inmon
 - Desarrollo e implementación por medio de iteraciones
 - A partir de un Depósito de Datos Operativo
- Ralph Kimball
 - Desarrollo por medio de Data Marts
 - Ciclo de Vida del Data Warehouse

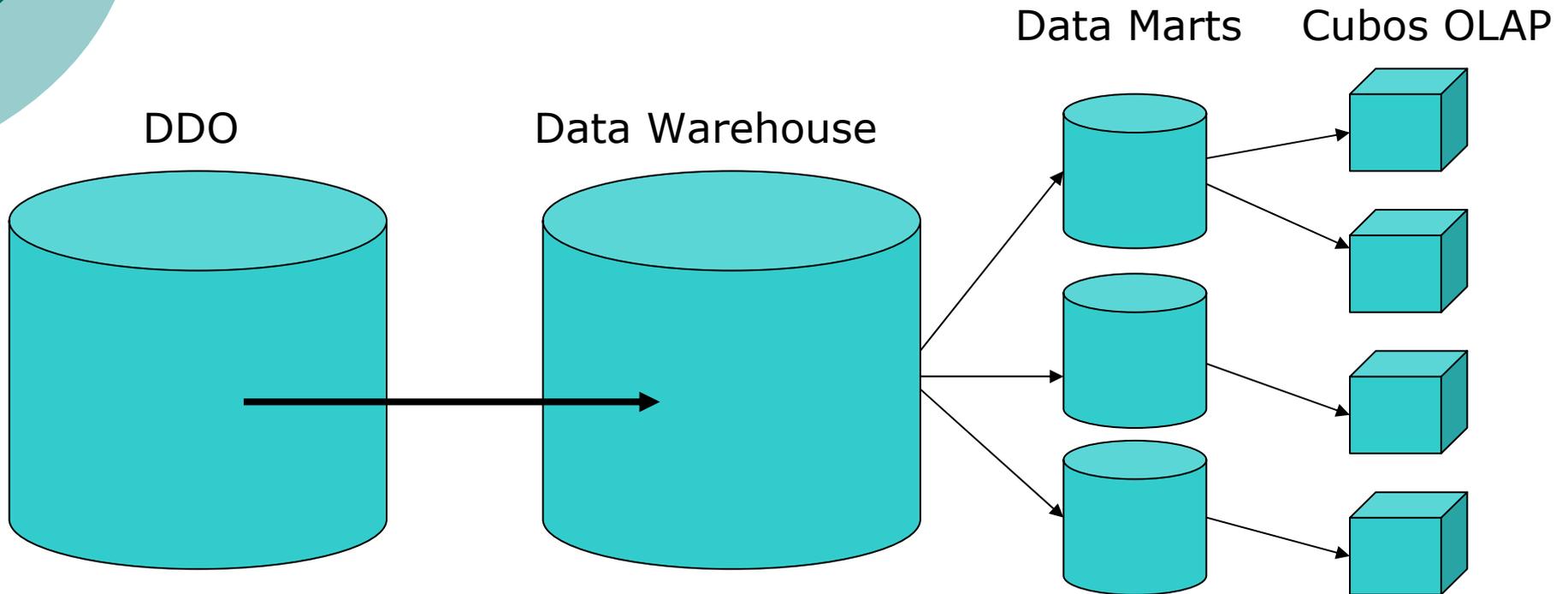
El Ciclo de Vida de Ralph Kimball

○ Data Warehouse: arquitectura de Bus



Enfoque de Inmon: Data Warehouse Centralizados

- Data Marts dependientes del DW





Algunos Factores Críticos de Éxito

- No se puede trabajar sin una metodología
- Utilización de una metodología de trabajo consolidada
- Evaluación de los desarrolladores
- Evaluación de opciones de desarrollo: a la medida, en forma iterativa, por medio de un Bus de Data Marts, etc.
- Confiabilidad de las especificaciones del sistema. Deben existir reglas de negocio, fórmulas, mapeos de orígenes de datos, etc., que sean correctos, claros y concisos para que la extracción de datos sea confiable.



Algunos Factores Críticos de Éxito

- Datos básicos confiables. Los datos de las fuentes primarias de información deben ser correctos y consistentes.
- Disponibilidad de personal técnico y de negocios para el proyecto.
- La evangelización lleva a generar expectativas adecuadas de los usuarios.
- La experiencia en IN es fundamental: La IN es un área de especialización en la TI, existe una gran acumulación de conocimientos, experiencias y técnicas específicos para desarrollar los sistemas de IN.

Data Warehouse: Metodología de desarrollo e implementación

- Ciclo de Vida (Ralph Kimball)
 - **Planeación del proyecto:** Se define el DW, se establecen sus limitaciones y sus alcances. Se establece un objetivo global. Constituye la columna vertebral del proyecto.
 - **Administración del proyecto:** Como en cualquier otro tipo de proyecto es requerida su administración en forma centralizada.
 - **Definición de requerimientos de negocios:** El DW responde a una necesidad de información particular por parte de los tomadores de decisiones de alto nivel. Los desarrolladores de DW deben estar en capacidad de realizar una interpretación correcta de la información obtenida de sus usuarios.
 - **Modelamiento Multidimensional:** se convierten los requerimientos en medidas y dimensiones, se generan Modelos Conceptuales y Lógicos del sistema.
 - **Diseño físico:** Es la construcción física de las bases de datos a partir de los Modelos Lógicos.

Data Warehouse: Metodología de desarrollo e implementación

- Ciclo de Vida (Ralph Kimball)
 - **Preparación de datos, diseño, desarrollo e implementación:** En esta etapa se realiza la extracción y transformación de datos hacia el DW.
 - **Diseño de la arquitectura técnica:** Se establece la arquitectura y estructura del sistema por desarrollar. Se determina qué sistemas operativos se van a utilizar, se toman decisiones en materia de Bases de Datos, aplicaciones de usuario final.
 - **Aplicaciones de usuario final:** Las aplicaciones de usuario final se deben especificar a partir de las necesidades y requerimientos de los usuarios.
 - **Mantenimiento y crecimiento:** Los DW tienen una naturaleza dinámica, pues evolucionan al ritmo de los negocios de la empresa. Si un DW cambia, esto es una buena señal.
 - **Capacitación** técnica y de usuario, “evangelización” en materia de Business Intelligence, Data Warehouse, pensamiento multidimensional, etc.
 - **Divulgación de resultados:** Los buenos resultados se deben “mostrar” y promover en la empresa, el proyecto se debe vender internamente.

Otras áreas de la IN: Minería de Datos

- DW y tecnología OLAP muestran las fotografías “crudas” de los resultados finales de los negocios: ventas, costos, producción, ...
- En un proceso de Minería de Datos se trata de llegar al descubrimiento de información valiosa utilizando distintas técnicas y algoritmos de diversa naturaleza: matemática, estadística, ...
- Se aplican técnicas como:
 - Árboles de Decisión
 - Redes Neuronales
 - Técnicas de Agrupamiento

Otras áreas de la IN: Balanced Scorecard

- La frase “no importa el juego, el concepto siempre es el mismo: si no llevas control del puntaje, cómo puedes saber si vas ganando”
- El proceso de medición y valoración debe ser ordenado, coherente y total, no se debe descuidar nada
- Kaplan y Norton: BS es un sistema de administración, basado en el análisis de medidas, financieras y no financieras, que resumen la gestión en todas las áreas y en todo momento
- En el BS las medidas se organizan en dos niveles:
 - Medidas de los resultados pasados (efecto)
 - Medidas que generan estos resultados (causa)
- Además se vigilan las metas constantemente para controlar el accionar de la empresa

Integración de la IN

- Diferentes niveles de la información estratégica:
 - **Data Warehouse** se enfoca en aspectos históricos y resumidos, centralización de la información e integración de sistemas
 - **Minería de Datos**: análisis estadístico matemático de la información, generación de información mas allá del sentido común y/o la experiencia
 - **Balanced Scorecard**: sistema de medición y control constante del *performance* (rendimiento) de la empresa.
 - **CRM**: administración de la relación con los clientes
 - ...