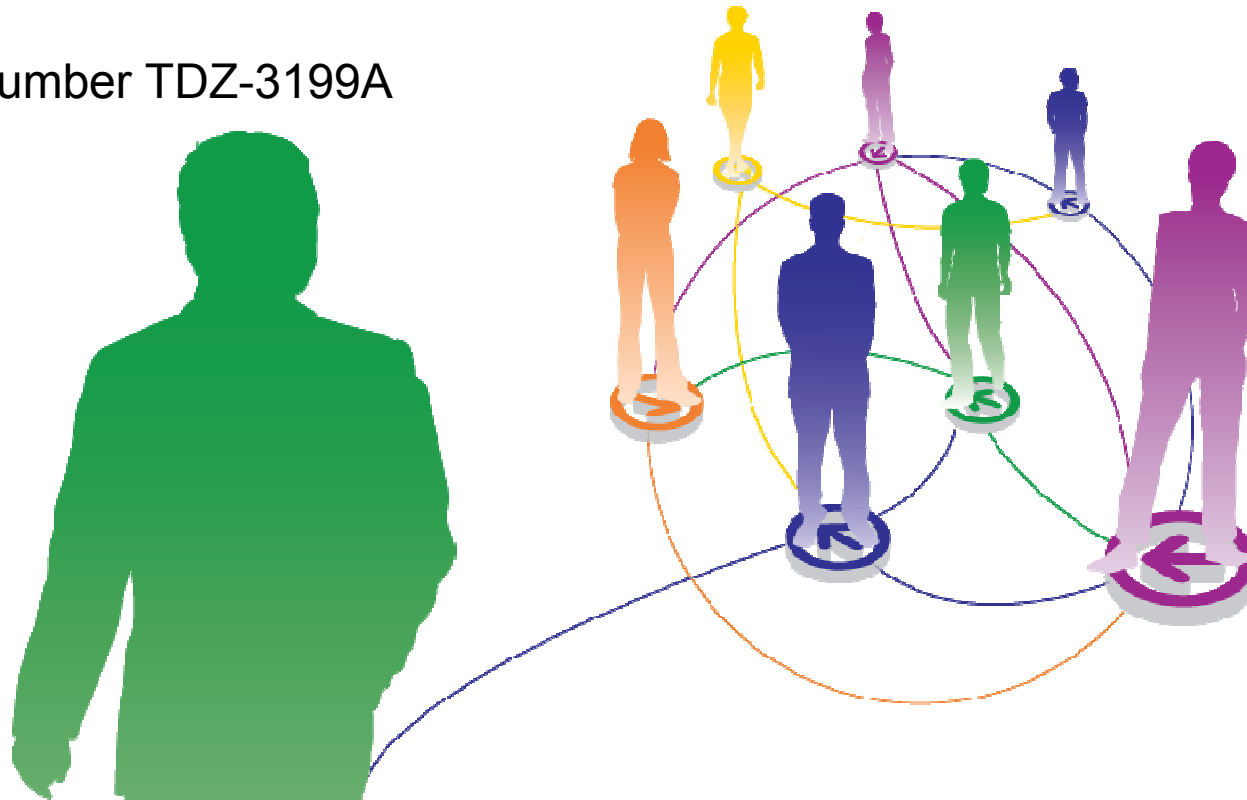# DB2 9 for z/OS Migration Planning & Experience Part 2

John J. Campbell, Distinguished Engineer, IBM DB2 for z/OS Development

campbelj@uk.ibm.com

Session Number TDZ-3199A

Objectives

→ Share lessons learned, surprises, pitfalls

→ Provide hints and tips

→ Address some myths

→ Provide additional planning information

→ Information on new enhancements

Agenda

➜ Quick hits

  – Preparing for the migration

➜ Migration

  – Overview

  – Plan stability

  – Converged TEMP space

➜ What to expect?

  – DB2 9 CPU performance

  – DBM1 Virtual Storage relief below the 2GB bar

➜ More quick hits

  – New Functions

DB2 9 CPU Performance

→ The target for DB2 9 CPU performance is to be roughly equivalent or marginally better relative to V8

  – Assumes z890, z990, z9 or z10

  – Assumes no access path regression

  – Assumes REBIND on the new release

→ Mileage in terms of reduced CPU reduction will vary

→ Do not spend the benefits until you see them

→ Customers running DB2 9 on old hardware (z800/z900) will definitely see CPU regression - may be 10-15% increase

→ Data sharing customers running on DB2 9 (NFM) may see significant savings from reduced LC19 contention and less spin to get unique LRSN
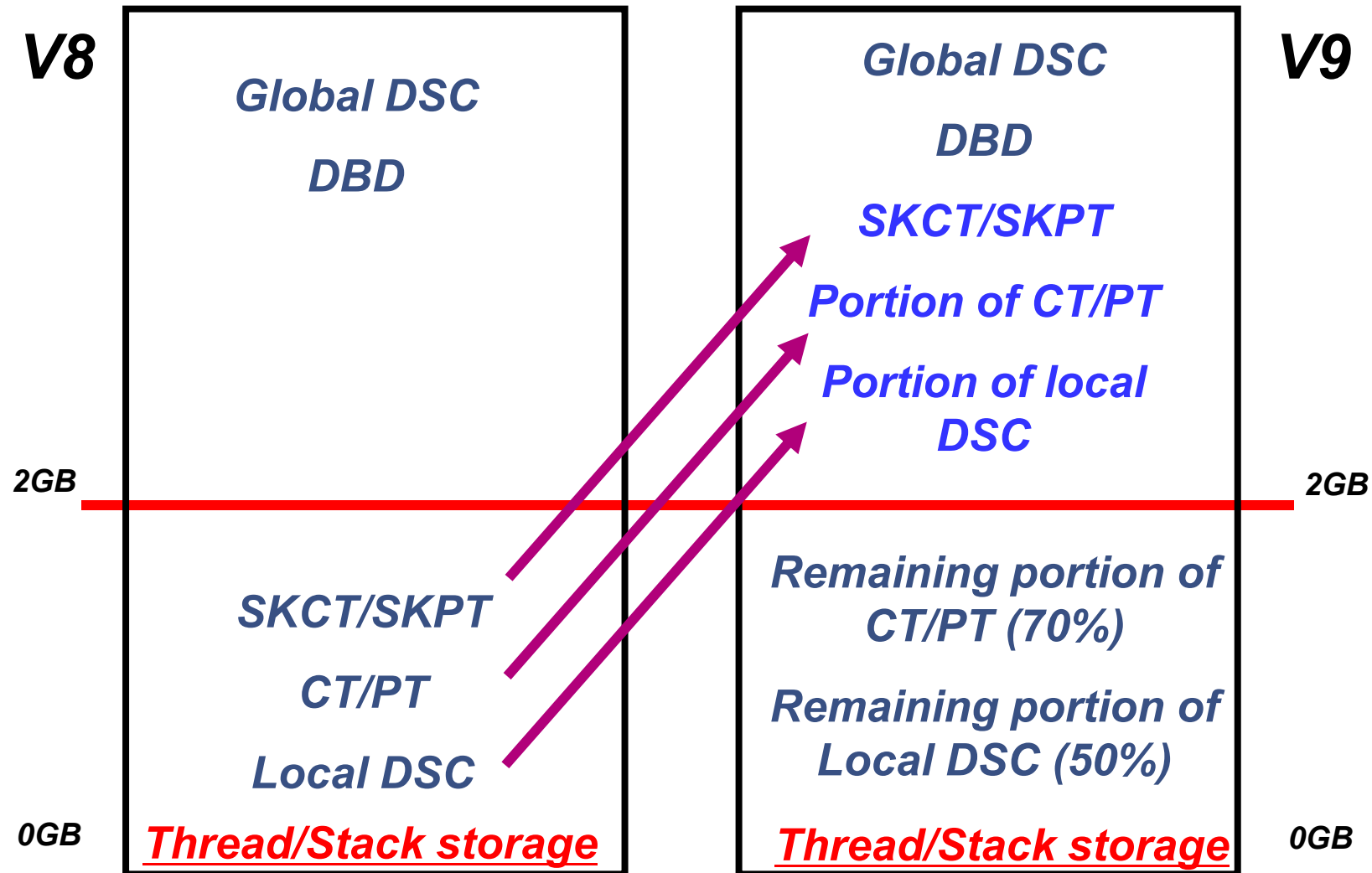
DB2 9 CPU Performance ...

→ Dynamic prefetch replaces all sequential prefetch in SQL calls (CM), except in tablespace scan

- No rebind is required to switch to dynamic prefetch as this change is transparent to Optimizer

- In V8 when pages are read in via dynamic prefetch they are marked as sequential and when they are subsequently touched by random getpage they are reclassified as random

- In DB2 9 the pages are not re-classified following subsequent touch by random getpage

DBM1 Virtual Storage below 2GB

V8

Global DSC

DBD

2GB

SKCT/SKPT

CT/PT

Local DSC

0GB **Thread/Stack storage**

Global DSC

DBD

SKCT/SKPT

Portion of CT/PT

Portion of local DSC

V9

2GB

Remaining portion of CT/PT (70%)

Remaining portion of Local DSC (50%)

Thread/Stack storage 0GB

DBM1 Virtual Storage below 2GB …

→ DBM1 Virtual Storage Constraint Relief for static SQL users

- EDM pool in DB2 9 – need REBIND

  - SKCT/SKPT moved above 2GB

  - A portion (close to 30%) of CT/PT moved above 2GB

  - Average estimated reduction of 60% but wide fluctuation from 20 to 90%

  - Only non-stealable components (CTs/PTs) are left in the EDM pool below 2GB

  - Plan on at least keeping 2 * 0.7 * CT/PT (V8 peak) for EDM Pool below 2GB

    - There are no stealable EDM components now below 2GB
    - Increase in package size after REBIND
    - Increased likelihood of hitting EDM Pool full condition
    - Intent of 2x factor is to provide some headroom

→ DBM1 VSCR for dynamic SQL users

- Local dynamic statement cache (KEEPDYNAMIC=YES)

  - Rough estimation of V9 = 50% of V8

DBM1 Virtual Storage below 2GB …

→ User thread storage, System thread storage, Stack storage

- Current expectation of less than 10% difference overall

→ Potential reduction can range from 0 to 300MB depending on thread/stack storage usage

- Very few installations will see as much as a 200-300MB reduction!
  - Especially if you are an IMS/TM customer who has very large ECSA and a small Extended Private Region (i.e., 1GB)
  - No VSCR for dynamic SQL users unless they are using local dynamic statement cache as a result of using BIND option KEEPDYNAMIC(YES)
  - VSCR from local dynamic statement cache savings will be small if the V8 size is heavily constrained by low MAXKEEPD

DBM1 Virtual Storage below 2GB …

→ Introduced with V8 APAR PK20800 8/07

- -DISPLAY THREAD(*) SERVICE(STORAGE)

- Produces DSNV492I message that can be used by DB2 service for diagnostics

  - V91A   N *   0 003.RCRSC 02 SYSOPR   0067   0
  - V490-SUSPENDED 07213-09:59:18.02 DSNRCRSC +00000230 01.51
  - V492-LONG 252 K VLONG 40 K 64 1028 K

- Includes Agent Local Non-System Storage usage

- Does not include Getmained Stack Storage usage

- The key values are the Long storage pool and the Very Long storage pool values. These reflects virtual storage consumption below the 2GB bar. They may be used to identify poorly behaved applications or DB2 code issues.

More Quick Hits

→ Object level RECOVER from system level backups is available as soon as DB2 9 (CM), whereas object level RECOVER to PIT with consistency is not available until DB2 9 (NFM)

→ Nice package of DB2 9 COPY enhancements

– MRU to avoid throwing away useful random pages

– Automatic implicit CHECKPAGE on table space pages and no setting of COPYPEND when damage found

– All with less CPU

More Quick Hits …

→ Pros and cons of large index page size (8K, 16K, 32K)

- Reducing the number of index levels, reducing the number of getpages for index traversal, there by reducing CPU resource consumption

- Reducing the number of index page splits which are painful especially for GBP-dependent index (data sharing)

- Page size greater than 4K is an essential requirement for index compression
  - Note index compression is NOT a performance improvement
  - On the other hand large page size may possibly lead to either
    - Wasted space within page to maintain ability to compress down to 4K CI

- Large page size can aggravate
  - Index buffer pool hit ratio for random access
  - Page P-lock contention for GBP-dependent indexes if heavy update

More Quick Hits …

→ OPTIOWGT & SYSDUMMY1

- Execute RUNSTATS against DSNDB06.SYSEBCDC, so that Optimiser knows it is a 1-row table
- This will avoid CPU query parallelism being chosen on a SYSIBM.SYSDUMMY1 query after OPTIOWGT is enabled

→ GROUP BY in different order

- Potential problem area is SQL query without an ORDER BY expecting order to be preserved from a GROUP BY
- Relational theory states that order is ONLY guaranteed with an ORDER BY
- DB2 9 can reorder the GROUP BY columns to match an index and thus allow a sort to be avoided
- Applications that have relied on the order from GROUP BY will be exposed

→ General sort avoidance

- As DB2 provides enhanced sort avoidance, applications written depending on cursor materialisation will be exposed
- For example - Application FETCHes a row, and then issues a stand-alone UPDATE which moves the row forward, then it is possible that the row will be seen a 2nd time as the FETCH processes the rows

More Quick Hits …

→ Reordered Row Format (RRF)

– Potential for significant reduction in CPU resource consumption when accessing many rows with many varying length columns

- Implements CPU tuning recommendation to place fixed length columns ahead of varying length columns

- Provides for direct access to each varying length column

- In some few cases may lead to increased logging volume

– On by default in DB2 9 (NFM) and applies to all tablespace types

– A pageset will be automatically converted from Basic Row Format (BRF) to RRF when pageset is re-allocated (e.g., REORG, LOAD REPLACE)

– Application transparent – even for SELECT *

– REORG and LOAD REPLACE utilities override KEEPDICTIONARY during first time migration where data compression of variable-length rows

More Quick Hits …

➡ Reordered Row Format (RRF) …

– Be careful when using DSN1COPY to export/import data during the transition period when all DB2 systems and data sharing groups have not yet been migrated to DB2 9 (NFM)

- Should consider disabling RRF on a temporary basis before migrating first system or data sharing group to DB2 9 (NFM), and re-enabling after all the systems and data sharing groups have been migrated to DB2 9 (NFM)

– How to disable RRF

- Find &SPRMRRF parameter in the macro DSN6SPRC in SDSNMACS

- Look for "&SPRMRRF SETC '1' 1=Enable High Perf Row Option"

- Edit the macro and set &SPRMRRF to '0' and rerun DSNTIJUZ job or similar

- Disabling RRF with SPRMRRF has the following effect:

  – Any tablespaces created in DB2 9 NFM will be created as BRF

  – When a BRF tablespace is reorganised it will remain in BRF

- Note carefully pageset/partitions already in RRF will remain in RRF

- REORG will not convert to BRF

More Quick Hits …

→ Reordered Row Format (RRF) …

   – Important DB2 maintenance

      • APAR PK78958 disables RRF conversion for compressed pagesets

      • APAR PK78959 adds new DIAGNOSE options and will also disable RRF for CREATEs of compressed pagesets

      • APAR PK79127 stops inadvertent conversion to RRF during CM mode when support is not there

More Quick Hits …

→ Reordered Row Format (RRF) …

– APAR PK87348 Enable BRF for UTS PBR/PBG

  – Hidden zparm SPRMRRF now converted to opaque zparm

  – SPRMRRF=<u>ENABLE</u>|DISABLE controls row format for new tablespaces

  – Online changeable

  – Set SPRMRRF=DISABLE to get BRF

  – SPRMRRF setting is ignored for

    – Existing tablespace with EDITPROC i.e., stays BRF

    – XML tablespace always RRF

      – New PART of existing XML tablespace will always be RRF

  – REORG and LOAD REPLACE will honour SPRMRRF setting

  – If SPRMRRF=ENABLE

    – Get BRF -> RRF conversion

    – COMPRESS YES|NO is no longer a factor

    – New PART via ALTER TABLE ADD PART or UTS PBG growth will be RRF

  – SPRMRRF=DISABLE

    – No BRF -> RRF conversion

    – If already RRF then stay as RRF

    – New PART will be BRF

More Quick Hits …

➡ Reordered Row Format (RRF) …

  – APAR PK85881 LOAD/REORG ROWFORMAT support

    – New keyword ROWFORMAT for LOAD REPLACE and REORG

    – Determines output row format

    – ROWFORMAT BRF | RRF

  – SPRMRRF and LOAD/REORG ROWFORMAT not honoured for CLONE
     tables relationship

    – UTS, LOB or XML

    – No conversion from BRF -> RRF or RRF -> BRF

More Quick Hits …

→ Universal Table Space (UTS)

- LOCKSIZE ROW (default) is not a recommendation

- SEGSIZE 4 (default) is not a recommendation

- No MEMBER CLUSTER support

- Hard prerequisite for CLONE TABLE

- Migration path: UNLOAD/DROP/CREATE/LOAD

More Quick Hits …

→ Effectiveness of Asymmetric leaf page split function

- Design point is to provide performance relief for classic sequential index key problem

- Asymmetric split information is tracked in the actual pages that are inserted into, so it is effective across multiple threads across DB2 members

- Prior to APAR PK62214, DB2 only remembered the last insert position and a counter

- APAR PK62214 introduces changes to the tracking and detection logic, and it should work much better for data sharing

- The new approach remembers an insert 'range' and tolerates entries being slightly out of order

- It may still not be effective for large key sizes (hundreds of bytes), or if entries come in very bad order (i.e., they do not look sequential)

- But for simple cases like 3, 2, 1, 6, 5, 4, 9, 8, 7, 12, 11, 10 ... DB2 will be able to determine that the inserted entries are ascending

More Quick Hits …

→ Enhanced index look-aside

    – In DB2 9 potential for extra usage of index look-aside during INSERT, DELETE, UPDATE processing

    – Applies to non-clustering indexes where CLUSTERRATIO is equal to or greater than 80%

→ Object level point-in-time RECOVER from System Level Backup (SLB) will fail if dataset has moved away to a different volume since the backup was taken

    – For example

        • PIT recovery to a point before REORG with inline copy

        • Storage administrator decides to move datasets around to defrag volumes

        • Migrate to new DASD

    – Consider use of Recovery Expert Tool, or ISV tool which provides similar function

More Quick Hits …

→ BACKUP SYSTEM, RESTORE SYSTEM, object level RECOVERY from SLB

- XRC

  - Object level RECOVER

    - Will use the normal copy (i.e., not FlashCopy) when restoring data from a CopyPool backup

  - RESTORE SYSTEM

    - Cannot use FlashCopy to restore the entire DB2 system from a CopyPool backup
    - But can use a system level backup on tape
    - To use FlashCopy to restore the entire DB2 system to a PIT, need to disable XRC before running the RESTORE SYSTEM utility

- PPRC

  - Same limitations as XRC unless use the new DS8000 Remote Pair FlashCopy (Preserve Mirror) function and the associated list of pre-reqs

  - Must set the preserve mirror attribute for the DFSMShsm copypools via DFSMShsm commands outside of DB2

  - RECOVER and RESTORE SYSTEM will benefit from the preserve mirror support

    - FlashCopy will be used for restore
    - PPRC volume pair will be maintained and not marked as duplex-pending

More Quick Hits …

→ Optimistic Locking is not just for WebSphere!

→ Real Time Statistics (RTS) always enabled

- In DB2 9 the RTS tables are now in the DB2 Catalog

- In-memory statistics are always externalised to the RTS tables

- Whereas in V8 the process had to be explicitly started

→ MODIFY RECOVERY AGE n

- Will set COPY PENDING for tablespaces which are defined with DEFINE NO and where SYSIBM.SYSTABLEPART.SPACE = -1

  - i.e., underlying VSAM pageset does not exist

- Circumvention START DB(name) SPACENAM(name) ACCESS(FORCE)

- Fixing APAR PK69427

More Quick Hits …

→ Declared Global Temporary Tables (DGTTs)

- Up to 20x increase possible in both CPU resource consumption and elapsed time without fixing maintenance

- APAR PK67301 to reduce excessive logging for sequences of insert and mass delete within the same commit scope
  - With more frequent commits, the performance degradation is less and PK67301 makes it even less

- APAR PK62009 to reclaim the space emptied by mass delete after commit for ON COMMIT DELETE DGTT
  - However, if you have a loop of Inserts/Mass Delete without commit, then Insert performance will steadily deteriorate and that is where DELETE WHERE can help

- APAR PK70347 to reduce DBM1 31-bit storage usage when Performance Trace Class 8 active

More Quick Hits …

→ Declared Global Temporary Tables (DGTTs) …

- APAR PK70060 to

  - For sort workfile, DB2 will first go after tablespaces with zero secondary quantity to find space and if none is available, will try to find space in the first 2GB region of tablespaces with non-zero secondary quantity and even if that is not possible, will use any available tablespace

  - For DGTT, DB2 will first try to find space in table space with non zero secondary quantity and if none is available will use any available table space

- To achieve separation for DGTT and Sort Workfile use, the recommendation will be to have some table spaces with zero secondary quantity and some with non-zero secondary quantity

- Maximum primary quantity of workfile tablespace is 2GB. But workfile tablespace can grow up to 64GB like any classic segmented table space if secondary quantity is not zero

More Quick Hits …

➔ SQL TRUNCATE

– Efficient mass delete of data rows only with classic segmented or universal tablespace

– No efficient mass delete of index entries

– Performance is ugly with classic partitioned tablespace and/or many indexes

➔ REOPT(AUTO)

– Only applies to dynamic SQL

More Quick Hits …

→ Do I still need QUIESCE points after DB2 9(NFM)?

- PIT recovery will still be faster to a quiesce point assuming you can get one!

- If you are using ISV tools which may not handle URs but do handle QUIESCE points

  - For example to allow a consistent PIT recovery to a different target

- To force changed pages to DASD for whatever reason

- Record a log RBA/LRSN in SYSCOPY for whatever reason

More Quick Hits …

→ Different Group Attach behavior when running multiple DB2 members of the same data sharing group on the same LPAR

– DB2 in V8 would attach to first available member based on IEFSSNxx parmlib specification

– Usermod available on V8 to change the behaviour and spread connections across available DB2 members on the same LPAR

– DB2 in V9 will spread connections across available DB2 members on the same LPAR

– Until now there was no way to re-instate the old V8 behaviour

– New co-requisite APARs PK79228 and PK79327 will provide such an option

  • Enhancement to the ZPARM macro (DSN6GRP)

  • Can specify new option RANDOMATT=NO

  • Activate via –SET SYSPARM RELOAD

  • From that point forward, the DB2 member will be considered "ineligible" for Random Group Attach

  • DB2 member will only be selected if no other members are available

  • To completely restore V8 behaviour, must specify RANDOMATT=NO for every member of the data sharing group

More Quick Hits …

→ Native SQL Procedures

- – Potential for significant reduction in CPU resource consumption by avoiding

  - • Overhead of stored procedure invocation overhead

  - • Overhead of roundtrip between WLM and DBM1 address spaces for each SQL call

- – Short running SQL procedure could achieve up to an 40% ITR improvement

- – But little or no improvement for long-running SQL procedure

- – When invoked from DRDA connection over TCP/IP

  - • zIIP eligible

  - • As it runs in DBM1 address space under DDF enclave SRB

- – Easy to code, develop and manage

- – Selective application re-engineering may be required when migrating to Native SQL Procedures

More Quick Hits …

→ Stored Procedures - Performance of different languages

– Environment Configuration

- z/OS 1.9

- DB2 9 for z/OS

- Universal Driver 3.52.76

- JDK 1.4.2 (SQLJ/JDBC stored procedures)

- 3 CP's

- 2 zIIP's

- 2 zAAP's

- IRWW OLTP workload

More Quick Hits …

→ Stored Procedures - Performance of different languages …

Language/API	Base CPU/tran Cost	Billable CPU/tran Cost after zIIP and/or zAAP redirect
COBOL stored proc	1X (BASE)	.88x
C stored proc	.95x	.83x
SQLJ stored proc	1.7x	1.15x (zIIP + zAAP)
JDBC stored proc	2.95x	1.76x (zIIP + zAAP)
External SQL stored proc	1.62x	1.49x
Native SQL stored proc	1.14x	.65x
Remote SQLJ	1.78x	1.06x

More Quick Hits …

→ Stored Procedures - Performance of different languages …
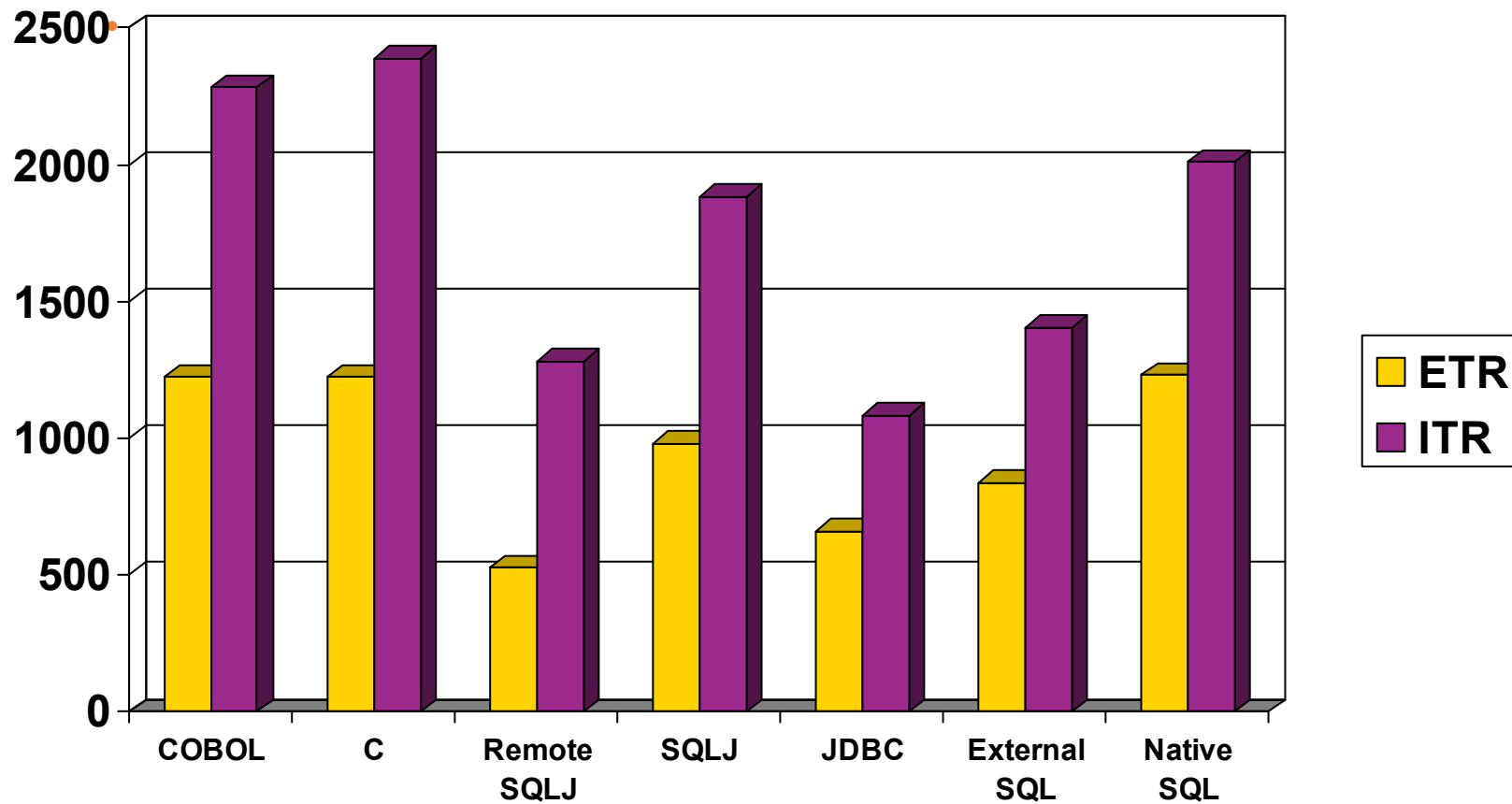
More Quick Hits …

→ Checklist for XML-related Configurations

- Basic XML parsing requires z/OS XMLSS: z/OS 1.8 or z/OS 1.7 with APAR OA16303

- XML schemas requires IBM 31-bit SDK for z/OS, Java 2 Technology Edition, V5 (5655-N98), SDK V1.5. And Java stored procedure setup

- zparms for virtual storage:

  - XMLVALA and XMLVALS
    - Default: 200MB and 10GB
  - Also LOBVALA and LOBVALS impact bind-in and bind-out of XML

- Buffer pool for XML tables (default BP16K0), requires authorisation for users who create or alter tables with XML columns

Questions

Data Management Communities for DB2

→ IDUG – the worldwide community of DB2 users

  – Membership is FREE – join today! www.idug.org

→ Data Management Community – share and interact with peers around the world

  – www.ibm.com/software/data/management/community.html

→ Information Champions – recognizes individuals who have made the most outstanding contributions to the Information Management community

  – www.ibm.com/software/data/champion

Disclaimer

IBM Information On Demand 2009