

# Asynchronous DB2 data sharing locks not counted

Judy Ruby-Brown - ATS Dallas Systems Center

## Introduction

In all current versions of DB2 for z/OS (V7/V8/V9), performance monitors may calculate data sharing global lock contention rates incorrectly if those locks are converted from synchronous requests and are granted asynchronously. Customers likely to encounter this situation have any of these conditions:

- A coupling facility (CF) is separated from the requesting z/OS image by distance
- Technology mismatch between requesting z/OS image and the CF with the DB2 lock structure (example: System z9 image with z990//z900 CF)
- Matched technology between z/OS requester and CF, but slow links (or overcommitted links)

The RMF Coupling Facility Activity Report provides accurate values.

## How global locks become asynchronous

In any of the above conditions the Cross System Extended Services (XES) component of z/OS may convert synchronous requests to asynchronous using a capability of z/OS 1.2 called the "heuristic algorithm". The goal is to spare the faster z/OS host from spinning while it is waiting for the synchronous request to complete. The processor is made available for other workload when the lock (or other request) is performed asynchronously. Even though the request is processed asynchronously, there is enough CPU overhead in converting the request that it is not done for all conditions. The thresholds for conversion have been revised upwards with z/OS APAR OA21635, in order to more accurately reflect the cost tradeoff between the CPU cost of processing a CF request synchronously vs. asynchronously.

You can find details on the heuristic algorithm in

- *z/OS MVS Setting Up a Sysplex*, SA22-7625-14, Chapter 6, "Tuning a sysplex", under "Tuning coupling facilities")
- <http://www.ibm.com/support/techdocs/atmastr.nsf/WebIndex/FLASH10159> for Flash10159 "z/OS Performance: New, heuristic algorithm for managing CF request conversion"

## DB2's Global Lock Contention calculation

Tivoli Omegamon XE for DB2 PM for z/OS (hereafter OMPE) provides calculations of the global contention rate. The current data sharing locking statistics look like this snippet:

DATA SHARING LOCKING QUANTITY

-----  
GLOBAL CONTENTION RATE (%) 0.16  
P/L-LOCKS XES RATE (%) 66.36

LOCK REQUESTS (P-LOCKS) 408.3K  
UNLOCK REQUESTS (P-LOCKS) 408.2K  
CHANGE REQUESTS (P-LOCKS) 2.00

SYNCH.XES - LOCK REQUESTS 2654.8K                    A            QTGSLSLM

```

SYNCH.XES - CHANGE REQUESTS 955.3K      B      QTGSCSLM
SYNCH.XES - UNLOCK REQUESTS 2381.7K     C      QTGSUSLM
ASYNCH.XES - RESOURCES 0.00              D      QTGSKIDS

SUSPENDS - IRLM GLOBAL CONT 4954.00     E      QTGSIGLO
SUSPENDS - XES GLOBAL CONT. 23.00        F      QTGSSGLO
SUSPENDS - FALSE CONTENTION 4338.00     G      QTGSFLMG
INCOMPATIBLE RETAINED LOCK 0.00

```

See *Tivoli Omegamon XE for DB2 PM on z/OS Report Reference*, SC18-9984-01, for detail on the Data Sharing Locking section of the Statistics Long report/trace.

Beside some of the fields, there is a letter and a field name, neither of which shows in the report. The current formula used to determine GLOBAL CONTENTION RATE is:  $(E+F+G)/(A + B + C + E + F + G)$ . That formula is described in *Data Sharing: Planning and Administration* for all current releases of DB2, Chapter 6, under the sections, "Monitoring DB2 locking" and "using the DB2 statistics trace". Note that the letter designations and meanings can be different, but the same fields are used to calculation the global lock contention. The field name is from IFCID0002, found in SDSNIVPD(DSNWMSG) under section QTGS.

The global contention is calculated only for locks that are granted synchronously. The field title, SYNCH.XES - LOCK REQUESTS, does not mean the number of Sync. XES requests, but the number that were granted synchronously.

Heuristic conversion is one way that lock requests can be processed asynchronously, and are, naturally, not included in the "synch." field. There is no field that includes locks processed asynchronously as a result of heuristic conversion. If 95% of locks from a member are converted from synchronous to asynchronous, they are not included in the calculation and the resulting statistics are misleading when a significant amount of conversion occurs.

### The calculation should be corrected in future releases of DB2 and OMPE.

There will be a new field for asynchronous locks, which would include the heuristics. The current field, ASYNCH.XES - RESOURCES, only counts child lock propagations.

### Are you affected?

You can tell if the DB2 calculations are inaccurate by looking at the RMF Coupling Facility Structure Activity section of the report. A brief snippet shows significant ASYNC counts,

```

STRUCTURE NAME = DB2M_LOCK1      TYPE = LOCK
# REQ ----- REQUESTS
SYSTEM   TOTAL      #      % OF
NAME     AVG/SEC     REQ   ALL
-----
ZOSA     2650K   SYNC  1211K  13.6
        2944   ASYNC 1439K  16.2
        CHNGD    0
. . . etc. . . other z/OS images
-----
TOTAL    8887K   SYNC  6095K  68.6   65.0
        9874   ASYNC 2792K  31.4   121.5
        CHNGD    0     0.0

```

This shows large ASYNC counts (in yellow) for the one image whose lock totals are shown (for brevity, others were omitted) for the total for the entire Lock Structure:

In this case, only the RMF CF Activity Report should be used for global and false contentions.

For a peak locking interval, create a spreadsheet similar to this:

RMF CF Activity Reports		29-May-07	14:45	15:00			Lock Structure				
DSNDB0G_LOCK1	Req Total (all KB)	Req Defer (actual)	Contention	False Cont	Cont %	False Cont %	Sync Svc Time	Aync Svc Time	No. Aync Requests	% Requests converted to Aync	
XXXA	2481	120000	46000	1697	4.72%	0.07%	51.4	101	442000	27.33%	
XXXB	6490	52000	46000	4400	0.78%	0.07%	57	189	248000	5.73%	
XXXC	2159	53000	35000	1505	2.40%	0.07%	55	118	1483000	95.92%	
XXXD	2596	37000	36000	1489	1.39%	0.06%	88	255	61000	4.29%	
Total	14000	262000	163000	9092	1.83%	0.06%	62.3	126	2233000	25.06%	
9902/sec											

These results were obtained by looking down the right hand side of the Activity report under the structure named DSNDB0G\_LOCK1 and copying the columns approximately (rounding to even thousands etc.). The purpose was to calculate the contention and false contention, always of interest to DB2 staff. For the purpose of this discussion, the critical column is the “% Requests converted to Aync”. Note that XXXC has over 95% of its locks converted, while XXXA has 27% sync-async conversions. For those members, the OMPE Data sharing locking results will not be accurate, and attempts to reduce lock contention will not be properly reflected.

Several more calculations should be performed for validation. While the above measurement occurred during the “online” day, it is a good idea to check a peak batch time, as is shown in the following table:

		31-May-07	100	115							
DSNDB0G_LOCK1	Req Total (all KB)	Req Defer (actual)	Contention	False Cont	Cont %	False Cont %	Sync Svc Time	Aync Svc Time	No. Aync Requests	% Requests converted to Aync	
XXXA	801	149000	148000	44000	18.17%	5.36%	50	119	64000	8.07%	
XXXB	1386	115000	115000	7668	8.10%	0.54%	53	799	7977	0.58%	
XXXC	740	181000	145000	30000	23.89%	3.96%	49	97	739000	94.26%	
XXXD	16	1842	1842	995	11.24%	6.07%	77	287	558	4.83%	
Total	2943	447000	410000	83000	14.83%	2.75%	52	106	811000	27.41%	
3473/sec											

It can be seen that XXXC still has almost 95% of locks sync-async, but XXXA does not. It is possible that one of the other members might have significant percentages converted to async, but that did not occur.

## How can you calculate Lock Contention now?

If you do have a different monitor or other versions of the named products, the RMF Coupling Facility Activity report has capabilities that you can use:

- Accurate Global Lock Contention
- Accurate False Contention
- Accurate XES plus IRLM contention

In simple terms, IRLM contention is experienced when more than one member of the group wants the same page/row lock. RMF does not distinguish between XES contention and IRLM contention, and you always receive the total.

If you are in DB2 V8 (NFM), you can use the new locking protocol (enabled by a group stop and start in NFM). You should then experience greatly reduced XES lock contention (and likely significantly reduced false contention) that results in reduced global lock contention compared to prior DB2 releases. You can then deduce the RMF Contention is due to IRLM contention.

### **Why do I care?**

As it was stated before, the conversion process in XES is not free. There is increased CPU in order to change the lock to asynchronous, and for the lock structure, it is borne by IRLM and the XCFAS address spaces.

The experience of one customer may illustrate. The customer was in DB2 V8 in NFM, so should be receiving excellent contention relief. Unfortunately, the IRLM CPU was increasing greatly. Upon further analysis, it was determined that over a 3 week period, the z/OS image with a DB2 that was local to the coupling facility had been moved temporarily to another location 5KM away. It was during that time the IRLM CPU went as high as 8 times its former value. When the z/OS image was returned to its original site, local to the CF, the IRLM CPU returned to normal. What was occurring at the second site was that due to distance, almost all locks were converted to asynchronous. When the image was returned to "home", most of the locks were granted synchronously, reducing IRLM CPU.

Note that IRLM CPU time is normally small.