



# IWA

## Informix Warehouse Accelerator for Real-Time Analytics



# Agenda

- Tendencias de Mercado y Posicionamiento
- ¿Qué es IWA?
- Experiencias de Clientes
- Ediciones IWA
- Seminario IWA



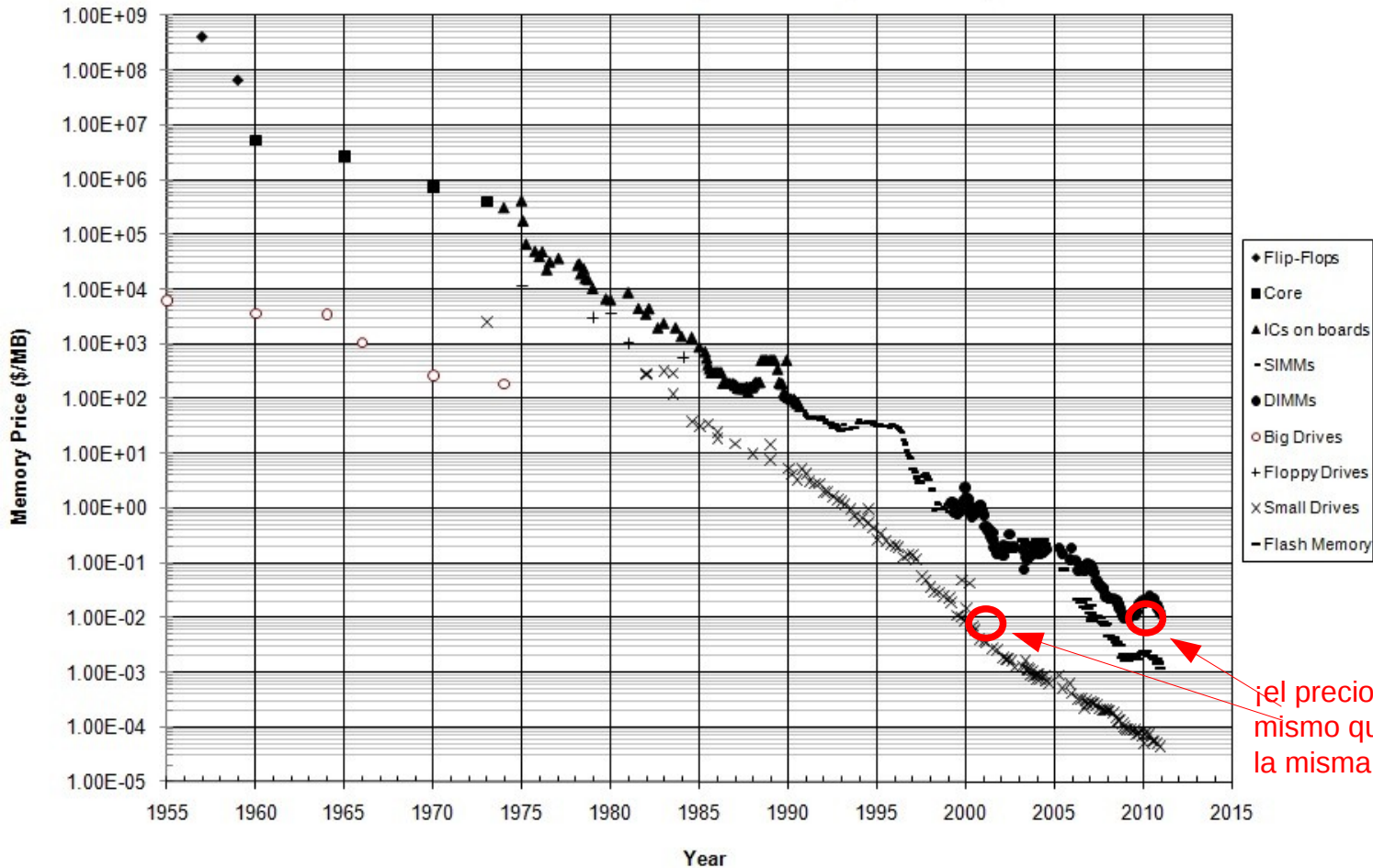


# Tendencias de Mercado y Posicionamiento



# Tendencias de Precios Para Memoria y Disco

Historical Cost of Computer Memory and Storage



¡el precio de la memoria hoy es el mismo que el precio del disco para la misma capacidad hace 10 años!





# Precio de la Memoria

- Precios diarios de Memoria en <http://www.simmtester.com/page/memory/memprice.asp>

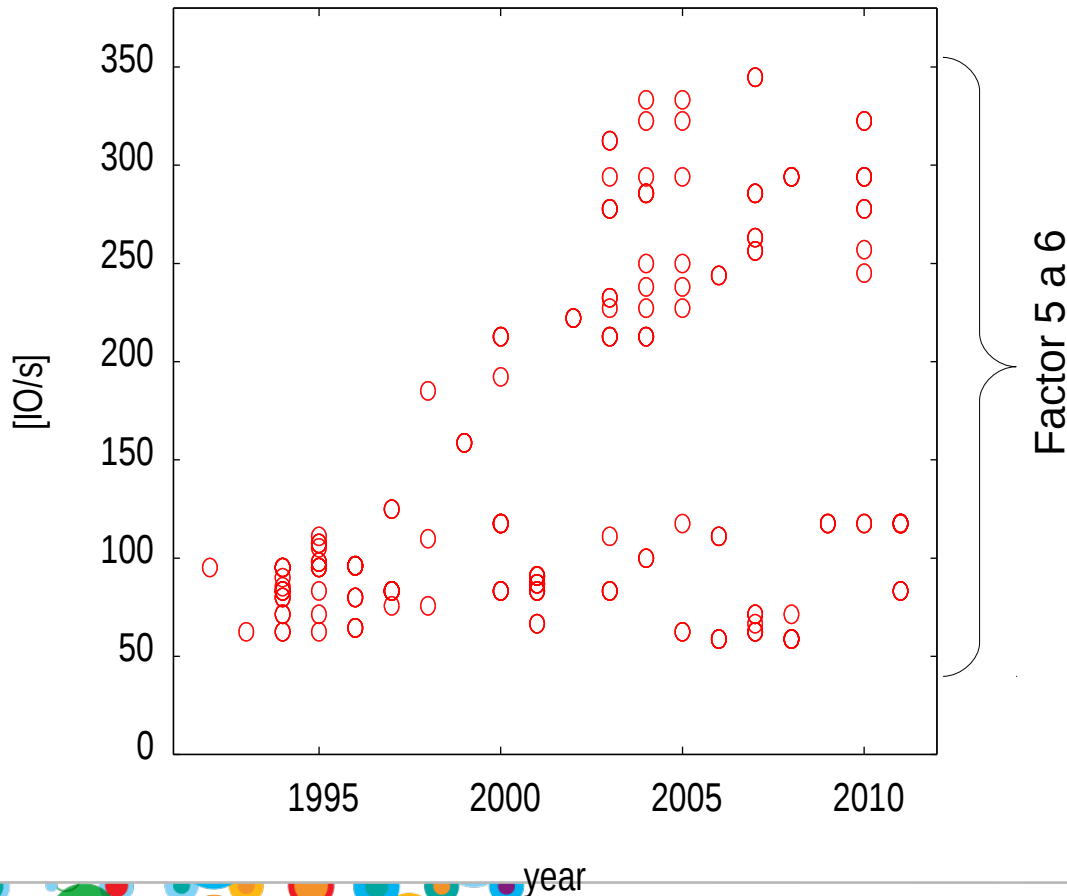
Memory Size	Price for Memory Chips
1 GB	\$ 20
1 TB	\$ 20,000
10 TB	\$ 200,000
100 TB	\$ 2,000,000

- ¡Ya es asequible instalar grandes cantidades de memoria!



# Tendencias del Rendimiento de Accesos Random en los últimos 20 años.

Número de IO/s por disco duro en los pasados 20 años



- Número de I/Os por disco se ha incrementado por 5 o 6 veces en los últimos 20 años.
- El rendimiento del acceso random a la memoria RAM se ha incrementado unas 25 veces en los últimos 20 años.
- La diferencia en rendimiento entre memoria RAM y disco se amplía.



# Tendencias Hardware

- El precio en \$/MB es hoy el mismo para la memoria que lo era hace 10 años para el disco
- Acceso aleatorio ('Random'):
  - el tiempo medio de posicionamiento (average seek time) para los discos ha mejorado lentamente a lo largo de los últimos 20 años
  - la latencia de acceso a memoria ha mejorado mucho más rápidamente
  - la diferencia en rendimiento entre memoria y disco es cada vez mayor y el precio menor
- Incremento en el paralelismo de CPUs:
  - más paralelismo SIMD
  - más cores

IWA tiene en cuenta todas estas tendencias hardware





# La Tercera Generación en Tecnología de BBDD

According to IDC's Article (Carl Olofson) – Feb. 2010

## 1ª Generación:

- Bases de datos propietarias IMS, IDMS, Datacom

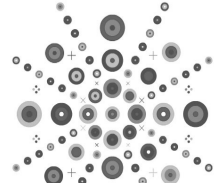
## 2ª Generación:

- RDBMS para Open Systems, flexibles en modelo y lenguaje SQL, dependientes en equilibrio disco/RAM

## 3ª Generación: IDC prevé que en 5 años:

- La mayoría de los data warehouses se almacenarán en formato 'por columna'
- La mayoría de las bases de datos OLTP serán potenciadas con tecnología 'en-memoria' (IMDB) o se almacenarán enteras en memoria
- Los grandes sistemas servidores tendrán un crecimiento horizontal mediante 'clustering'.

DATABASE



IWA propone todas estas características de un DBMS  
de tercera generación ¡HOY!







# Historia de Informix en Warehousing



- Informix tiene 3 productos de base de datos:
  - XPS para MPP Data Warehousing.
  - Red Brick para Star Schema data marts/data warehousing.
  - Informix Dynamic Server (IDS) para OLTP & Data Warehousing.
  
- Nuestros clientes Informix necesitan una mejor plataforma para warehousing que se adapte a sus necesidades y tamaño.
  
- La Solución:
  - Informix Ultimate Warehouse Edition e
  - Informix Growth Warehouse Editions
  
- Informix Warehouse Accelerator es la clave.





# Características de Informix para Warehousing

- Particionamiento – Fragmentación de Datos
  - PDQ y eliminación de particiones
  - Gestión cíclica y automática de particiones
- Joins sobre modelos en Estrella: Star and snowflake hash joins
- Multi-index scans
- Tablas Externas – fast bulk loader
- Tablas sin logging - Raw tables
- Áreas privadas de memoria (VIRTSHM) para procesos PDQ
  - PDQPRIORITY
  - Light scans ahora mejorados (VARCHAR)
  - Ordenaciones y agrupaciones
  - Potente indexación
- Actualización de Estadísticas Inteligente
- Informix-Warehouse ELT
- **Informix Warehouse Accelerator – para entornos analíticos**



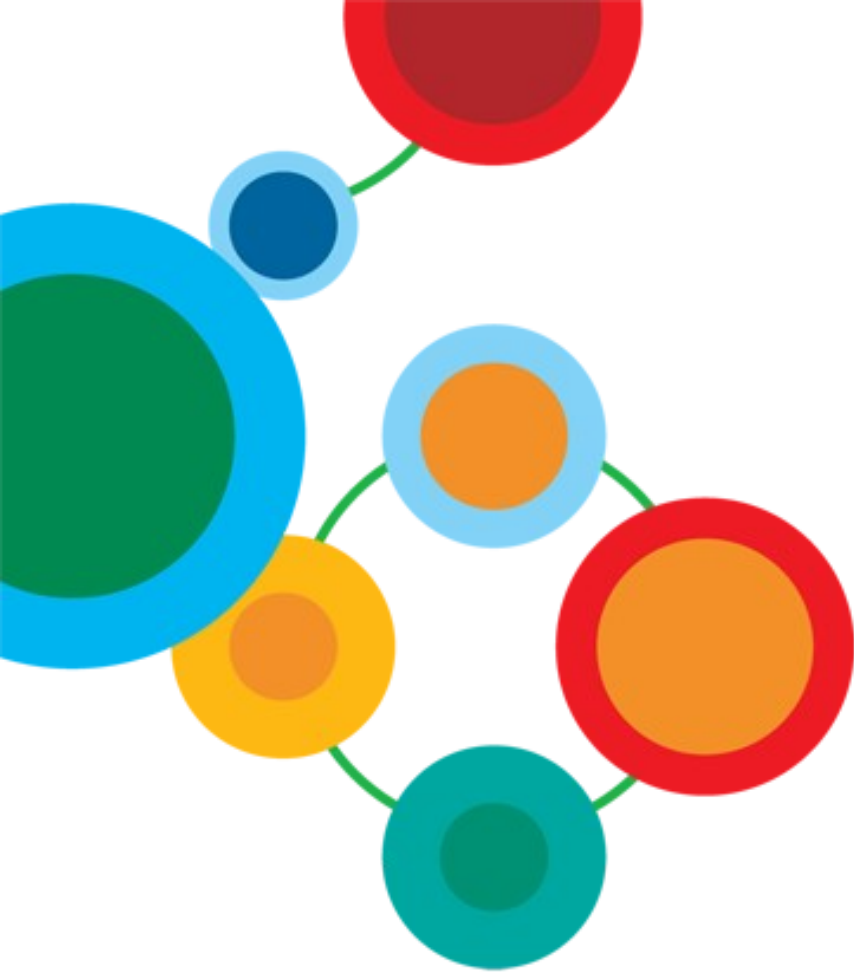


# Informix Warehousing Mirando al Futuro



- El objetivo es proveer una plataforma de warehousing completa y competitiva
  - Incorporando las mejores características de XPS y Red Brick en IDS para OLTP/Warehousing y para Cargas-Mixtas.
  - Utilizando tecnología punta en:
    - Disponibilidad Continua y Replicación de Datos
    - Data Warehouse Accelerator para acelerar consultas pesadas
  - Integración del stack IBM para BI
    - Cognos
      - Express Manager
      - Cognos Reporter
      - Cognos Advisor
      - Advisor Server
        - » Data Advisor
        - » Xclerator
      - El gestor de contenidos de estos productos Cognos es Informix Dynamic Server
    - SPSS

¿Qué es IWA?





# Informix Ultimate Warehouse Data Warehouse Accelerator

## ¿Qué es?

*El Informix Warehouse Accelerator (IWA) es un optimizador de consultas, como un appliance, añadido a IDS, que permite la integración de procesos analíticos e informes de negociación con el sistema operacional. Acelera consultas 'SELECT', con tiempos de respuesta sin precedente.*



IWA está basado en la tecnología BLINK desarrollada por IBM Research en Almaden

## ¿Cómo Trabaja?

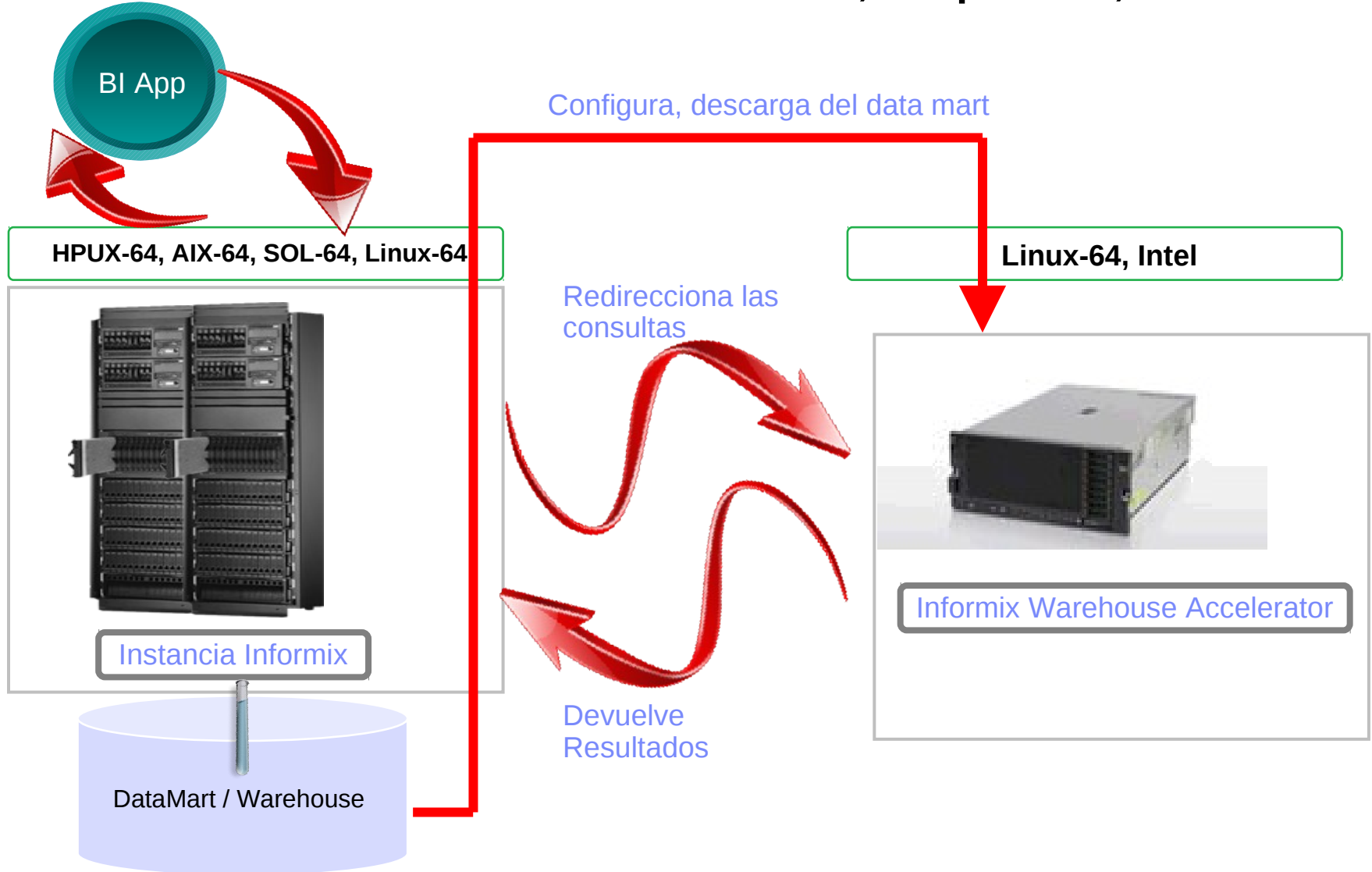
- **Rendimiento:** tiempos de respuesta más rápidos medidos en órdenes de magnitud (x10 - x100).
- **Integración:** se conecta a IDS permitiendo el acceso transparente a las aplicaciones.
- **Auto-gestión de la carga:** las consultas se ejecutan de la manera más eficiente.
- **Transparente:** las aplicaciones se conectan a IDS, no son conscientes del WA
- **Sencilla Administración :** como un appliance, no requiere intervención, eliminando muchas tareas de ajuste.

**Tecnología Innovadora que Genera Nuevas  
Oportunidades Para Su Negocio**



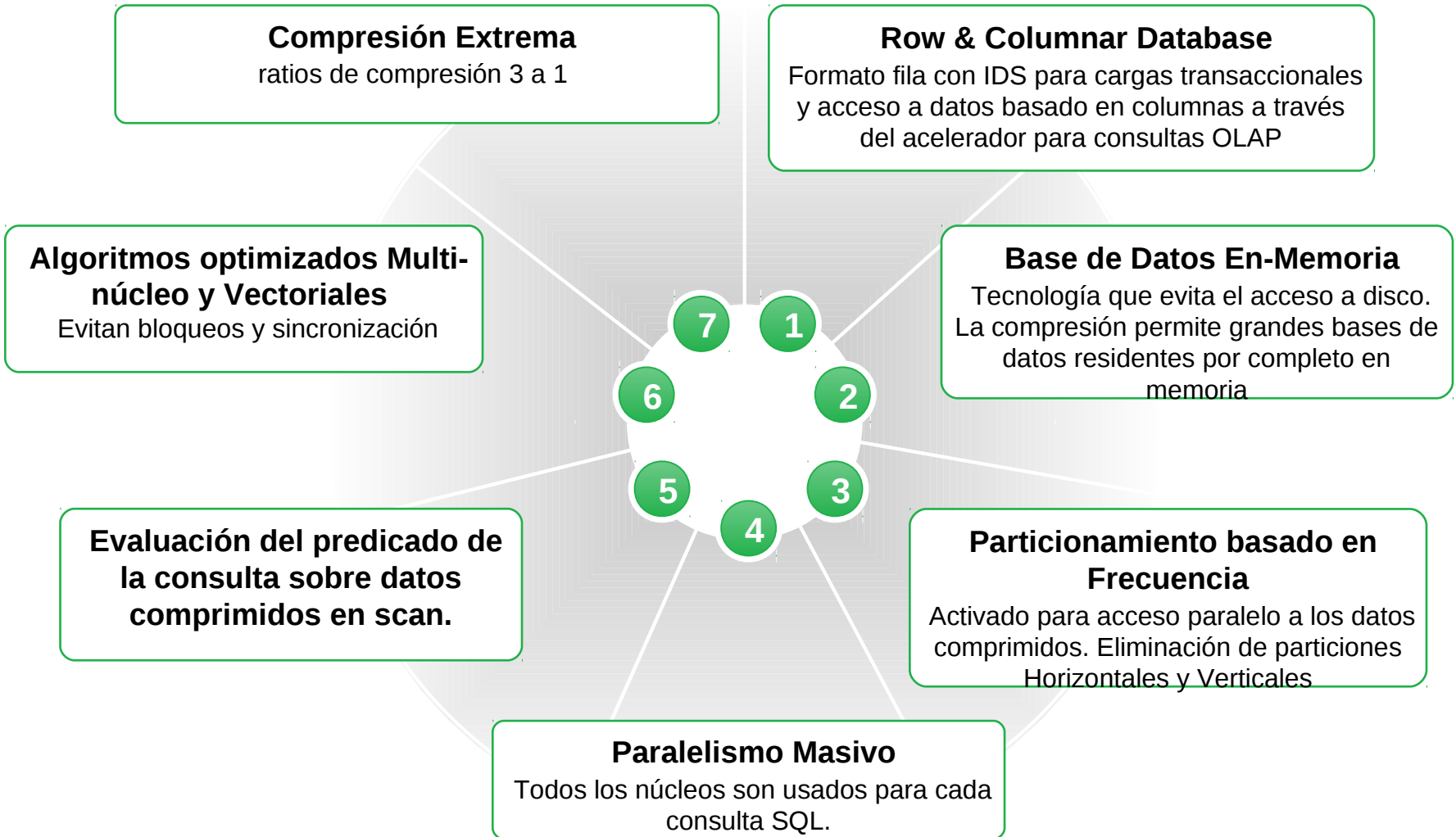


# Informix Ultimate Warehouse – Rendimiento, Simplicidad, Invisible





# Informix Ultimate Warehouse – Máximo Rendimiento





# IBM Smart Analytics Studio

The screenshot displays the IBM Smart Analytics Studio interface. The main window shows a data mart diagram for 'dm\_ast\_201101' with an estimated size of 0.4 GB. The diagram includes a source table 'tf\_ast\_201...' and several target tables such as 'td\_dwh\_sta...', 'td\_dwh\_ges...', 'td\_dwh\_beru...', 'td\_dwh\_ges...', 'td\_dwh\_alter...', 'td\_dwh\_alte...', 'td\_dwh\_arb...', 'td\_dwh\_ber...', and 'td\_dwh\_mo...'. The 'Data Project Explorer' on the left shows the project structure, and the 'Data Source Explorer' shows the database connections. The bottom panel shows the configuration for the 'DEMO' accelerator, including the name, shared memory available (0.3 GB of 0.4 GB), and trace profile (DEFAULT).

Database Development - demo/dm\_ast\_201101/default.mart\_diagram - IBM Smart Analytics Optimizer Studio

File Edit Diagram Navigate Search Project Run Data Mart Report Advise Window Help

Optim Solutions

Sans 9 B I A

Data Project Explorer

demo

Data Marts

dm\_ast\_201101 (demo)

\*dm\_ast\_201101 sel.sql

Estimated size: 0.4 GB

tf\_ast\_201... 0.4 GB

mon\_id  
uniq\_id  
bws\_id  
sex\_id  
age\_id  
gem\_id  
gst\_id  
ber\_id  
dim\_10\_1  
dim\_100\_1  
dim\_1000\_1  
dim\_10000\_1  
dim\_100000\_1  
anzahl

td\_dwh\_sta... 1.0 KB  
bws\_id  
bws\_bez

td\_dwh\_ges... 0.3 MB  
gst\_id  
aam\_id  
gst\_bez

td\_dwh\_beru... 2.6 MB  
ber\_id  
beg\_id  
ber\_bez

td\_dwh\_ges... 0.5 KB  
sex\_id  
sex\_bez

td\_dwh\_alter... 13 KB  
age\_id

td\_dwh\_alte... 1.5 KB  
agr\_id

td\_dwh\_arb... 26 KB  
aam\_id  
aam\_bez

td\_dwh\_ber... 26 KB  
beg\_id  
beg\_bez

td\_dwh\_mo... 6.0 KB  
mon\_id  
mon\_bez

Palette

Data Mart

Table

1:n  
n:m

SQL Results Execution Plan Properties

<Accelerator> DEMO

General Start Stop Configure Trace... Save Trace... Change Authentication Token

Data Marts

Software Version

Name: DEMO (Fully Operational)

Shared Memory available: 0.3 GB of 0.4 GB

Trace Profile: DEFAULT

Cluster status (1 of 1 coordinators and 1 workers active)

1 items selected

Herramienta gráfica fácil de utilizar para configurar el data mart



¡Increíble Rendimiento!



¡Acepta un 'NO' por respuesta!

¡No necesita mantenimiento!

~~Configuración del almacenamiento/particionamiento~~

¡No reorgs!!

~~Mantenimiento de índices~~



~~Mantenimiento de estadísticas~~

**NO**



~~Cambios de Aplicación~~



~~Mantenimiento de Cubos o tablas de agregados~~

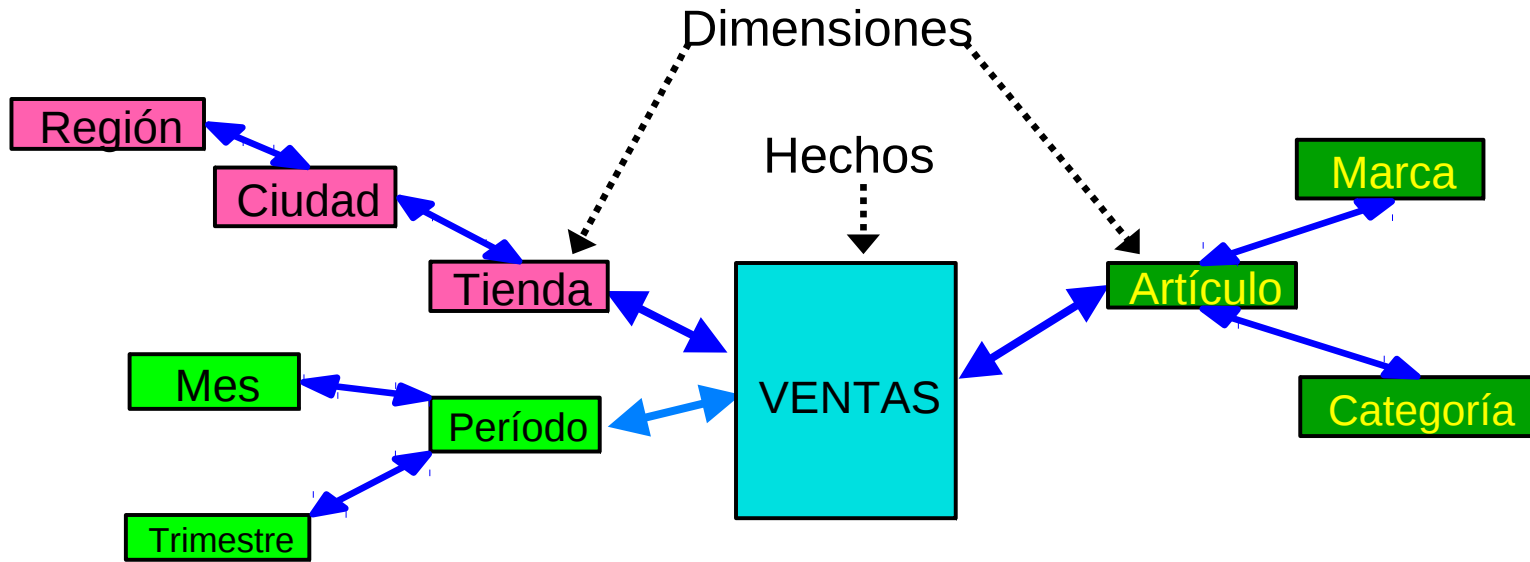
# IWA - Características

- Un sistema SMP dedicado (Linux en Intel x86\_64)
- No hay que cambiar las aplicaciones
  - Las aplicaciones se conectan a IDS.
  - Cuando una consulta es candidata para ser 'acelerada' IDS se la pasa al acelerador de manera transparente para las aplicaciones
- Mejoras medidas en órdenes de magnitud
- Reducción las necesidades de ajustar IDS (particionamiento, índices, etc.)
- Es como un appliance
  - No requiere administración sólo una mínima configuración
- Mejora significativa en precio/rendimiento y TCO por efecto de:
  - Aceleración de consultas analíticas complejas y consumidoras de recursos
  - Mejora el rendimiento en órdenes de magnitud
  - Ahorro del tiempo de DBA en continuos ajustes



# Mercado Objetivo: Business Intelligence (BI)

- Caracterizado por:
  - Modelos tipo “estrella” o “copo-de-nieve”:



## Consultas complejas, generadas dinámicamente

- Buscan tendencias, excepciones para toma de decisiones
- Acceden a grandes subconjuntos de datos
- Ordenan, agrupan, y agregan (ej., COUNT, SUM, AVG,...)
- **Haga su analítica empresarial con IWA**

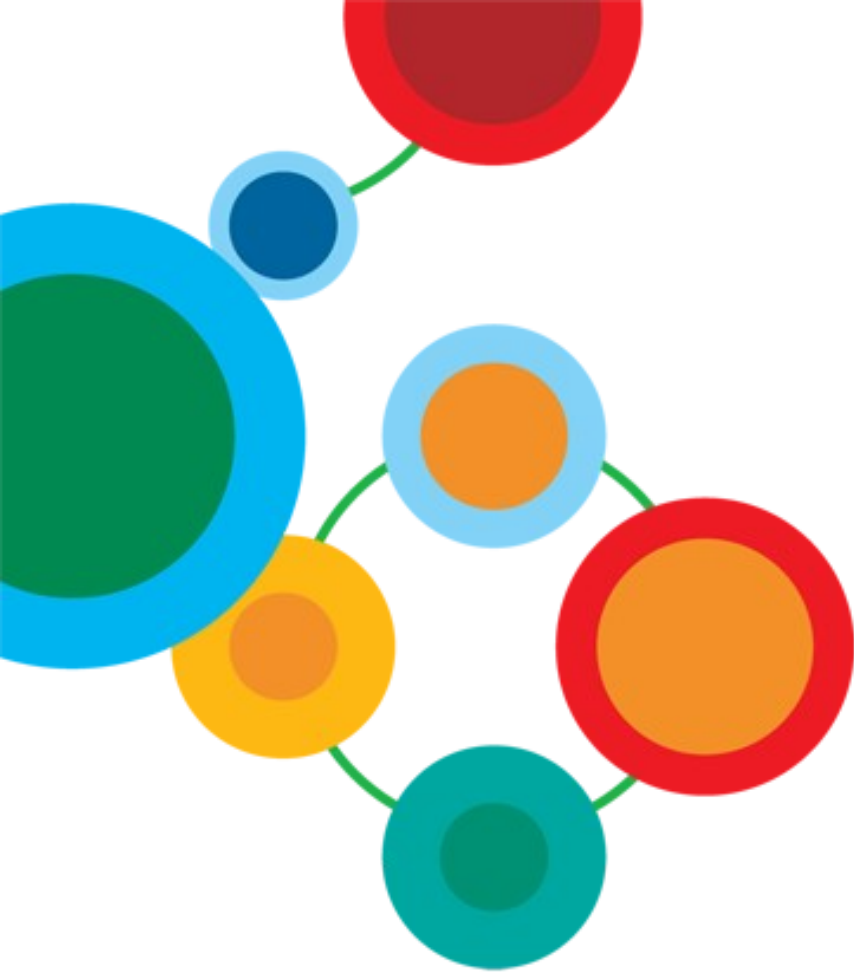


# Para Qué Está Diseñado IWA:

- Scans masivos o selectivos de grandes tablas de hechos
- Joins con tablas de dimensión más pequeñas
- Consultas de tipo OLAP contra modelos en estrella con agrupaciones y ordenaciones

```
SELECT articulo, region, SUM(facturacion)
FROM ventas V, productos P, clientes C, tiempo T
    INNER JOIN DIM_PRODUCT P ON V.FKP = P.PK
    INNER JOIN DIM_REGION R ON V.FKR = C.PK
    LEFT OUTER JOIN DIM_TIME T ON V.FKT = T.PK
WHERE T.año = 2009
    AND C.pais = 'Francia'
    AND P.tipo = 3
GROUP BY articulo, region
```





# Experiencias de Clientes

# Escenarios de Pruebas en Clientes

## Federal Agency - Germany

- 537 queries executed over a 30 GB data mart
- 432 Microstrategy queries accelerated via IWA
- Performance with IDS+IWA was **90 times faster than** the current system

## Federal Agency - USA

- Sample data set from 2 TB warehouse
- 8 representative queries ran **127 times faster**

## Large Shoe Retailer

- BI Queries on 150 GB data set
- Fact table - 1 billion rows
- Queries against Informix - 22 minutes
- Queries against Informix + IWA : **4 secs, 330 times faster**

## Large Global Retailer

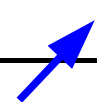
- 10 GB data mart sample
- Fact table - 25 million rows
- Queries against Informix + IWA: **4 secs**
- **5 to 275 times faster than** Teradata



# Sample Customer Result – Sketchers

The numbers say it all!

Query	Informix (secs)	Informix w/ IWA (secs)	Improvement
1	1320	4	<b>330 times</b>
2	63	2	31 times
3	220	2	110 times
4	1800	4	<b>450 times</b>
5	120	2	60 times
6	1800	2	<b>900 times</b>
7	2700	2	<b>1350 times</b>



much smaller variation in run times  
 ⇒ more predictable response times



# Paper with Intel for Skechers



SPONSORED SECTION: PERFORMANCE-OPTIMIZED INFORMATION MANAGEMENT



SPONSORED SECTION: PERFORMANCE-OPTIMIZED INFORMATION MANAGEMENT

## Maximize Data Analytics Performance

Rethink the data warehousing equation and empower data-driven business decisions

Skechers USA Inc. is a global leader in lifestyle footwear with more than US\$2 billion in annual revenue. Retail daily sales and inventory data is loaded each night into their data warehouse system. About 15 business analysts use this data to understand retail sales and market trends, and to decide on product marketing and promotions.

With more than a billion rows in fact tables, analytical queries could take anywhere from a few minutes to 45 minutes to complete. Time lost waiting for answers is a productivity loss because it limits the number of business analytical queries being executed. This puts more pressure on DBAs to tune each query.

Skechers isn't the only company facing these hurdles. We live in a data-rich world, and the opportunities that come with unlimited data are epic. Retailers have access to fine-grain demographics, buying history, social media identities, online community involvement, information from public records, and more. Manufacturers record and track supplier metrics in exquisite detail, from performance against delivery deadlines to breakage and discount levels. Law enforcement agencies rely on enormous amounts of data and information, locally, nationally, and internationally. Criminal and scientific data that is fast and easy to access allows law enforcement to nail down airtight cases.

When it's used in a timely manner to drive critical business decisions, data analysis equals competitive advantage. The key phrase is "when it's used in a timely manner."

Companies are finding that as data volumes grow, they overwhelm storage systems and the analytic engines that turn data into game-changing insight. In this article, we'll look at how and why the tremendous potential of unlimited data can become a huge challenge, and how Intel and IBM are partnering to help organizations structure data delivery and data management processes that turn unlimited data into unlimited business value opportunity.

### Companies strive for rapid analysis

Trying to keep up with unlimited data puts a huge strain on an entire organization, starting with slow analytics response times. As organizations amass more and more data, even basic queries can take a substantial amount of time. More sophisticated queries—particularly those generated by business analytics applications—take even more time. In extreme cases, the query may return results already out of date.

Businesses and public agencies are devoting growing amounts of their IT budgets to handle these vast quantities of data. Data tools and infrastructures have become quite efficient, but even the most advanced technologies cannot keep pace with unlimited data. Creating schemas, indexing, building cubes and summary tables, and creating partitioning strategies all require extensive planning and human effort. Where a single database administrator (DBA) once managed a centralized data center for the entire company, there may now be many DBAs



Finally, the accelerator is designed to run on commodity hardware and to easily scale with a growing infrastructure. The accelerator runs on all hardware platforms supported by Informix, and significant data warehouse acceleration can be achieved while minimizing hardware costs.

Informix Warehouse Accelerator is already delivering real-world results. "Complex inventory and sales analysis queries on the enterprise warehouse took anywhere from a few minutes to 45 minutes to run," says Ashutosh Khunta, manager of data management services for Skechers USA. "When we ran those same queries using Informix Warehouse Accelerator, they finished in 2 to 4 seconds. That means they ran from 60 to 1,400 times as quickly, with an average acceleration factor of more than 450—all without any index or cube building, query tuning, or application changes!"

### Informix Warehouse Accelerator runs best on the Intel Xeon processor E7 family

Transforming traditional disk I/O-based query execution to in-memory execution requires a new level of processor support. The Intel Xeon processor E7 family delivers the high processor performance needed to accelerate even the most demanding warehouse queries, and supports large quantities of system RAM, enabling organizations to maximize the benefits of the Informix Warehouse Accelerator in-memory technology.

The Intel Xeon processor E7 family is engineered to meet demands for rapidly increasing compute capacity, faster performance, and high availability. It provides up to 10 cores, up to 20 hyper-threads, and 30 MB of last-level cache per processor, all in a four-socket server. The Intel Xeon processor E7 family supports 32 GB dual in-line memory modules (DIMMs), enabling application workloads that require larger memory capacities. Together with Informix Warehouse Accelerator, the Intel Xeon processor E7

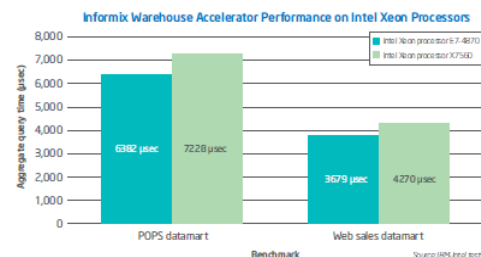


Figure 1: The Intel Xeon processor E7-4870 works with IBM Informix Warehouse Accelerator to deliver increased performance over previous generations of processors for datamarts of widely varying sizes. The proof of performance and scalability (POPS) datamart consists of two fact tables, one with 350 million rows and one with 1 billion rows. The web sales datamart is relatively small, consisting of a single fact table with 7 million rows.

family eliminates the need for disk I/O and dramatically improves the speed of access to complex data sets.

Also, with the Intel Xeon processor E7 family, organizations do not need to introduce additional infrastructure security capabilities. Protection against hardware errors and malicious software attacks is automatically managed by advanced, built-in features. For example, Machine Check Architecture (MCA) enables the CPU to coordinate with the OS to gracefully recover from memory errors.

Finally, advanced compression capabilities in Informix Warehouse Accelerator enable the Intel Xeon processor E7 family to handle more data in the same memory configurations without sacrificing performance. With Intel® Turbo Boost Technology, processor performance can be enhanced beyond normal operating constraints when thermal, power, and current limits have not been exceeded.

The Intel Xeon processor E7 family—together with in-memory software solutions—demonstrates innovation that can lead to a more efficient performance experience for your data warehousing implementations. Extensive testing by Intel and IBM shows that the Intel Xeon processor E7 family provides a significant performance boost (see Figure 1).

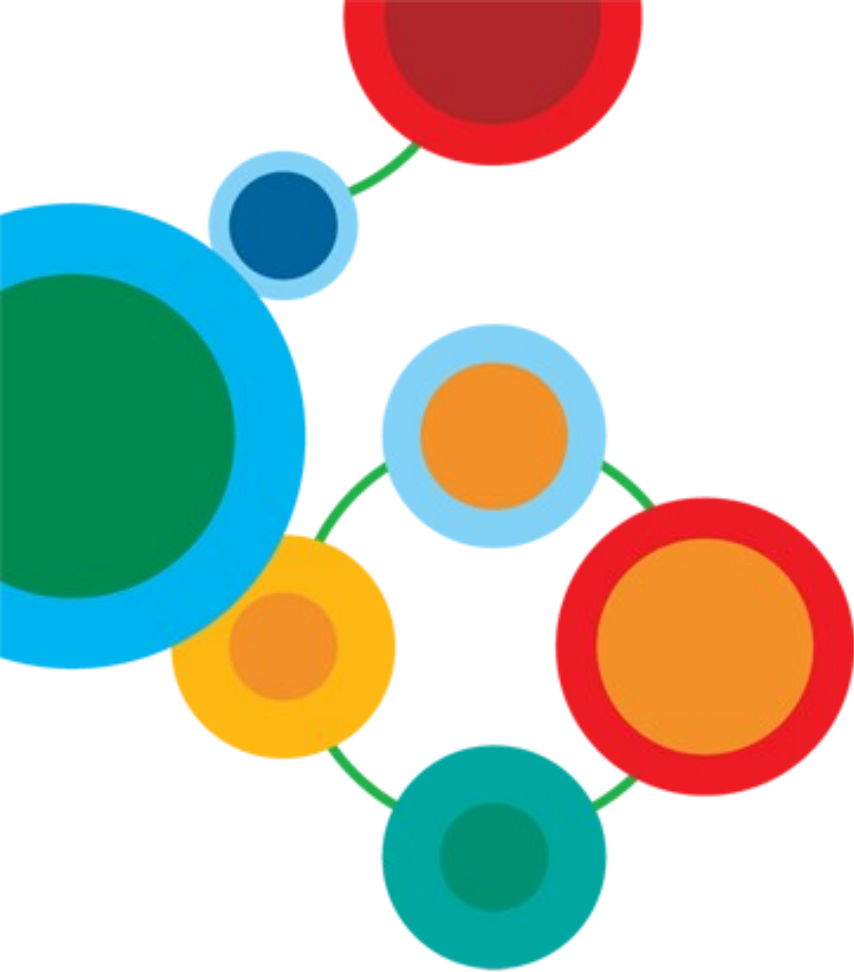




# Datamart at a Government Agency

- Microstrategy report was run, which generates
  - 667 SQL statements of which 537 were Select statements
- Datamart for this report has 250 Tables and 30 GB Data size
- Original report on XPS and Sun Sparc M9000 took 90 mins
- With IDS 11.7 on Linux Intel box, it took 40 mins
- With IWA, it took 67 seconds:
  - 60 seconds within Microstrategy.
  - and 7 seconds within IWA.





## Ediciones IWA



# Informix Ultimate Warehouse Edition (IUWE) vs. Informix Growth Warehouse Edition (IGWE)

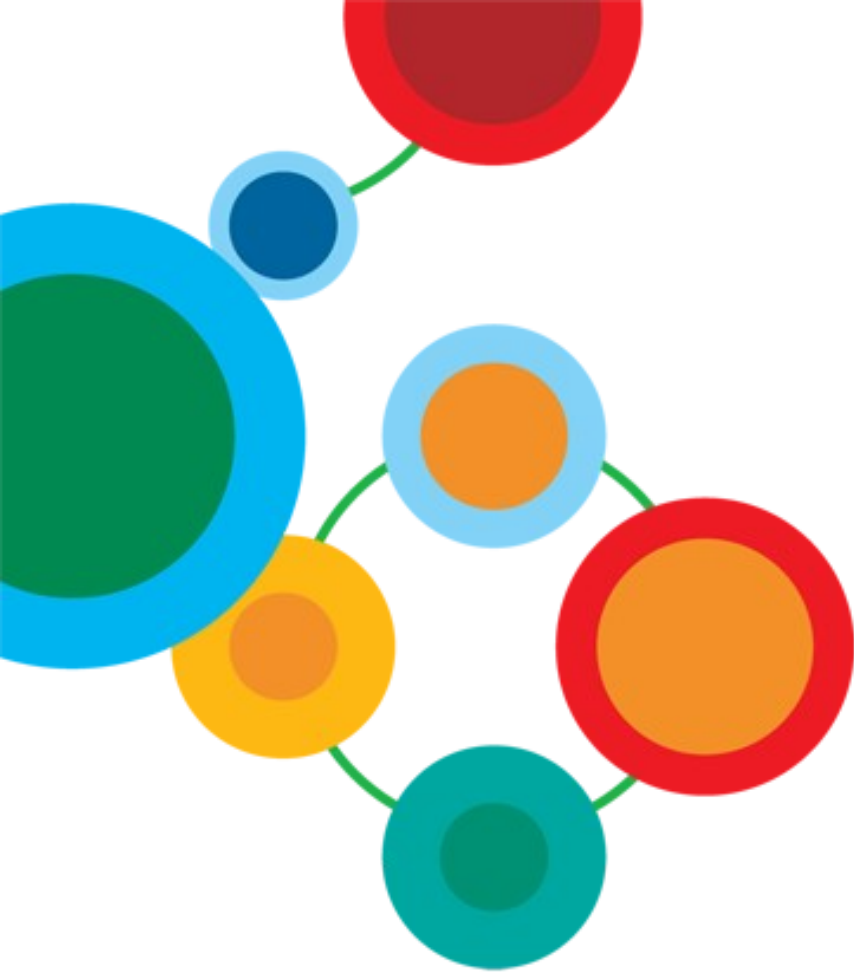
	IUWE	IGWE
Components	Informix Ultimate Edition Compression IWA ISAO Studio	Informix Growth Edition IWA ISAO Studio
Limits	Max memory available No core limits Informix on 4 platforms: AIX64, Sol64, HPUX64, Linux-ntel64 IWA on Linux-Intel 64	Informix Growth <i>16 cores, 16 GB Memory max</i> Informix on 4 platforms: AIX64, Sol64, HPUX64, Linux-Intel64 IWA on Linux-Intel 64 <i>48 GB Max, 1600 PVU limit</i>
Target	> 300 GB of raw data	< 250 GB of raw data





# Componentes de IWA

- Linux en Intel x86\_64 (RHEL 5 o SUSE SLES 11).
- IDS 11.70 + IWA, software incluye IDS Stored Procedures.
- ISAO Studio Plug-in – definición del data mart mediante herramienta gráfica
- OnIWA – Utilidades para monitorizar IWA.



## Hardware de Referencia

# IWA Referenced Hardware Configuration

Intel(R) Xeon(R) CPU	X7560 @ 2.27GH 4 X 8
Memory	512G
6 disks	300 GB SAS hard disk drives each



## Options:

- 4-processor, 4U rack-optimized enterprise server with Intel® Xeon® processors.
- 8-core, 6-core and 4-core processor options with up to 2.26 GHz (8-core), 2.66 GHz (six-core) and 1.86 GHz (four-core) speeds with up to 16 MB L3 cache
- Scalable from 4 sockets and 64 DIMMs to 8 sockets and 128 DIMMs
- Optional MAX5 32-DIMM memory expansion
- 16x 1.8" SAS SSDs with eXFlash or 8x 2.5" SAS HDDs



# IWA Hardware Platform for Clustering (11.70xC4)

- Referenced platform for clustering is BladeCenterS.
- Also supports 64 bit INTEL based platforms supporting Linux (RHEL & SLES).





# 11.70xC4 – Blade Server Support

Product	Business use	Processor (Max)	Memory (Min/Max)	Internal storage (Max)
 <p>IBM BladeCenter HX5</p> <p>→ More details → Browse &amp; Buy</p>	Ideal applications include database, virtualization, business intelligence, modeling and simulation and other enterprise applications.	2- and 4-socket, six-core, eight-core or ten-core Intel® Xeon®	Up to 256 GB per 2-socket HX5; scalable up to 640 GB using MAX5 expansion blade	Up to 100 GB per 2-socket HX5 using solid-state drives; scalable to 200 GB per 4-socket HX5 (with 50 GB SSDs)
 <p>IBM BladeCenter HS22V</p> <p>→ More details → Browse &amp; Buy</p>	High-density, high-performance blade server with expanded memory ideal for optimizing performance for virtualization and memory-intensive applications	2-socket, six-core or quad-core Intel® Xeon®	Up to 288 GB of VLP DDR-3 memory	Up to 100 GB total internal storage
 <p>IBM BladeCenter HS22</p> <p>→ More details → Browse &amp; Buy</p>	Versatile, easy-to-use blade optimized for performance, power and cooling - ideal for most enterprise applications including virtualization, hosted client and SAP	2-socket, six-core or quad-core Intel® Xeon®	Up to 192 GB of VLP DDR-3 memory	Up to 1.0TB with 2 hot-swap SAS, SATA or solid-state drives

## Hardware:

- HS22 192GB/blade, up to 640GB per blade.
- Up to 14 blade slots.
- Stackable blade chassis.

## Software:

- Fact table divided among nodes.
- Dimension tables replicated to each node.
- Multiple coordinator and worker nodes.