DB2 for z/OS: Data Sharing Technical Deep Dive

Session 17006

Mark Rader
IBM – DB2 for z/OS
March 5, 2015

#SHAREorg

SHARE is an independent volunteer-run information technology association that provides **education**, professional **networking** and industry **influence.**

Acknowledgements and Disclaimers

Availability. References in this presentation to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates.

The workshops, sessions and materials have been prepared by IBM or the session speakers and reflect their own views. They are provided for informational purposes only, and are neither intended to, nor shall have the effect of being, legal or other guidance or advice to any participant. While efforts were made to verify the completeness and accuracy of the information contained in this presentation, it is provided AS-IS without warranty of any kind, express or implied. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this presentation or any other materials. Nothing contained in this presentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer. Nothing contained in these materials is intended to, nor shall have the effect of, stating or implying that any activities undertaken by you will result in any specific sales, revenue growth or other results.

Assumptions

- This is a technical discussion of DB2 for z/OS Data Sharing topics
- The audience should be familiar with DB2 for z/OS Data Sharing concepts, behavior and benefits, based on
  - Experience with a data sharing environment
  - Recent data sharing education
  - DB2 for z/OS publications or Redbooks®
  - Flexible capacity
  - Scalability
  - High availability
  - Dynamic workload balancing

Acronyms

- CF – Coupling Facility LPAR
  - ICF – Integrated CF, aka Internal CF
- CFRM – CF Resource Management, definitions in CFRM policy
- CFCC – CF Control Code
- CF Links – connectivity between CF LPAR and 'host' CECs
  - ISC – fiber links, medium to long distance
  - ICB – copper links, very short distance
  - PSIFB - InfiniBand® links, short (12X IB) to long (1X IB) distance
  - IC – internal, microcode links for ICFs
- XCF – Cross-System Coupling Facility – communication between CECs
- XES – Cross-System Extended Services, z/OS component that manages CFs

Agenda

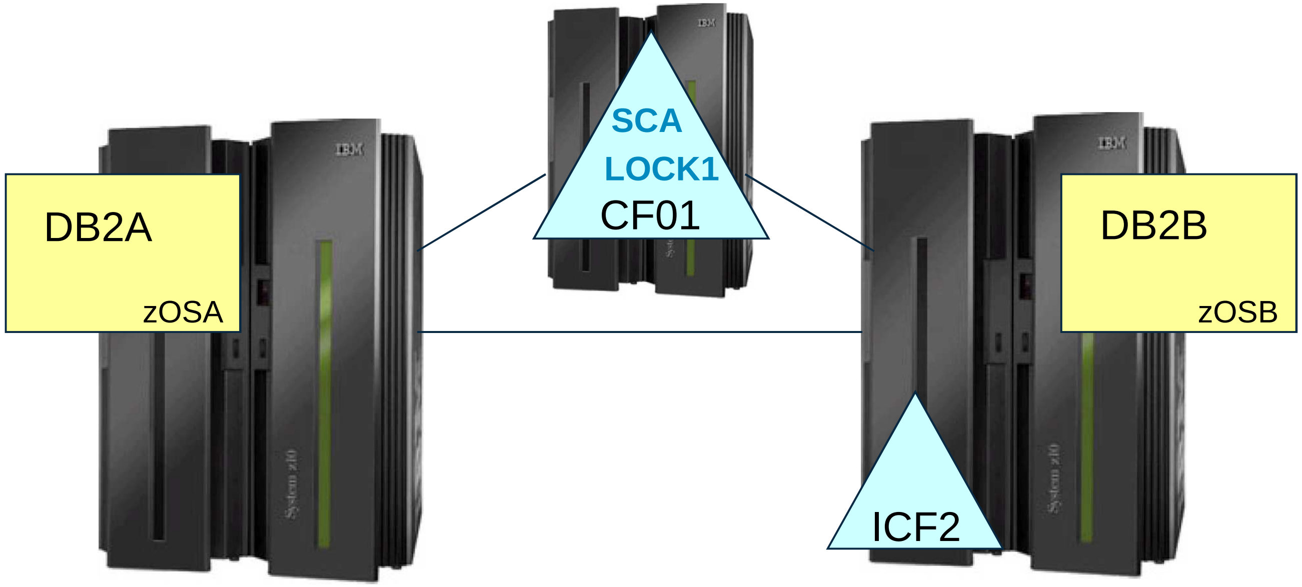
- DB2 Data Sharing
  - Configurations
  - Standard CF interaction
  - Performance monitoring
  - Auto Alter
- Workload growth
  - Lock structure
  - GBPs
  - Changes in configuration
    - CF considerations
- What's New in DB2 10 and DB2 11

DB2 Data Sharing Starting Configuration

- Starting with DB2 V4

DB2 Data Sharing: Usual Configuration

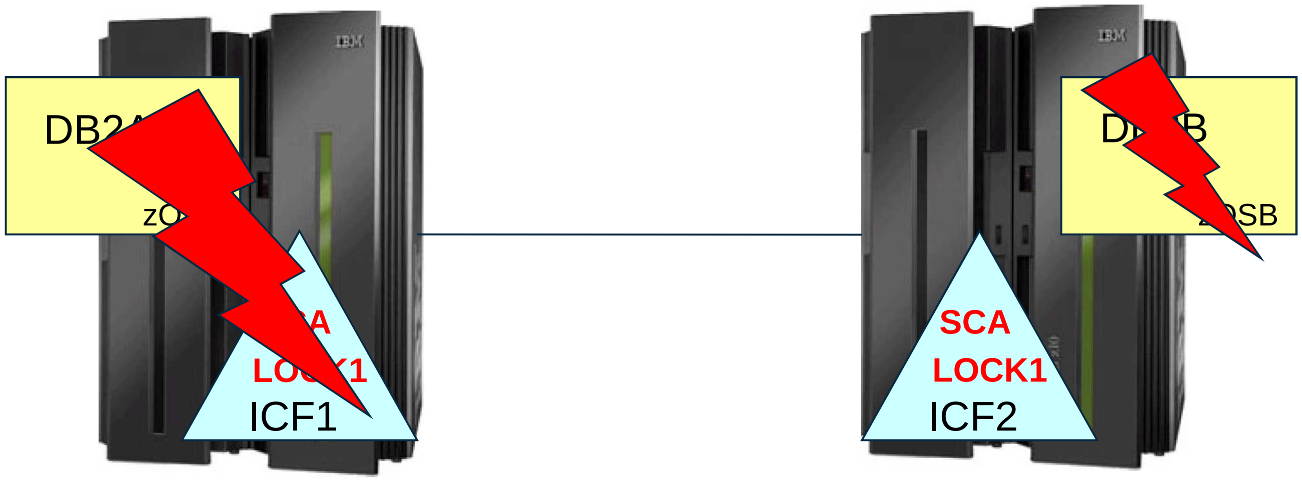
- Introduction of ICF



  - SCA and LOCK1 on external CF; isolated from DB2 and IRLM members
  - Duplexed GBPs spread across CF01 and ICF2

DB2 Data Sharing: 2-ICF Configuration

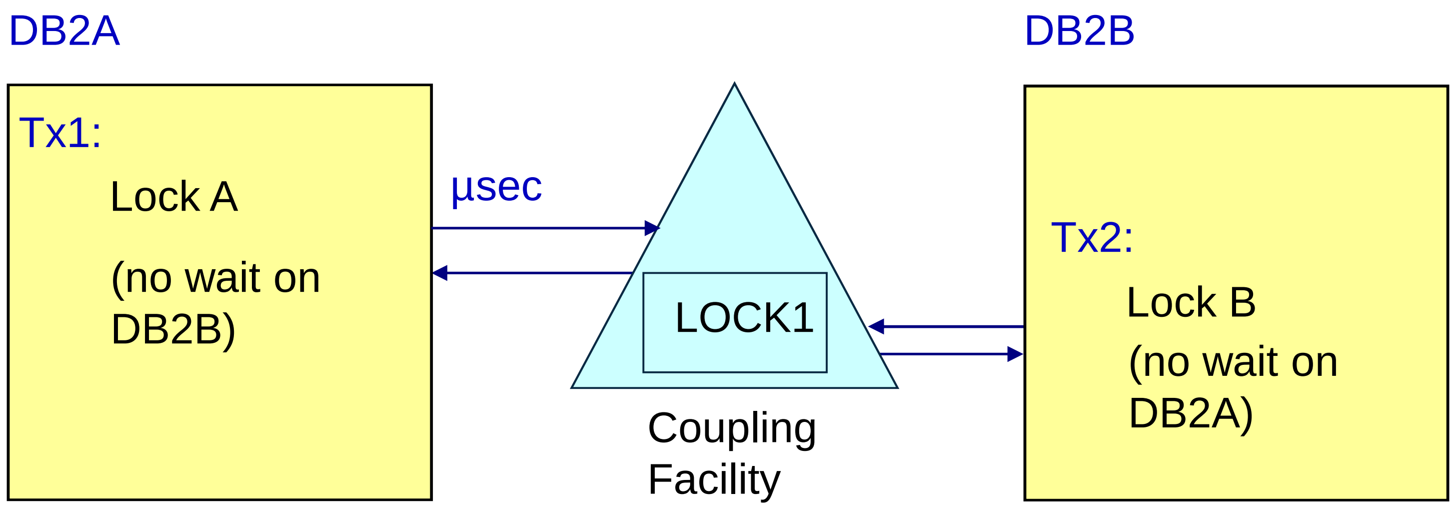
- Reduced number of CEC footprints
- Risk of 'double failure': DB2 and SCA, IRLM and LOCK1
  - If structure and exploiter fail, other members fail, too.



- Duplexed SCA and LOCK1 strongly recommended in this configuration
  - DB2B remains active, even if CEC on left is lost
  - Additional cost: host CPU, CF CPU, and CF link busy

Data Sharing: Locking

- Global locking using Parallel Sysplex® coupling technology
  – Inter-system concurrency control



  – Cost of obtaining lock does not increase when adding 3rd through nth members
  – This example assumes no contention

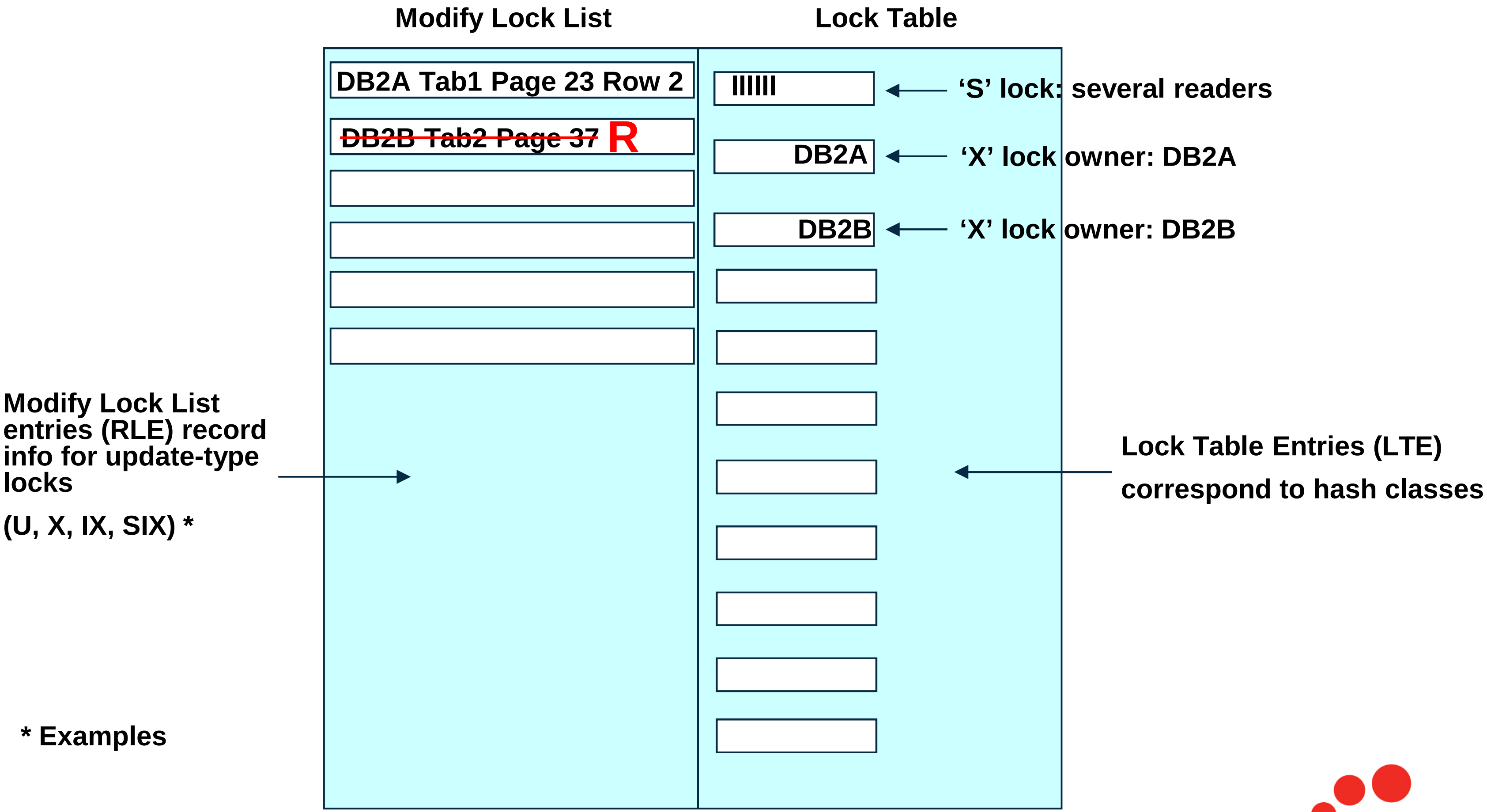
Notes: Lock Structure (LOCK1)

- Used by IRLM to manage global locking

- Holds L-locks and P-locks
  - L-locks to track concurrency
    - Parent L-locks: e.g. table space intent locks
    - Child L-locks: page or row locks
    - Others…
  - P-locks to track coherency. Examples:
    - Page set P-locks: table space, partition, index, index partition
    - Page P-locks: data page (RLL), index leaf page, space map page
    - Others…
- Consists of a lock table (hash table) and a modify lock list
  - Lock table controls access to resources
    - One entry can record multiple readers and one updater (owner)
  - Modify lock list contains detailed information for update-type locks
    - Entries become retained locks in case of an IRLM or DB2 failure

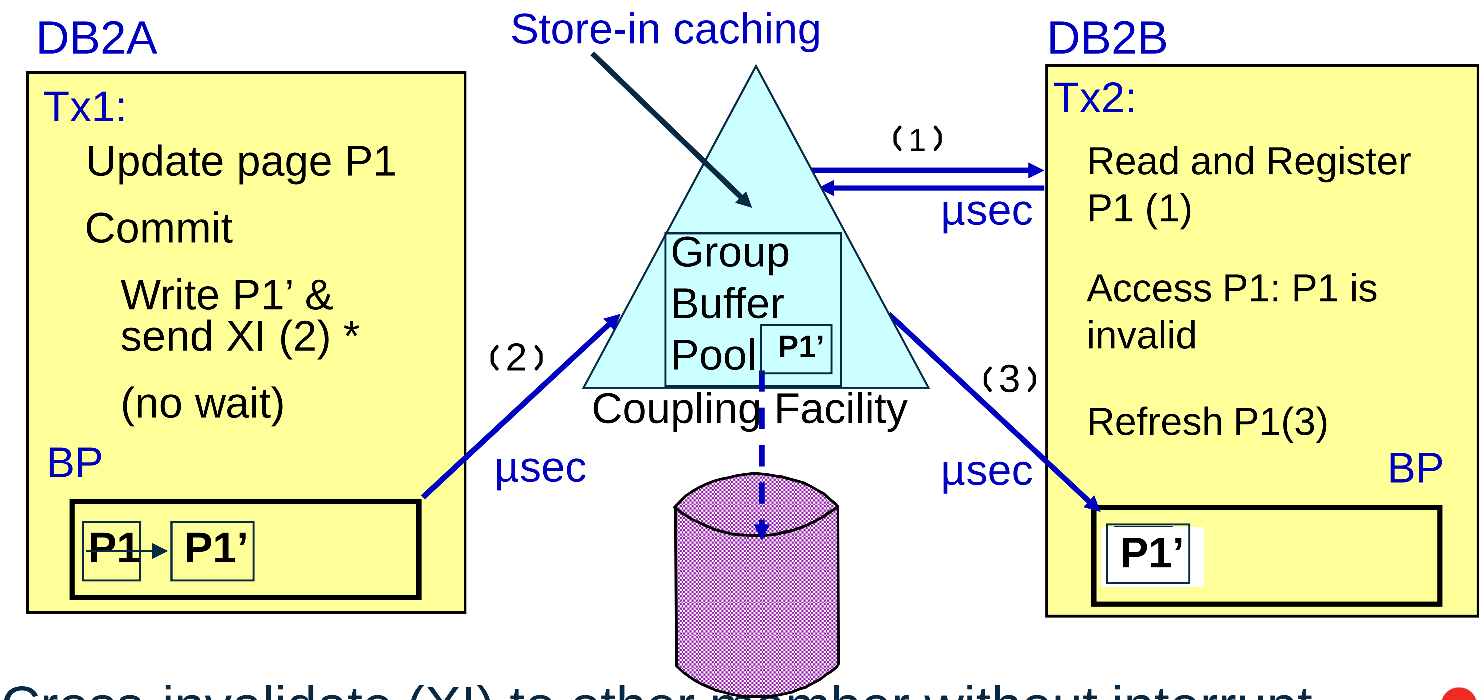
Lock Structure (LOCK1)

- Simplified view



Data Sharing: Managing changed data

- Inter-system buffer coherency control
  - Example: DB2A has write interest in the table space, and page P1 is in DB2A's buffer pool



- * Cross-invalidate (XI) to other member without interrupt

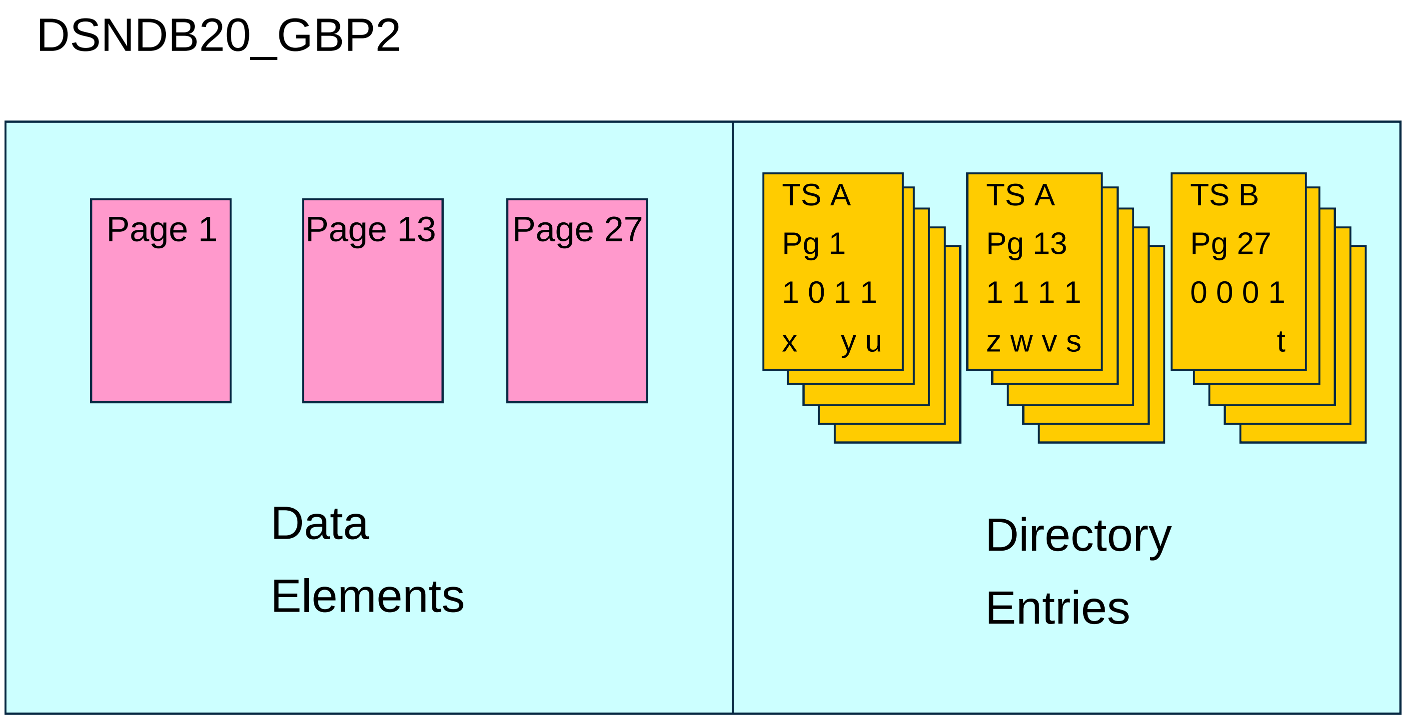
Notes: Group Buffer Pools (GBPs)

- DB2 uses GBPs to
  - Manage buffer coherency
  - Cache changed pages
    - Optionally cache read-only pages

- GBP consists of directory entries and data elements
  - Directory entries manage coherency by tracking interest in a data or index page by any DB2 member in the data sharing group
    - There is one directory entry for each page in the aggregate pool, no matter how many DB2 members have a copy of that page
  - Data elements are the cached pages that a DB2 member changed
  - In GBP duplexing, data elements exist in both the primary and secondary GBP
    - Directory entries in secondary GBP only exist for the changed pages

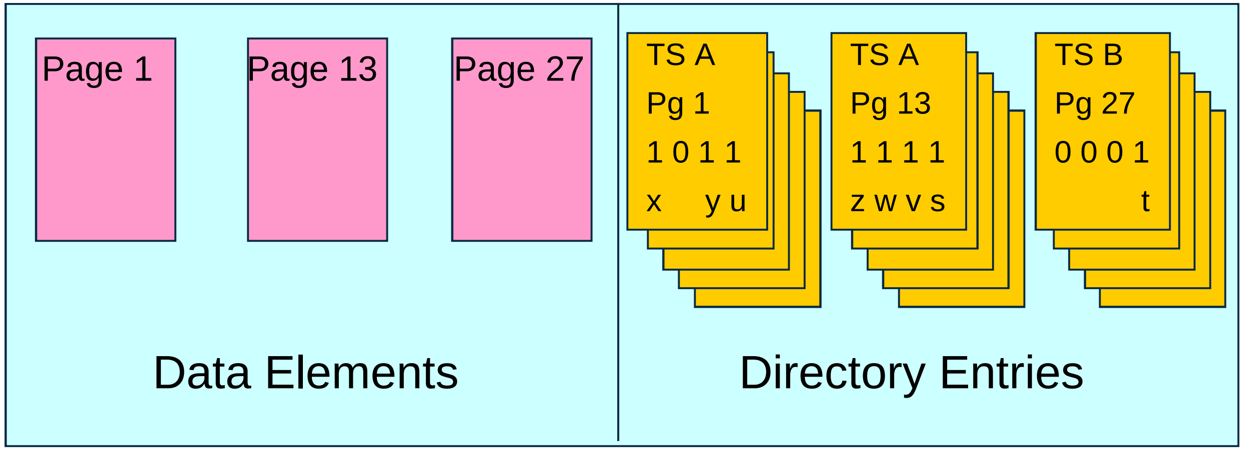
Group Buffer Pool (GBP)

- Simplified view



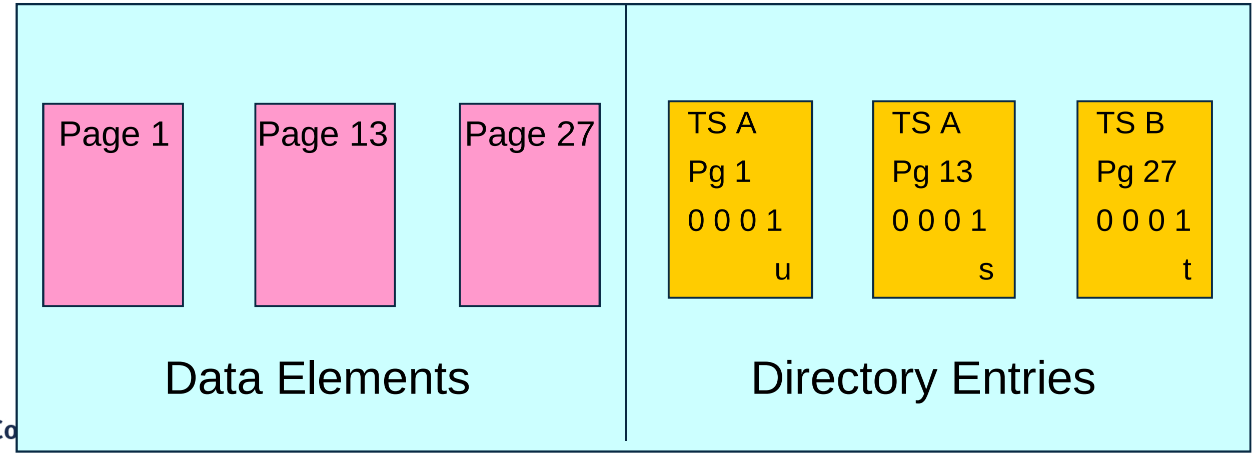
GBP Duplexing

- DSNDB20_GBP2 – Primary; “Old” on CF01



-Cache changed pages
-Register interest
-Search if local buffer miss
-Castout

- DSNDB20_GBP2 – Secondary; “New” on ICF2



-Cache changed pages

Monitoring LOCK1

- RMF CF Activity Report
  - Structure Summary

```
1                                          C O U P L I N G   F A C I L I T Y   A C T I V I T Y
                                                                                                          PAGE    1
         z/OS V1R6             SYSPLEX ****              DATE 09/16/2008            INTERVAL 015.00.489
                          CONVERTED TO z/OS V1R9 RMF     TIME 08.59.00              CYCLE 10.000 SECONDS

    ------------------------------------------------------------------------------------------------------------------------------
     COUPLING FACILITY NAME = CFP01
     TOTAL SAMPLES(AVG) =    90   (MAX) =    90   (MIN) =    89
    ------------------------------------------------------------------------------------------------------------------------------
                                                  COUPLING  FACILITY  USAGE  SUMMARY
    ------------------------------------------------------------------------------------------------------------------------------
     STRUCTURE SUMMARY
    ------------------------------------------------------------------------------------------------------------------------------

                                                      % OF              % OF   % OF    AVG     LST/DIR  DATA      LOCK     DIR REC/
            STRUCTURE                         ALLOC   CF          #     ALL    CF      REQ/    ENTRIES  ELEMENTS  ENTRIES  DIR REC
     TYPE   NAME              STATUS CHG      SIZE    STOR      REQ     REQ    UTIL    SEC     TOT/CUR  TOT/CUR   TOT/CUR  XI'S

     LOCK   DSNDB2B_LOCK1     ACTIVE           16M     0.1        0     0.0     0.0    0.00      24K         0     4194K      N/A
                                                                                                  32         0      7381      N/A
            DSNDB2P_LOCK1     ACTIVE           64M     0.5    1483K    10.5     0.0  1646.5     100K         0       17M      N/A
                                                                                                2121         0      207K      N/A
            DSNDB2Q_LOCK1     ACTIVE           16M     0.1        0     0.0     0.0    0.00      24K         0     4194K      N/A
                                                                                                 272         0       48K      N/A
            DSNDB2R_LOCK1     ACTIVE           16M     0.1        0     0.0     0.0    0.00      24K         0     4194K      N/A
                                                                                                  13         0      5717      N/A
```

Notes: Key Points – LOCK1 Structure Summary

- Size can be an issue
  - Determines the number of Lock Table entries (LTE) and space for Modify Lock List entries (RLE)
- Requests per second is important
  - "Busy" is relative; < 5K/sec is not very busy
  - Observed: 166K/sec – very busy
- LIST/DIR ENTRIES = Modify Lock List entries (RLE)
- LOCK ENTRIES = 2-byte Lock Table entries (LTE)
  - May be 4- or 8-byte entries if > 7 members in the data sharing group
  - IRLM automatically rebuilds the lock structure when the 8th member (4-byte entries) or 23rd member (8-byte entries) joins the data sharing group

Monitoring LOCK1, cont.

- RMF CF Activity Report
  - Structure Activity

```
STRUCTURE NAME = DSNDB2P_LOCK1     TYPE = LOCK   STATUS = ACTIVE
              # REQ    -------------- REQUESTS -------------   -------------- DELAYED REQUESTS ------------
              TOTAL              #     % OF  -SERV TIME(MIC)-    REASON     #   % OF  ---- AVG                  EXTERNAL REQUEST
    NAME      AVG/SEC           REQ    ALL      AVG    STD_DEV              REQ   REQ   /DEL                    CONTENTIONS

    SYSA        567K   SYNC     523K   35.3    44.6      64.3    NO SCH   316    0.1   16.8       94.5    0.0   REQ TOTAL      784K
              630.1    ASYNC     44K    3.0   150.0     325.8    PR WT      0    0.0    0.0        0.0    0.0   REQ DEFERRED  4634
                       CHNGD      0     0.0  INCLUDED IN ASYNC   PR CMP 3016    0.5  643.6       1418    3.4   -CONT         4198
                                                                                                               -FALSE CONT    742

    SYSB        916K   SYNC     853K   57.6    43.5      85.3    NO SCH    49    0.0   80.7      184.8    0.0   REQ TOTAL     1256K
                1017   ASYNC     62K    4.2   147.7     259.6    PR WT      0    0.0    0.0        0.0    0.0   REQ DEFERRED  5437
                       CHNGD      0     0.0  INCLUDED IN ASYNC   PR CMP     0    0.0    0.0        0.0    0.0   -CONT         4703
                                                                                                               -FALSE CONT    705

  ------------------------------------------------------------------------------------------------------------------------------
                1483K  SYNC    1376K   92.8    43.9      78.0    NO SCH   365    0.0   25.3                    REQ TOTAL     2040K
                1647   ASYNC    106K    7.2   148.6     288.9    PR WT      0    0.0    0.0        0.0    0.   REQ DEFERRED   10K
                       CHNGD      0     0.0                      PR CMP 3016    0.2  933.9       1597    1.9   -CONT         8901
                                                                                                               -FALSE CONT   1447
```

Complete your session evaluations online at www.SHARE.org/Seattle-Eval

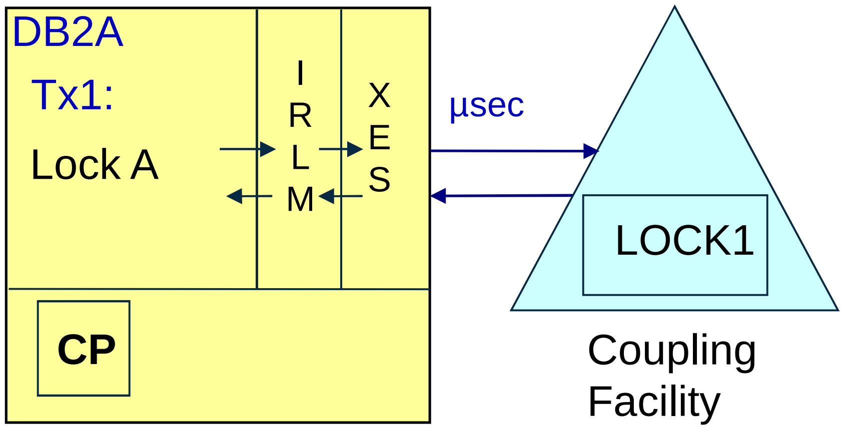
SHARE in Seattle 2015

Notes: Key Points – LOCK1 Structure Activity

- # REQ TOTAL
  - These are requests on the subchannel
    - Compare with EXTERNAL REQUEST CONTENTIONS: REQ TOTAL, which reflects API requests to XES and should be the higher number
- SERV TIME(MIC) – service time in microseconds
  - SYNC is key metric – ‘good’ number is relative to CF configuration
    - If ASYNC is non-zero it could be ‘block unlock’, or some requests were converted, either due to subchannel busy or heuristic algorithm
- CONT and FALSE CONT
  - Contention - recommend: CONT/REQ TOTAL < 2%
  - False Contention - recommend: FALSE CONT/REQ TOTAL < 1%
    - If higher, adjust size of LOCK1 to double size of Hash Table

Heuristic Algorithm and LOCK1

- Most LOCK1 requests are synchronous CF requests
  - Synchronous CF request means host CP is busy for duration of request
  - Long synchronous service times = high host CPU overhead



- XES can convert synchronous request to asynchronous
  - Heuristic algorithm based on measured lock service times
  - Host CP can now do other work during CF request
  - There is some host CP cost to setting up asynchronous request
  - Also elapsed time impact on lock requests

Monitoring GBPs

- RMF CF Activity Report
  - Structure Summary

```
1                        C O U P L I N G   F A C I L I T Y   A C T I V I T Y
                                                                                                        PAGE    1
     z/OS V1R6               SYSPLEX ****                DATE 09/16/2008            INTERVAL 015.00.489
                       CONVERTED TO z/OS V1R9 RMF        TIME 08.59.00              CYCLE 10.000 SECONDS

 ------------------------------------------------------------------------------------------------------------------------------
 COUPLING FACILITY NAME = CFP01
 TOTAL SAMPLES(AVG) =     90   (MAX) =     90   (MIN) =     89
 ------------------------------------------------------------------------------------------------------------------------------
                                          COUPLING  FACILITY  USAGE  SUMMARY
 ------------------------------------------------------------------------------------------------------------------------------
 STRUCTURE SUMMARY
 ------------------------------------------------------------------------------------------------------------------------------
```

TYPE	STRUCTURE NAME	STATUS CHG	ALLOC SIZE	% OF CF STOR	# REQ	% OF ALL REQ	% OF CF UTIL	AVG REQ/ SEC	LST/DIR ENTRIES TOT/CUR	DATA ELEMENTS TOT/CUR	LOCK ENTRIES TOT/CUR	DIR REC/ DIR REC XI'S
CACHE	DSNDB2P_GBP0	ACTIVE	34M	0.3	529	0.0	0.0	0.59	29K	5732	N/A	0
									19	10	N/A	0
	DSNDB2P_GBP1	ACTIVE	501M	4.1	18380	0.1	0.0	20.41	494K	82K	N/A	0
									5406	5406	N/A	0
	DSNDB2P_GBP16K0	ACTIVE	8M	0.1	120	0.0	0.0	0.13	1596	1270	N/A	0
									0	0	N/A	0
	DSNDB2P_GBP16K1	ACTIVE	32M	0.3	42641	0.3	0.0	47.35	50K	2876	N/A	0
									18K	1980	N/A	0
	DSNDB2P_GBP2	ACTIVE	2G	16.8	8681	0.1	0.0	9.64	1236K	412K	N/A	0
									844	844	N/A	0
	DSNDB2P_GBP3	ACTIVE	8M	0.1	94	0.0	0.0	0.10	6008	1201	N/A	0
									0	0	N/A	0
	DSNDB2P_GBP32K	ACTIVE	10M	0.1	132	0.0	0.0	0.15	840	1344	N/A	0
									2	16	N/A	0
	DSNDB2P_GBP32K1	ACTIVE	16M	0.1	120	0.0	0.0	0.13	1438	2862	N/A	0
									1	0	N/A	0
	DSNDB2P_GBP5	ACTIVE	256M	2.1	358	0.0	0.0	0.40	521K	13K	N/A	0
									7	7	N/A	0

Notes: Key Points – GBPs Structure Summary

- Size and requests per second important
- LIST/DIR ENTRIES = directory entries
- DATA ELEMENTS = data pages
  - If current directory entries = current data pages, probably secondary GBP (GBP duplexing)
    - Could also be the effect of Auto Alter
- DIR REC/DIR REC XI'S = directory reclaims / cross-invalidations (XI's) due to directory reclaims
  - Should be zero! Investigate if non-zero, especially XI's
    - If DIR REC XI'S non-zero, potential performance impact
  - CF report does not have directory reclaim details
    - Use –DIS GBPOOL GDETAIL

Monitoring GBPs, cont.

- RMF CF Activity Report
  - Structure Activity

```
STRUCTURE NAME = DSNDB2P_GBP1      TYPE = CACHE   STATUS = ACTIVE
                # REQ     -------------- REQUESTS -------------    -------------- DELAYED REQUESTS -------------
     SYSTEM     TOTAL              #      % OF  -SERV TIME(MIC)-    REASON     #    % OF   ---- AVG TIME(MIC) -----
     NAME       AVG/SEC           REQ     ALL      AVG    STD_DEV             REQ    REQ    /DEL     STD_DEV    /ALL

                          SYNC     19      0.1     12.0      6.8    NO SCH    34     0.3  853.4       837.8     2.3
                          ASYNC   12K     67.6     73.2    297.3    PR WT      0     0.0    0.0         0.0     0.0
                          CHNGD     0      0.0  INCLUDED IN ASYNC   PR CMP     0     0.0    0.0         0.0     0.0
                                                                    DUMP       0     0.0    0.0         0.0     0.0

                          SYNC     44      0.2     54.9    246.0    NO SCH     0     0.0    0.0         0.0     0.0
                          ASYNC  5887     32.0     54.7    110.9    PR WT      0     0.0    0.0         0.0     0.0
                          CHNGD     0      0.0  INCLUDED IN ASYNC   PR CMP     0     0.0    0.0         0.0     0.0
                                                                    DUMP       0     0.0    0.0         0.0     0.0

    ---------------------------------------------------------------------------------------------------------------------------
     TOTAL      18380     SYNC     63      0.3     41.9    205.9    NO SCH    34     0.2  853.4       837.8      .6    -- DATA ACCESS ---
                20.42     ASYNC   18K      100     67.3    253.0    PR WT      0     0.0                                READS            0
                          CHNGD     0      0.0                      PR CMP     0     0.0                                WRITES       29183
                                                                    DUMP       0     0.0                                CASTOUTS         0
                                                                                                                        XI'S         11542
```

Secondary GBP

Monitoring GBPs, cont.

- RMF CF Activity Report
  - Structure Activity

```
STRUCTURE NAME = DSNDB2P_GBP1      TYPE = CACHE  STATUS = ACTIVE
               # REQ    -------------- REQUESTS -------------   -------------- DELAYED REQUESTS -------------
     SYSTEM    TOTAL            #     % OF   -SERV TIME(MIC)-    REASON    #    % OF   ---- AVG TIME(MIC) -----
     NAME      AVG/SEC         REQ    ALL     AVG    STD_DEV              REQ    REQ    /DEL     STD_DEV   /ALL

                        SYNC    183K   30.5    18.0     47.5     NO SCH   819    0.4    1341      2830     5.6
                        ASYNC    14K    2.3   152.7    369.3     PR WT      0    0.0     0.0       0.0     0.0
                        CHNGD   818     0.1   INCLUDED IN ASYNC  PR CMP     0    0.0     0.0       0.0     0.0
                                                                 DUMP       0    0.0     0.0       0.0     0.0

                        SYNC    363K   60.5    15.9     74.5     NO SCH   124    0.0   770.1     912.6     0.2
                        ASYNC    40K    6.6   143.7    354.7     PR WT      0    0.0     0.0       0.0     0.0
                        CHNGD   124     0.0   INCLUDED IN ASYNC  PR CMP     0    0.0     0.0       0.0     0.0
                                                                 DUMP       0    0.0     0.0       0.0     0.0

    ----------------------------------------------------------------------------------------------------------------------------------
     TOTAL        599K  SYNC    545K   90.9    16.6     66.7     NO SCH   943    0.2    1266      2665            -- DATA ACCESS ---
               665.9    ASYNC    53K    8.9   146.0    358.5     PR WT      0    0.0                              READS        16712
                        CHNGD   942     0.2                      PR CMP     0    0.0                              WRITES       29186
                                                                 DUMP       0    0.0                              CASTOUTS     12172
                                                                                                                  XI'S         20401
```

Primary GBP

Notes: Key Points – GBPs Structure Activity

- SERV TIME(MIC)
  - SYNC is key metric – ‘good’ number is relative to CF configuration
    - If REQ/SEC < 100, variations in service time probably not significant
  - ASYNC requests are expected, especially in secondary GBPs
- XI’s in lower right are not necessarily reclaims
  - Most likely business as usual

Monitoring GBPs: -DIS GBPOOL

- –DIS GBPOOL(*) TYPE(GCONN) GDETAIL(*)
  - Contains status and definition information as well as statistics
  - Reports statistics since GBP allocation
- –DIS GBPOOL(*) TYPE(GCONN) GDETAIL(INTERVAL)
  - To monitor an interval, execute this command before and after the desired interval.
  - Output messages from second command will show GBP statistics for the interval
- Typical problems due to incorrectly defined GBP
  - Directory entry reclaims
  - XIs due to directory entry reclaims
  - Writes failed due to lack of storage

–DIS GBPOOL(*) TYPE(GCONN) GDETAIL(*)

```
07.57.32 STC34822  DSNB784I  -DB2A GROUP DETAIL STATISTICS  362
   362                          READS
   362                            DATA RETURNED                           = 1842830
07.57.32 STC34822  DSNB785I  -DB2A     DATA NOT RETURNED  363
   363                              DIRECTORY ENTRY EXISTED               = 1490516
   363                              DIRECTORY ENTRY CREATED               = 9995482
   363                              DIRECTORY ENTRY NOT CREATED           = 26712646, 0
07.57.32 STC34822  DSNB786I  -DB2A   WRITES  364
   364                            CHANGED PAGES                           = 50473770
   364                            CLEAN PAGES                             = 3408467
   364                            FAILED DUE TO LACK OF STC               = 48
   364                          CHANGED PAGES SNAPSHOT VALUE              = 5568
07.57.32 STC34822  DSNB787I  -DB2A   RECLAIMS  365
   365                            FOR DIRECTORY ENTRIES                   = 80726
   365                            FOR DATA ENTRIES                        = 28878053
   365                          CASTOUTS                                  = 28679918
07.57.32 STC34822  DSNB788I  -DB2A   CROSS INVALIDATIONS  366
   366                            DUE TO DIRECTORY RECLAIMS               = 56680
   366                            DUE TO WRITES                           = 2666240
   366                            EXPLICIT                                = 0
07.57.32 STC34822  DSNB762I  -DB2A DUPLEXING STATISTICS FOR GBP11-SEC  367
   367                          WRITES
   367                            CHANGED PAGES                           = 50072797
   367                            FAILED DUE TO LACK OF STORAGE           = 48
   367                          CHANGED PAGES SNAPSHOT VALUE              = 5568
```

Notes: Sizing CF Structures

- http://www.ibm.com/systems/support/z/cfsizer
  - CF Structure Sizer Tool
- *DB2 Version 9.1 for z/OS Installation Guide*, GC18-9846
- *DB2 10 for z/OS Installation and Migration Guide*, GC19-2974
- *DB2 11 for z/OS Installation and Migration Guide*, GC19-4056
  - Knowledge Center: cf sizing for DB2 10 or 11
- Rule of thumb for GBPs
  - Start with CFSizer INITSIZE
  - Round up
  - Make that result INITSIZE; make SIZE at least 20% larger than INITSIZE
    - But no larger than 2 times INITSIZE
  - Use Auto Alter

Auto Alter – What is it?

- Autonomic effort by XES to avoid filling up any kind of structure. For GBPs:
  - If all data elements (pages) are changed, writes cannot occur
  - If all directory entries are marked changed, new pages cannot be registered
- Auto Alter has algorithms that
  - can increase or decrease number of entries and/or elements to avoid structure full conditions
  - can increase or decrease the size of the structure
- Can alter, dynamically, the precise directory to data ratio for GBPs
- Design point is for gradual growth, not spikes

Auto Alter and DB2

- DB2 Structures support Auto Alter
- LOCK1 – effective on Modify Lock List entries (RLEs)
  - Lock Table entries (LTE) cannot be changed without a rebuild
- SCA – can be increased
- Main value is for Group Buffer Pools (GBPs). Why?
  - People tend not to tune GBPs
    - Organizational division of labor
      - DB2 DBAs responsible for local BPs – may forget about GBPs
      - z/OS responsible for GBPs – and they own the CFRM Policy
  - DB2 needs ?? more directory entries than data page elements
  - Each –ALTER to change directory entries means manual GBP rebuild
- Works for duplexed GBPs

Auto Alter – When not to use it

- CF available storage is <10%
  - Auto Alter reduces the size of "alterable" structures below INITSIZE (to MINSIZE), attempting to get 10% available storage in the CF
- Not enough storage for size of structure, especially in Test environments
  - XES reaches SIZE quickly
  - Reclaim avoidance results in constant XES attempts to increase directory entries and reduce data pages
    - Reclaim avoidance alone does not allow structure size increase
  - Attempts usually fruitless - produce alarming console messages
  - Hint: test one structure, correctly sized, instead of all

Workload Growth

- Increased transaction, batch and/or query volumes
- New applications
- Mergers
- New business opportunities
- Regulatory compliance
- Technology advances

Workload Changes and LOCK1

- Increased lock requests may lead to
  - Higher CF CPU busy
  - Higher synchronous service time, and host CPU cost
  - Higher transaction or query elapsed time, higher job run time
- New applications may not follow standards
  - Less lock avoidance by new applications may mean more locking for existing applications
  - Long commit scopes hold Modify Lock List entries (RLEs) longer
  - Row level locking increases demand for RLEs
- False contention could increase, requiring more Lock Table Entries (LTEs)

Workload Changes and LOCK1

- Possible solutions
  - Increase CF CPU capacity
    - More CPs and/or faster CPs
  - Increase the number of RLEs
    - SETXCF START,ALTER,strnm=&,SIZE=&
      to increase the size of LOCK1
      - Assumes allocation < SIZE in CFRM policy
      - Else change CFRM policy definition, rebuild structure
  - Increase the number of LTEs
    - Requires a structure rebuild with larger allocation
      - CFRM policy change required if allocation already = SIZE
  - CF storage increase unlikely to be necessary in a production environment

LOCK1 Example

- RMF CF Activity Report
  - Structure Activity

```
STRUCTURE NAME = DSN****_LOCK1     TYPE = LOCK   STATUS = ACTIVE
              # REQ    -------------- REQUESTS -------------    -------------- DELAYED REQUESTS -------------
    SYSTEM    TOTAL              #     % OF  -SERV TIME(MIC)-    REASON     #    % OF  ---- AVG TIME(MIC) -----   EXTERNAL REQUEST
    NAME      AVG/SEC           REQ     ALL      AVG    STD_DEV              REQ    REQ    /DEL     STD_DEV   /ALL   CONTENTIONS

    S***        232M   SYNC     232M   38.8     11.0      5.8    NO SCH   88K    0.0   30.8       236.2    0.0   REQ TOTAL       133M
              64403    ASYNC    750     0.0     94.2    413.5    PR WT     0     0.0    0.0         0.0    0.0   REQ DEFERRED    239K
                       CHNGD      0     0.0   I       ED IN ASYNC  PR CMP    0     0.0    0.0         0.0    0.0   -CONT           239K
                                                                                                           -FALSE CONT     162K

    S***        187M   SYNC     187M   31.2               5.5    NO SCH   14K    0.0   15.0        67.2    0.0   REQ TOTAL        99M
              51870    ASYNC      0     0.0               0.0    PR WT           0.0    0.0         0.0    0.0   REQ DEFERRED    218K
                       CHNGD      0     0.0   I       ED IN ASYNC  PR CMP          0.0    0.0         0.0    0.0   -CONT           218K
                                                                                                           -FALSE CONT     111K

    S***        179M   SYNC     179M   30.0               5.9    NO SCH 5        0.0   19.7        37.5    0.0   REQ TOTAL       106M
              49841    ASYNC   1500     0.0              87.7    PR WT           0.0    0.0         0.0    0.0   REQ DEFERRED    333K
                       CHNGD      0     0.0              IN ASYNC  PR CMP          0.0    0.0         0.0    0.0   -CONT           333K
                                                                                                           -FALSE CONT     242K

              ----------------------------------------------------------------------------------------------------------------------
                598M   SYNC     598M    100     10.7      5.7    NO SCH  108K    0.0   28.2       215.6                          337M
              166.1K   ASYNC   2250     0.0     80.2    249.4    PR WT     0     0.0    0.0         0.0    0.0   REQ DE  ERRED   790K
                       CHNGD      0     0.0                      PR CMP    0     0.0    0.0         0.0    0.0   -CONT           790K
                                                                                                           -FALSE CONT     515K
```

Workload Changes and GBPs

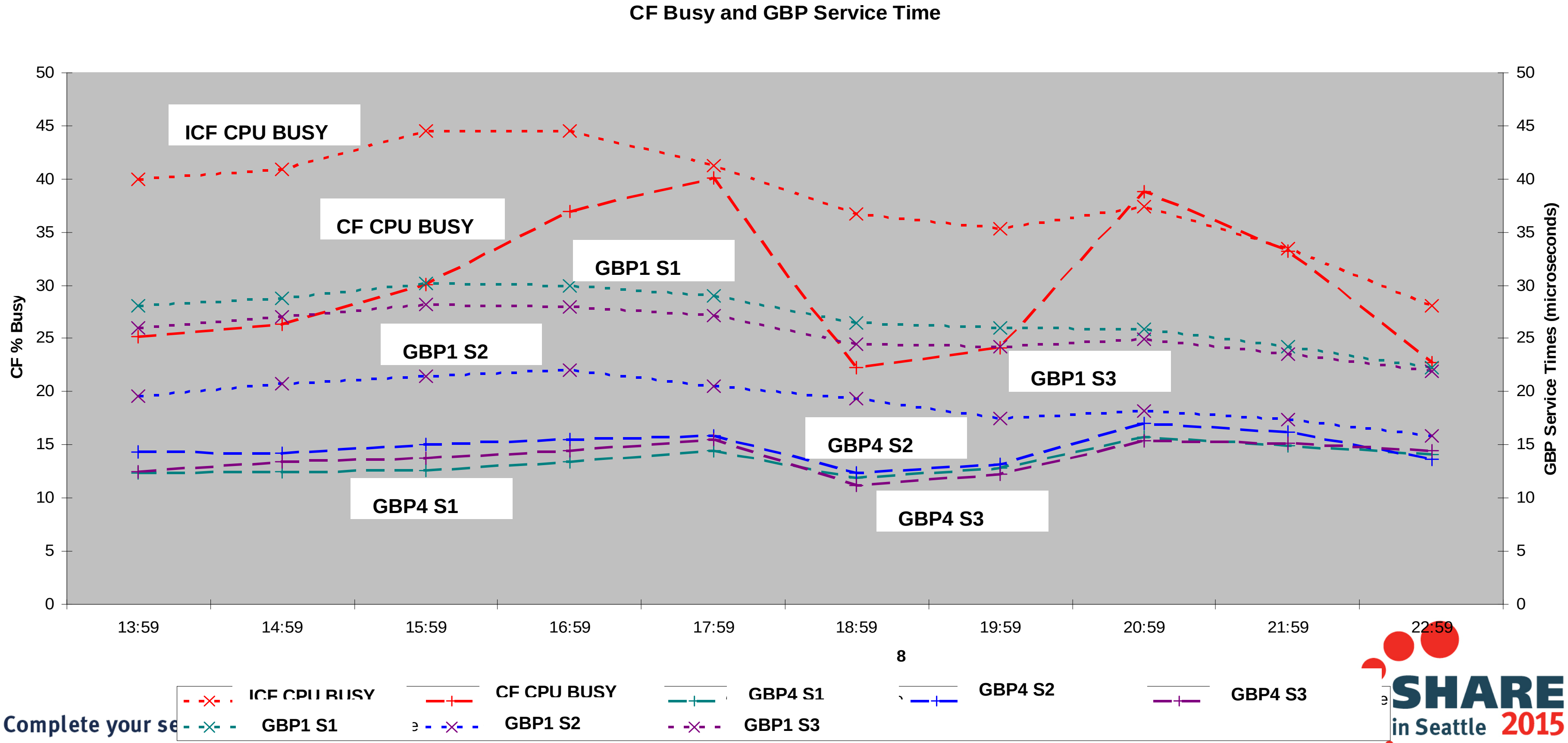
- Increased GBP requests may lead to
  – Higher CF CPU busy
  – Higher synchronous service time, and host CPU cost
  – Higher transaction or query elapsed time, higher job run time
- New applications may
  – Change access patterns of existing tables or indexes
  – Add tables and indexes to existing buffer pools
- Local buffer pool allocations may increase
  – GBPs might be forgotten

Workload Changes and GBPs

- Possible solutions
  - Increase CF CPU capacity
    - More CPs and/or faster CPs
  - Increase the size of the GBPs
  - Tune local buffer pool thresholds and GBP thresholds
  - CF storage increase may be necessary

GBPs and Impact of CF Busy

- ICF has two CPs
- CF (external) has three CPs



GBPs and Impact of CF Busy, cont.

- GBP1 on ICF was very busy over a 10 hour interval
  - S1: 178 M synchronous requests
  - S3: 144.5 M synchronous requests
  - If 10 µsec saved from each request, over 300 CPU seconds per hour of ‘host effect’ could be saved from GBP1 alone
- How could 10 µsec be saved?
  - Increase number of CPs on ICF to reduce CF busy and improve service time
  - Upgrade CEC with ICF to reduce CF service times

Workload Changes and SCA

- New applications or new workloads may add tables and indexes
- New clients may require additional databases
- Auto Alter may be able to handle most of the increase
- Use CF Structure Sizer Tool to validate CFRM policy definition

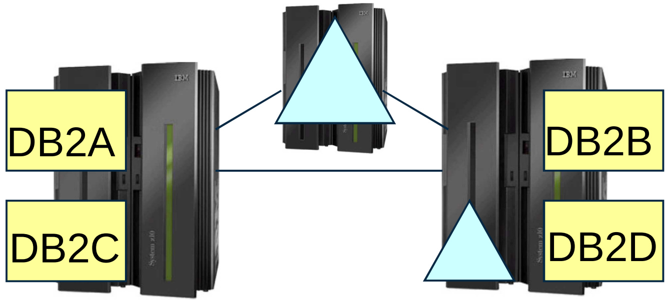
When New Members Join the Data Sharing Group

- GBP – review sizes
  - Increased demand for directory entries and data elements
  - Auto Alter alone may not be sufficient to handle multiple new members
- LOCK1
  - 4-byte LTEs required when 8th member joins the group
    - Automatic rebuild will normally result in half as many LTEs, so false contention will increase
    - Prepare for larger LTEs before adding 8th member
  - 8-byte LTEs required when 23rd member joins the group
    - Automatic rebuild has same considerations

Configuration Changes

- CF Considerations
  - Balanced performance: CF technology = CEC technology
  - Unbalanced configuration examples:
    - zEC12 CF and z196 CEC – good for the CEC
    - z196 CF and zEC12 CEC – more Host Effect cost to CEC
    - z10 CF and zEC12 CEC – ‘heuristic algorithm’ likely to convert many synchronous requests to asynchronous
      - Algorithm represents tradeoff of host effect versus cost of conversion
      - Elapsed times, contention, and time outs likely to increase
  - Increase in distance between CF and CEC can have similar effect
    - Asynchronous conversion frequently observed as distance between CEC and CF increases

Creative Use of CF Storage



- As more DB2 members join the group
  - Consider GBPCACHE ALL
    - Each page is read into GBP on first access
    - Only one member incurs I/O cost for each page
    - Local buffers can be smaller – GBP acts as very fast cache
- If large objects with very random access and minimal page re-reference
  - Consider GBPCACHE NONE
    - Saves GBP access on local page miss
    - Enforces 'store through cache': synchronous writes to disk at commit
      - Modern cache controllers minimize negative impact

DB2 10 for z/OS

Data Sharing Highlights

DB2 10 for z/OS and Data Sharing

- Deleting member of data sharing group
  - Offline utility
- Deleting structures during group restart
  - DEL_CFSTRUCTS_ON_RESTART - DSNZPARM for DR
- Sub-group attach
- DDF Restart Light – handle indoubts
- MEMBER CLUSTER for UTS
- -MODIFY DDF – online changes for LOCATION ALIAS

- LRSN spin avoidance
- IFCID 359 – index split
- GBP DELETE_NAME processing
- BP scan avoidance

DB2 11 for z/OS

Data Sharing Highlights

DB2 11 for z/OS Data Sharing Enhancements

- Castout enhancements: New CLASST setting – similar to VDWQT
- RESTART LIGHT Enhancements
- Buffer pool enhancements

- GBP Write-around
  - If GBP / CF busy, write new pages to directly to disk
  - Reduce impact of flood of new pages on rest of GBP
- Automatic LPL or GRECP recovery
- CF DELETE_NAME
- Locking enhancements
- Index split performance
- LRSN spin avoidance – extended LRSN

Additional Resources

- ***Data Sharing: Planning and Administration***
  - DB2 9 for z/OS: SC18-9845
  - DB2 10 for z/OS: SC19-2973
  - DB2 11 for z/OS: SC19-4055
  - KC db2 data sharing planning

Questions?

mrader@us.ibm.com

Thank you!