Rule DB2-216: Page fault rate was high for individual buffer pool

Finding: The page fault rate for read and write I/O was high for an individual buffer pool.

Impact: This finding can have a MEDIUM IMPACT, or HIGH IMPACT on the performance of the DB2 subsystem and of the MVS system.

Discussion: DB2 buffer pools are allocated and managed by DB2. Pages in the buffer pools normally reside in real (central) storage. The buffer pool management algorithms that DB2 uses are very efficient in managing large virtual buffer pools. Locating pages in large virtual buffer pools does not use any more of the processor's resources than locating pages in smaller pools. Additionally, the parameters that guide DB2's algorithms provide a significant amount of user control over how the buffer pools are managed.

From an internal DB2 view, the buffer pools and the pages in the buffer pools are managed by DB2's algorithms. However, DB2 operates in a larger MVS context. MVS might need pages of central storage because of real storage demands from other applications. Consequently, MVS might *steal* the buffer pool pages[1] to satisfy the overall real memory requirements.

If buffer pool pages are stolen from central storage, MVS will send the pages to expanded storage[2], or send the pages to auxiliary storage on DASD. Subsequent access to these pages results in a *page fault* and MVS must initiate action to bring the buffer pool pages back into central storage. There is little delay if the page fault is resolved from expanded storage, as these page faults are resolved in microseconds. DB2 performance can be **seriously** degraded if DB2 buffer pool pages are sent to DASD, as these delays are in milliseconds.

Additionally, a high page fault rate can cause performance problems for the entire MVS system. In general, any DB2 buffer pool paging indicates that there is a problem with DB2's use of buffer pools, a problem with the amount of central storage available, or (before z/Architecture) a problem with the amount of expanded storage available.

[1] Please see the referenced papers at www.cpexpert.com

[2] Expanded storage is not available with z/OS running on z/Architecture systems, so the pages are sent to auxiliary storage on DASD.

©Copyright 1999, Computer Management Sciences, Inc.

To help prevent paging problems (both for DB2 and for the entire system), DB2 limits the total amount of storage that is allocated for virtual buffer pools to approximately twice the amount of real storage.

CPExpert sums the QBSTRPI (the number of page-ins required for read I/O) and QBSTWPI (the number of page-ins required for write I/O) in DB2STATB. The sum is divided by the DB2 statistics interval to yield a page-in rate per second. CPExpert then compares this page-in rate with the **PAGERATE** guidance variable in USOURCE(DB2GUIDE). CPExpert produces Rule DB2-216 when the number of page-ins for read I/O and write I/O exceeds the value specified by the **PAGERATE** guidance variable.

Prior to DB2 UDB for z/OS Version 8, the default value for the **PAGERATE** guidance variable is 25, indicating that CPExpert should produce Rule DB2-216 whenever page faults for read I/O and write I/O for a particular buffer pool exceed 25 page-ins per second. This default value was selected because most paging would be directed to expanded storage rather than to DASD.

With DB2 UDB for z/OS Version 8, the default value for the **PAGERATE** guidance variable is 0, because page faults are resolved from DASD. IBM strongly recommends that page faults not be resolved from DASD, since DB2 performance can be seriously degraded if page faults must be resolved from auxiliary storage.

Please note that an acceptable page fault rate generally is dependent on the overall workload executing on the system and on the importance of the DB2 work.

The following example illustrates the output from Rule DB2-216:

```
RULE DB2-216: PAGE FAULT RATE WAS HIGH FOR INDIVIDUAL BUFFER POOL

   Buffer Pool 1: The page fault rate for read and write I/O was larger
   than 25 pages per second. This situation occurred for Buffer Pool
   during the intervals shown below:

                             BUFFERS     PAGE-IN FOR   PAGE-IN FOR  PAGE
   MEASUREMENT INTERVAL     ALLOCATED     READ I/O     WRITE I/O   RATE
    8:46- 9:16, 16SEP1999    25,000        39,315          69      43.8
   10:45-11:15, 16SEP1999    25,000        25,707         226      28.8
   11:15-11:45, 16SEP1999    25,000        36,904         195      41.2
   11:45-12:15, 16SEP1999    25,000        30,892         563      35.0
   12:45-13:15, 16SEP1999    25,000        23,890         170      26.7
```

Suggestion: The basic problem is that the amount of central (real) storage is insufficient for the demands of the workload (demands from DB2 and demands from

other work executing concurrently with DB2). As a result, MVS is stealing pages from DB2 buffer pools.

If Rule DB2-216 is produced more than occasionally, you should consider the following alternatives:

- Reduce the size of buffer pools so that they will not require as much central (real) storage. This will cause DB2 to manage the pages in the buffer pool, rather than allowing the MVS page fault process to manage the pages. DB2 normally will more efficiently manage pages in the buffer pool than the MVS page fault process.

  IBM strongly recommends that you set the total buffer pool size to less than the real storage that is available to DB2.

- If you are operating in MVS Compatibility Mode, verify whether storage isolation is specified for any performance group periods **other** than DB2. This specification will be reflected in the IEAIPSxx member of SYS1.PARMLIB (see the PWSS keyword for other performance group periods).

  If storage isolation is specified for any performance periods **other** than DB2, you should either remove the storage isolation specification of the other performance group periods, or specify storage isolation for the DB2 address spaces and the IRLM. You can use the PWSS keyword to specify storage isolation for the DB2 performance group and the IRLM.

- If you are operating in MVS Goal Mode, verify that the Goal Importance for DB2 (and that of allied agents) is sufficiently high (compared with other service class periods) that the Workload Manager will adequately manage processor (central and expanded) storage for DB2.

- If you are operating in MVS Goal Mode with OS/390 Version 2 Release 10 or later, verify that Protective Processor Protection has not been specified for other service classes operating on the system.

- If you are operating in MVS Goal Mode with OS/390 Version 2 Release 10 or later, consider specifying Protective Processor Protection for the service class assigned to DB2.

- Assess other applications running concurrently with DB2 to determine whether their real storage requirements can be reduced. Reducing the real storage requirements of these applications could make more real storage available to DB2.

©Copyright 1999, Computer Management Sciences, Inc.

- Determine whether other workloads running on the system can be rescheduled to periods when DB2 is not running (or not running with a heavy application volume), or consider moving the other workloads to a system in which DB2 is not running.

- You can alter CPExpert's analysis by modifying the **BUFFPGRT** guidance variable in USOURCE(DB2GUIDE).

The following additional alternatives do not apply with DB2 UDB for z/OS Version 8.

- If you have the necessary hardware, consider using hiperpools so that movement of the pages in the buffer pool between central storage and expanded storage will be managed by DB2 rather than by MVS. DB2 normally manages pages in virtual pools and hiperpools better than the normal page fault resolution process in MVS.

  If you are already using a hiperpool for this buffer pool, consider increasing the size of the hiperpool[3].

- Review the Expanded Storage Control Table (ESCT) parameters in the IEAOPTxx member of SYS1.PARMLIB. These parameters guide MVS in sending stolen pages to expanded storage or auxiliary storage. With proper settings, you might be able to guide the system in sending stolen pages to expanded storage rather than to auxiliary storage. Page fault resolution from expanded storage is significantly faster than page fault resolution from auxiliary storage.

- Consider using the CASTOUT=NO option for the buffer pool if you have established hiperpools. This alternative should be considered **only** if the DB2 system is extremely important from a management objective view. Specifying CASTOUT=NO should not **normally** be specified.

Reference: DB2 for OS/390 Version 3: Administration Guide
Section 7.5.1.1 (Buffer Pools and Hiperpools)

DB2 for OS/390 Version 4: Administration Guide
Section 5.4.1.1 (Buffer Pools and Hiperpools)

DB2 for OS/390 Version 5: Administration Guide
Section 5.4.1.1 (Buffer Pools and Hiperpools)

[3] However, depending on the workload executing on the system and the amount of expanded storage on the system, CPExpert could begin producing Rule DB2-206 or Rule DB2-207 if the hiperpool is increased. These rules would indicated that DB2 is experiencing problems with the availability of expanded storage.

©Copyright 1999, Computer Management Sciences, Inc.

DB2 for OS/390 Version 6: Administration Guide
Section 5.4.1.1.1 (Buffer Pools and Hiperpools)

DB2 UDB for OS/390 and z/OS, Version 7: Administration Guide
Section 5.4.1.1.1 (Buffer Pools and Hiperpools)

DB2 UDB for z/OS Version 8: Administration Guide
Chapter 26. Tuning DB2 buffer, EDM, RID, and sort pools

DB2 UDB for z/OS Version 9: Performance Monitoring and Tuning Guide
Chapter 4. Tuning DB2 buffer, EDM, RID, and sort pools

DB2 10 for z/OS: Managing Performance
Chapter 7. Configuring storage for performance

DB2 11 for z/OS: Managing Performance
Chapter 6. Configuring storage for performance

"Central Storage Management with MVS/ESA" by Deese, Donald R., available at **www.cpexpert.com**.

"Expanded Storage Management with MVS/ESA" by Deese, Donald R., available at **www.cpexpert.com**.

"DB2 Buffer Pool Tuning - Top Down or Bottom Up", Joel Goldstein, *CMG98 Conference Proceedings*, pages 51-62.

MVS Planning: Workload Management
Section 12 (Defining Special Protection Options for Critical Work)

©Copyright 1999, Computer Management Sciences, Inc.

©Copyright 1999, Computer Management Sciences, Inc.