IBM System z Technology Summit

DB2 10 – The Secrets of Scalability

Agenda

- **Introduction**

- **The Need for Scalability**

- **DB2 10 Scalability Enhancements**

  - Virtual storage constraint relief

  - Latch contention reduction

  - Catalog concurrency enhancements

  - SMF compression

  - Other scalability enhancements

    - Concurrency
    - Index Enhancement
    - Multirow Operations
    - LOBs
    - Hash Organization

    - Optimization
    - Additional zIIP Exploitation

- **Summary & Questions**

The Need for Scalability

- **IT volumes continue to increase**
  - More applications
  - More data
  - More transactions

- **Performance is ever more important**
  - Customers need to support workload growth without a drop-off in performance

- **Availability is ever more important**
  - Pressure to reduce both planned and unplanned outages

- **End result: each DB2 environment is being asked to work harder, with less downtime**

- **Every DB2 release attempts to push back these boundaries, but major progress has been made in DB2 10**

DB2 10 Scalability Enhancements

Savings … right out of the box

IBM DB2 10 for z/OS delivers faster queries and reduced cost with optimized technology

Proven
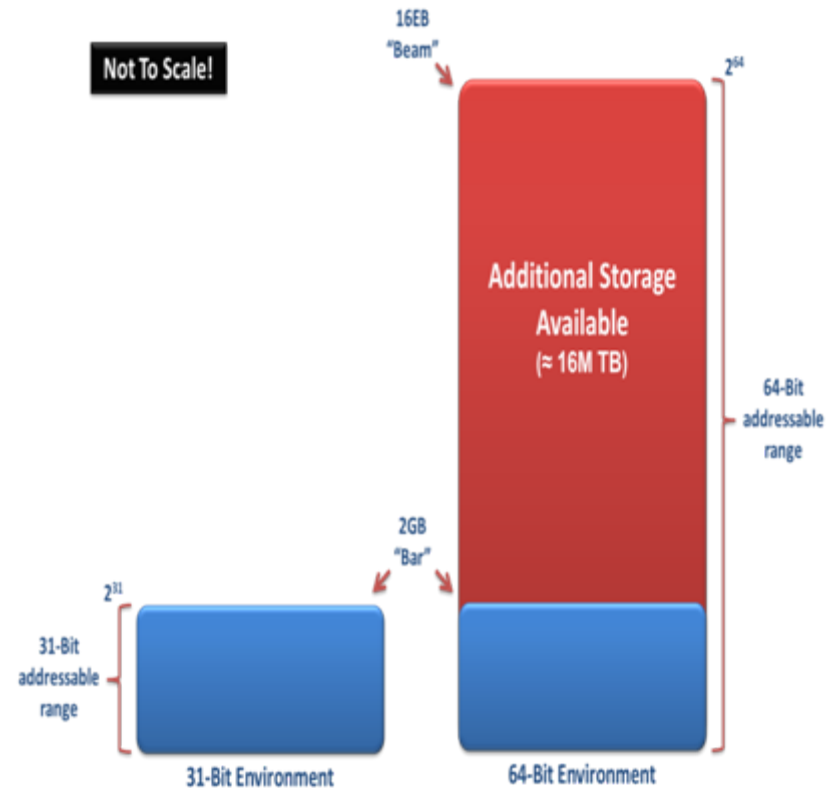Secure
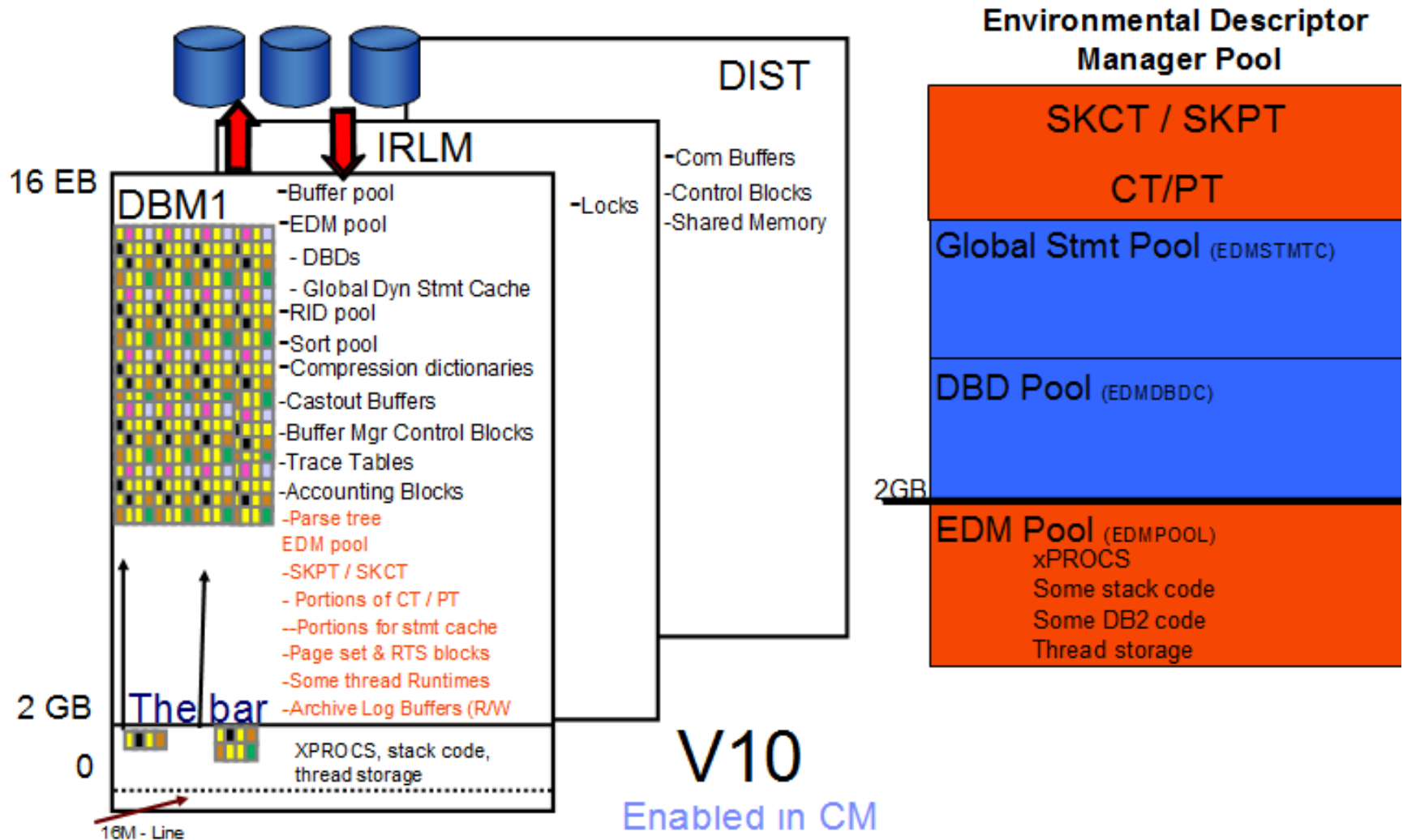Simple
Cuts costs
Innovative

Top New Features

- **CPU/Performance Improvements**
- **Virtual Storage Enhancements**
- **Security Extensions**
- **Improved Catalog Concurrency**
- **Temporal Data**
- **Access Path Management**
- **pureXML enhancements**
- **Currently Committed semantics**
- **Automated statistics**
- **Dynamic schema change enhancements**
- **In-memory object support**

- **Optimiser enhancements**
- **MEMBER CLUSTER for UTS**
- **Backup and recovery enhancements**
- **Enhanced audit**
- **Include additional index columns**
- **Enhanced SQL OLAP functions**
- **Skip Migration**
- **And many more….**

Virtual Storage Enhancements

- **V8 began a major project to transform DB2 into a 64-bit RDBMS**

  - Laid the groundwork and provided some scalability improvements but a lot of DBM1 objects remained below the 2GB bar

- **DB2 9 improved things a little, but only by another 10-15% for most customers**

  - Practical limit of 300-500 threads per DB2 subsystem

- **DB2 10 moves 80-90% of the remaining objects above the bar, resulting in 5-10x improvement in threads per subsystem (CM)**

Not To Scale!

16EB "Beam" — 2^{64}

Additional Storage Available (≈ 16M TB)

64-Bit addressable range

2GB "Bar"

2^{31}

31-Bit addressable range

31-Bit Environment

64-Bit Environment

64 bit Evolution (Virtual Storage Relief)

Environmental Descriptor Manager Pool

16 EB

DBM1

- Buffer pool
- EDM pool
  - DBDs
  - Global Dyn Stmt Cache
- RID pool
- Sort pool
- Compression dictionaries
- Castout Buffers
- Buffer Mgr Control Blocks
- Trace Tables
- Accounting Blocks
- Parse tree
  EDM pool
- SKPT / SKCT
- Portions of CT / PT
-- Portions for stmt cache
- Page set & RTS blocks
- Some thread Runtimes
- Archive Log Buffers (R/W

IRLM

- Locks

DIST

- Com Buffers
- Control Blocks
- Shared Memory

SKCT / SKPT
CT / PT

Global Stmt Pool (EDMSTMTC)

DBD Pool (EDMDBDC)

2GB

EDM Pool (EDMPOOL)
xPROCS
Some stack code
Some DB2 code
Thread storage

2 GB

The bar

0

XPROCS, stack code, thread storage

V10

Enabled in CM

16M - Line

© 2012 IBM Corporation

Virtual Storage Enhancements

- **Possibility for fewer DB2 subsystems (and possibly fewer LPARs) in a data sharing environment**

  – Lower data sharing overhead

  – Fewer systems to manage / maintain

  – Delete of data sharing member is coming! See:

    - PM31003
    - PM31009
    - PM54873
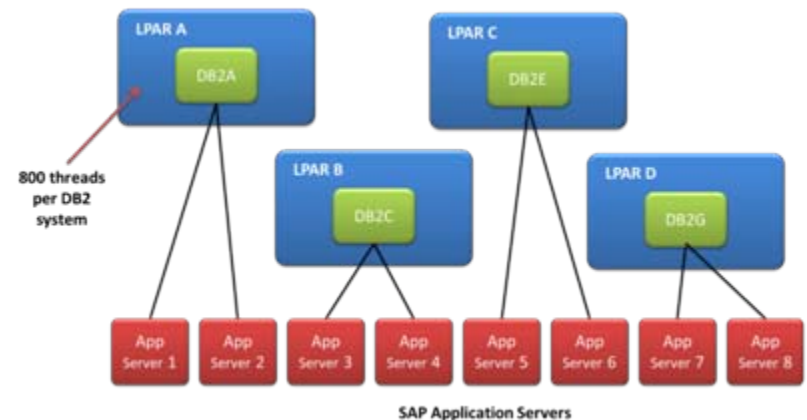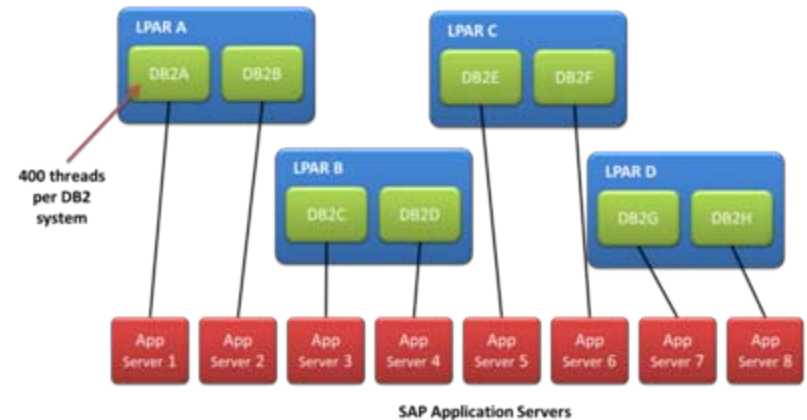
Virtual Storage Enhancements

- **More space for performance critical storage objects such as dynamic statement cache**

- **Potential to reduce legacy OLTP CPU cost through**

  - More use of CICS protected entry threads

  - More use of RELEASE(DEALLOCATE) with persistent threads (with trade-off on concurrency)

  - DB2 10 High-Performance DBATs
    - -**MODIFY DDF PKGREL(BINDOPT | COMMIT)**

- **Other limiting factors on vertical scalability still remain**

  - Real storage
    - Plan on additional 10-30% real memory following migration

  - ESQA/ECSA (31-bit) storage

  - Active log write

Real Storage Enhancements

- **For prior releases, z/OS always managed DB2 bufferpool pages as 4K frames**

- **Move to 64-bit architecture made much larger buffer pools viable**

  – Bufferpools can use many millions of pages

  – Increased z/OS overhead for page management

DB2 9 Buffer Pool

4K Pages

4K Pages
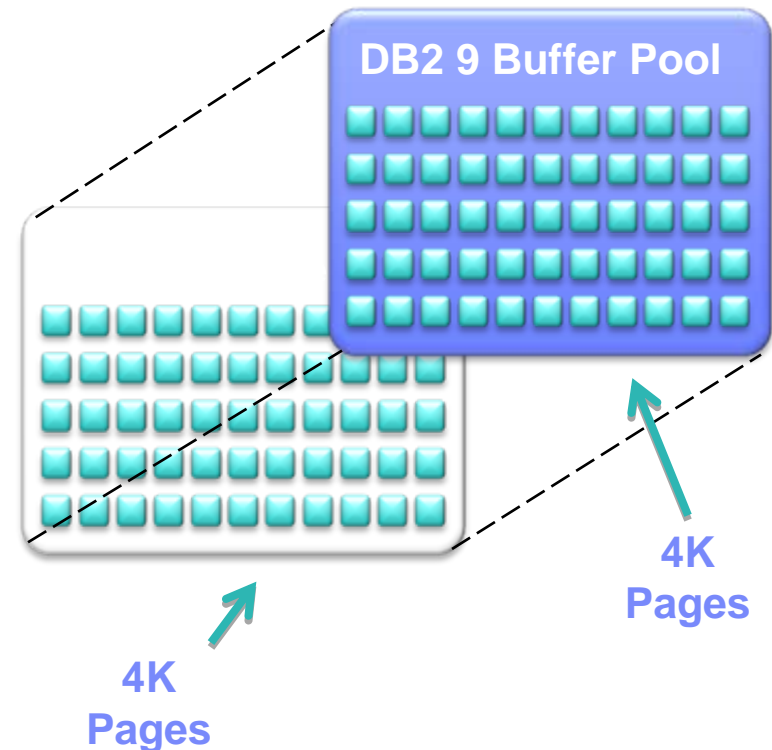
Real Storage Enhancements

- **DB2 10 introduces support for 1MB pages to reduce z/OS page management overhead**

  – Needs z10 or newer z196 server

  – Needs bufferpool to be defined with PGFIX=YES

    • Buffer Pool intensity can help determine

    • PGREADS + PGWRITES / # of Pages

  – z/OS Sysprogs must partition real storage between 1K and 1MB frames (LFAREA in IESYSnn / PARMLIB)

- **Customer testing during beta program showed CPU reductions of 0-6% with this feature enabled**

DB2 9 Buffer Pool

4K Pages

1MB Pages
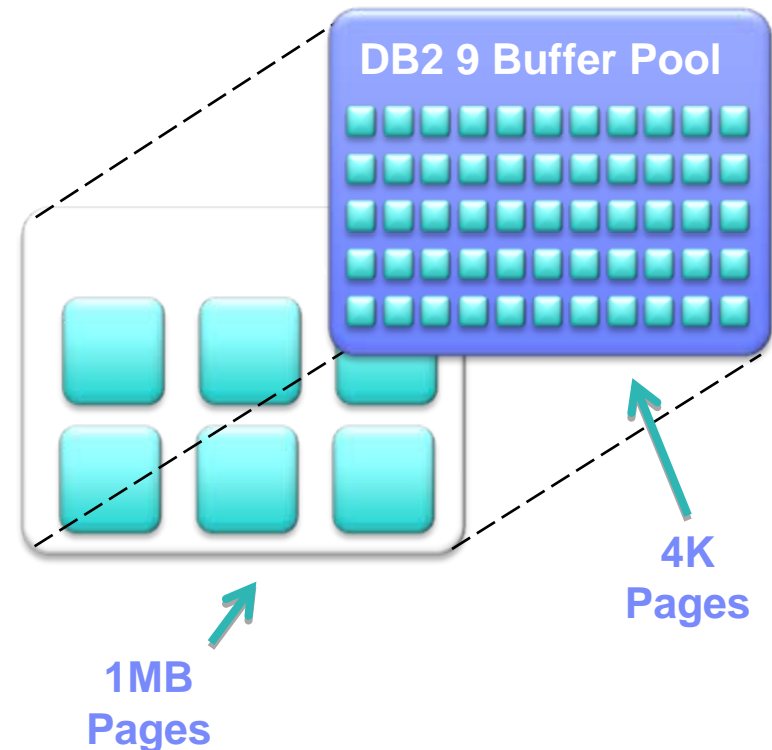
Storage Enhancements

- **Remains critical to ensure that there is no paging in DB2 address spaces**

- **Focus changes from virtual memory constraints & monitoring to real memory constraints & monitoring**
  - See (PM24723) for real storage monitoring and contraction enhancements – advised not to go into production without this!

- **Ensure use of PGFIX=YES to exploit 1MB real storage frames**
  - Many customers still haven't exploited this feature in their DB2 8 and DB2 9 systems – significant CPU savings!
  - Support for 1MB non page-fixed bufferpools in future release

- **Ensure you are up to date on z/OS maintenance before using 1MB pages**
  - OA31116, OA33702, OA33529, check for others.

Latch contention reduction

- **Latch: DB2 mechanism for controlling concurrent events or the use of system resources**

  - Reported in accounting and statistics traces

  - Latch wait time can be significant for high-volume environments

- **DB2 10 reduces latch contention for a large number of situations, including:**

  - LC12: Global Transaction ID serialization

  - LC14: Buffer Manager serialization

  - LC19: Log write in both data sharing and non data sharing

  - LC24: EDM thread storage serialization

  - LC24: Buffer Manager serialization

  - LC27: WLM serialization latch for stored procedures and UDF

  - LC32: Storage Manager serialization

  - IRLM: IRLM hash contention

  - CML: z/OS Cross Memory Local suspend lock

  - UTSERIAL: Utility serialization lock for SYSLGRNX (removed in NFM)

Catalog Concurrency

- **Contention on DB2 catalog is a major ongoing pain for most large DB2 customers**

- **DB2 10 introduces UTS PBG format for catalog tablespaces in NFM**
  - Internal hashes and links are removed during ENFM processing
  - Use of row-level locking and reordered row format
  - Use of new currently committed semantics and other lock avoidance techniques
  - SMS pre-reqs for migration

- **Greatly improves access to catalog/directory**
  - REORG SHRLEVEL(CHANGE) for complete catalog/directory
  - BIND concurrency much improved, but more work required in future releases – especially with heavy parallel DDL against different databases

Catalog Contention Issues

- **Be prepared for some short-term degradation on entry to CM for single-thread BIND/REBIND processes, until you get to NFM**

  - PLANMGMT=EXTENDED the default, so multiple copies of access plan kept in catalog

  - New indexes are defined, in preparation for hash links to be removed in NFM

  - No concurrency improvement until catalog restructure in ENFM

  - Redbook testing showed REBIND elapsed time increases CPU increases of 50-70%

SMF compression

- **High transaction volume usually means high SMF volume, which can become a limiting factor**

- **Some customers forced to switch off useful accounting data, or resort to SMF rollup (via ACCUMACC ZPARM)**

- **New SMF compression feature can provide increased throughput due to I/O efficiency improvement**

  - Uses z/OS compression service to deliver approx. 60%-90% compression for 1% CPU cost

SMF compression

- **Enabled via new SMFCOMP DSNZPARM (member scope)**

- **All data after SMF header is compressed**

- **Needs vendor support to allow compressed SMF records to be processed**

- **New sample DSNTSMFD application to uncompress SMF data (via PM27872)**

- **Can be used in conjunction with accounting rollup to achieve up to 99% reduction**

Other scalability enhancements

- **Others enhancements in DB2 10**
  - SPT01 restructured: split into several pieces with LOBs used for larger package sections
    - Inlined BLOB with PM27811 for performance
    - Reported in **–DIS GROUP DETAIL**
  - Workfile enhancements: support for spanned records to increase maximum record length, better use of in-memory workfiles, use of PBG tablespaces
  - Support for Extended Address Volumes: EAVs theoretically allow up to 221TB per volume (223GB in z/OS 1.10)
  - Decrease dataset allocation/deallocation times: using new function in z/OS 1.12 DB2 startup/shutdown times can be improved

Other scalability enhancements

- **Concurrency**

  - Access Currently Committed
    - With INSERT / DELETE workloads
  - XML storage restructure (Multi-Versioned Format)

- **Index Enhancements**

  - Parallel Index Maintenance
    - For tables with 3 or more indexes
    - Measures 50% ET reduction during the beta
      - Table with 6 indexes, during 2000 INSERTs
  - INCLUDE Columns
    - Added columns to unique indexes
      - Not part of the unique constraint
      - Enhance index only access
      - Eliminate some indexes
  - Sequential INSERT improvements in DB2 9 & 10

Other scalability enhancements

- **Multirow Operations for Data Sharing**

  - LRSN changes for Multirow Insert

- **LOBs**

  - Streaming & INLINE(ing)

- **Hash Organization**

  - Use caution in determining the right table candidates

  - Significant savings given the correct workload

    - Performance Redbook reports 13 – 37% CPU savings, slightly less if the application was getting index only access

- **Optimization / Operation**

  - Literal Replacement

  - Dynamic Index List Prefetch for disorganized indexes

Other scalability enhancements

- **Additional zIIP Exploitation**
  - Buffer operations
    - Prefetch
    - Deferred Write
  - RUNSTATS
    - Simple statements
  - DB2 9 added
    - Native SQL Procedures called remotely
    - zIIP exploitation for DFSort for Utilities
    - XML Serialization / Parsing
  - REORG(UNLOAD) in DB2 maintenance stream

Summary

- **DB2 10 delivers some very significant enhancements for increasing throughput, supporting more users and reducing planned downtime**

  – Many of these enhancements available in Conversion Mode (CM)

- **Remember that sufficient real storage is needed to back any increase in virtual**

- **If you are still on DB2 V8, remember that support ended in April 2012**

  – DB2 9 EOS announced for June 2014

DB2 for z/OS Exchange

DB2 for z/OS Best Practices