

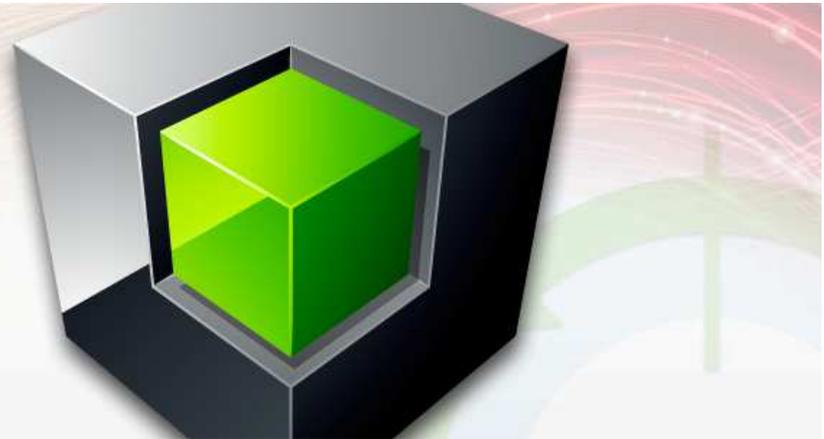
# IBM Informix 12.1. Simplemente poderoso.



# Informix Warehouse Accelerator 12.1

Vicente Salvador – DEISTER

“Gracias a Fred Ho por su docu”



# ¿Qué es IWA?

- *Informix Warehouse Accelerator (IWA) es un sistema de tratamiento de datos optimizado que actúa como una "caja negra" o un complemento y que permite el análisis de información y la integración de los aspectos de negocio en los procesos operativos para impulsar estrategias ganadoras. Acelera las consultas con tiempos de respuesta sin precedentes.*



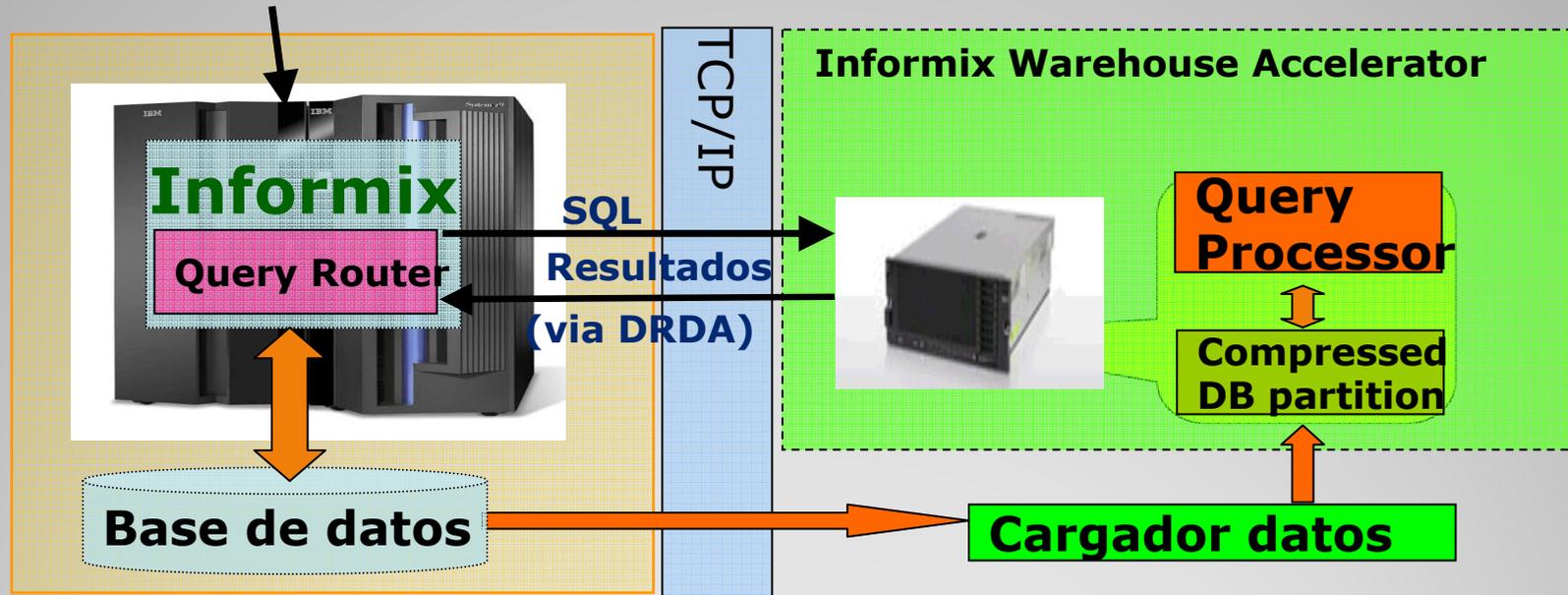
## Características Clave:

- **Rendimiento:** Tiempos de respuesta sin precedentes que permiten realizar análisis de información que normalmente no son posibles dados los elevados tiempos de respuesta
- **Integración:** Se conecta a Informix de forma totalmente integrada, permitiendo ser utilizado por todas las aplicaciones sin cambios
- **Autogestionado:** Las consultas son siempre ejecutadas de la forma más eficiente
- **Transparente:** Las aplicaciones se conectan a Informix y no necesitan conocer nada acerca de IWA
- **Administración simplificada:** Es un complemento que no requiere de tareas administrativas para su mantenimiento



# ¿Cómo funciona IWA?

## Consultas SQL



### Informix:

- Enruta las consultas
- **No necesita cambiar SQL ni aplicaciones**
- Siempre se puede ejecutar el SQL en el motor IDS

### Informix Warehouse Accelerator:

- Conecta a Informix via
- Analiza, comprime y carga
- Procesa las sentencias SQL y retorna los resultados a Informix



# Administración Simplificada

- No** – Query Tuning & Optimizer Hints
- No** – Ajustes de la base de datos
- No** – Creación o actualización de índices
- No** – Update Statistics
- No** – Particionamiento/Fragmentación
- No** – Gestión de discos o configuración de tamaños de página
- No** – Cambios en la BB.DD o en el esquema
- No** – Cambios en las aplicaciones
- No** – Tablas de agregación/Materialized Views
- No** – Compra de hardware más caro
- No** – Cambios en las expectativas



# Tecnologías de vanguardia

## Compresión Extrema

Necesaria porque la memoria impone el limite

## Row & Columnar Database

Formato en filas en Informix para operaciones transacciones y acceso via columnas para consultas OLAP

## Algoritmos Multi-core & Vector Optimized

Eliminando bloqueos o sincronización

## Base de datos en memoria

Tecnología de base de datos de 3ª generación que elimina la necesidad de I/O. Las bases de datos se mantienen completamente en memoria

## Evaluación de predicados en datos comprimidos

Permite búsquedas con/sin decompresión

## Paralelismo masivo

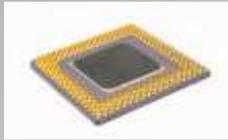
Se utilizan simultáneamente todos los cores para las consultas

## Particionamiento por frecuencia

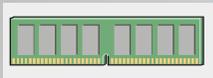
Acceso en paralelo a los datos comprimidos para el escaneo de datos. Eliminación de particiones Vertical y Horizontal.



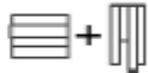
# Otras tecnologías clave



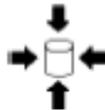
64-bit processor



RAM in TB



Row and Column Store



Compression



No Aggregate Tables

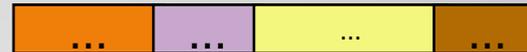
Number of Occurrences



Particionamiento por frecuencia

Common Values

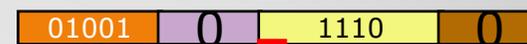
Rare values



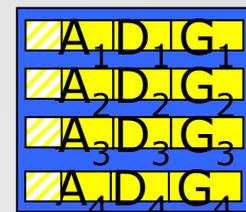
&



=



Evaluación de predicados comprimidos



SIMD

# Evaluación de predicados

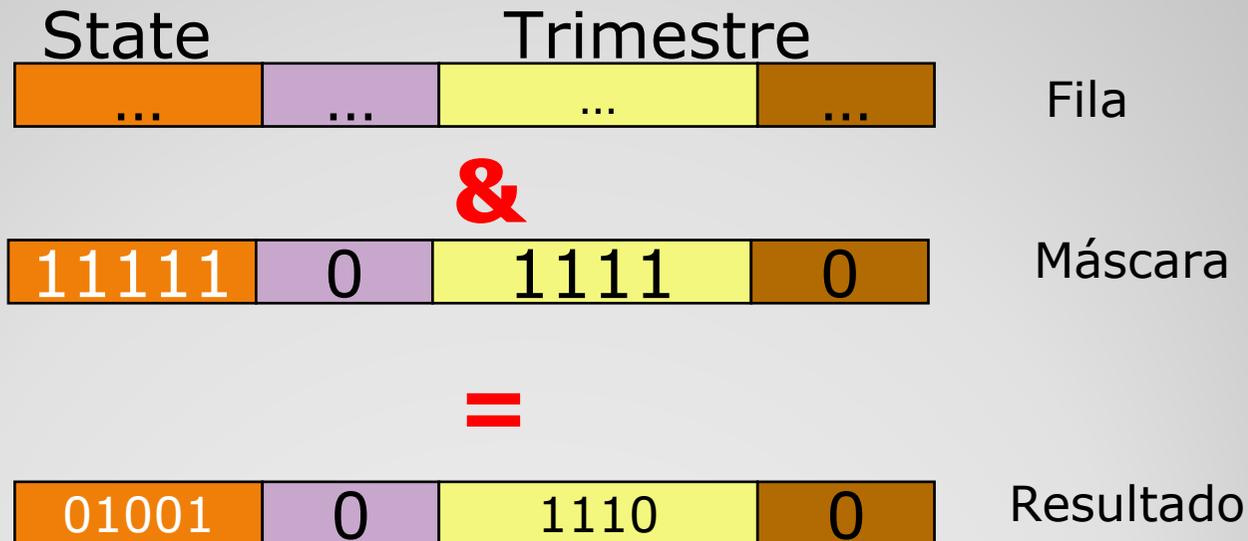
- CPU opera en bloques de 128b
  - Muchos campos caben en 128b
- Aplica el filtro a todas las columnas simultáneamente!

State=='CA' && Quarter == '4T'



State==01001 && Quarter==1110

Transforma el valor de la consulta en códigos internos

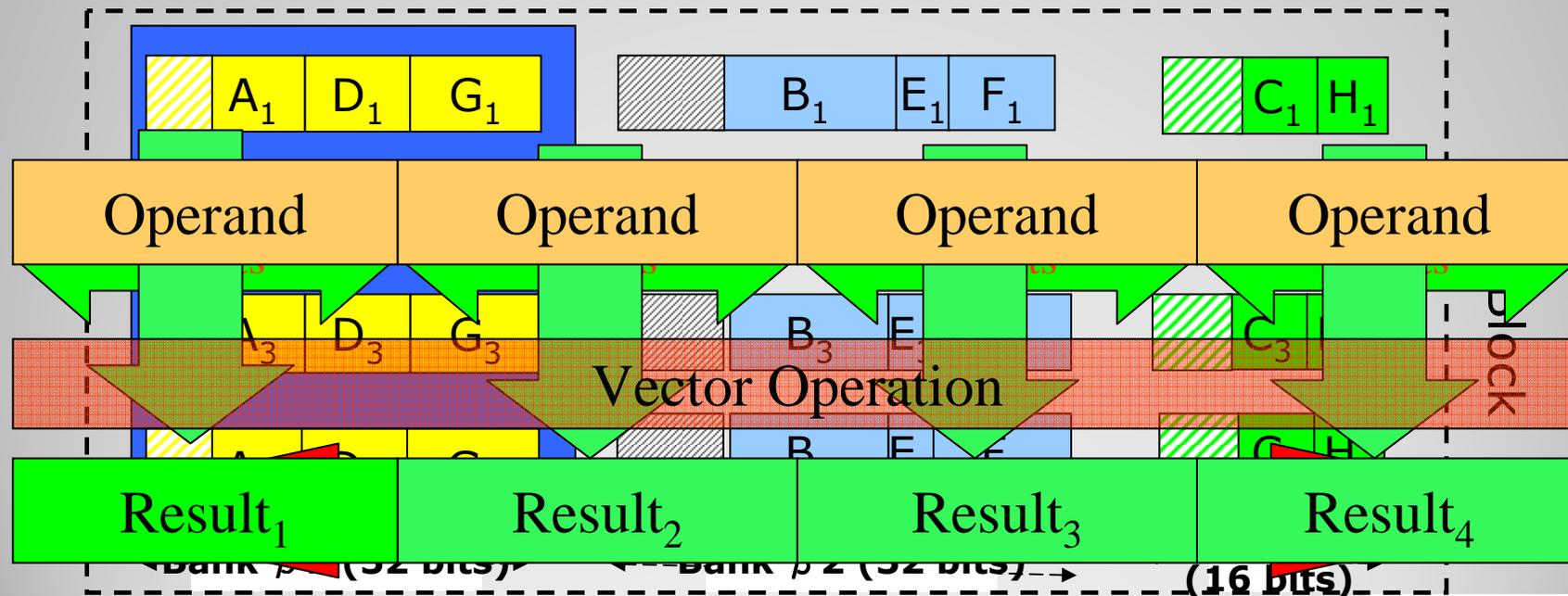


# Paralelismo SIMD

- Accede solo a los bancos utilizados en la consulta:

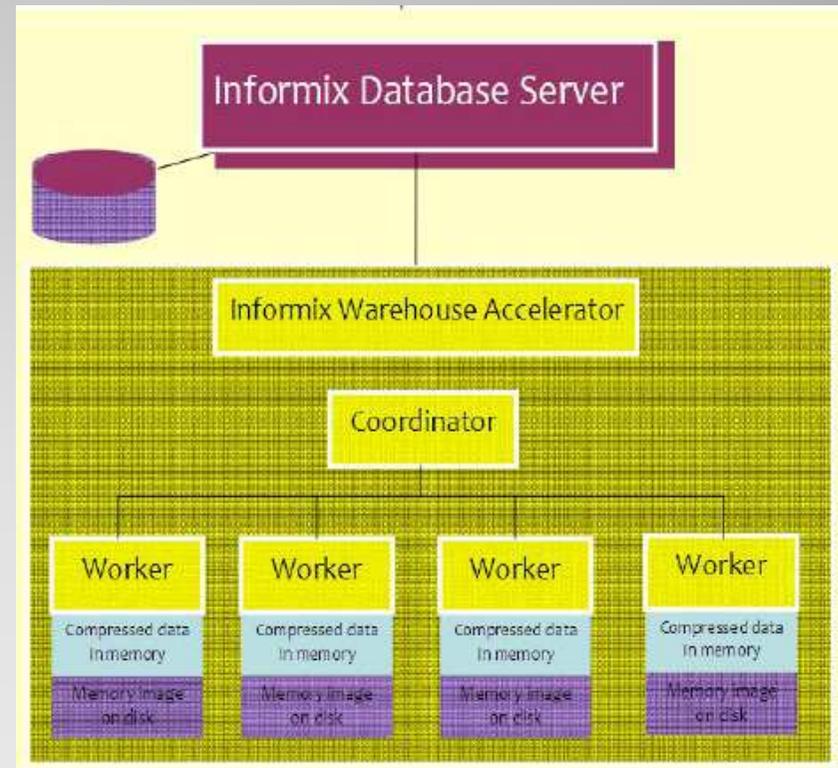
```
SELECT SUM (T.G)
FROM T
WHERE T.A > 5
GROUP BY T.D
```

- Empaqueta múltiples filas del mismo banco en un registro de 128 bits
- Permite otro mecanismo de paralelismo: SIMD (Single-Instruction, Multiple-Data)!



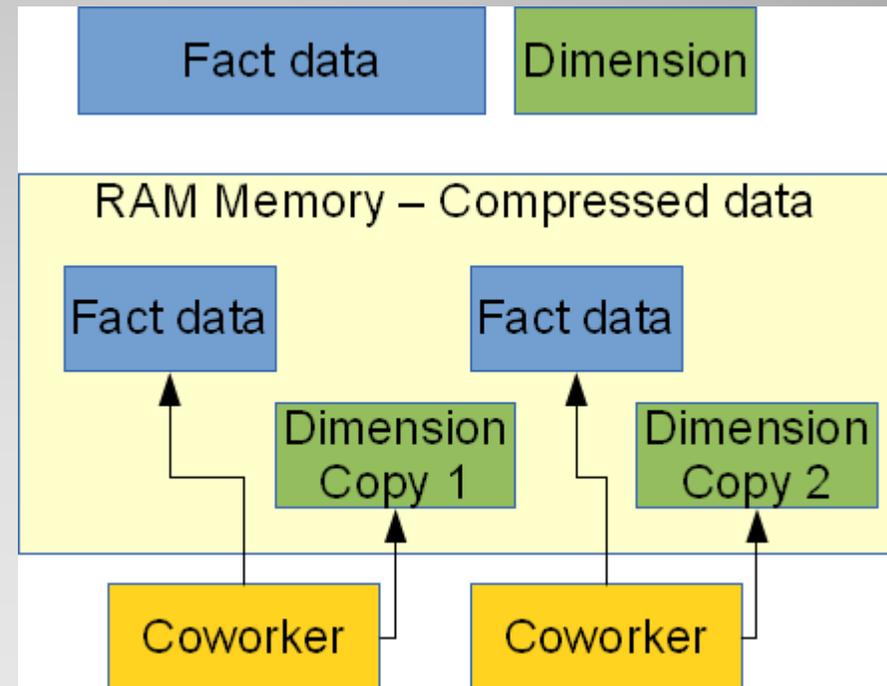
# ¿Cómo funciona?

- Si una sentencia SQL puede ser acelerada, IDS transforma el SQL y lo envía a IWA
- Cada procesador recorre su segmento de memoria, realiza las operaciones y devuelve los datos al coordinador
- El coordinador, recibe los datos y realiza las operaciones de fusionado y ordenación: OPNQR operations.
- IWA devuelve el resultado al servidor Informix
- IDS envía los resultados del SQL a la aplicación cliente como un resultset.



# ¿Cómo se utiliza la memoria?

- Los datos se comprimen y se reparten entre cada coprocesador
- Los ratios de compresión suelen estar en torno a relaciones 4 a 1 o 5 a 1
- Existe un proceso coordinador que se encarga de recoger los resultados escaneados por los coprocesadores y construir el resultado final.



# Resumen de características

- Tecnología revolucionaria que permite realizar consultas analíticas a la velocidad del pensamiento
- Tecnología de primer nivel que utiliza un sistema de almacenamiento columnar en memoria
- Elimina las operaciones E/S y las operaciones de mantenimiento como indexación, particionamiento, ajustes de SQL, update statistics, etc.
- Puede ejecutarse en equipos Intel convencionales capaces de almacenar varios Terabytes de memoria
- IWA se ejecuta como un acelerador de las instalaciones Informix existentes, permitiendo el uso simultáneo de operaciones analíticas y de OLTP



# Nuevas características IWA 12.10

- Trickle Feed y mejoras en refresco de datos
- Permite el uso de operadores OLAP en sentencias SQL
- Evaluación de sentencias concatenadas mediante UNION
- Mejora en los algoritmos de ordenación
- Aceleración de Timeseries mediante IWA





# Carga inicial de datos

- El conjunto total de datos que componen el datamart ocupa 550 Gbytes (tablas de hechos + tablas de dimensiones)
- 585 millones de registros en las tablas de hechos
- Una vez cargados los datos en IWA ocupan 115 Gbytes de RAM
- El proceso inicial de carga tarda 2 H
- 2 entornos distintos de test:
  - 16 core Intel E5 512 GB RAM
  - 16 core Intel E5 512 GB RAM (Virtualizado)



# Proceso de cliente

- Proceso de cliente de generación de reportes que utiliza cursores de datos y realiza consultas SQL complejas
  - Sin IWA: 3 H 56 min (Utilizando agg tables)
  - Entorno 1: 4 min 40 sec (Usando tabla de hechos)
  - Entorno 2: 5 min 10 sec (Usando tabla de hechos)
  - IWA en x-Series (40 CPU 2T RAM): 4 min: 15 sec (test done on IBM San Francisco Labs)
- Impacto en el rendimiento del # CPUs



# Datamart completo

- Se ejecutan 746 operaciones SQL que consideran todas las posibles combinaciones entre las dimensiones y las tablas de hechos
  - Sin IWA: las primeras 14 sentencias estuvieron ejecutandose durante 24 h
  - Entorno 1: 12 h 25 m
  - Entorno 2: 15 h 05 m
- En el peor de los casos se obtiene una aceleración de los procesos de 100 a 1



# Estadísticas de bajo nivel

- En cualquiera de las operaciones sobre los 585 millones de registros, la operación de escaneo y filtrado tarda siempre 53 segundos.
- Las operaciones de agrupación y ordenación tienen un tiempo de respuesta variable que oscila entre 0 y 40 segundos
- **Puede obtenerse el resultado de cualquier consulta en menos de 1 minuto 30 segundos**



# Gracias por su atención

¿Preguntas?

Vicente Salvador

[vsalvador@deister.es](mailto:vsalvador@deister.es)

+34 688 92 19 19



**Informix**<sup>®</sup>

# Gracias



[@ibmanalytics\\_es](https://twitter.com/ibmanalytics_es)

