IBM Software Group

DB2 Utility Update

DB2 Information Management Software

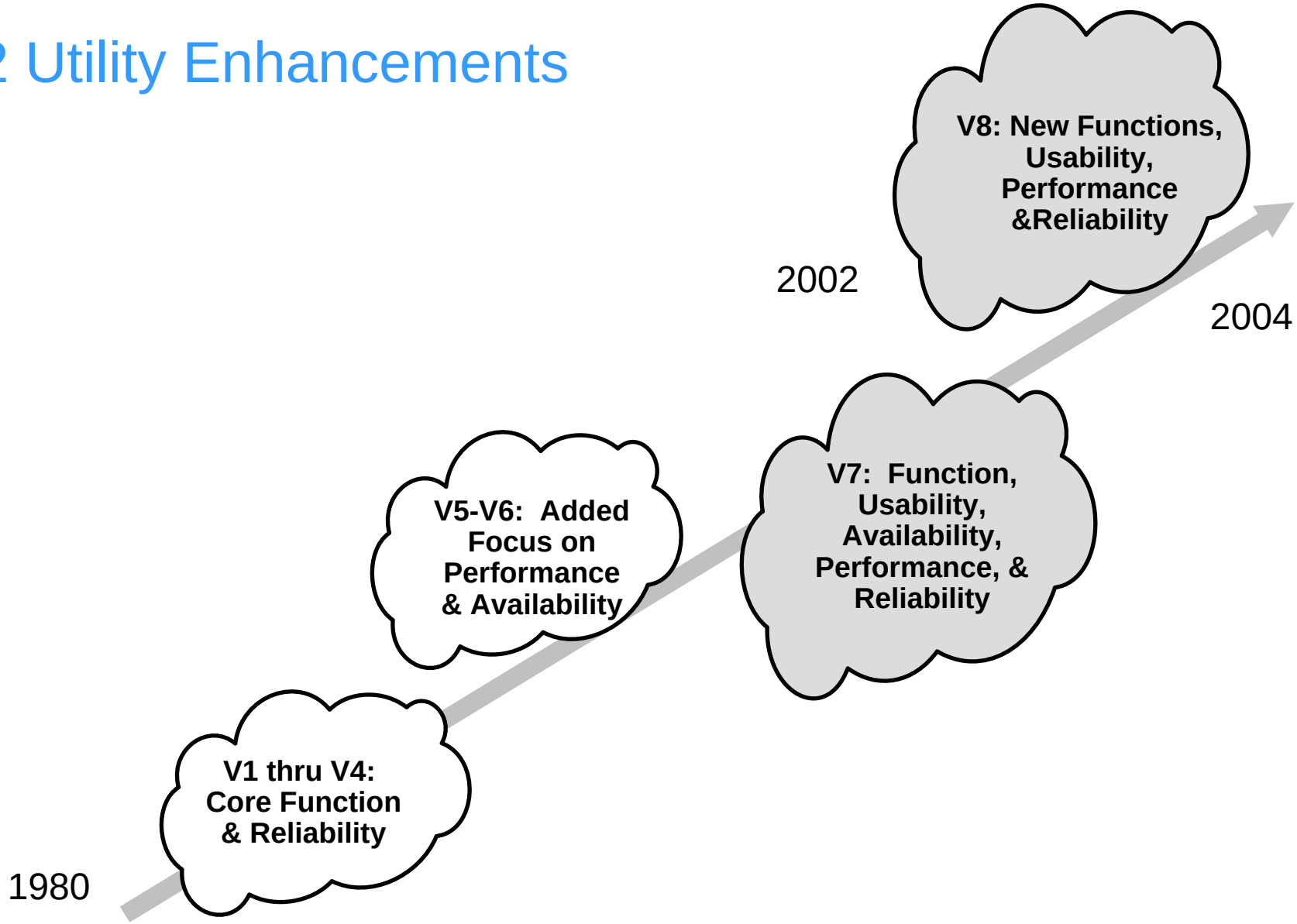
ON DEMAND BUSINESS

Winston A. Tobias
Technical Specialist, DB2 for zTools
wtobias@us.ibm.com

Agenda

- DB2 Utility Enhancement History
- DB2 Utilities V7 (function, usability, and availability)
- DB2 Utilities V8
- Best practices

DB2 Utility Enhancements



Version 7 Summary

- Object Wildcarding/Dynamic Allocation
- Online REORG Improvements
  - FASTSWITCH
  - Parallel BUILD2
  - Online Reorg Time-out & Retry options
- LOAD Enhancements
  - LOAD Partition Parallelism
  - Online LOAD RESUME
  - DB2 Family Cross Loader
- COPYTOCOPY Utility
- COPY/RECOVER Parallelism for tape
- UNLOAD Utility
- Statistics History
- Improved utility restart (APAR PQ72337)

Not covered in following slides

Object Wildcarding/Dynamic Allocation

- Utilities invoked against one or more database objects
  - Table spaces and/or Index spaces
- Multiple objects generally require
  - Run multiple utility invocations
  - Specify an itemized list of objects
- Now allow a pattern-matching expression
  - Generate a list of objects
  - Passed to one or more specific utilities for processing
- Dynamically allocate data sets based on data set template
  - GDG Base Creation
  - Intelligent data set sizing
- Used together, these two facilities utility job streams are developed more quickly with less future maintenance as the underlying list of database objects change

Example V6 and V7

- V6 and before

```
//DDn DD DSN=..,UNIT=..,VOL=SER=..,DISP=
  ...
//SYSIN DD *
 QUIESCE TABLESPACE DBA.X TABLESPACE DBA.Y
                TABLESPACE DBA.Z
 COPY TABLESPACE DBA.X COPYDDN (DD1,DD2)
           TABLESPACE DBA.Y COPYDDN (DD3,DD4)
           TABLESPACE DBA.Z COPYDDN (DD5,DD6)
```

- With V7

```
//SYSIN DD *
 OPTIONS PREVIEW    -- Preview list expansion and dataset names
 LISTDEF X INCLUDE TABLESPACE DBA.*
 TEMPLATE A DSNAME(&DB..&TS..D&JDATE..&PRIBAC&TIME.)
 OPTIONS OFF            -- Turn off preview to run following utilities

 QUIESCE LIST X
 COPY LIST X COPYDDN(A,A)
```

Fast Switch

- User option for REORG SHRLEVEL REFERENCE or CHANGE
- FASTSWITCH YES or NO
  - Default set by ZPARM SPRMURNM
    - IBM V7 default is FASTSWITCH NO
  - Catalog and directory is only FASTSWITCH NO
- V5 and V6 use Access Method Services (IDCAMS) for rename during SWITCH
- Version 7 allows two naming conventions for DB2 tablespace and index dataset names.
- The "instance: node of the DB2 dataset can be either I0001 or J0001
  - DSNC710.DSNDBC.DBDV0701.TBDV0701.I0001.A0001
  - DSNC710.DSNDBC.DBDV0701.TBDV0701.J0001.A0001
- Up to 12 times faster Elapsed Time

Build2 Parallelism

- Elapsed time improvement to the BUILD2 phase of REORG SHRLEVEL REFERENCE or CHANGE
  - REORG PART m
  - REORG PART m:n
- Logical partitions of non-partitioning indexes are updated using parallel subtasks.
- Availability improvement.
- (80% reduction in elapsed time for 5 NPIs).

REORG Timeout & Retry

- Reorg Drain with minimum application disruption
  - Tolerates less well behaved applications
  - Time-out on aggregate of partitions
  - Handle unpredictable access spikes
- Adjust so Reorg time-out occurs before application time-out
- Retry after a specified delay period
- Example: time-out in 20 seconds, but retry 6 times after waiting a minute

REORG TABLESPACE SHRLEVEL CHANGE

DRAIN_WAIT 20 RETRY 6 RETRY_DELAY 60

Load Partition Parallelism

- When a single LOAD can't be used:
  - Too much data to load, takes too long
  - Batch window tightly constrained
- How we currently support it:
  - Multiple load jobs, one per partition
  - NPI contention is a problem
- Uses multiple tasks in a single job to load partitions in parallel
- Easier to use:
  - Single job submission with multiple input data files
- Performance:
  - Load phase is faster due to parallelism
  - Eliminates contention on NPI, so build phase is faster
  - Up to 30% faster Elapsed Time

Online Load Resume

- Classic LOAD drains all access to tablespace
- Customers write Insert programs to avoid drain and allow availability to data
- Maintaining hundreds of Insert programs expensive - wish they could use LOAD instead
- Willing to trade performance for availability
  - Especially in data warehouse
- Add SHRLEVEL NONE | CHANGE syntax
  - Default SHRLEVEL NONE is Classic LOAD RESUME
- Online LOAD operates like an SQL INSERT program
  - Claims instead of drains
  - Data manager insert
  - Tries to maintain clustering order of data
  - Fire triggers
  - LOG YES only

DB2 Family Cross Loader

- High speed transfer of data from one table to another
  - Local or remote
- Combines the power and performance of
  - SQL
    - Including DataJoiner or Relational Connect
  - DRDA
  - LOAD Utility
- Single step process instead of
  - Unload or Export
  - File transfer
  - Load or Import

The Cross Loader was introduced in DB2 V7 after GA with PTF UQ55541 for APAR PQ45268 and PTF UQ55542 for APAR PQ46759

COPYTOCOPY Utility

- COPY, LOAD, and REORG can all make two local and two recovery site copies
  - Customers constrained by number of tape drives
  - Recovery site copies via remote attached tape drives impacts data availability
- Requirements:
  - Make asynchronous copies of copies
  - Register in SYSCOPY
- COPYTOCOPY can make up to three copies of a copy
- Tablespaces, indexes, indexspaces, lists
- choice of
  - **FROMLASTCOPY**
  - FROMLASTFULLCOPY
  - FROMLASTINCRCOPY
  - FROMCOPY ***dsn***

COPY/RECOVER Parallelism

- PARALLEL keyword introduced in V6
- Image copy objects to DASD and Tape in parallel.
  - stacking of multiple copies on tape
- Restore in parallel from DASD and Tape
- New keyword, TAPEUNITS indicates total number of tape units that can be dynamically allocated in processing this statement.
- PQ56293/PQ56295/PQ56296

Version 8

- New utilities BACKUP SYSTEM and RESTORE SYSTEM
- Delimited data support for LOAD and UNLOAD
- New defaults for better "out of the box" performance
- REORG SHRLEVEL NONE/REFERENCE allow REBALANCE
- Online Schema Support (e.g., REPAIR VERSIONS)
- Non-uniform statistics and on non-indexed columns
- HISTORY statistics without updating main statistics
- REORG SHRLEVEL CHANGE allow DISCARD
- REORG SHRLEVEL REFERENCE catalog tables with links
- Online Concurrent Copy support for 32K pages

Not covered in following slides

System Level Point in Time Recovery

- Easier, more flexible, less disruptive, faster recovery
- Handle large numbers of table spaces & indexes
- Two new utilities are introduced
  - BACKUP SYSTEM: Fast volume-level backups
    - DB2 databases and logs
    - Data sharing group scope
    - z/OS V1R5 required for new COPYPOOL function
  - RESTORE SYSTEM
    - To an arbitrary point-in-time
    - Handles creates, drops, LOG NO events

Delimited Data Support for LOAD and UNLOAD

- LOAD FORMAT DELIMITED COLDEL x CHARDEL y DECPT z
- UNLOAD DELIMITED COLDEL x CHARDEL y DECPT z

- DELIMITED- BSAM file with column and character data string delimiters
- COLDEL - column delimiter (default comma ,)
- CHARDEL - character data string delimiter (default quote ")
- DECPT - decimal point (default period .)

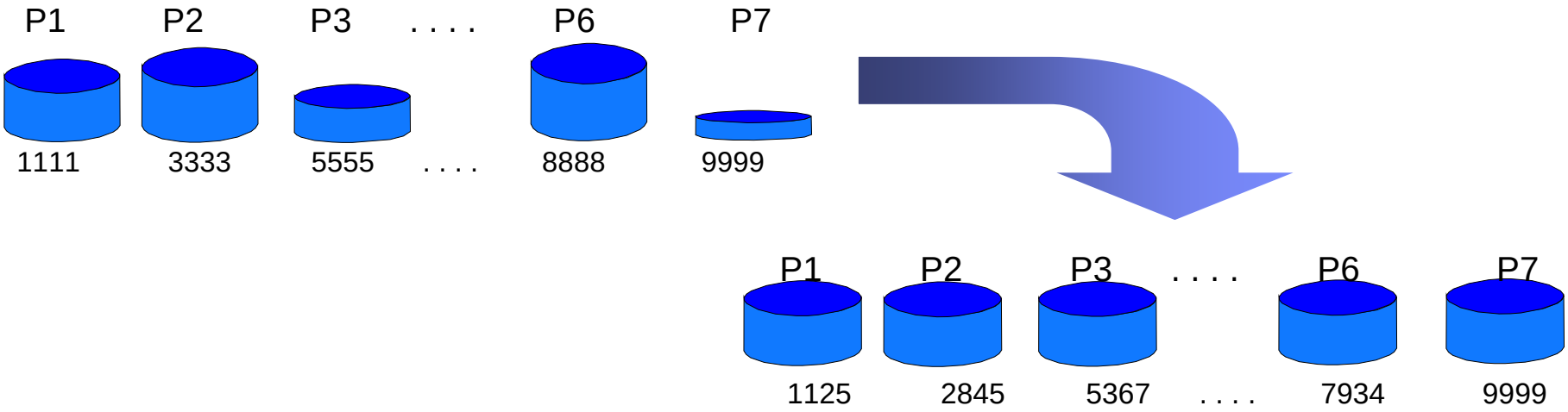
"Smith, Bob",4973,15.46
"Jones, Bill",12345,16.34
"Williams, Sam",452,193.78

Defaults for Better Performance

- SORTKEYS for LOAD/REORG/REBUILD
- SORTDATA for REORG
  - SORTDATA now allowed for 32K records with DFSORT
- REORG will use implicit clustering index

REORG REBALANCE

- REORG TABLESPACE SHRLEVEL NONE or REFERENCE
- Relative balancing of pages across page range or entire table space
- Useful to provide better space utilization across partitions
- Query parallelism benefits from balanced I/O across partitions
- DBA does not have to perform tedious analysis to determine partition boundaries



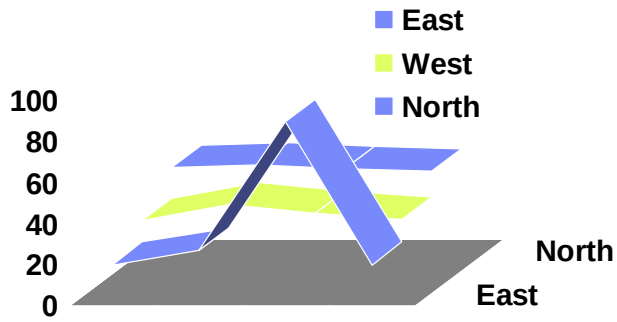
Online Schema Support

- Improved usability for objects placed in pending states
  - SCOPE PENDING for REORG and REBUILD
  - REORG SCOPE PENDING will operate only on objects in
    - REORP, AREO*
  - REBUILD SCOPE PENDING will operate only on objects in
    - RBDP, AREO*, RECP

- REPAIR VERSIONS - Updates the versions in the catalog and directory from the information in the table space or index. Use when you perform the following tasks:
  - When you use the OBIDXLAT option of DSN1COPY to move objects from one system to another.
  - As part of version number management for objects that do not use the IBM REORG utility.

RUNSTATS Distribution Statistics Enhanced

- Non-uniform distribution statistics on non-index columns
- RUNSTATS improvement that allows optimizer to consider non-uniform distribution statistics on columns that aren't part of an index
- Current technique is separate DSTATS program
- Significant performance improvements possible
- Collected with the FREQVAL keyword on a specified group of columns (COLGROUP)
- Most or least frequently occurring values can also be collected



V8 Post GA

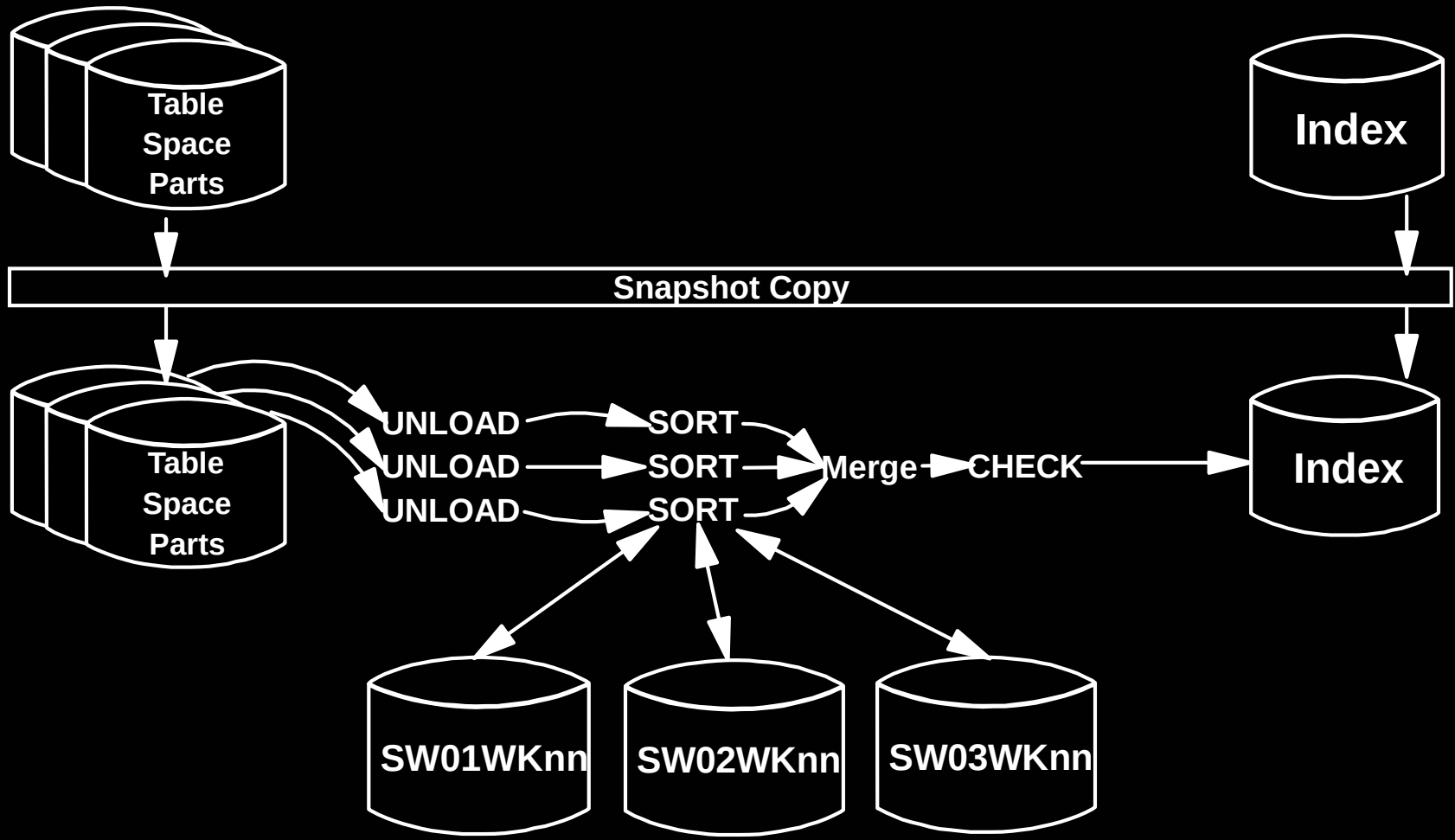
- Online CHECK INDEX
  - APARs PQ92749 (DB2 base) and PQ96956 (Utility Suite)
- Cross Loader support for > 32K LOBs
  - APAR PQ90263 (PTF available now for V7 and V8)
- LOAD/UNLOAD support for very large LOBs
  - Running prototype
  - Production code in unit test
  - APAR PK10278 for V7 and V8
- Data first claiming/draining
  - Greatly reduces the chances of a deadlock between SQL and utilities – now SQL and utilities will always claim/drain the data first, and then the index. This doesn't prevent any deadlocks -- if SQL accesses partition M, then tries for partition N while utilities does the reverse, there is still a potential deadlock among data partitions. ZPARM CLAIMDTA (default is NO)
  - APAR PK09781

Covered in following slides

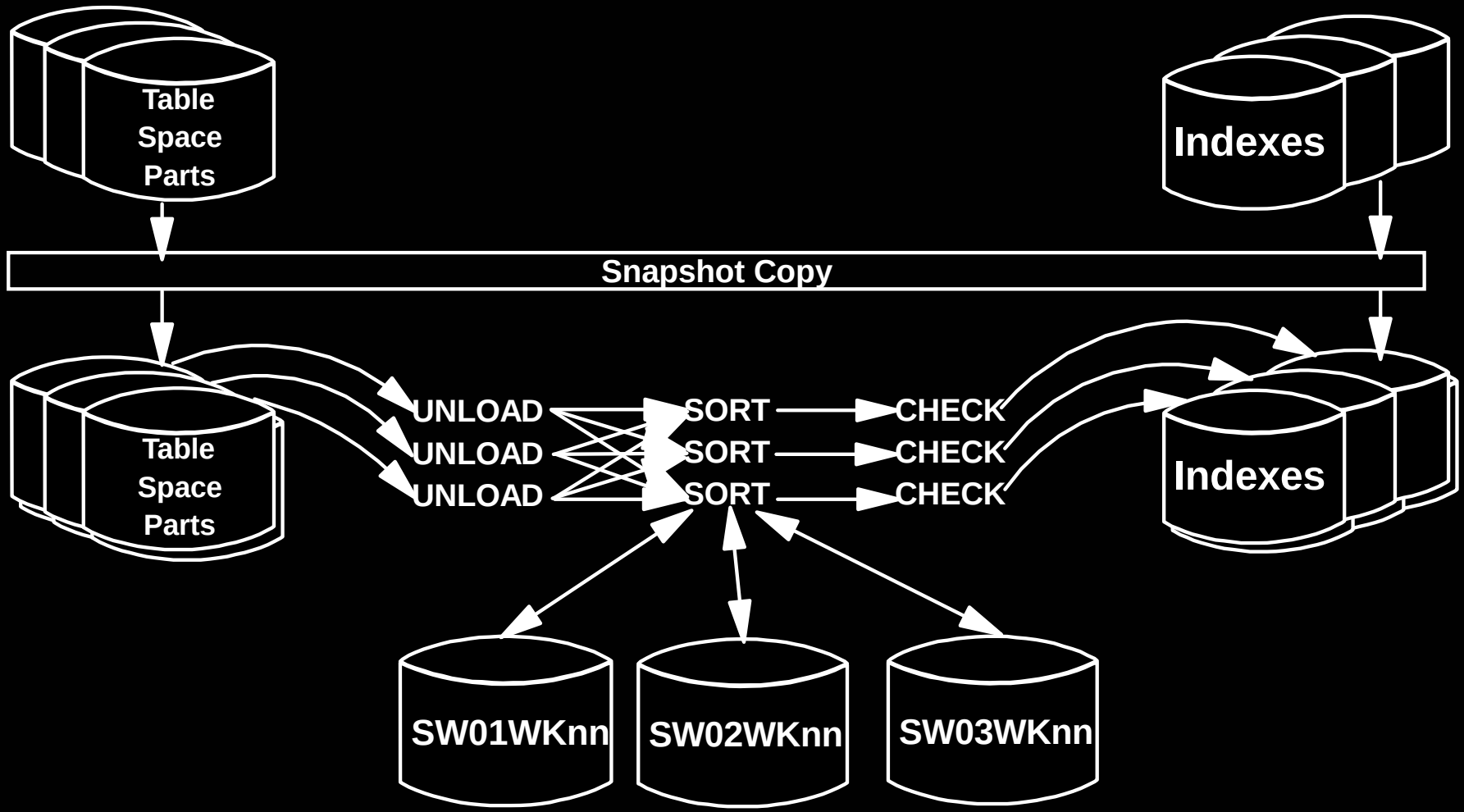
Online CHECK INDEX

- Current CHECK INDEX causes data and indexes are unavailable for update for the duration
- Online CHECK INDEX different design than Online REORG
- Claim as reader for target data and indexes
- Create shadow datasets
  - same dataset naming convention as Online REORG
  - cannot run Online CHECK INDEX on two logical parts of NPI
- Drain writers for target data and indexes
- Flash data and indexes from target to shadows
- After copy logically complete for ALL,
  - dedrain target data and indexes
  - run parallel check index on shadow data and indexes
    - same parallel design as REBUILD INDEX
- At utilterm delete shadow datasets when DB2 managed

Single NPI on partitioned



All indexes on partitioned



Cross Loader support for > 32K (rows with) LOBs

- Architectural limits within LOAD/UNLOAD did not allow for a record greater than 32K to be loaded or unloaded
- New buffering scheme for LOB values to bypass the 32K limit
- Will be constrained by region size
- Cross Load of 2GB LOBs will still not be possible
- Cross Loader will also allow for conversion between CLOBs and DBCLOBs
  - not currently supported when loaded from file

LOAD/UNLOAD support for very large LOBs

- Requirement is to move LOBs from one z/OS system to another z/OS system
- Need to support millions of rows
- Typical LOB sizes are 25K, 200K, 1MB
- Need to allow user to limit LOAD at target with WHEN clause
- LOB column values will be stored as separate PDS member, PDS/E member, or HFS directory member.
- LOB column values from each row will have identical member names in each PDS, PDS/E, or HFS
- Data set name stored in output record
- Design fits well with File Reference Variables where LOB values are in individual datasets

Best Practices

COPY/RECOVER/QUIESCE Best Practices

- COPY
  - PARALLEL keyword provides parallelism for lists of objects
  - CHECKPAGE YES
- RECOVER
  - PARALLEL keyword provides parallelism for lists of objects
  - Enable Fast Log Apply (which can use dual-copy logs)
- QUIESCE
  - WRITE NO is less disruptive
  - Use TABLESPACESET

- Large BUFNO
  - Anecdotal evidence of improved performance with a large BUFNO (e.g., BUFNO=100) but we have not seen this in our benchmarks – we suspect that this helped in cases where I/O configuration was not well tuned

LOAD Best Practices

- LOAD
  - LOG NO reduces log volume, must be followed by a copy
  - KEEPDICTIONARY (track dictionary effectiveness with history statistics PAGESAVE)
  - Load Partition Parallelism (V7)
    - not individual LOAD part level jobs
  - Inline COPY & Inline STATISTICS
  - Index parallelism (SORTKEYS)
    - remove SORTWKxx and use SORTDEVT/SORTNUM
  - When using DISCARD, try to avoid having the input on tape
    - input is re-read to discard the errant records

REORG Best Practices

- REORG
  - LOG NO reduces log volume; requires an image copy (inline is a good choice)
  - KEEPDICTIONARY (track dictionary effectiveness with history statistics PAGESAVE)
  - On V7, SORTDATA to use table space scan and then sort
  - NOSYSREC to avoid I/O (always used for SHRLEVEL REFERENCE and CHANGE)
    - Use only if taking full image copy before REORG
  - Inline COPY & Inline STATISTICS
  - Index parallelism (SORTKEYS)
    - remove SORTWKxx and use SORTDEVT/SORTNUM

Online REORG Specific Best Practices

- REORG SHRLEVEL CHANGE (sometimes called online REORG)
  - TIMEOUT TERM frees up the objects if timeouts occur in getting drains
  - DRAIN ALL
    - Some customers have better success draining users if they drain readers and writers at once rather than writers first and then readers later
  - MAXRO = lock timeout ZPARM/2 (30 seconds by default)
  - DRAIN_WAIT = lock timeout ZPARM/2
  - RETRY = utility lock timeout multiplier (6 by default)
  - RETRY_WAIT = DRAIN_WAIT*RETRY

REBUILD/CHECK/RUNSTATS Best Practices

- REBUILD
  - Index parallelism (SORTKEYS)
    - remove SORTWKxx and use SORTDEVT/SORTNUM
  - Inline STATISTICS
- CHECK DATA
  - If large volumes of delete data (e.g. after REORG DISCARD)
    - LOG NO to avoid log archive and log latch contention
    - Image COPY will be required
- CHECK INDEX
  - SHRLEVEL CHANGE and large region size to get parallelism (for availability, performance, and for additional checks on root and non-leaf pages (PQ90086)) with Flashcopy V2 or snapshot on RVA
- RUNSTATS
  - SHRLEVEL CHANGE for availability
  - SAMPLE reduces CPU time when gathering column stats

Sorting with DFSORT Best Practices

- Remove SORTWKxx and use SORTDEVT/SORTNUM
  - This will use dynamic allocation
  - To direct datasets to storage group, use ACS (see DFSMSrmm SMS ACS Support reference on *References* slide)
- DFSORT installation options (see APAR II14047)
  - Leave the default for SIZE set to MAX
  - Don't bother with changing TMAXLIM (initial storage for each sort)
  - The only knob to consider adjusting is DSA (Dynamic Size Adjustment)
    - R14 DFSORT default is 32M; V1R5 DFSORT default is 64M
    - You could set this to 128M, but then look to see if DFSORT ever uses this much
    - Follow DFSORT tuning recommendation to use hiperspaces, data spaces, etc. (if not on 64-bit LPAR)
- >64K track datasets for DFSORT supported in z/OS 1.7

References

- DB2 UDB for z/OS home page
  - http://www.ibm.com/software/data/db2/zos/index.html
- utilities@work
  - http://www.ibm.com/software/data/db2imstools/db2tools/db2utilsuite8.html
- DB2 UDB for z/OS and OS/390 Version 7 Performance Topics, SG24-6129
- DB2 UDB for z/OS and OS/390 Version 7: Using the Utilities Suite, SG24-6289
- DB2 Magazine Fall 1998 - DB2 OS/390 Online Reorganization
  - http://www.db2mag.com/db_area/archives/1998/q3/98fextra.shtml
- DB2 Magazine Quarter 2, 2003 - Programmer's Only - Programs vs. Utilities
  - http://www.db2mag.com/db_area/archives/2003/q2/programmers.shtml
- Implementing Online Reorg in a Production Environment
  - http://www.ibm.com/software/data/db2/os390/pdf/oreorg.pdf
- Moving Data Across the DB2 Family, SG24-6905
- Recommendations for Tuning Large DFSORT Tasks
  - http://www.ibm.com/servers/storage/support/software/sort/mvs/tuning/index.html
- DFSMSrmm SMS ACS Support
  - http://www.redbooks.ibm.com/abstracts/TIPS0530.html?Open

DB2 UDB for z/OS information resources

- Information center
  http://publib.boulder.ibm.com/infocenter/dzichelp/index.jsp
- Information roadmap
  ibm.com/software/db2zos/roadmap.html
- DB2 UDB for z/OS library page
  ibm.com/software/db2zos/library.html
- Examples trading post
  ibm.com/software/db2zos/exHome.html
- DB2 for z/OS support
  ibm.com/software/db2zos/support.html
- Official Introduction to DB2 for z/OS
  ibm.com/software/data/education/bookstore

Disclaimers & Trademarks*

Information in this presentation about IBM's future plans reflect current thinking and is subject to change at IBM's business discretion.

You should not rely on such information to make business plans. Any discussion of OEM products is based upon information which has been publicly available and is subject to change.

The following terms are trademarks or registered trademarks of the IBM Corporation in the United States and/or other countries: AIX, AS/400, DATABASE 2, DB2, OS/390, OS/400, ES/9000, MVS/ESA, Netfinity, RISC, RISC SYSTEM/6000, SYSTEM/390, SQL/DS, VM/ESA, IBM, Lotus, NOTES. The following terms are trademarks or registered trademarks of the MICROSOFT Corporation in the United States and/or other countries: MICROSOFT, WINDOWS, ODBC