Willie Favero
Dynamic Warehouse on System z Swat Team
IBM Silicon Valley Laboratory

Busted...

Myth Busting DB2's Parallelism

Willie Favero
Dynamic Warehouse on System z Swat Team
IBM Silicon Valley Laboratory

Busted...

Taking a Peak at DB2's Parallelism

August 1, 2010

Disclaimer

The information contained in this presentation has not been submitted to any formal IBM review and is distributed on an "As Is" basis without any warranty either expressed or implied. The use of this information is a customer responsibility.

The materials in this presentation are also subject to

enhancements at some future date,
a new release of DB2, or
a Programming Temporary Fix (PTF)

IBM MAY HAVE PATENTS OR PENDING PATENT APPLICATIONS COVERING SUBJECT MATTER IN THIS DOCUMENT. THE FURNISHING OF THIS DOCUMENT DOES NOT IMPLY GIVING LICENSE TO THESE PATENTS.

TRADEMARKS: 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, e-business logo, Enterprise Storage Server, ESCON, FICON, OS/390, OS/400, ES/9000, MVS/ESA, Netfinity, RISC, RISC SYSTEM/6000, iSeries, pSeries, xSeries, SYSTEM/390, IBM, Lotus, NOTES, WebSphere, z/Architecture, z/OS, zSeries, System z.

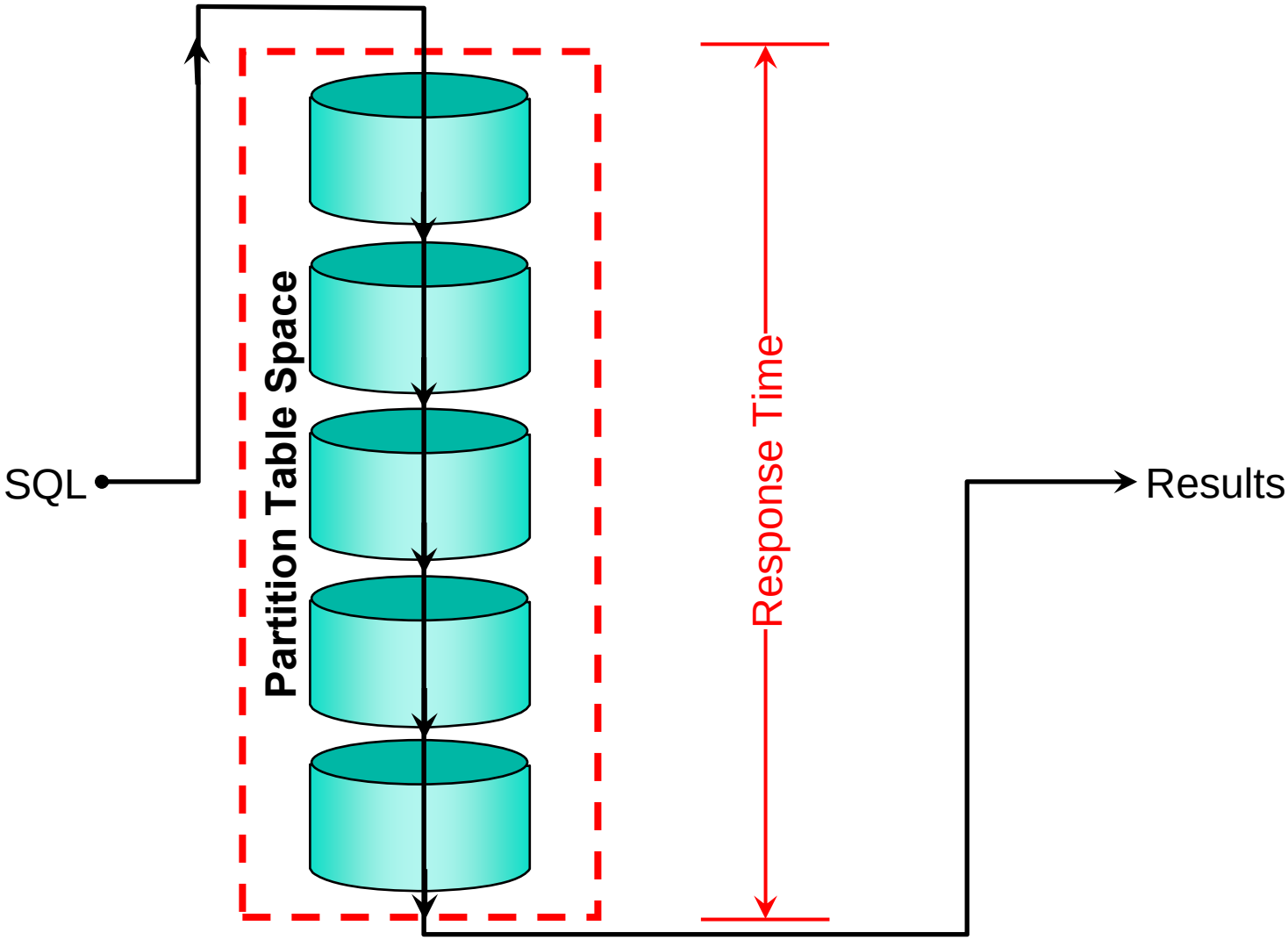
THE FOLLOWING TERMS ARE TRADEMARKS OR REGISTERED TRADEMARKS OF THE MICROSOFT CORPORATION IN THE UNITED STATES AND/OR OTHER COUNTRIES: MICROSOFT, WINDOWS, WINDOWS NT, ODBC and WINDOWS 95.

For additional information visit the URL
http://www.ibm.com/legal/copytrade.phtml ***for* "Copyright and trademark information"**

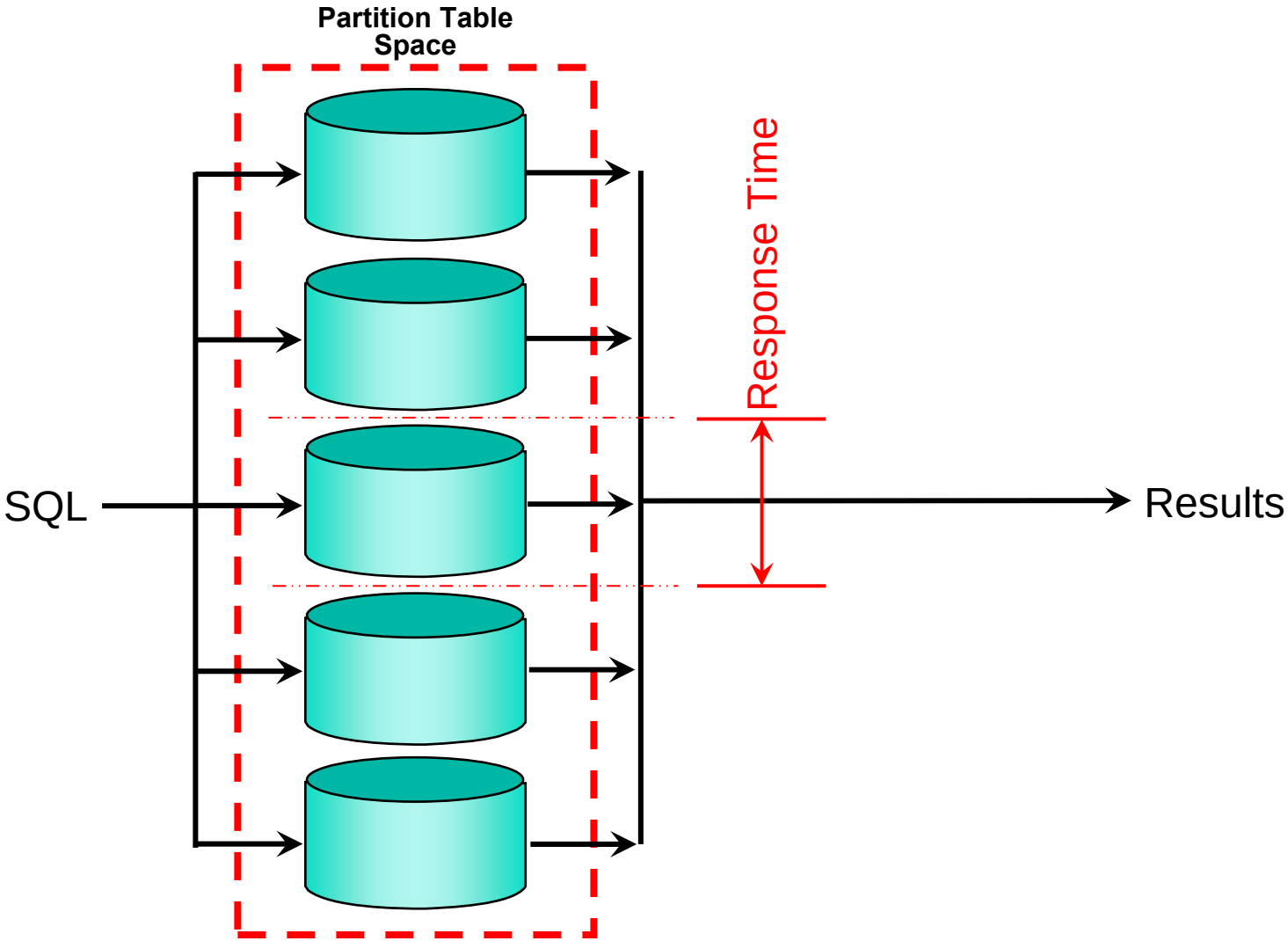
The History of DB2's Parallelism

(on the mainframe of course)

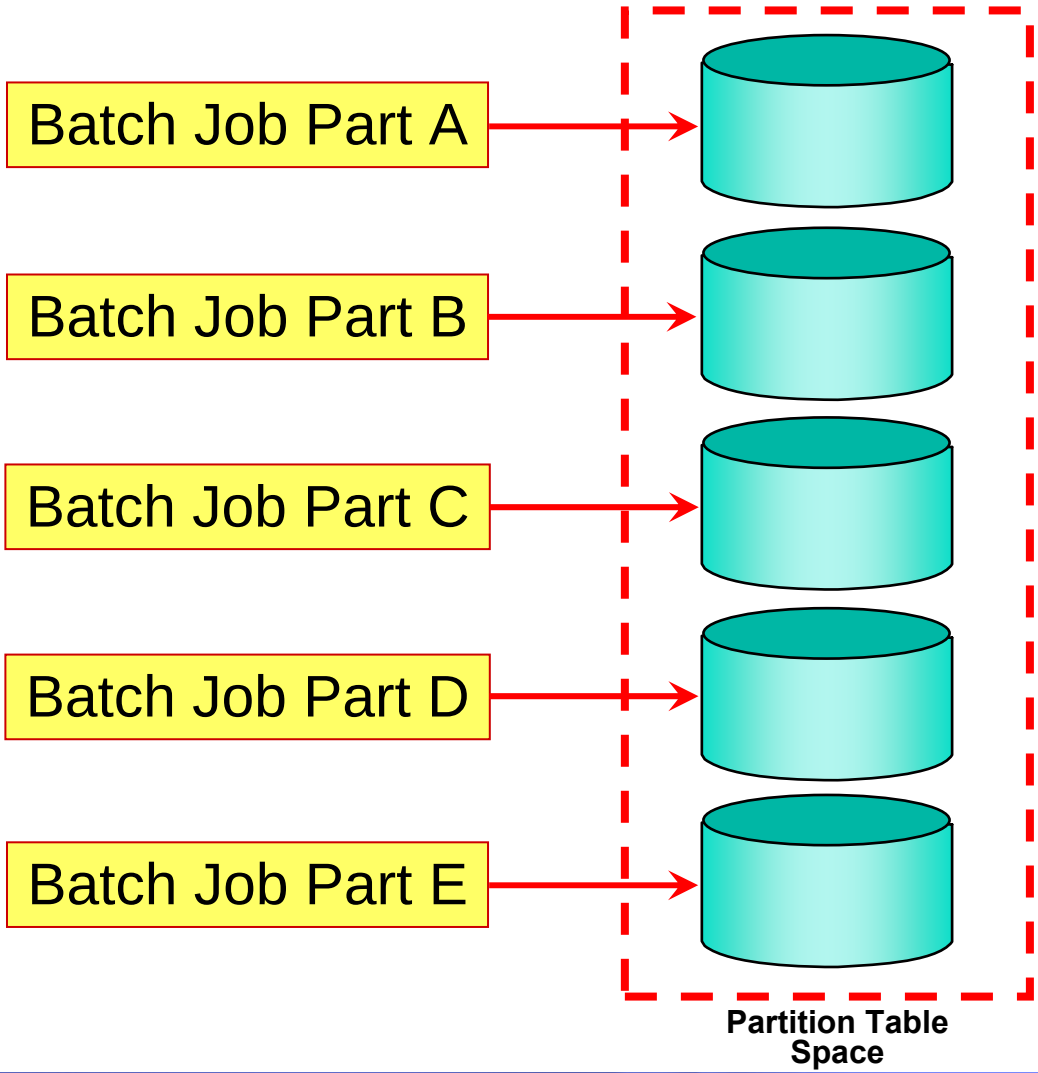
Query Without Parallelism



Query With Parallelism (potentially)



Batch Program Parallelism



Part A = 1 -10
Part B = 11- 20
Part C = 21- 30
Part D = 31- 40
Part E = 40 - ?

1/5 the time,
≈1/5 the contention
…in theory

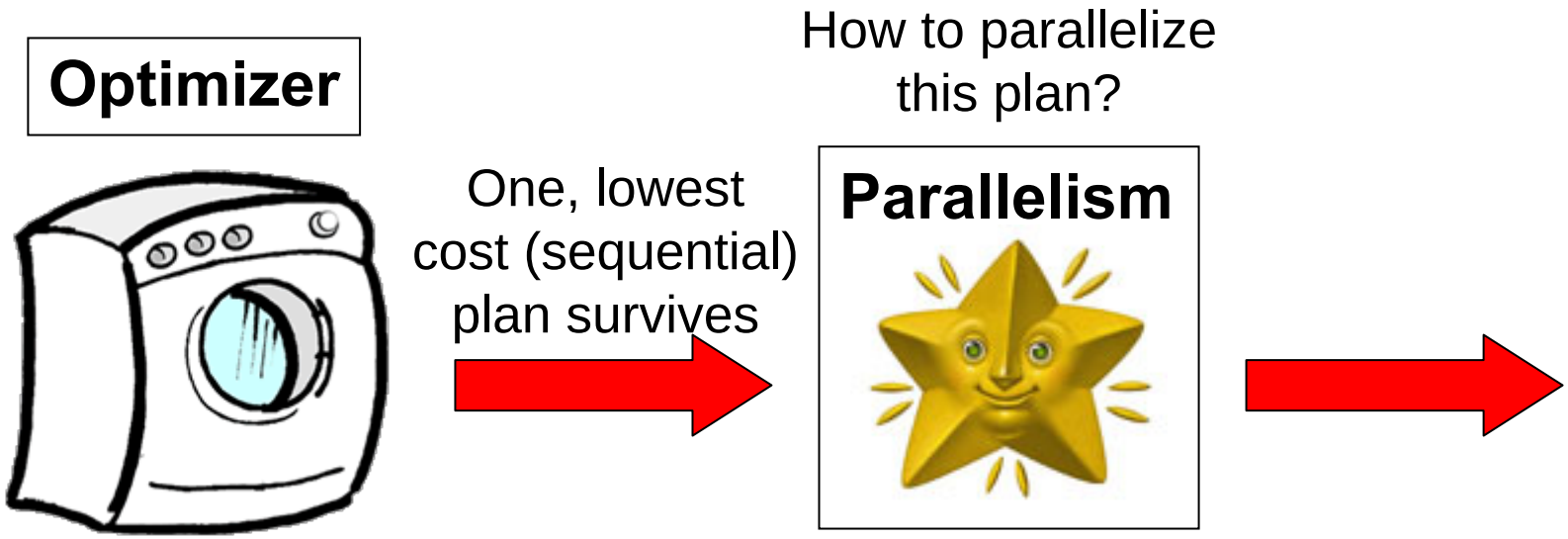
Parallelism modes

- In priority order
  - Sysplex
    - Eligible for zIIP redirect
    - EXPLAIN's PARALLELISM_MODE='X'
  - CPU 

    - zIIP eligible
    - EXPLAIN's PARALLELISM_MODE='C'
  - I/O
    - Not eligible for zIIP redirect
    - EXPLAIN's PARALLELISM_MODE='I'

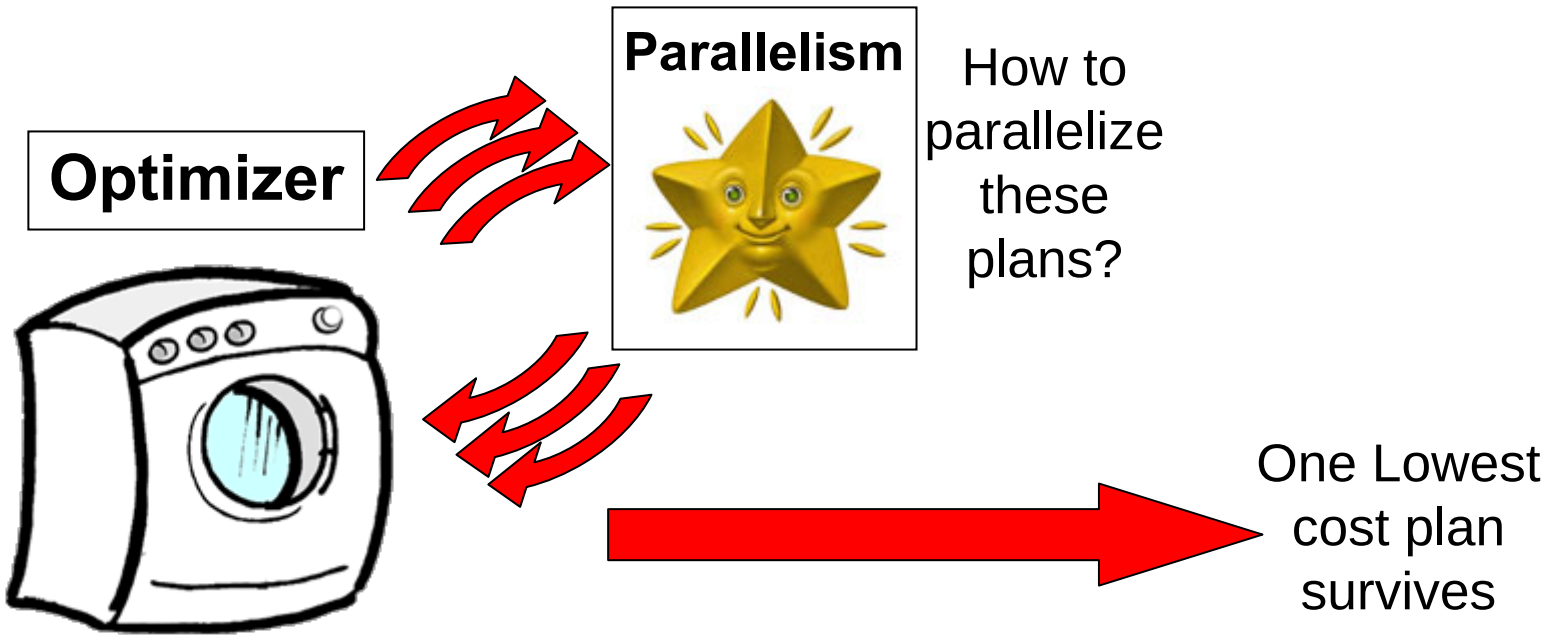
How is parallelism plan determined?

- Parallelism is considered for SELECT only
- In V8 and prior
  - Optimizer chooses the lowest cost sequential plan
    - And then determines what operations of the winning plan can be executed in parallel



How is parallelism plan determined?

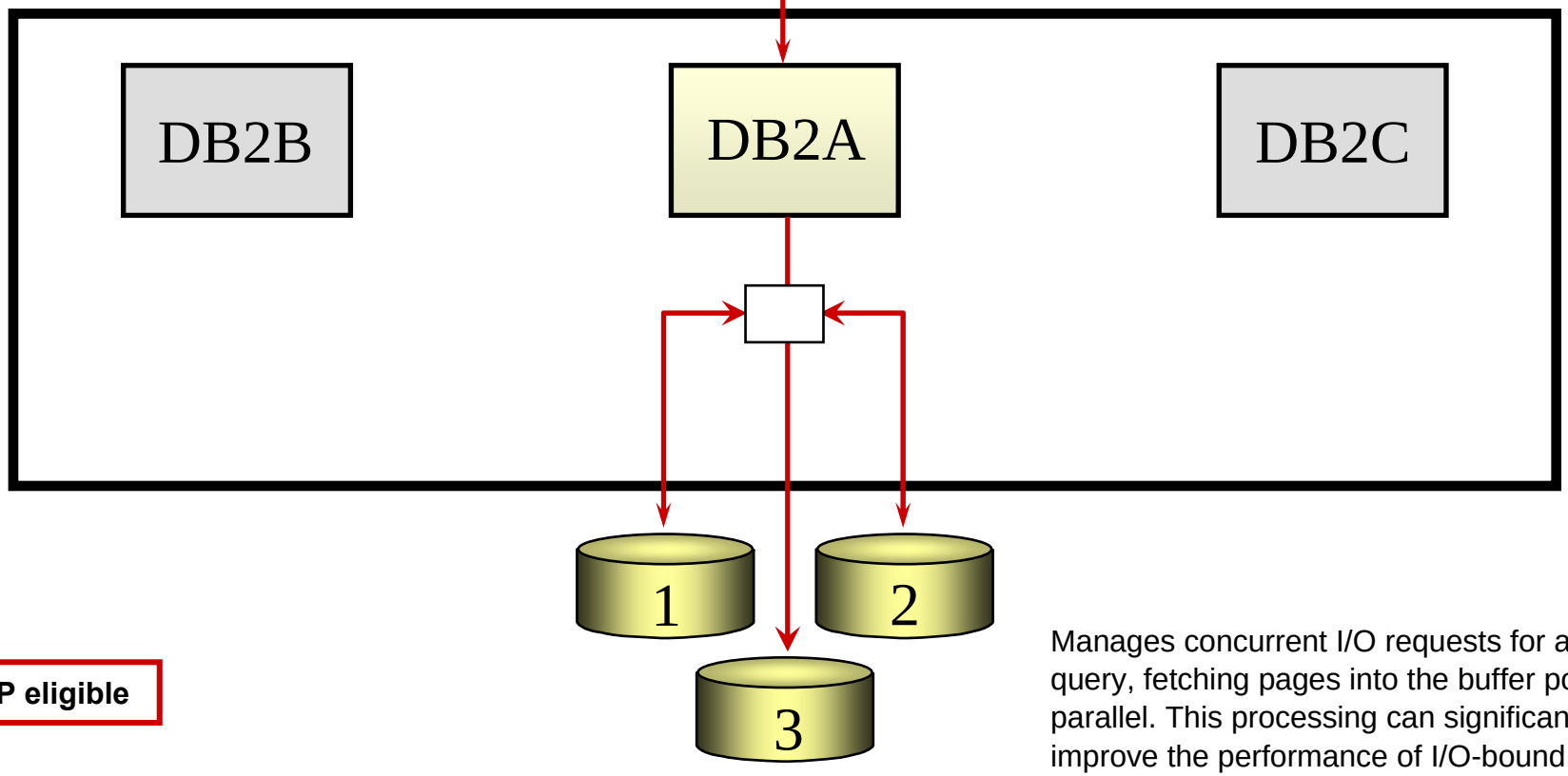
- In DB2 9
  - Lowest cost is AFTER parallelism
    - Only a subset of plans are considered for parallelism



I/O Parallelism

```
SELECT * FROM TAB1 A, TAB2 B
   WHERE A.COLA=B.COLZ
      ORDER BY A.COLM
```

Introduced in DB2 V3



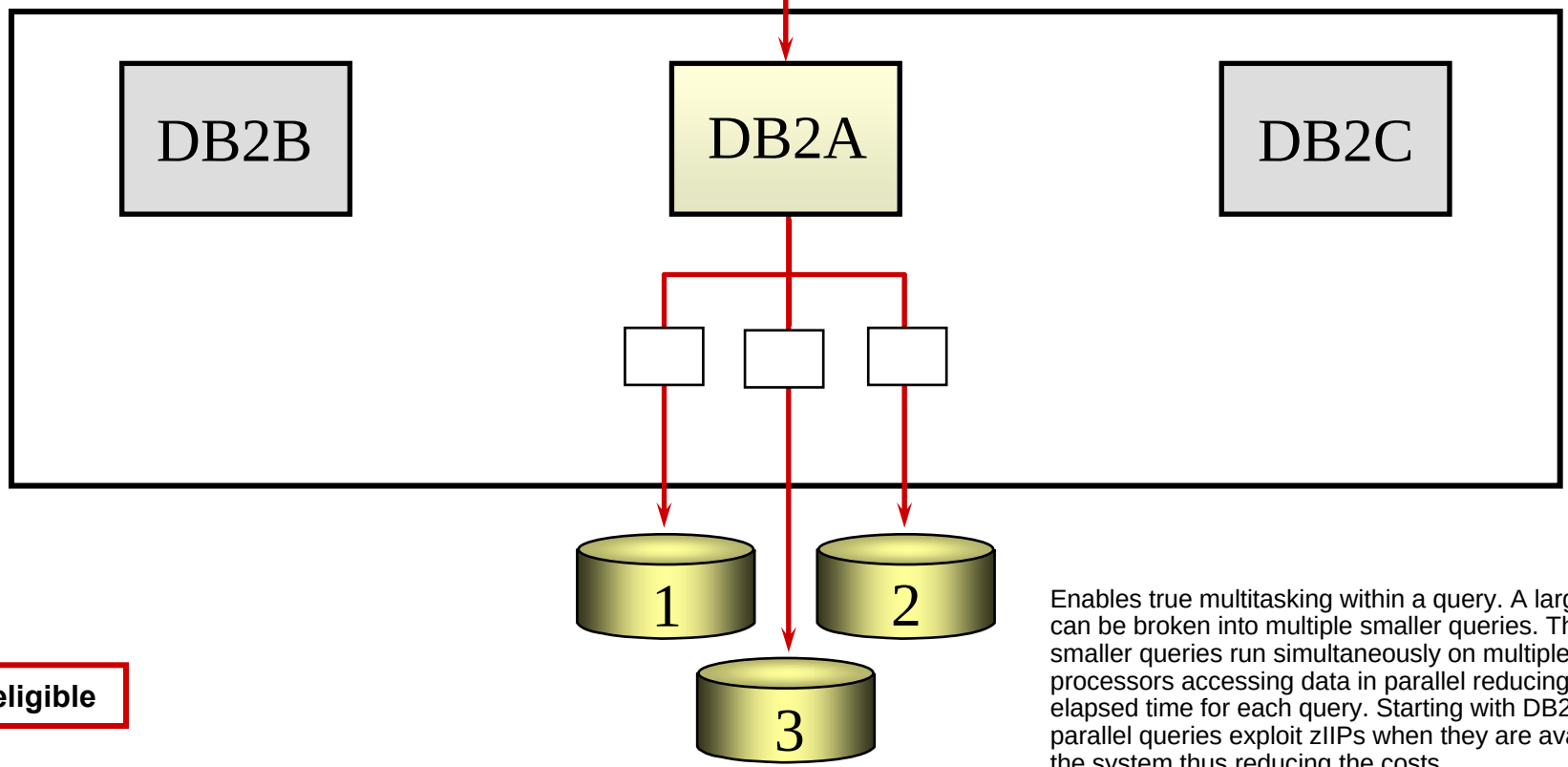
Not zIIP eligible

Manages concurrent I/O requests for a single query, fetching pages into the buffer pool in parallel. This processing can significantly improve the performance of I/O-bound queries.

CP Parallelism

```
SELECT * FROM TAB1 A, TAB2 B
   WHERE A.COLA=B.COLZ
      ORDER BY A.COLM
```

Introduced in DB2 V4.1

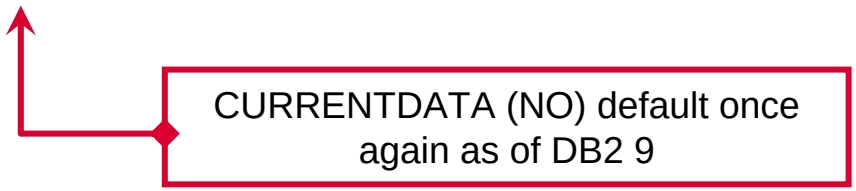


zIIP eligible

Enables true multitasking within a query. A large query can be broken into multiple smaller queries. These smaller queries run simultaneously on multiple processors accessing data in parallel reducing the elapsed time for each query. Starting with DB2 V8, the parallel queries exploit zIIPs when they are available on the system thus reducing the costs.

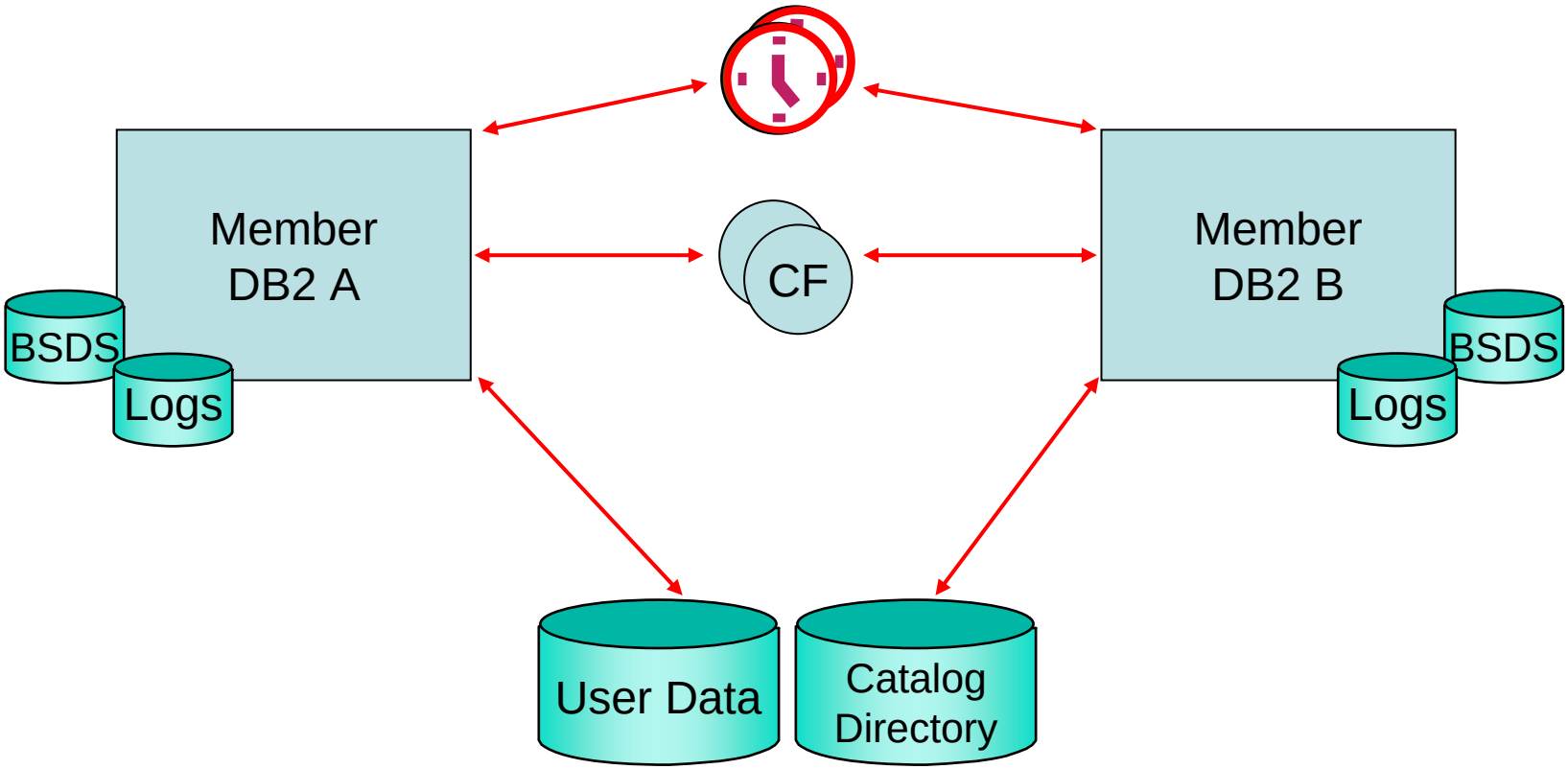
Parallelism Modes - CPU

- CPU parallelism is not considered if:
  - Only one CPU
  - CPU parallelism disabled by RLF
    - RLFFUNC='4' in RLF table (for plan/package/authid)
  - Cursor declared WITH HOLD and RR or RS isolation
  - No MVS enclave service
  - VPPSEQT=0
  - Ambiguous cursor with
    - CURRENTDATA(YES) and ISOLATION(CS)



Data Sharing

A really high level view!

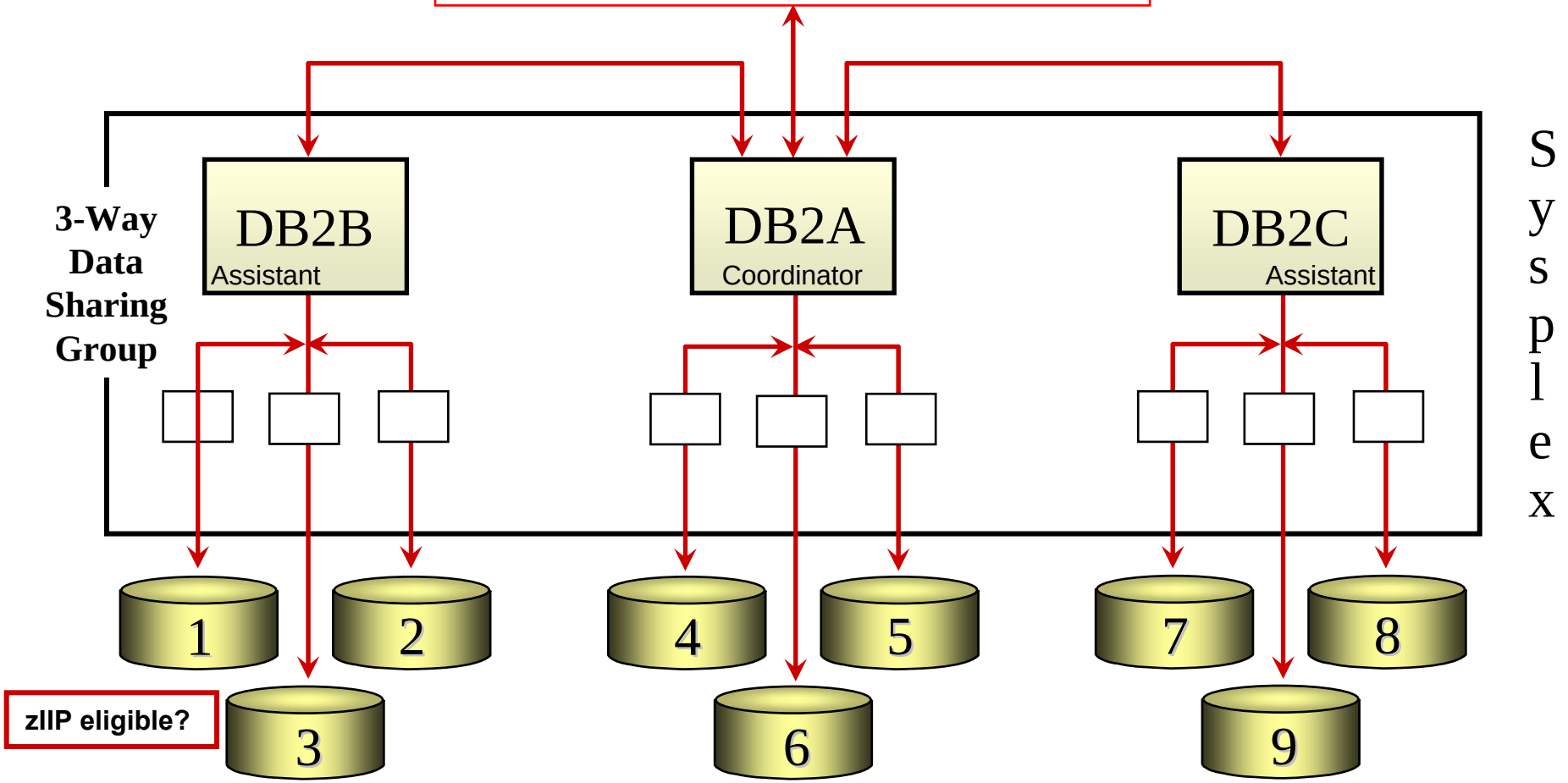


Sysplex Query Parallelism

DB2 can split a large query across different DB2 members in a data sharing group, known as Sysplex query parallelism.

```
SELECT * FROM TAB1 A, TAB2 B
  WHERE A.COLA=B.COLZ
  ORDER BY A.COLM
```

Introduced in DB2 V5

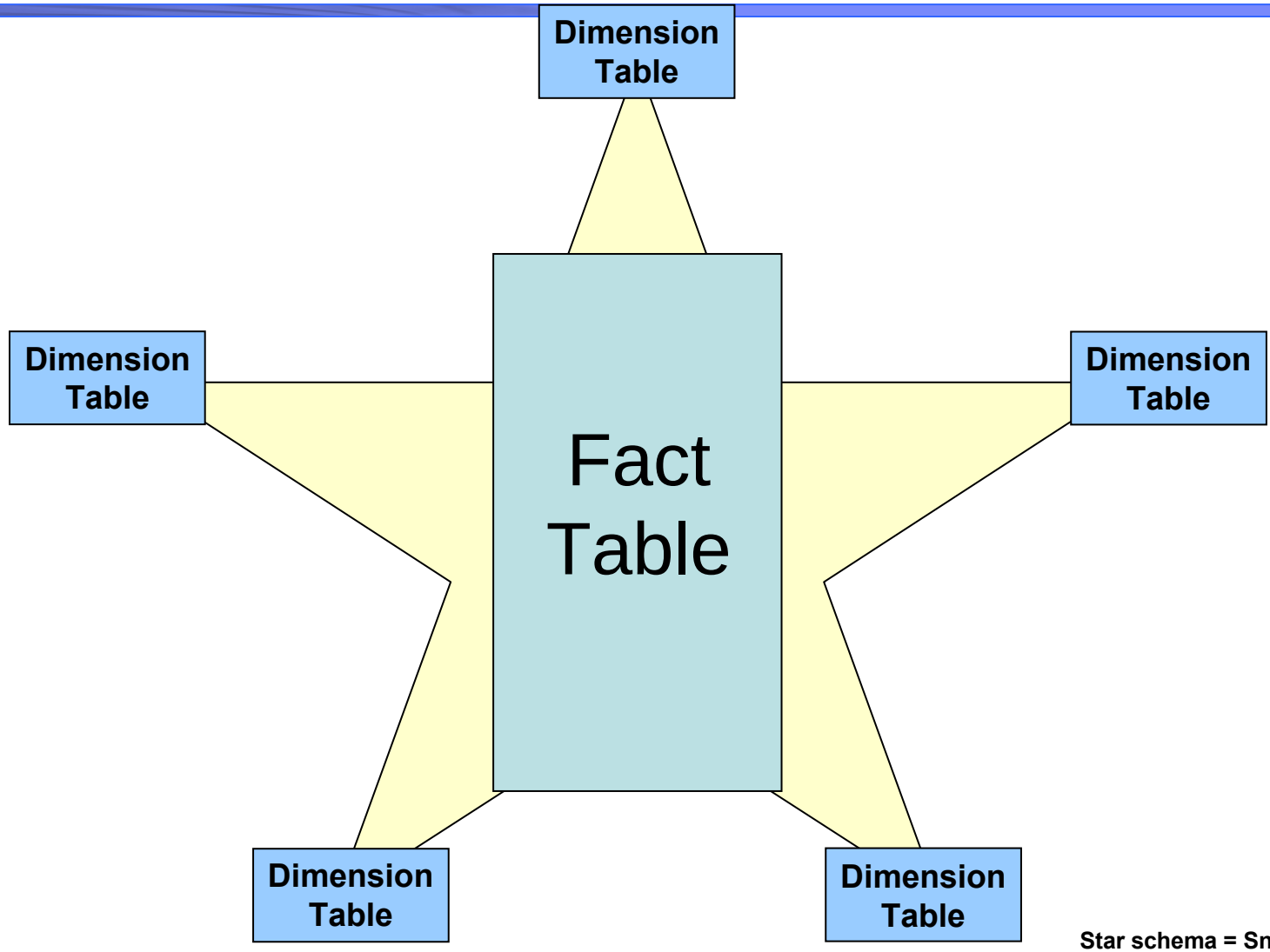


Sysplex Parallelism Considerations

- Sysplex parallelism is not considered if:
  - Data sharing not supported
  - Sysplex parallelism disabled by RLF
    - RLFFUNC='5‘ in RLF table (for plan/package/authid)
  - RR or RS isolation specified

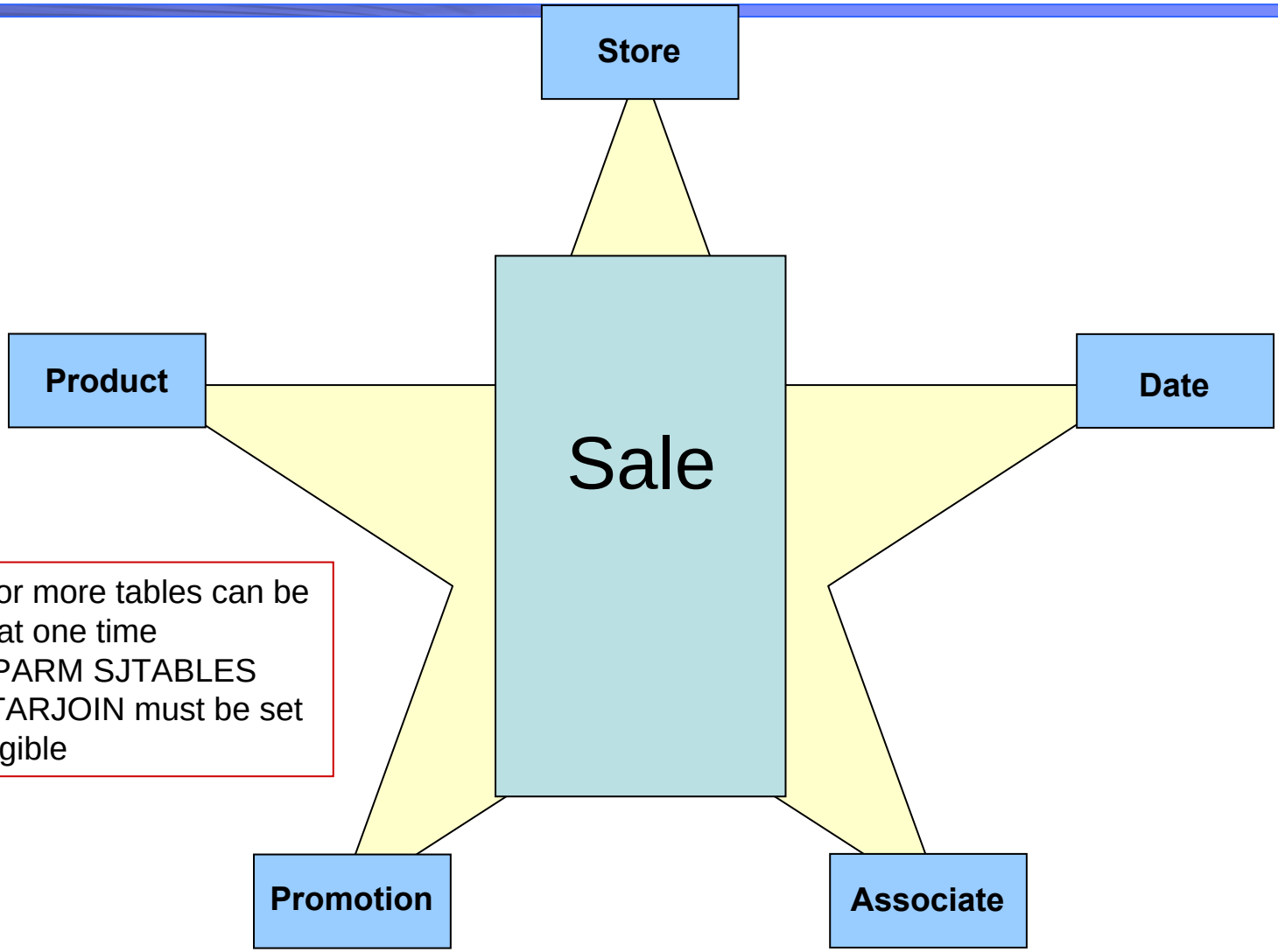
- Sysplex parallelism degrades to CPU parallelism if:
  - Star join query
  - RID access or IN-list parallelism
  - Sparse index used

Star Schema



Star schema = Snowflake schema

Star Schema



- Three or more tables can be joined at one time
- DSNZPARM SJTABLES and STARJOIN must be set
- zIIP eligible

Star schema = Snowflake schema

Pause for Questions

Myths?

- Parallelism doesn't work correctly.
- IBM recommends "*x*".
  - Corollary 1: PARAMDEG should be…
  - Corollary 2: Never exceed x degrees of parallelism
  - Corollary 3: Always use degree of 0 (zero)
- No documentation exist for parallelism.
- Parallelism is difficult to control?

Pause

simply for effect

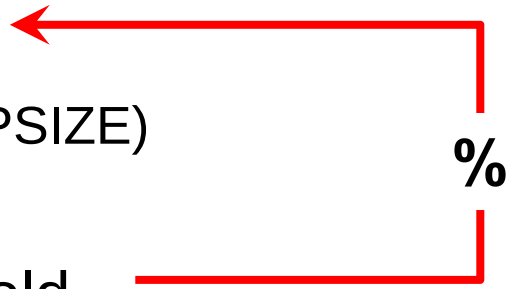
Enabling Parallelism – Part 1

- BIND - DEGREE ANY (not 1, the default)
- CDSSRDEF in macro DSN6SPRM (CURRENT DEGREE on install panel)
  - 1 is default, Valid values are 1 or ANY, Can update online
    - Consider taking the default

Enabling Parallelism – Part 2

Buffer Pool Affect on Parallelism

- VPSEQT - sequential steal threshold
  - Percentage of the total buffer pool size (VPSIZE)
  - Default 80%
- VPPSEQT - parallel sequential threshold
  - Percentage of the VPSEQT
  - Default 50%
  - VPPSEQT shuts down parallelism
- VPXPSEQT - assisting parallel sequential threshold
  - Percentage of the VPPSEQT
  - Default 0%



Enabling Parallelism – Part 3

Set Parallelism's Max. Degree

- Macro DSN6SPRM, keyword PARAMDEG
  - MAX DEGREE field on install DSNTIP8 (DB2 9)
  - Default 0
  - Valid range 0 to 254
    - Online Updatable (YES)
  - Set the maximum degree between the # of CPs and the # of partitions
  - CPU intensive queries - closer to the # of CPs
  - I/O intensive queries - closer to the # of partitions
  - Data skew can reduce # of degrees

Disabling Parallelism

- Disabling parallelism:
  - For dynamic SQL:
    - SET CURRENT DEGREE = '1'
    - Install value CURRENT DEGREE sets special register
      - For DSNZPARM, CDSSRDEF in macro DSN6SPRM
  - For static SQL:
    - DEGREE(1) at BIND PACKAGE or BIND PLAN
  - ALTER BUFFERPOOL(BPx) VPPSEQT(0)
  - Resource Limit Facility (RLST)
  - Add row for plan, package, or authid with RLFFUNC
    - 3 disables I/O parallelism
    - 4 disables CP parallelism
    - 5 disables Sysplex query parallelism
    - Need row for each type to disable all types

EXPLAIN

- **PARALLELISM_MODE='I', 'C', or 'X'**
  - I - parallel I/O operations
  - C - parallel CP operations
  - X - Sysplex query parallelism

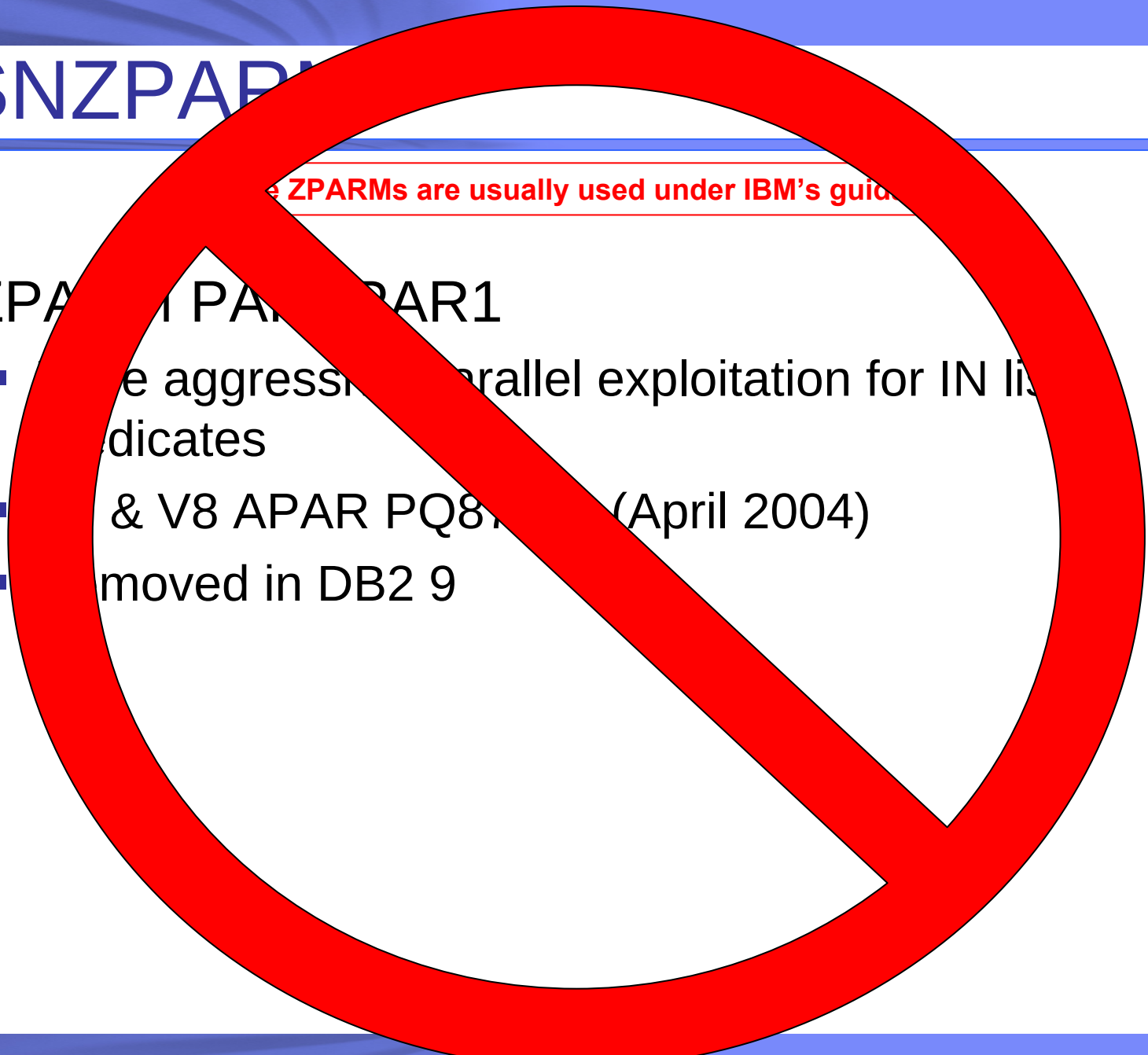
Parallelism and DSNZPARM

- PTASKROL in macro DSN6SYSP
  - YES is default, Valid values are YES or NO, Can update online
  - Accounting trace rollup
  - Performance verses granularity

DSNZPAR

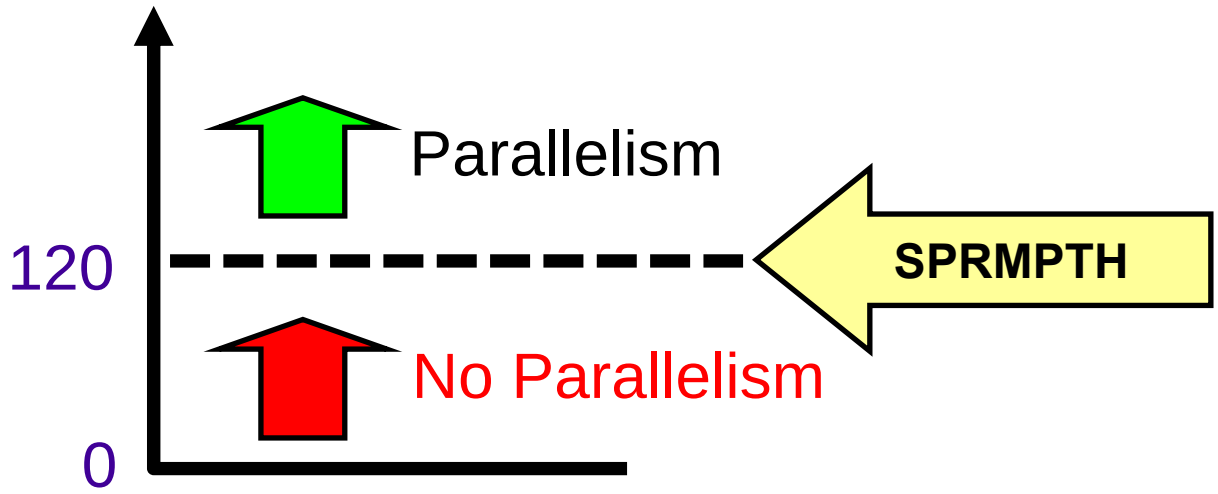
e ZPARMs are usually used under IBM's guid

- ZPA PA AR1
  - e aggressi rallel exploitation for IN li dicates
  - & V8 APAR PQ87 (April 2004)
  - moved in DB2 9



Lower Bound

- To avoid parallelism for short running queries
  - DSNZPARM SPRMPTH on macro
    - Query block with a cost lower than SPRMPTH will not choose parallelism
      - Default = 120 msec



APAR PQ25135 (V6),
APAR PQ45820

Cost-Base[illegible]arallel Se[illegible]

- DB2 V8 i[illegible]ces cost-based consideration[illegible]determine if parallel[illegible]ngle or multiple (composite) t[illegible]s should be dis[illegible]ed
  - S[illegible]ata size < 2[illegible]500 pages)
  - S[illegible]ata per parall[illegible]e < 100 KB (25 pages)
- **Hid[illegible] DSNZPARM C[illegible]PSE**
  - [illegible]rol - default is ON
  - [illegible] cost-based parallel sort [illegible]erations for both sing[illegible]ble a[illegible]multi-table
  - C[illegible] not cost-based - sort parallel[illegible]havior same a[illegible], i.e.
    - [illegible]gle (composite) table : parallel sort
    - [illegible]le tables : sequential sort only
- Consid[illegible]tions:
  - Elapse[illegible]e improvement
  - More usa[illegible]workfiles and virtual storage

- DB2 will use a higher degree of parallelism for read-only queries, batch jobs, and utilities when batch jobs are run in parallel and each job goes after different partitions. Parallel processing is most efficient when you spread the partitions over different disk volumes and allow each I/O stream to operate on a separate channel

Parallelism Helps Fully Utilize zIIP

- SQL could be tuned to increase parallelism
- Parallel child task will run on zIIP
  - Threshold exist that control when child goes to zIIP
- Parallel child tasks get higher % redirect than DRDA
  - Applies to local or distributed
    - Local non-parallel obtains 0% redirect
    - Distributed non-parallel obtains x% redirect
    - Parallel obtains x++% redirect
      - Except first “y” milliseconds
- If you want to reduce TCO
  - Exploit your zIIP as much as possible
    - Key is to increase parallelism for your queries

Tips to Optimize Parallelism

- Partition size does matter; the closer in size the partitions, the better
- The more CPs (and zIIPs) available, the better the degree of parallelism that might be achieved
- Separating partitions across as many I/O devices & I/O paths as possible can improve parallelism's performance
- For CPU intensive queries, try to make the number of partitions and number of CPs as close as possible
- Place the partitions (table spaces & indexes) on separate disk volumes and separate control units if possible to minimize contention
- Execute RUNSTATS on a regular basis to maintain accurate partition level stats
- Monitor buffer pool sizes and thresholds to make sure they are adequate to maximize parallelism’s performance

The Latest News

- **PM06953**
  - **Single Enclave Support For CP Query Parallelism**
- **PM16020**
  - **New feature to adjust parallelism reduction**
  - DSN6SPRM PARA_EFF
    - 0 – 100
    - Default 100

Final Disclaimer

- Customers can design to encourage parallelism
  - However, there is no guarantee that parallelism will be achieved
- Some restrictions on parallelism still exist
  - DB2 Optimizer Development are working to reduce restrictions
    - Substantial improvements in DB2 9
    - More improvements in DB2 10
    - And the future?????

Questions

More References

- My Blog:

  “Getting the Most out of DB2 for z/OS and System z”

  http://blogs.ittoolbox.com/database/db2zos

Willie Favero

Senior Certified Consulting IT Software Specialist

Dynamic Warehousing on System z Swat Team

IBM Silicon Valley Laboratory

IBM Academic Initiative Ambassador for System z

IBM Certified Database Administrator - DB2 Universal Database V8.1 for z/OS

IBM Certified Database Administrator – DB2 9 for z/OS

IBM zChampion

wfavero@attglobal.net