Improve Performance with DB2 V9 for zOS

By David Beulke

Dave@DaveBeulke@com

Dave Beulke & Associates
3213 Duke Street
Suite 805
Alexandria, VA 22314
703 798-3283

Dave@DaveBeulke.com - 703 798-3283

- Member of the inaugural IBM DB2 Data Champions
- One of 45 IBM DB2 Gold Consultant Worldwide
- Past President of International DB2 Users Group - IDUG
- Best speaker at CMG conference & former TDWI instructor

- Co-Author of certification tests
  - DB2 V8 & V7 DBA certification test
  - IBM Business Intelligence certification test

BLOG:
www.DaveBeulke.com

- Columnist for DB2 Magazine
  - Former editor of the IDUG Solutions Journal

- Consulting
  - CPU Demand Reduction Guaranteed
  - DB2 Performance Review
  - Database Design Review
  - Security Audit & Assessments
  - Migration Assistance
- Teaching Educational Seminars
  - DB2 Version 9 Transition
  - DB2 Version 9 Java Performance
  - Data Warehouse Performance Seminar
  - SOA Architecture Review
  - DB2 Business Intelligence

- Extensive experience in performance of large systems, databases and DW systems
  - Working with DB2 on z/OS since V1.2
  - Working with DB2 on LUW since OS/2 Extended Edition

- Author of Syspedia – Find, understand and integrate your data faster! - www.Syspedia.com

DB2 9 for z/OS Enhancements

SHRLEVEL(REFERENCE) for REORG of LOB table spaces	APPEND option at insert
Online RENAME COLUMN	Autonomic index page split
Online RENAME INDEX	Index page sizes 8K, 16K, 32K
Online CHECK DATA and CHECK LOB	Support for optimistic locking
Faster REORG by intra-REORG parallelism	Faster and more automatic DB2 restart
More online REORG by eliminating BUILD2 phase	MODIFY RECOVERY enhancements
LOB Lock reduction	RLF improvements for remote application servers such as SAP
Skipping locked rows option	Preserving consistency when recovering individual objects to a prior POT
Online REBUILD INDEX	
Change SCHEMA & VCAT	DECIMAL FLOAT, BIGINT
Tape support for BACKUP and RESTORE SYSTEM utilities	VARBINARY, BINARY
Recovery of individual tablespaces and indexes from volume level backups	TRUNCATE TABLE statement
	MERGE statement
Enhanced STOGROUP definition	FETCH CONTINUE
Utility TEMPLATE switching	ORDER BY and FETCH FIRST n ROWS in sub-select and full-select
Conditional restart: automatic search for appropriate checkpoi	ORDER OF extension to ORDER BY
CLONE Table: fast replacement of one table with another	Various scalar functions
Buffer management by WLM	XML support in DB2 engine
Global query optimization	Native SQL Stored Procedures, able to use zIIP
Generalizing sparse index and in-memory data caching metho	SELECT FROM UPDATE/DELETE/MERGE
Optimization Service Center	Enhanced CURRENT SCHEMA
Autonomic re-optimization	IPv6 support
Logging enhancements	Unified Debugger
LOBs Network Flow Optimization	Network Trusted Context
Faster operations for variable length rows	Database ROLEs
NOT LOGGED table spaces	Automatic creation of database objects
Index on expressions	Modify early code without requiring an IPL
Universal Table spaces	Utilities CPU reduction
Partition-by-growth table spaces	Temporary space consolidation

 Dave@DaveBeulke.com

DB2 9 – Many new settings & controls

- MAXTEMPS - online-updateable
  - Work file temp storage allowed for single process, 0=unlimitied
  - Important: Declared Global Temporary Tables use temp storage
    - No longer uses TEMP database DASD
    - Now use temp storage MEMORY!
      - **GROUP BY or HAVING (without index)**
      - **IN (subselect)**
      - **ORDER BY (without index)**
      - **ANY (subselect)**
      - **DISTINCT (without index)**
      - **UNION (except UNION ALL)**
      - **EXISTS (subselect) Some joins**
      - **Created Global Temp Tables**
      - **SORTs needed for processing SQL statements**
- No longer uses DASD now it uses faster memory!
  - Verify that its not over allocated

New Utility Controls

- Nine new zParms for controlling utilities

```
DSNTIP6 INSTALL DB2 - DB2 UTILITIES PARAMETERS
===>
Enter system-level backup options for RESTORE SYSTEM and RECOVER below:
1 SYSTEM-LEVEL BACKUPS          ===> YES             As a recovery base: YES or NO
2 RESTORE/RECOVER               ===> NO              From dump (NO or YES)
3 DUMP CLASS NAME               ===>                 For RESTORE/RECOVER from dump
4 MAXIMUM TAPE UNITS                        ===> NOLIMIT        For RESTORE SYSTEM: NOLIMIT or
1-255
Enter other DB2 Utilities options below:
5 TEMPORARY DS UNIT NAME        ===> SYSDA           Device for temporary utility data sets
6 UTILITY CACHE OPTION          ===> NO              3990 storage for DB2 utility IO
7 STATISTICS HISTORY                        ===> NONE           Default for collection of stats
history
8 STATISTICS ROLLUP             ===> NO              Allow statistics aggregation. NO or YES
9 UTILITY TIMEOUT               ===> 6               Utility wait time multiplier
```

- Important for online utilities especially reorgs
  - System restore options are new in DB2 9

Dave@DaveBeulke.com

Utility CPU Reductions

- 0 to 15% Copy Tablespace
- 5 to 20% in Recover index, Rebuild Index, and Reorg Tablespace/Partition
- 5 to 30% in Load
- 20 to 60% in Check Index
- 35% in Load Partition
- 30 to 50% in Runstats Index
- 40 to 50% in Reorg Index
- Up to 70% in Load Replace Partition with NPIs and dummy input

- **Eliminating the BUILD2 phase provides a dramatic improvement in availability, but can increase the CPU time and the elapsed time when one or a few partitions are reorganized.**

DB2 9 New Settings & Controls

- REOPT AUTOMATICALLY ==> **YES**
  - Star Join parameters externalized
- DECFLOAT ROUNDING MODE===> **ROUND_HALF_EVEN**
- DB2 Version 9 introduces a second RLST table
  - DSNRLMTnn, and a unique index DSNMRLnn
    - Provides more control for governing applications
    - Don’t have empty DSNRLMTnn tables - just RLF overhead
- DB2 Private Protocol is finally gone!
  - Installation process verification

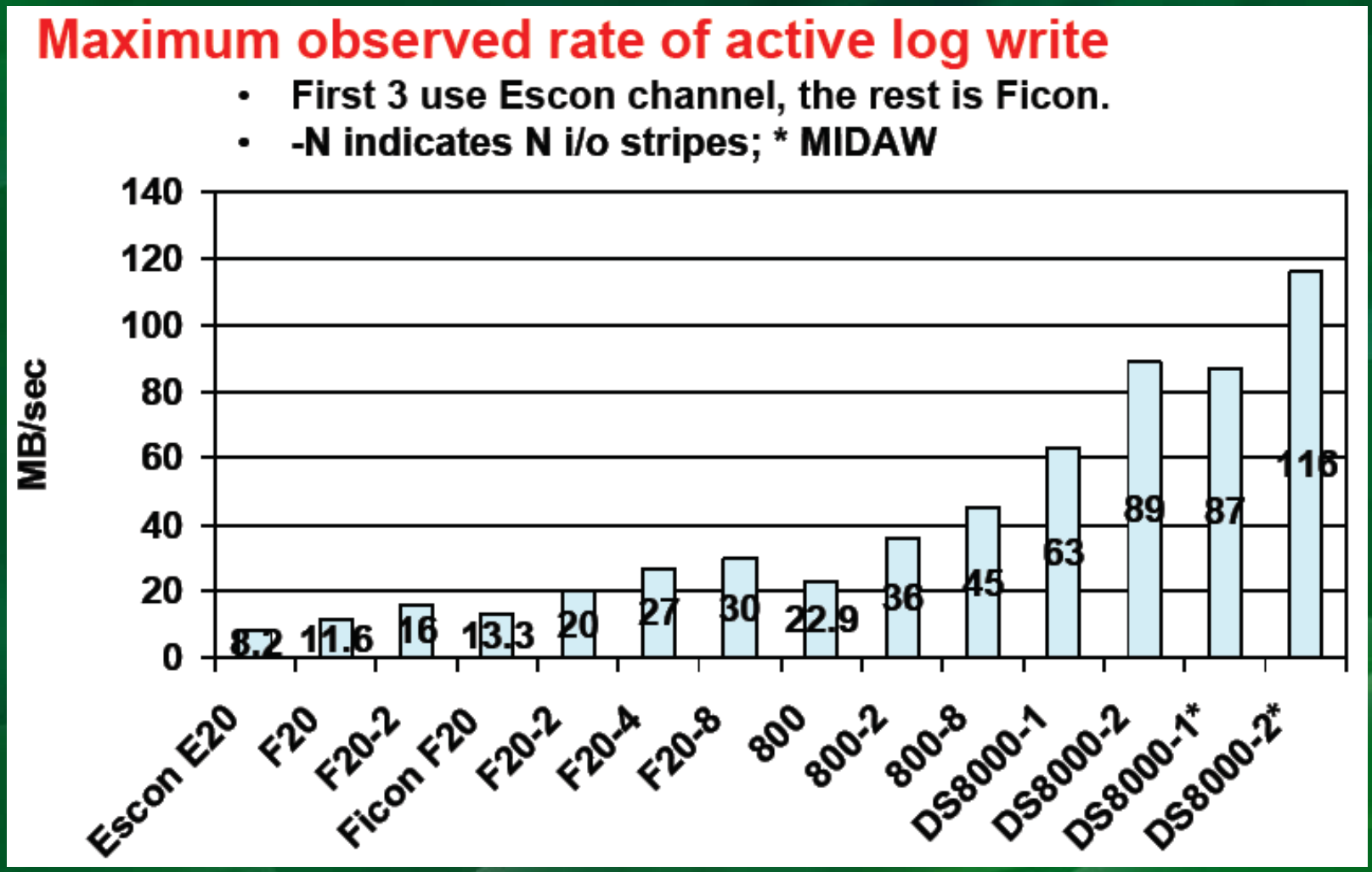
Dave@DaveBeulke.com

Disk Network – I/O x 2!

- Then use of the z9 and MIDAW improved that data rate to 109 MB / second.
- With two stripes,that configuration can reach 138 MB / second.
- DB2 9 changes in read quantity, write quantity and pre-format quantity allow the same hardware to deliver 183 MB / second in reading and a similar speed for writing.
- With the MIDAW, the performance gap between Extended Format (EF) data sets and non EF data sets is practically eliminated.

Dave@DaveBeulke.com

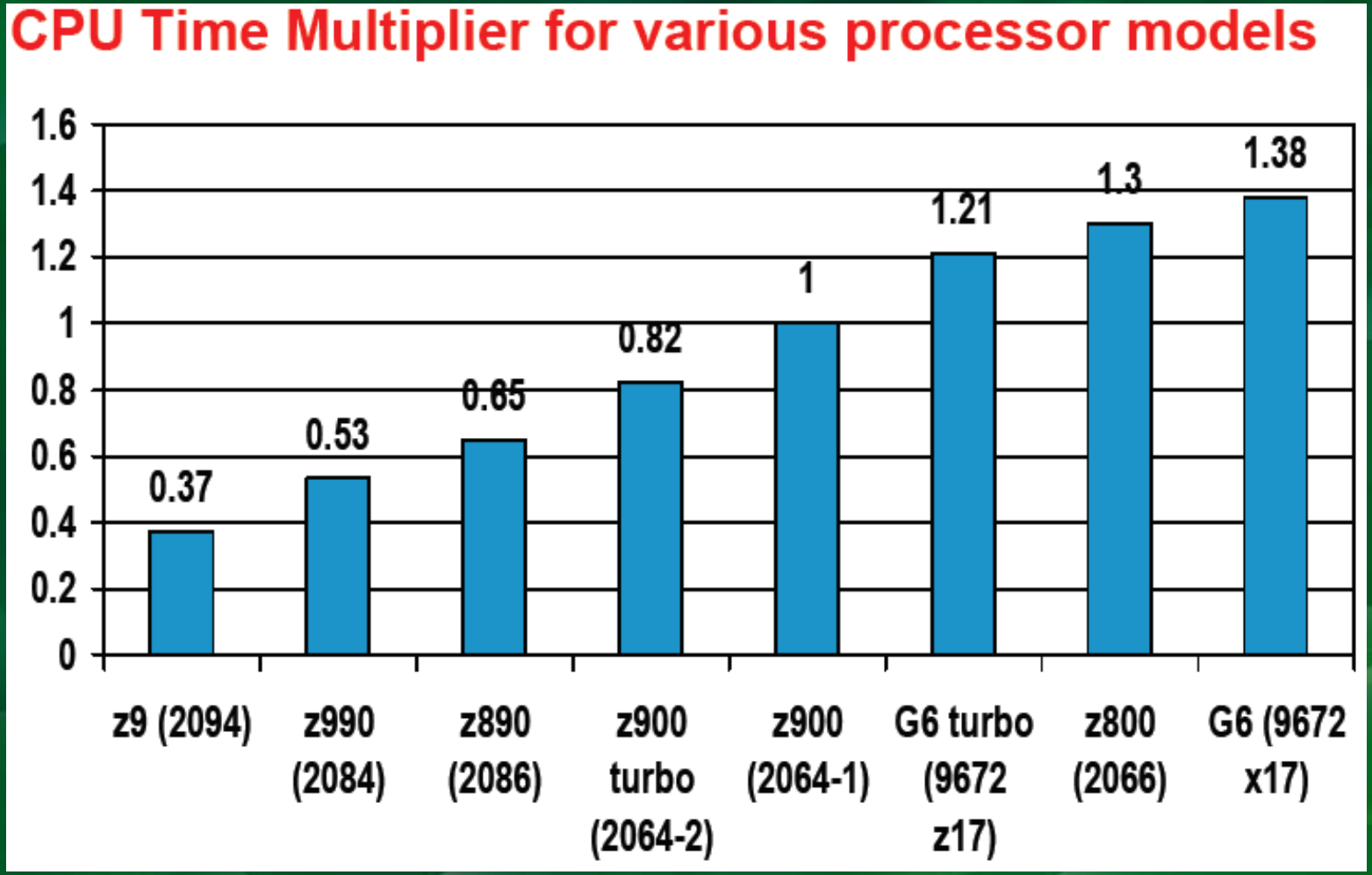
I/O and Network speeds over the years



Information from "DB2 9 for z/OS Performance Preview" by Akira Shibamiya

 Dave@DaveBeulke.com

CPU Processor Speed over the Years



Information from "DB2 9 for z/OS Performance Preview" by Akira Shibamiya

Dave@DaveBeulke.com

Version 9 Universal Tablespace

- Partitioned – many possibilities
- Each Partition has Space Map page(s)
  - Tracks the row density of the database tablespace pages

DSSIZE	4K Page	8K Page	16K Page	32K Page
1-4 GB DSSIZE	4096	4096	4096	4096
8 GB DSSIZE	2048	4096	4096	4096
16 GB DSSIZE	1024	2048	4096	4096
32 GB DSSIZE	512	1024	2048	4096
64 GB DSSIZE	256	512	1024	2048

- Partition by Growth
  - Defined with no partitioning key
  - Starts out as single partition
  - Dynamically adds partitions
- Up to 128TB in one table!

Dave@DaveBeulke.com

Universal TS and Clones

- Partition by Range
  - Defined with partitioning key
  - Uses keywords SEGSIZE and NUMPARTS
  - Can create ranges just like range partitioning

- Clones Tables in V9 NFM only
  - Defined with Universal Tablespace
  - Also initiate the creation of multiple additional clone objects such as indexes, constraints, and triggers
  - Many restrictions but great for availability issues

V9 TS APPEND, LOG NO & CATMAINT options

- Tablespace has an APPEND option
  - With some restrictions
    - Cannot be specified for tables created in LOB, XML or Work file table spaces
  - Alter APPEND YES can be done
    - Places rows into the table disregarding the clustering sequence
    - Rows are appended at the end of the table or appropriate partition
- Tablespaces can now have their logging turned off
  - Create or alter table to turn off logging
  - ***Beware of the recovery dangers!***
  - ***TRUNCATE***
- New CATMAINT facilities for VCAT, SCHEMA & OWNER
  - Previously done offline now can be handled within DB2 processes

Dave@DaveBeulke.com

V9 XML Tablespaces

- DB2 adds pureXML as a distinct storage mechanism
  - Mix and match XML with your regular DB2 tables
    - XMLDocument Data type
  - Serialized XML data column is limited to 2 GB
  - XML schema repository (XSR) validates an XML document on Insert or Update

- The major differences between XML data and relational data are:
  - An XML document contains information about the relationship of data items to each other in the form of the hierarchy.
  - With the relational model, the only types of relationships that can be defined are parent table and dependent table relationships. *From the XML Guide and Reference

V9 XML Tablespaces

- XML data is self-describing; relational data is not.
  - An XML document contains not only the data, but also tagging for the data that explains what it is. A single document can have different types of data.
  - With the relational model, the content of the data is defined by its column definition. All data in a column must have the same type of data.
- XML data has inherent ordering; relational data does not.
  - For an XML document, the order in which data items are specified is assumed to be the order of the data in the document. There is often no other way to specify order within the document.
  - For relational data, the order of the rows is not guaranteed unless you specify an ORDER BY clause on one or more columns.
    - *From the XML Guide and Reference

V9 XML Tablespace – When should I use XML

- Need performance - use DB2 relational
  - There is overhead serializing and interpreting XML data
  - Also can only update/replace full documents no update of small pieces
- Where else is the data used?
  - Is the data used in mainly relational application or XML data streams
  - XML decomposition can be done to store data in relational columns and then referenced via XML
- Example:
  - `CREATE TABLE DAVEBTAB (CustID BIGINT, INFO XML)#`
  - `CREATE UNIQUE INDEX MYCUT_CID_XMLIDX ON DAVEBTAB(INFO) GENERATE KEY USING XMLPATTERN 'declare default element namespace "http://syspedia.com"; /davebtab/@ CustID' AS SQL DECFLOAT#`

Dave@DaveBeulke.com

V9 XML Tablespace - When should I use XML

- Do you need Referential Integrity (RI)?
  - XML does not support RI of pieces of the XML structures
- Is the data understandable outside of its XML structure?
  - The XML values or the order within the hierarchy denotes its meaning
- Example:
  - `INSERT INTO DAVEBTAB (CustID, INFO) VALUES (1000, '<davebtab xmlns="http://syspedia.com" CustID="1000"> <name>Andy Clark</name> <addr country="USA"> <street>5 Buddy lane</street> <city>Rosebud</city> <prov-state>AR</prov-state> <pcode-zip>87616</pcode-zip> </addr> <phone type="work">703-798-3283</phone> </ davebtab>')#`

Many XML resources

- IBM XML Guide and Reference
  - Good Examples and explanations.
- DB2 9: pureXML – Overview and Fast Start
  - Great with DB2-C free download version of DB2 LUW
- XML and Java: Developing Web Applications
  - pureQuery
  - second edition already!
  - Great XML examples XSLT, DOM, SOAP, WSDL etc..
- Many more to come……

DB2 9 Index Compression

- DB2 9 different index page sizes are available
  - 4k, 8k, 16k and 32k page size
  - Index compression
    - Different than row compression
    - No compression dictionary
- Improves the number of index entries per page
  - Calculate the number of entries per index page
    - Make sure a proper number fit on a page
    - Choose the bigger page size for active indexes
      - Reduces the number of page splits
      - Reduces the data sharing issues
- Compressed on the index page
  - Expanded in the buffer pool

Dave@DaveBeulke.com

DB2 9 Index Improvements

- Index Access Improvements
  - DB2 keeps track of additional index entries
  - Index Lookaside caches index entries accessed
- Index Lookaside is now used for Delete process
  - Improves Delete processing by dramatically reducing I/O
    - Up to **100 time** reduction through improved caching
- Defaults changed
  - DB2 9 BIND ISOLATION defaults to CS
    - Not changed for distributed, REBIND
  - DB2 9 BIND CURRENTDATA defaults to NO
    - Not changed for REBIND

Non Partitioning Indexes - NPIs

- Any index not used to partition the data
  - Zero to many NPIs per table
- Spans the entire tablespace
  - Needs all the entries to build index
- Can be a problem for utility processing
  - Build2 Phase of utility processing eliminated
    - Online reorgs in Version 9 for any table

Index Enhancements with DB2 9

- Better comparison capabilities
  - Unlike Column and Host variable definitions
  - Character comparisons of unequal length
  - Decimal and integer comparisons
    - Especially important for Java, C#, .NET apps
- Better local and remote join operations
  - WHERE host= x + z
- EBCDIC and UNICODE Joins index-able
- Varchar padding or not
  - More index entries per page fewer pages/levels
  - Index only access available

Dave@DaveBeulke.com

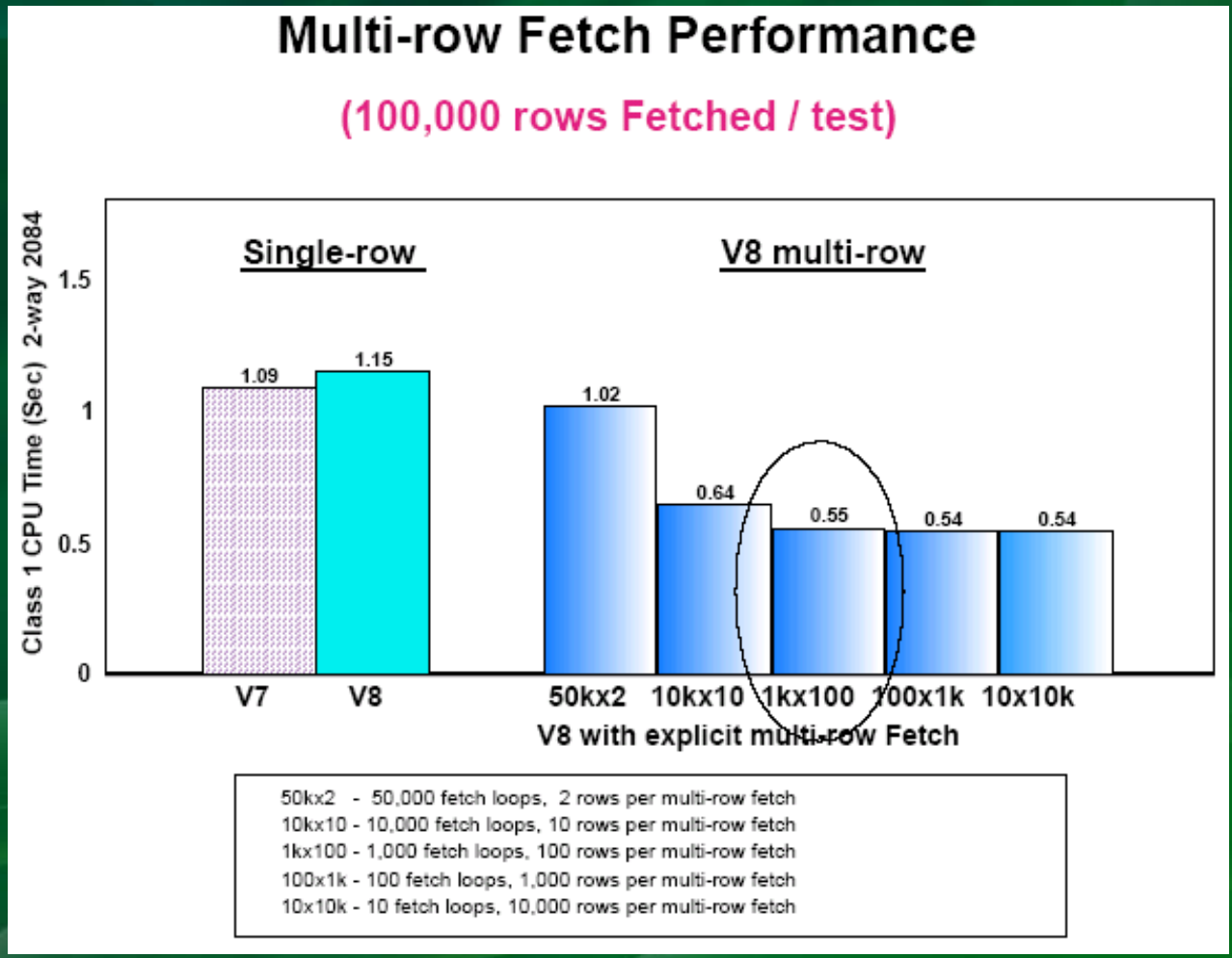
Online Profiles - Websphere

- Distributed application can connect several ways
  - ODBC – definition parameters
  - JDBC – Type 2 or Type 4 (pure Java) drivers
  - SQLJ – Pre-compile, bind and execute
  - DDF is now 64-bit
- Websphere applications – Server options
  - Connection concentrator
  - Many driver types available
  - JNDI – Java Naming & Directory Interface
  - DAO and .Net
- System Performance Characteristics
  - Security, object verification, and access authorization
  - Caching of resources and authorization

Block to 100x cuts costs almost half

- 10x gives 35% improvement
- 100x give 45% improvement
- Great for batch processing!

* From DB2 Performance Topics redbook-sg24646



Dave@DaveBeulke.com

Catalog V8 to V9

- **V9 CM to V9 NFM takes the next highest time, although much faster than V8 CM to V8 NFM, as online reorg of 2 catalog table spaces SYSPKAGE and SYSOBJ takes place.**

Large Buffer Pools

- In order to maximally exploit the performance characteristics of newer I/O hardware, both pre-fetch quantity and deferred write quantity are doubled resulting in 30 to 50% I/O throughput improvement.

- Now can have WLM controller Buffer Pools
  - Dynamically adjust Buffer pool storage allocations
  - Storage will be a problem..
  - BE CAREFUL!!!!!!!!!

Many Access Path Types

- Table Access
  - Tablespace scan – segmented or partitioned
- Index Access
  - One Fetch
  - Nested loop join
  - Merge Scan join
  - Star Join – Cartesian or Pair-wise
  - Outer join – Left/Right
  - Hybrid join – 2 Types: C or N
  - Multiple index access
  - Matching Index access
  - Non-Matching Index Access
  - List Pre-fetch
  - Index Lookaside

- **Join tables based on criteria**
- **Qualify rows – filtering data**
- **Aggregate data - grouping**
- **Ordering data – Sorts**

- **How much CPU and I/O does your access path require?**

New V9 Application Access Improvements

- Global Query Optimization
- Histogram Statistics
- Generalized Sparse Index
- In Memory Data Caching

New V9 Application Access Improvements

- GROUP BY/DISTINCT Optimization
- Pair-wise Star Join
- Query Parallelism Enhancements
- Page Range Index Screening Against Partitioned Table Indexes
- Index on Expression

Dave@DaveBeulke.com

DB2 9 for z/OS Enhancements

SHRLEVEL(REFERENCE) for REORG of LOB table spaces	APPEND option at insert
Online RENAME COLUMN	Autonomic index page split
Online RENAME INDEX	Index page sizes 8K, 16K, 32K
Online CHECK DATA and CHECK LOB	Support for optimistic locking
Faster REORG by intra-REORG parallelism	Faster and more automatic DB2 restart
More online REORG by eliminating BUILD2 phase	MODIFY RECOVERY enhancements
LOB Lock reduction	RLF improvements for remote application servers such as SAP
Skipping locked rows option	Preserving consistency when recovering individual objects to a prior POT
Online REBUILD INDEX	
Change SCHEMA & VCAT	DECIMAL FLOAT, BIGINT
Tape support for BACKUP and RESTORE SYSTEM utilities	VARBINARY, BINARY
Recovery of individual tablespaces and indexes from volume-le	TRUNCATE TABLE statement
backups	MERGE statement
Enhanced STOGROUP definition	FETCH CONTINUE
Utility TEMPLATE switching	ORDER BY and FETCH FIRST n ROWS in sub-select and full-select
Conditional restart: automatic search for appropriate checkpoi	ORDER OF extension to ORDER BY
CLONE Table: fast replacement of one table with another	Various scalar functions
Buffer management by WLM	XML support in DB2 engine
Global query optimization	Native SQL Stored Procedures, able to use zIIP
Generalizing sparse index and in-memory data caching method	SELECT FROM UPDATE/DELETE/MERGE
Optimization Service Center	Enhanced CURRENT SCHEMA
Autonomic re-optimization	IPv6 support
Logging enhancements	Unified Debugger
LOBs Network Flow Optimization	Network Trusted Context
Faster operations for variable length rows	Database ROLEs
NOT LOGGED table spaces	Automatic creation of database objects
Index on expressions	Modify early code without requiring an IPL
Universal Table spaces	Utilities CPU reduction
Partition-by-growth table spaces	Temporary space consolidation

 Dave@DaveBeulke.com

Additional resources - Too many enhancements

- IDUG Conference CD
  - *DB2 9 for z/OS Performance Preview* by Akira Shibamiya
  - *Optimizing DB2...Through Statistics Trace* (Part 1 & 2) by John Campbell
  - *What's New in DB2 9 Buffer Pool Management* by James Teng
  - *Planning Your Migration to DB2 9 for z/OS* by Roger Miller
  - *DB2 z/OS Optimizer: What have you done for me lately?* By Terry Purcell
- IBM RedBooks
  - DB2 9 for z/OS Performance Topics - www.rebooks.ibm.com

Questions or Comments?

Thank you!

BLOG: www.DaveBeulke.com

Email: Dave@DaveBeulke.com

Phone: 703 798-3283