



DB2 V9 UNIX, Windows y Linux

Juan García Erviti
Luis Reina Juliá



DB2 Information Management Software

ON DEMAND BUSINESS™

© 2006 IBM Corporation



Agenda

- Compresión de Datos.
- Gestión Automática de la Memoria: Self-Tuning Memory Manager.
- Café y visita al Salón de Soluciones
- Datos Relacionales y XML Puro: SQL y XQuery.
- Nueva Característica de Seguridad: LBAC.
- Particionamiento con DB2: Particionamiento de Tablas, MDC y Hashing.
- Gestión Automática del Almacenamiento.
- Incremento de los Limites: Tablespaces, Indices, etc.
- Otras características.
- IBM DB2 9 para sistemas SAP.
- Cocktail y Visita al Salón de Soluciones



DB2 V9

COMPRESION DE DATOS



ON DEMAND BUSINESS™

© 2006 IBM Corporation



Evolución de la Compresión en DB2

- **V8 GA** – Compresión de Nulos y Valores por defecto

- **V8 GA** - Multidimensional Clustering (MDC)
 - ▶ Indices de tipo *block* permiten una reducción en el tamaño del índice
 - Una clave para miles de registros (versus una clave por registro con índices tradicionales)

- **V8 FP4** – Compresión de Backups
 - ▶ Imágenes de backup más pequeñas.

- **DB2 9** – Compresión de Filas



Compresión de Filas: Conceptos I

- **Basado en un Diccionario: Tabla para comprimir/descomprimir registros**
 - ▶ Basado en el Algoritmo Lempel-Ziv (LZ) utilizando símbolos de 12 bits.
 - ▶ Un diccionario por tabla almacenado con la tabla (~75KB).
- Los datos están comprimidos en las páginas de DB2 (en disco y bufferpool)
 - ▶ Ahorro de Entrada/Salida.
 - ▶ Ahorro de Memoria (bufferpool).
 - ▶ Costes de CPU
 - Filas tienen que ser descomprimidas antes de ser procesadas.
- *Logs* de registros comprimidos también comprimidos.
- No comprime filas que no supongan un ahorro en almacenamiento para esa fila:
 - “red” -> ‘#1#2#3’ símbolo 1=“r” símbolo 2=“e” símbolo 3=“d”
 - “red” == 3 bytes ‘símbolos1,2,3’ == 12bits+12bits+12bits (4.5 bytes)



Compresión de Filas: Conceptos II

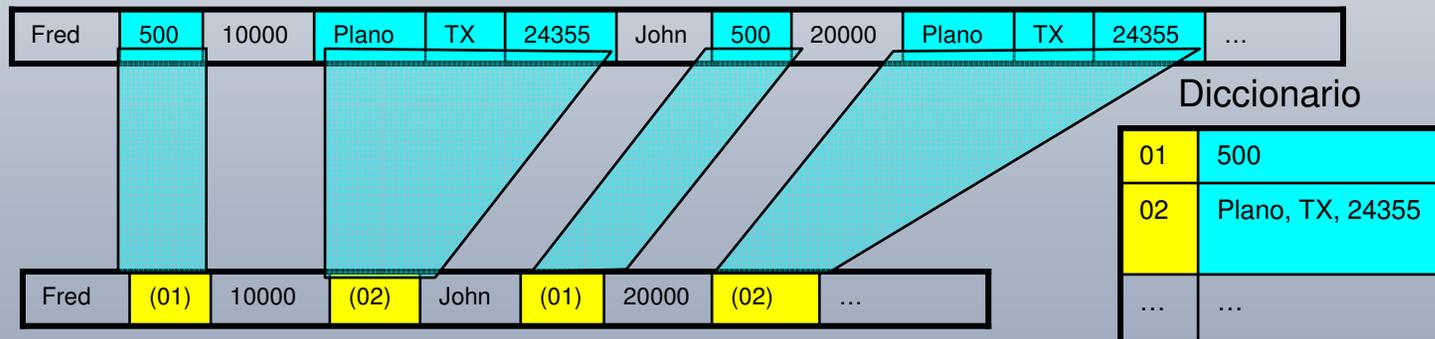
- Se trata de sustituir: Reemplazar cadenas con símbolos de 12 bits
 - “En un lugar de la Mancha....Fin”
 - “En un” -> símbolo 0
- Usa un Diccionario Estático:
 - ▶ Lista del tipo: [Cadena, símbolo <n>]
 - ▶ Previamente creado y almacenado en la tabla
 - {(“En un”, 0),(“lugar”,1),(“de la”,2).....(“Fin”,4095)}
- Compresión es el proceso de encontrar una cadena del diccionario (la más larga) en una fila y sustituirla por el símbolo correspondiente
- Expansión o descompresión es más sencilla: Partiendo del símbolo de la fila buscar y reemplazar la cadena del diccionario



Compresión de Filas: Ejemplo

- Datos de texto tienden a comprimirse bien debido a la existencia de cadenas recurrentes así como datos con muchos caracteres repetidos y blancos al comienzo y al final.

Name	Dept	Salary	City	State	ZipCode
Fred	500	10000	Plano	TX	24355
John	500	20000	Plano	TX	24355



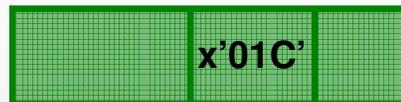


Compresión de Filas: Páginas

Fila Descomprimida

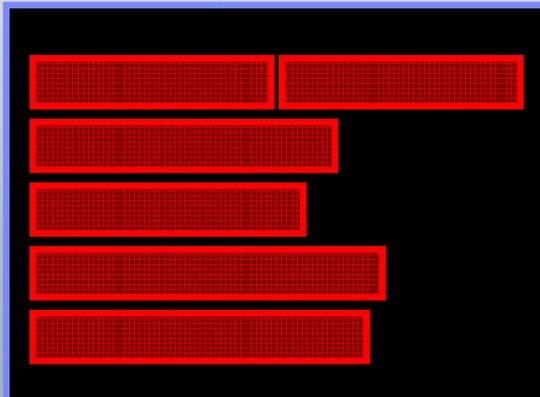


Fila Comprimida

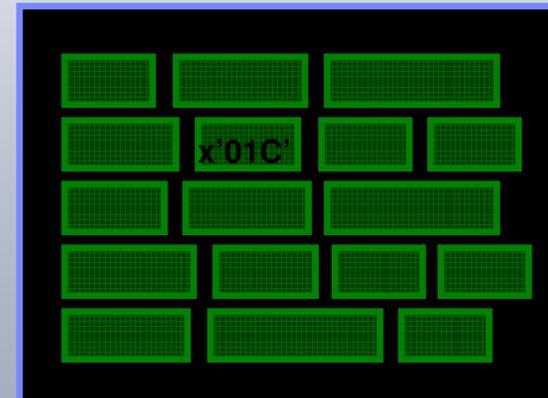


Símbolo de 12 bits

Páginas con Filas no comprimidas



Páginas con Filas comprimidas



“Más Filas en un Página”



Compresión de Filas: DDLs

- **CREATE TABLE <table name> --->**

|---COMPRESS NO---|

---+-----+--->

|---COMPRESS YES--|

- **ALTER TABLE <table name> --->**

---+-----+--->

|--COMPRESS--+YES--+--|

|--NO--|

- La compresión se habilita a nivel de tabla via CREATE o ALTER TABLE.
- La compresión será efectiva una vez se haya construido el diccionario



Compresión de Filas: Catálogo

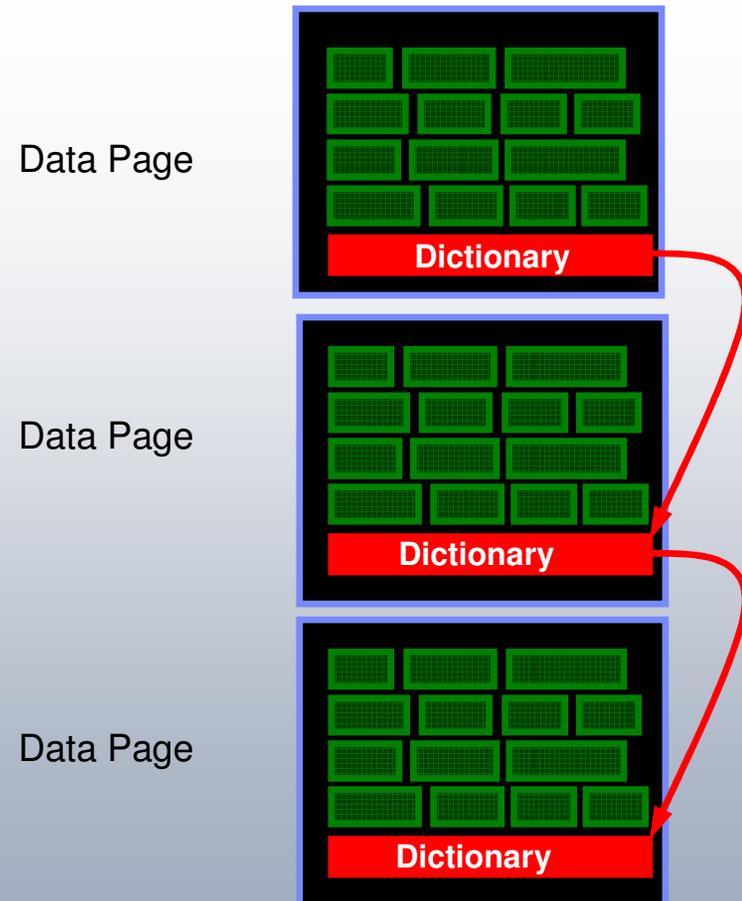
- Vista: SYSCAT.TABLES
- Nueva Columna: "COMPRESSION"

Valor	Significado
N	No hay compresión
V	Compresión de Valor por Defecto
R	Compresión de Fila
B	Ambos



Compresión de Filas: Diccionario

- Una tabla tiene que tener un diccionario de compresión antes de poder comprimir las filas.
- Se Almacena en la misma tabla:
 - ▶ En “filas” no seleccionables enlazadas entre sí.
 - ▶ Del orden de 75KB.
- Se puede crear de 2 formas:
 - ▶ REORG
 - ▶ INSPECT





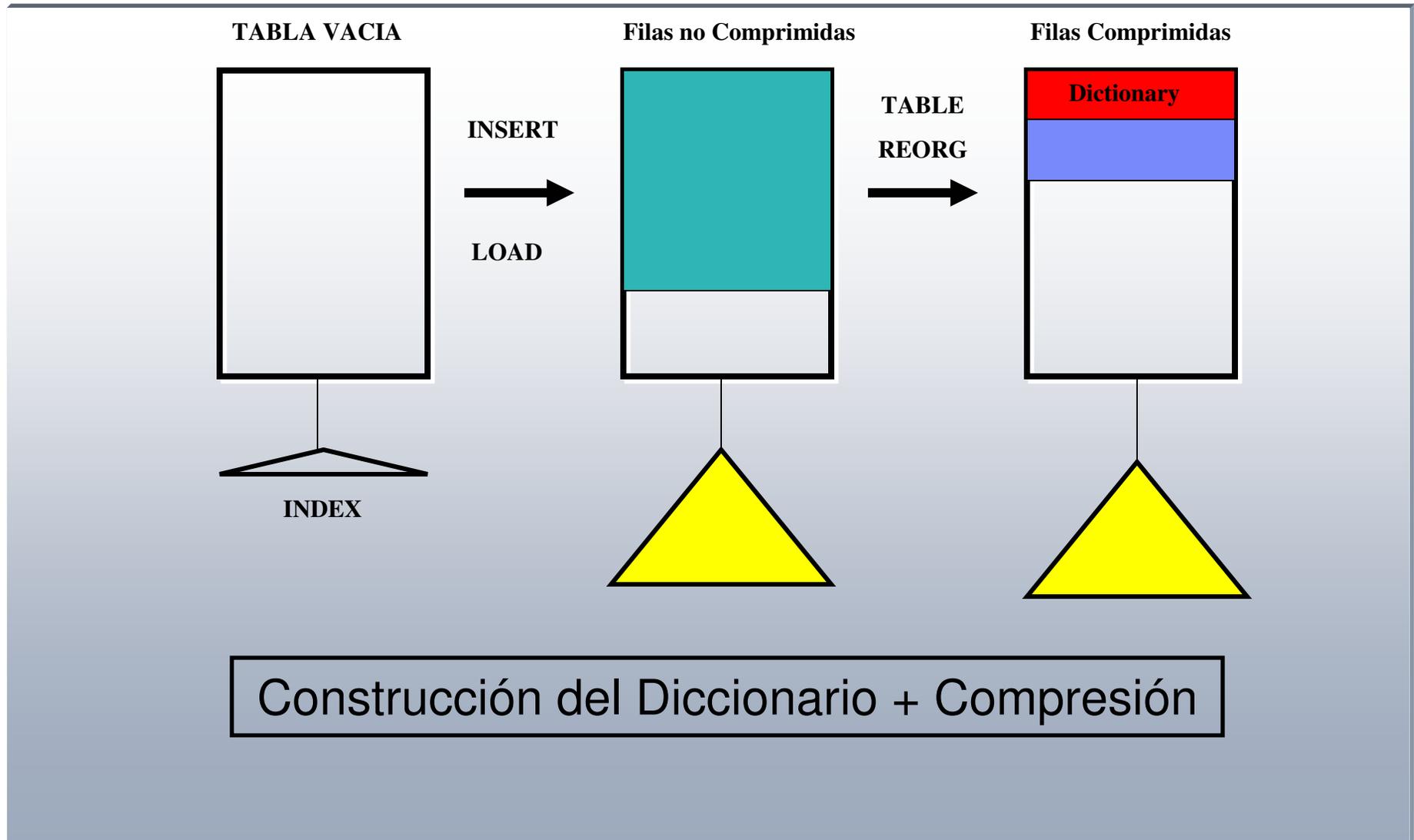
Creación del Diccionario : REORG OFFLINE

```
>--REORG--<table name>--+-----+----->
                                '--INDEX--<index name>--'
    .-ALLOW READ ACCESS-.
>--+-----+--+-----+--+-----+--+-----+-->
    '-ALLOW NO ACCESS---'  '-USE-<tbspace>-'  '-INDEXSCAN-'
                                .-KEEPDICTIONARY---.
>--+-----+--+-----+--+-----+--+----->
    '-LONGLOBDATA-'  '-RESETDICTIONARY---'
```

- Todas las filas de la tabla “participan” en la construcción del diccionario.
- Cuando el diccionario se construye se necesita un buffer de 10MB
 - ▶ Se toma esta memoria del “utility heap”



Creación del Diccionario : Ejemplo





Gestión del Diccionario con REORG

RESETDICTIONARY

COMPRESS Atributo?	Diccionario Existe?	Resultado
SI	SI	Construye Diccionario; filas comprimidas
SI	NO	Construye Diccionario; filas comprimidas
NO	SI	Elimina Diccionario; filas descomprimidas
NO	NO	Sin Efecto

KEEPDICTIONARY

COMPRESS Atributo?	Diccionario Existe?	Resultado
SI	SI	Preserva Diccionario; filas comprimidas
SI	NO	Construye Diccionario; filas comprimidas
NO	SI	Preserva Diccionario; filas descomprimidas
NO	NO	Sin Efecto



Creación del Diccionario : INSPECT I

```
>>-INSPECT+-| Check Clause |-----+----->
      '-| Row Compression Estimate Clause |-'
...
|--ROWCOMPESIMATE-TABLE--+-NAME--<table-name>--+-----+--+
|
|                                     '-SCHEMA-- <schema-name>-' |
| '-TBSpaceID-- <tbspc id> --OBJECTID-- <obj id>----`
```

- Crea el diccionario “ONLINE” y proporciona su tamaño (mirando todas las filas).
- Estimación del nivel de compresión (usando una muestra).
- Filas insertadas posteriormente serán comprimidas (si *compress* habilitado – alter o create).
- Filas existentes no son comprimidas -> Se precisa REORG



Creación del Diccionario : INSPECT II

- Salida de Inspect en el directorio db2dump
- Formatear salida con
 - ▶ `db2inspf <entrada> <salida>`
- <salida> contiene la estimación:

Acción: ROWCOMPESTIMATE TABLE

Nombre de esquema: SP29388

Nombre de tabla: LUIS

ID de espacio de tabla: 3 ID de objeto: 6

Nombre de archivo de resultados: entrada

Inicio de fase de tabla (ID con signo: 6, Sin signo: 6; ID de espacio de tabla: 3) : SP29388.LUIS

Inicio de fase de datos. Objeto: 6 Espacio de tabla: 3

Resultados del cálculo de la compresión de filas:

Porcentaje de páginas guardadas desde la compresión: 80

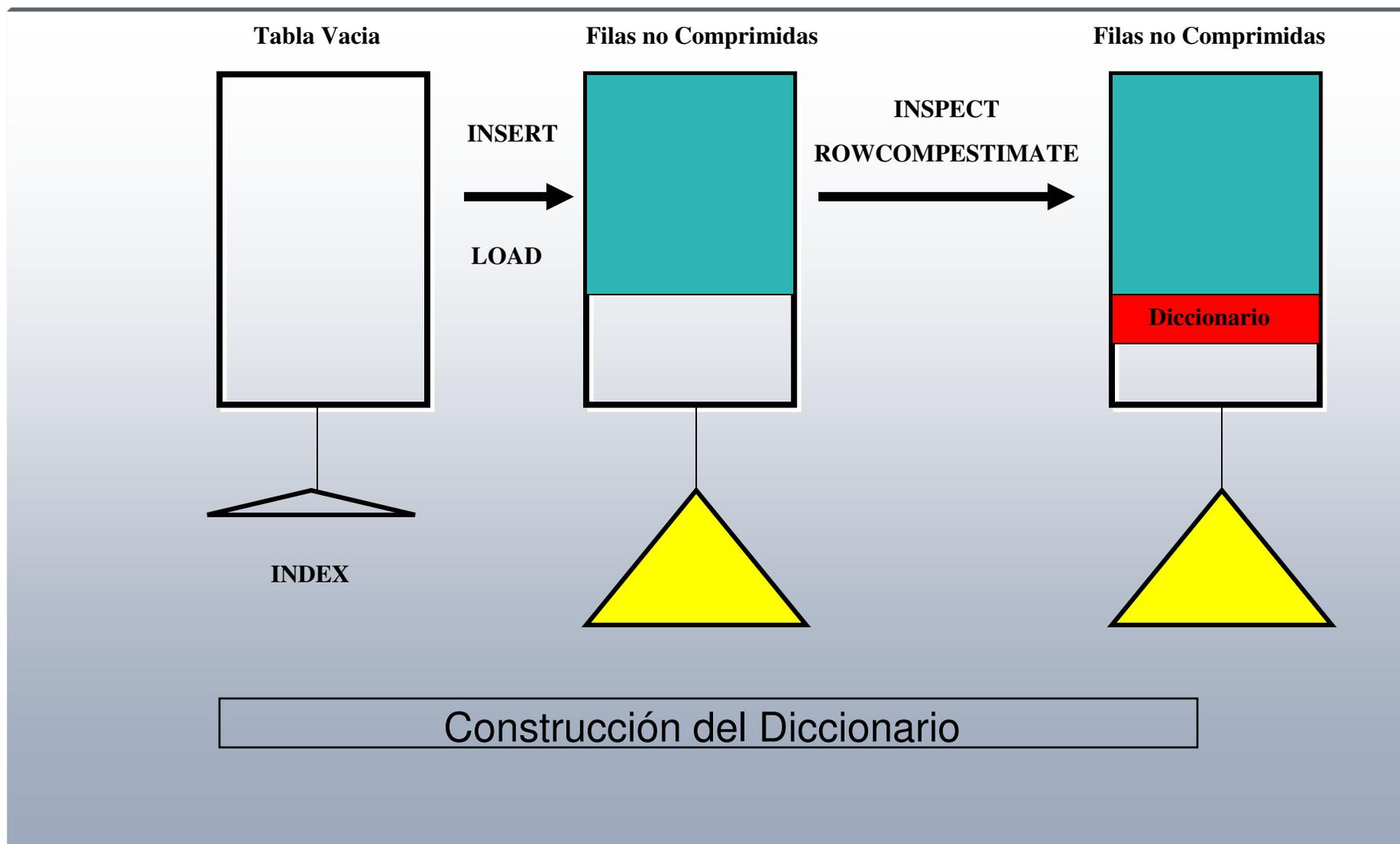
Porcentaje de bytes que se han ahorrado con la compresión: 80

Porcentaje de filas no elegibles para la compresión debido al pequeño tamaño de la fila: 0

Tamaño del diccionario de compresión: 5248 bytes.



Creación del Diccionario : Ejemplo





Ejemplo Completo

- Crea la tabla y habilita la compresión:
CREATE TABLE Sales (...) COMPRESS YES
- Carga de datos representativos:
LOAD FROM filesmall OF DEL REPLACE INTO Sales
- Crea el diccionario usando los datos cargados:
REORG TABLE Sales
- Carga nuevos datos usando el diccionario creado en el paso anterior:
LOAD FROM filerest OF DEL INSERT INTO Sales



Table Function para Administración

```
db2 describe "select * from
table(sysproc.admin_get_tab_info('MIKEW', 'STAFF')) as t"
```

sqltype	sqllen	sqlname.data	sqlname.length
449	VARCHAR	128 TABSCHEMA	9
449	VARCHAR	128 TABNAME	7
453	CHARACTER	1 TABTYPE	7
501	SMALLINT	2 DBPARTITIONNUM	14
497	INTEGER	4 DATA_PARTITION_ID	17
453	CHARACTER	1 AVAILABLE	9
493	BIGINT	8 DATA_OBJECT_L_SIZE	18
493	BIGINT	8 DATA_OBJECT_P_SIZE	18
493	BIGINT	8 INDEX_OBJECT_L_SIZE	19
493	BIGINT	8 INDEX_OBJECT_P_SIZE	19
493	BIGINT	8 LONG_OBJECT_L_SIZE	18
493	BIGINT	8 LONG_OBJECT_P_SIZE	18
493	BIGINT	8 LOB_OBJECT_L_SIZE	17
493	BIGINT	8 LOB_OBJECT_P_SIZE	17
493	BIGINT	8 XML_OBJECT_L_SIZE	17
493	BIGINT	8 XML_OBJECT_P_SIZE	17
501	SMALLINT	2 INDEX_TYPE	10
453	CHARACTER	1 REORG_PENDING	13
449	VARCHAR	10 INPLACE_REORG_STATUS	20
449	VARCHAR	12 LOAD_STATUS	11
453	CHARACTER	1 READ_ACCESS_ONLY	16
453	CHARACTER	1 NO_LOAD_RESTART	15
501	SMALLINT	2 NUM_REORG_REC_ALTERS	20
453	CHARACTER	1 INDEXES_REQUIRE_REBUILD	23
453	CHARACTER	1 LARGE_RIDS	10
453	CHARACTER	1 LARGE_SLOTS	11
493	BIGINT	8 DICTIONARY_SIZE	15



Estadísticas de Compresión

- SYSCAT.TABLES
- Nuevos campos de Medias y Porcentajes
- Usados por el optimizador para determinar costes.
 - ▶ AVGROWSIZE
 - ▶ PCTPAGESSAVED
 - ▶ PCTROWSCOMPRESSED
 - ▶ AVGROWCOMPRESSIONRATIO
 - ▶ AVGCOMPRESSEDROWSIZE



Compresión de Backup y Filas

Escenarios	Tiempo (segundos)	Páginas Usadas	Tamaño del Tablespace (GB)	Tamaño de la imagen de Backup (GB)
Sin Compresión	468	1.510.400	11.57	12
Sólo Compresión de Backup	1028	1.510.400	11.57	4.2
Sólo Compresión de Filas	198	610.816	4.68	4.7
Ambas Compresiones	662	610.816	4.68	4.2

¿Merece la Pena usar ambos tipos de Compresión?



Limitaciones en la Compresión de Filas

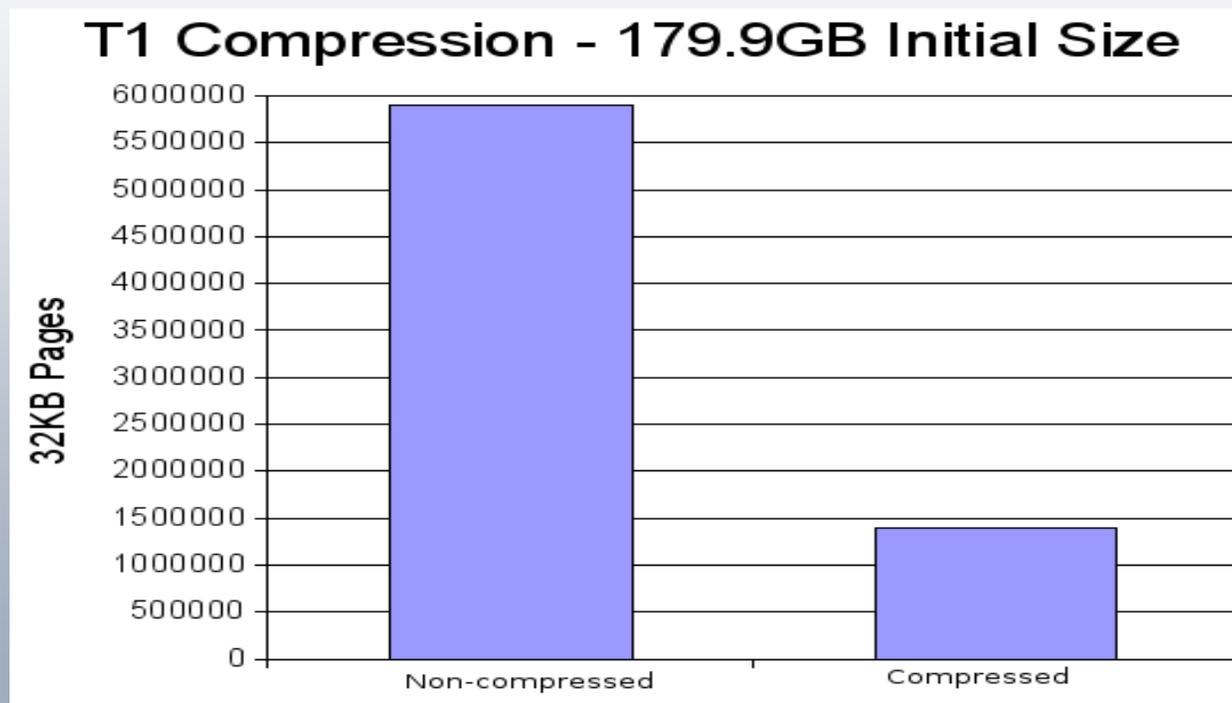
- Sólo se soporta compresión de tablas soportadas por REORG. No se soportan:
 - Tablas del Catálogo.
 - Declared global temporary tables.
 - System temporary tables.
 - Tablas RCT (range-clustered tables).
- Sólo se realiza sobre las filas del tablespace. No se comprime:
 - ▶ Datos XML (almacenamiento nativo)
 - ▶ LOBs (no van en las páginas de datos)
 - ▶ LONGs (no van en las páginas de datos)
- Compresión de filas y Replicación de datos no son compatibles
 - ▶ Opción DATA CAPTURE CHANGES no es compatible con la opción COMPRESS YES.
- Compresión no es práctica para tablas más pequeñas de 100KB
 - ▶ Debido a lo que ocupa el diccionario.



Ratios de Compresión I

<u>Compression Type</u>	<u>32KB Page Count</u>	<u>Space Required on Disk</u>
No compression	5893888	179.9GB
Row compression	1392446	42.5GB

% Pages Saved: 76.4%





Ratios de Compresión II

Table	% of Table Sampled	Number of Rows	Sample Size (GB)	Sample Size Compressed (GB)	Compression Ratio	Total Estimated Disk Savings (GB)
T1	100	159259747	71.49	13.23	81.5%	58
T2	10	77390826	21.48	5.23	75.7%	162
T3	100	191362309	141.46	36.18	74.4%	105
T4	20	19474261	11.80	3.05	74.2%	43
T5	10	23370458	25.78	8.54	66.9%	172
T6	10	39573969	23.24	6.32	72.8%	169

Reducción de 2.8 TB a 1.4TB del Data Warehouse



Conclusiones de la Compresión

- Se pueden obtener grandes reducciones en Almacenamiento.
- Rendimiento neto dependerá si la base de datos está CPU o *I/O bound*.
 - ▶ Mejora en E/S (20-30% mejor)
 - ▶ Coste de CPU (10% mayor consumo)
- Mejores que la competencia (Mejores Ratios)
 - ▶ SQL Server -> no tiene
 - ▶ Oracle -> diccionario a nivel de página y bulk load no soportado
 - ▶ Teradata -> A nivel de columna con límite de 255 valores por columna



DB2 V9

**GESTION AUTOMATICA
DE LA MEMORIA**
(Self Tunning Memory Manager - STMM)



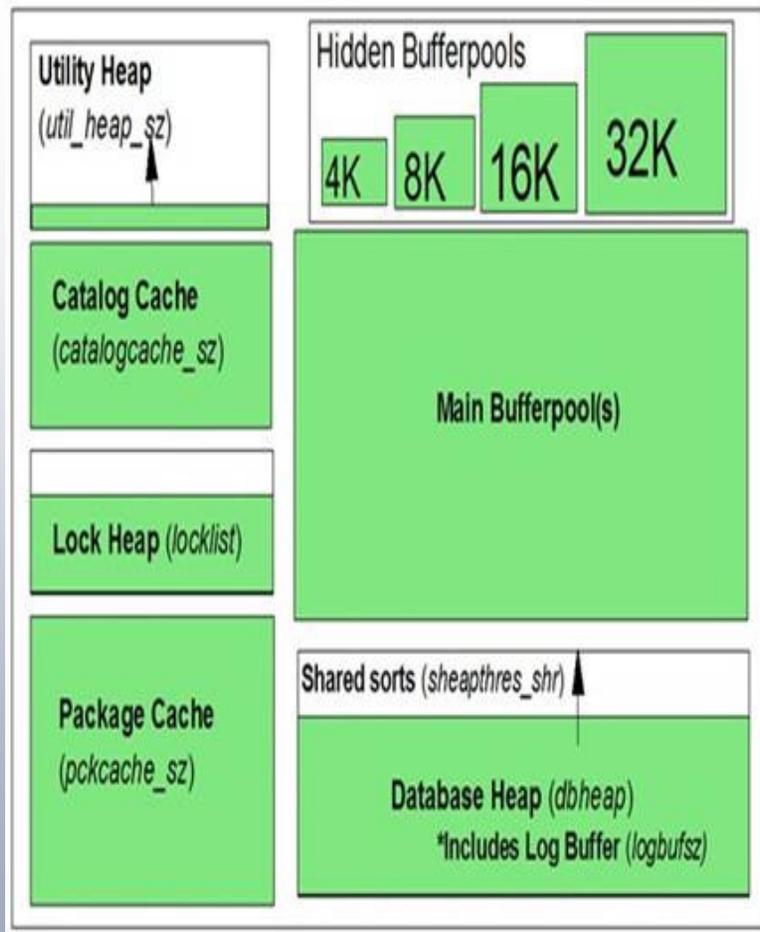
ON DEMAND BUSINESS™

© 2006 IBM Corporation



SELF_TUNING_MEMORY = ON

Modelo de La Memoria en DB2



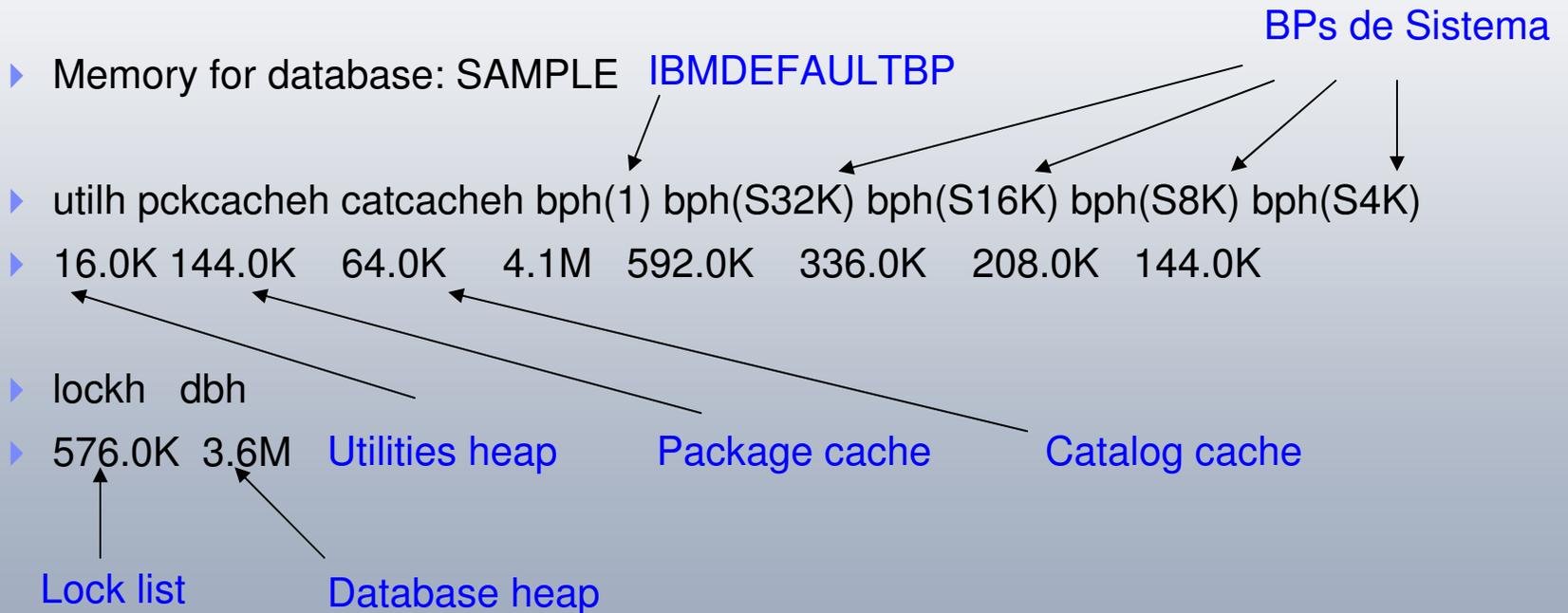
- Modelo de Memoria complejo
- 32 bit tamaño limitado
- 64 bits tamaño virtualmente ilimitado.
- Se divide en “heaps”
 - ▶ Bufferpool y locklist se establecen al activar la base de datos.
 - ▶ Otros heaps cuando se necesitan.



Visualización de la Memoria Usada

- Memory Tracker (db2mtrk)

- ▶ db2mtrk -d
- ▶ Tracking Memory on: 2006/06/29 at 11:08:29





Configuración de la memoria en DB2 V8

- La mejor forma es usar el *Configurator Advisor*
 - ▶ Comando `autoconfigure` desde la línea de comandos (CLP) ó
 - ▶ Forma Gráfica -> Centro de Control
 - ▶ Establece los parámetros de memoria más importantes.
 - ▶ Configura las cargas de trabajo bien.
 - En OLTP configuración similar a un experto.
- Consideraciones:
 - ▶ La configuración es estática
 - Si la carga de trabajo cambia la configuración puede ya no ser buena.
 - ▶ La configuración depende de “la descripción” que el Administrador “le diga” al *Configurator Advisor*
 - Si la carga luego no coincide, la configuración puede no ser buena.



Configuration Advisor desde el Centro de Control

Configure Performance Wizard

Review the performance configuration recommendations.

Based on your selections in this wizard, as well as the volume of data in the database, and system information, this wizard recommends the following values. Below the list, specify if you want to save the new values to a script, or apply them to the database immediately.

Parameter	Current value	Suggested value	DB2 Parameter
Application control heap size	128	128	app_ctl_heap_sz
Buffer pool size	250	35956	buffpage
Catalog cache size	32	262	catalogcache_sz
Changed pages threshold	60	60	chnpggs_thresh
Database heap size	600	912	dbheap
Default degree	1	1	dft_degree
Default prefetch size	16	32	dft_prefetch_sz
Maximum storage for lock list	50	329	locklist
Log buffer size	8	32	logbufsz
Log file size	250	250	logfilsiz
Number of primary log files	3	6	logprimary
Number of secondary log files	2	10	logsecond
Maximum number of active applications	40	84	maxappl
Maximum locks per application	22	13	maxlocks

Apply these recommendations immediately

Save these recommendations to the Script Center

Script name:

Script description:

◀ Back Finish Cancel



Self Tuning Memory Manager STMM



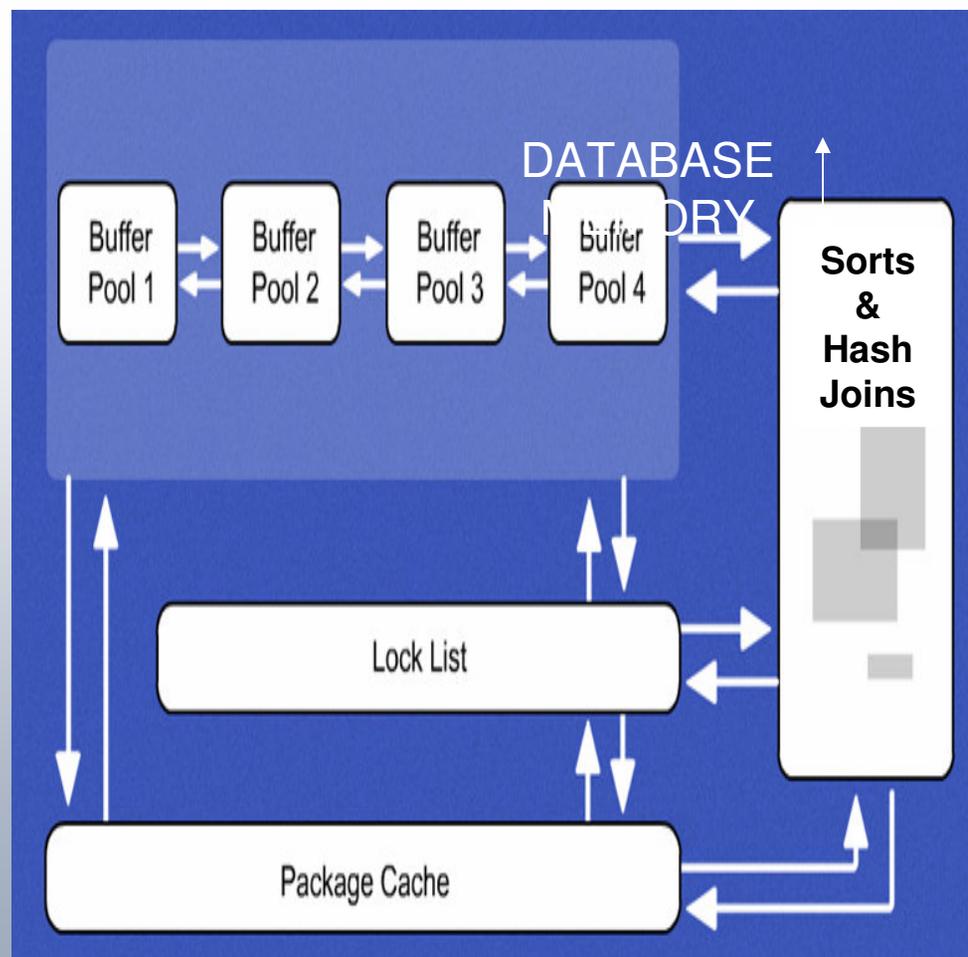


SELF-TUNING MEMORY MANAGER

- Revolucionario gestor automatico de la memoria
- Se encarga de dar y quitar memoria entre los distintos heaps.
- También libera y coge memoria del sistema operativo según se necesite.
- Sin intervención del DBA -> basado en la carga de trabajo actual.

STMM: Heaps Controladas

- Buffer pools
- Package cache
- Bloqueos: locklist y maxlocks
- Memoria de *sorts:sheaphres_shr* y *sortheap*
- Memoria *shared: database_memory*





Como Activar/Desactivar STMM

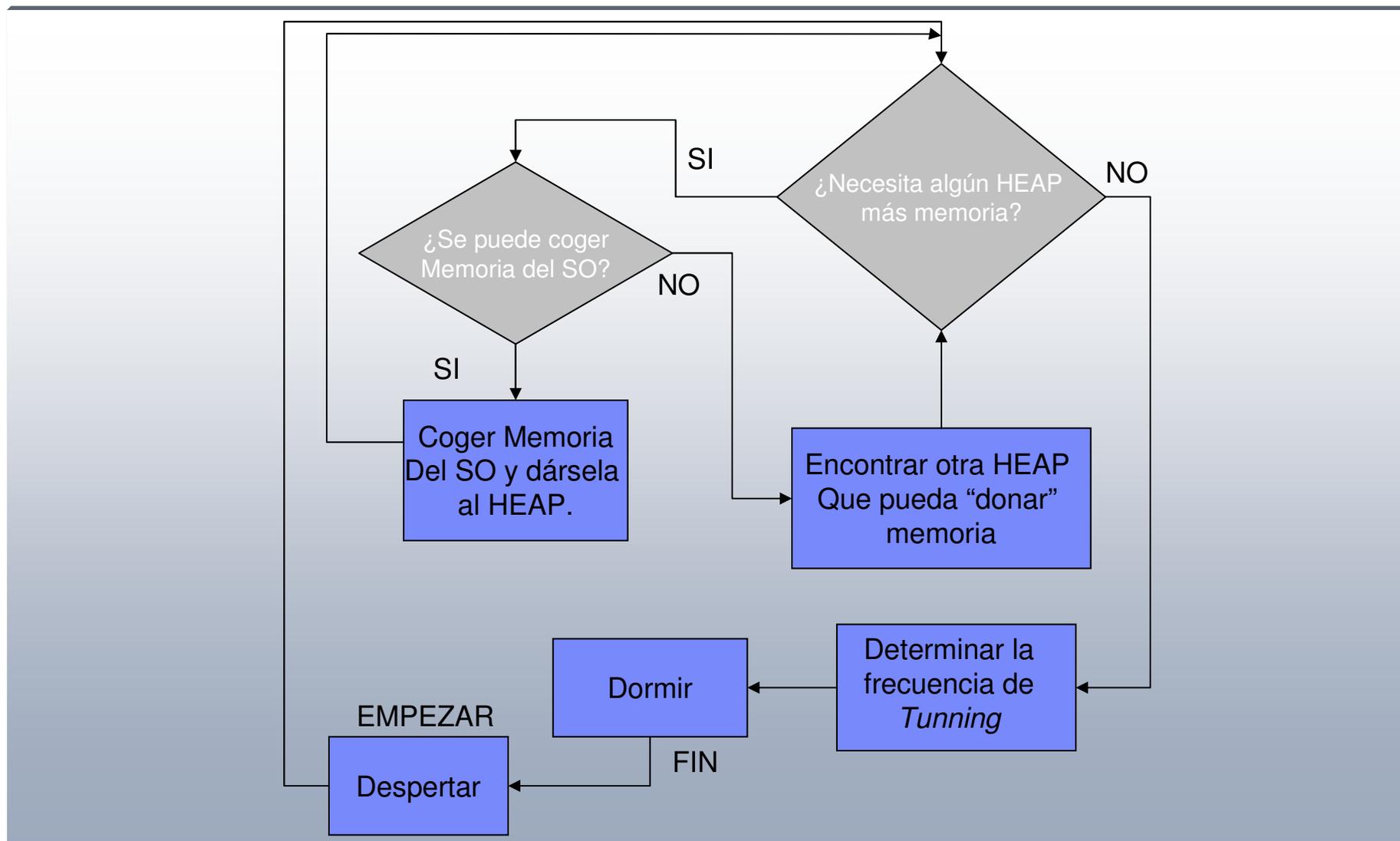
- SELF_TUNING_MEM ON (por defecto)
- Cada Heap se puede incluir o no (ONLINE): por defecto AUTOMATIC
 - ▶ DATABASE_MEMORY = AUTOMATIC
 - ▶ LOCKLIST = AUTOMATIC
 - ▶ MAXLOCKS = AUTOMATIC
 - ▶ PCKCACHESZ = AUTOMATIC
 - ▶ SHEAPTHRES_SHR= AUTOMATIC
 - ▶ SORTHEAP = AUTOMATIC
 - ▶ MAXLOCKS = AUTOMATIC
 - ▶ BUFFERPOOL:

Alter bufferpool bp1 size automatic (por defecto)

Create bufferpool bp1 size automatic (por defecto)



STMM: Como Funciona





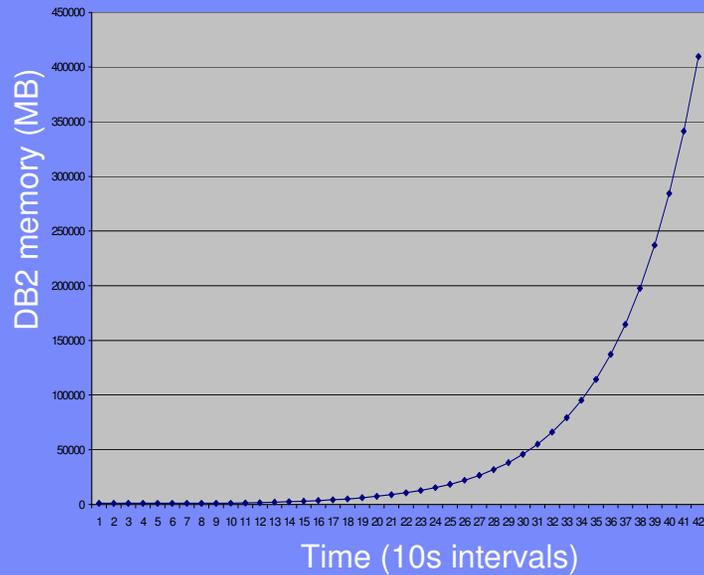
DPF y STMM

- Todos las Particiones se Configuran igual
 - ▶ Sólo un Memory Tunner para todas la particiones.
 - ▶ Misma configuración para todas las particiones.
 - ▶ Importante que los máquinas (nodos físicos) sean equivalentes.
 - ▶ Particiones sean consistentes.
 - ▶ Se configura una partición y la nueva configuración se envía al resto en cada intervalo.
 - ▶ Las particiones deben estar activas.
 - ▶ La partición de configuración se elige en la primera activación de la base de datos.
 - ▶ Existe un algoritmo de selección de partición para elegir una que sea representativa (media).
 - ▶ La partición que decide la configuración puede ser cambiada a través del SP Admin.
 - ▶ Se puede poner, para una partición, STMM=OFF
 - Catálogo
 - Coordinador (sin datos)
 - También a nivel de HEAP (valor distinto de AUTOMATIC)



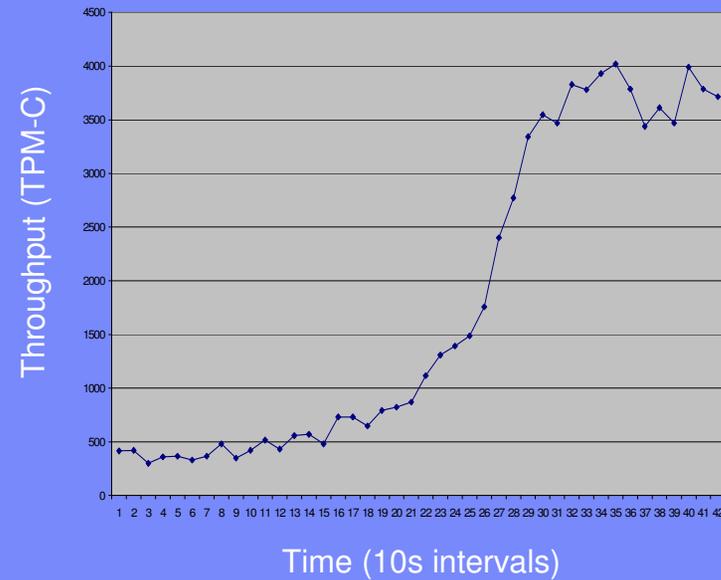
Ejemplos I

Automáticamente da más memoria a DB2



En rendimiento del Sistema mejora inmediatamente

10 veces





Ejemplos II

TPCH Query 21 - After drop index - Average times for the 10 streams

