IBM

Redbooks Paper

Jim Grauel
Frank Kyne

Performance Considerations and Measurements for CICS and System Logger

Introduction

This Redpaper describes the interactions between IBM® CICS® and z/OS® System Logger in order to achieve a well-tuned environment between the two subsystems. This paper describes the tools available to capture the performance data, how to interpret the data, and shows samples of different CICS and System Logger configurations, so you can actually see the different behaviors.

This Redpaper has been written in support of the IBM Redbook *System Programmer's Guide to: z/OS System Logger*, SG24-6898.

Performance tools

Obtaining the best performance is a matter of tuning for the optimum data residency time in interim storage. For CICS system log streams (DFHLOG and DFHSHUNT), the data required for backout and recovery should be available in interim storage for the duration of the unit of work. For non-system log streams, the data residency time should be such that a steady movement of data from interim storage to the offload data sets is observed, while requiring a minimal amount of storage to duplex the data within the System Logger data space.

The performance of CICS log streams is influenced by the sizing, type, and placement of the log streams. In addition, activity keypoint frequency (AKPFREQ) and the log defer interval (LGDFINT) can have a significant impact on log stream performance. It is important to remember that the IXGWRITE calls are issued under the CICS Quasi-Reentrant (QR) TCB, thus the CPU time is charged to the transaction.

Performance data is available from several sources:

- ► System Management Facilities (SMF) Type 88 records produced by the z/OS System Logger
- ► SMF Type 110 data produced by CICS

  Includes both statistical and performance data.

- ► SMF Type 70-78 data produced by the Resource Measurement Facility (RMF™) (or an equivalent product)

There are several tools that aid in the analysis of log stream performance problems:

- ► DFH0STAT: A CICS-supplied sample transaction that collects the CICS statistics and writes them to the CICS JES log.
- ► DFHSTUP (CICS Statistics Utility Program): Used to post-process the CICS SMF Type 110 subtype 2 statistical records.
- ► CICS Performance Analyzer (part of the CICS Tools portfolio): Used to post-process the CICS SMF 110 subtype 1 performance records and SMF Type 88 records.
- ► IXGRPT1 (sample program provided with z/OS): Formats and reports the SMF Type 88 data produced by the z/OS System Logger.
- ► Various RMF reports (Coupling Facility Activity reports, Workload Activity reports): Based on the data captured in the SMF 70 -78 records.
- ► The sample DFSORT™ jobs provided with the IBM Redbook *System Programmer's Guide to: z/OS System Logger*, SG24-6898: These reports provide similar information to the IXGRPT1 program; however, they enable you to summarize all the activity by log stream rather than by interval, and they are also more readable than the IXGRPT1 output.

DFH0STAT

This section describes the DFH0STAT sample transaction.

Use: Dynamically gather CICS statistics

DFH0STAT is supplied as a sample COBOL program in CICS.SDFHSAMP. It contains self-documenting source code to be compiled and run as a transaction to collect CICS statistics and write them to the JES spool. The output can then be viewed under TSO. The SIT parm SPOOL=YES is required. The information presented with DFH0STAT is also available in the CICS shutdown statistics. The interval reported is based on the CICS stats interval; to be precise, it is the data since the last stats reset, which is determined by the stats interval.

A compiled version is available in CICS Transaction Server Version 2.2.

A sample output is shown in Figure 1 on page 3. The MVS release is provided in the System Status section, along with the values for activity keypoint frequency (AKPFREQ) and log stream Deferred Force Interval (LGDFINT).

```
Applid IYOT1      Sysid JIM    Jobname IYOT1         Date 01/30/2000  Time 01:56:47                      CICS 5.3.0
_____

System Status
_____
  MVS Product Name. . . . :  MVS/SP6.1.0
  Activity Keypoint Frequency. . . . . . . :      4,000
  Logstream Deferred Force Interval. . . . :        30

                    Use                 Sys                          Max Block  DASD    Retention   Auto    Stream    Browse    Browse
Logstream Name     count   Status       Log   Structure Name         Length     Only    Period      Delete  Deletes   Starts    Reads
_____
IYOT1.DFHLOG         1      OK          YES                          65,532     YES          0       NO        3        34         0
IYOT1.DFHSHUNT       1      OK          YES   LOG_JG                  64,000     NO           0       NO        1         0         0
IYOT1.J02            1      OK          NO                            65,532     YES          0       NO      N/A       N/A       N/A

                    Write                         Average        Buffer        Buffer       Force    Current    Peak      Retry
Logstream Name     Requests     Bytes Written     Bytes          Appends       Full Waits   Waits    Waiters    Waiters   Errors
_____
IYOT1.DFHLOG         9,023        2,546,491         282            9,037           0          2         0          1         1
IYOT1.DFHSHUNT           0                0           0                0           0          0         0          0         0
IYOT1.J02               14          916,800      65,485            9,009           0          0         0          0         0
```

Figure 1   Sample DFH0STAT output

In the log stream statistics, we see each log stream connected to this CICS region. If the log stream resides in a Coupling Facility (CF) structure, the structure name is given; if it is a DASD-only log stream, the structure name is blank. In the example, IYOT1.DFHLOG and user journal J02 are DASD-only log streams, and IYOT1.DFHSHUNT is a CF-Structure log stream in structure LOG_JG.

Note that CICS allocates two buffers for each log stream, so CICS can process journal writes for one buffer while an IXGWRITE is in progress and an alternate buffer.

The values under Max Block Length are worth noting. This value originates in the log stream or structure definition and is returned to CICS when it connects to the log stream. For a CF-Structure log stream, the block size is specified as MAXBUFSIZE on the structure definition. The value specified in MAXBUFSIZE determines the element size for the log streams in the structure. If the MAXBUFSIZE is specified equal to or less than 65276, the element size is 256; if greater than 65276, the element size is set to 512.

For DASD-only log streams, MAXBUFSIZE can be specified on the log stream definition. MAXBUFSIZE defines the largest block that can be written to the log stream. The default value is 65532.

In either case, the MAXBUFSIZE value is returned to CICS and is used to determine the CICS log stream buffer size. Note, for user journals, unless the application uses the wait option, the IXGWRITE call is issued when the buffer fills. Refer to the Average Bytes column. This is the average buffer size written using IXGWRITE calls. For DFHJ02, the average size is 65,485. It is sometimes advantageous to reduce MAXBUFSIZE on a user journal to force the data into the log stream more frequently. This reduces the residency time and allows quicker access to the data from batch applications.

It's also worth noting in the case of a hard failure (MVS, CICS, or the hardware), where the data has not been written to the log stream, it will be lost. For DFHLOG and DFHSHUNT, when the region is emergency restarted, in-flight tasks would be backed out using the information that had already been written to the log.

In the case of a user journal, the data in the buffer would be lost. There is no transaction backout associated with data lost in a non-system log buffer.

The columns titled DASD Only, Retention Period, and Auto Delete reflect the attributes specified for the log stream in the System Logger policy.

The value given under Stream Deletes is the number of times CICS issued an IXGDELET call to the System Logger for log tail deletion.

The value under Browse Starts is a count of the number of times a browse start request is issued. Prior to CICS Transaction Server Version 2.2, a large value might be observed for DFHLOG in a low volume region, due to CICS using a Browse Start to verify that the System Logger is still operational. In CICS Transaction Server Version 2.2, a change was implemented to call the System Logger function CHECK_CONNECTION_STATUS.

The Browse Reads column reflects the number of times CICS reads the log stream to obtain information for transaction backout.

The number of Write Requests is the number of times CICS called the System Logger for an IXGWRITE. The number of Buffer Appends is a count of the times a record was added (appended) to the CICS log buffer. The number is normally larger than the number of Write Requests due to calls to the CICS logger domain that do not include the force option. For example, the records produced during activity keypoint processing are not written with the force option. In contrast, the before images for a recoverable VSAM file are written with the force option specified. Unfortunately, it is not possible to predict what percentage of writes will be with the force option; it depends on the work load. If the workload consists of mostly updates for a recoverable VSAM file, most of the records on DFHLOG will be forced. If the bulk of the records are the result of an activity keypoint, we simply append the records without force.

Buffer Full Waits is a count of the number of attempts to append a record to the current log stream buffer while the buffers were logically full. This situation arises when the current log stream buffer does not have sufficient space to accommodate the journal record, and IXGWRITE is already in progress for the alternate log stream buffer.

Force Waits is the total number of tasks suspended waiting for the IXGWRITE for their log records to complete.

Buffer Full Waits and Force Waits can be an indication that there is a delay in I/O processing. They can also be an indication that the log defer interval is too large. Additional items to check are the CF service times for a CF-Structure log stream or DASD I/O time for DASD-only log streams.

If Buffer Full Waits or Force Waits, or both, are seen consistently, verify that the value for LGDFINT is set to 5. Prior to CICS Transaction Server V2.2, the default value was 30. Normally, the recommendation is not to reduce the value lower than 5; however, in a very high-volume region, it might be advantageous to set it to 0 for CF-Structure log streams. In addition, CF service time (for CF-Structure log streams) or DASD I/O time should be investigated.

Setting LGDFINT to zero causes the data to be written to the log stream (through an IXGWRITE) immediately, without waiting for additional records to be added to the buffer.

The Peak Waiters field contains the high watermark of the number of tasks (UOWs) waiting for a buffer to be written. The Current Waiters field is the number of tasks (UOWs) waiting for a buffer to be written at the point in time when the statistics were collected.

Retry Errors is the number of requests to System Logger that resulted in a return code indicating an error and that were subsequently retried. An example would be 868 errors returned while the staging data set is being formatted. This can happen with DASD-only log streams, or if staging data set is used with a CF-Structure log stream. For example:

```
01.48.56 JOB07716 +DFHLG0777 IYOT1
A temporary error condition occurred during MVS logger operation IXGWRITE for log stream
IYOT1.DFHLOG. MVS logger codes: X'00000008', X'00000868'.
```

With CICS Transaction Server V2.2, DFH0STAT has been enhanced. The format has been changed for ease of reading, and information has been added to calculate (and display) the number of log stream writes per second. The samples in Figure 2 and Figure 3 on page 6 show the format of the new displays.

```
Applid IYOT1     Sysid IYO1  Jobname IYOT122     Date 08/11/2002  Time 13:50:13                      CICS 6.2.0

System Status

  MVS Product Name. . . . . . . :  MVS/SP7.0.4                      CICS Transaction Server Level . . :  02.02.00
  CICS Startup. . . . . . . . . :  INITIAL
  CICS Status . . . . . . . . . :  ACTIVE                           RLS Status. . . . . . . . . . . . :  RLS=NO
  Storage Protection. . . . . . :  ACTIVE                           RRMS/MVS Status . . . . . . . . . :  OPEN
  Transaction Isolation . . . . :  ACTIVE
  Reentrant Programs. . . . . . :  NOPROTECT                        VTAM Open Status. . . . . . . . . :  OPEN
                                                                    IRC Status. . . . . . . . . . . . :  OPEN
  Force Quasi-Reentrant . . . . :  No                               TCP/IP Status . . . . . . . . . . :  OPEN

  Program Autoinstall . . . . . :  ACTIVE                           Max IP Sockets. . . . . . . . . . :      255
  Terminal Autoinstall. . . . . :  ENABLED                          Active IP Sockets.. . . . . . . . :        0

  Activity Keypoint Frequency. . . . . . . :       1,000            WEB Garbage Collection Interval . :       60
  Logstream Deferred Force Interval. . . . :           0            Terminal Input Timeout Interval . :        5

  DB2 Connection Name . . . . . :  SYSD
```

Figure 2   DFH0STAT for CICS Transaction Server V2.2: Part 1

```
Applid IYOT1    Sysid IYO1  Jobname IYOT122     Date 08/11/2002  Time 13:50:13        CICS 6.2.0       PAGE    10
_____
Logstream - System Logs
_____

System log - DFHLOG
_____
  Logstream Name . . . . . . . . . . . . . :  IYOT1.CICS22.DFHLOG          Logstream Status . . . . :         OK
  DASD Only. . . . . . . . . . . . . . . . :  YES                         Retention Period (days). :          0
  Coupling Facility Structure Name . . . . :                              Auto Delete. . . . . . . :         NO
  Logstream Writes . . . . . . . . . . . . :       18,788                 Maximum Block Length . . :     65,532
  Logstream Writes per second. . . . . . . :        14.74
  Average Bytes per Logstream Write. . . . :        1,289
  Logstream Deletes (Tail Trims) . . . . . :          105
  Logstream Query Requests . . . . . . . . :            3
  Logstream Browse Starts. . . . . . . . . :            0
  Logstream Browse Reads . . . . . . . . . :            0
  Logstream Buffer Appends . . . . . . . . :      104,811
  Logstream Buffer Full Waits. . . . . . . :            0
  Logstream Force Waits. . . . . . . . . . :       93,764                 Logstream Current Force Waiters:    3
  Logstream Retry Errors . . . . . . . . . :            1                 Logstream Peak Force Waiters:      11


--------------------------------------------------------------------------------------------------------------------
Applid IYOT1    Sysid IYO1  Jobname IYOT122     Date 08/12/2002  Time 08:31:13        CICS 6.2.0       PAGE    11
_____
Logstreams - Resource
_____
                          Use           Sys                   Max Block DASD Retention  Auto   Stream  Browse  Browse
Logstream Name           Count  Status  Log  Structure Name    Length  Only   Period   Delete Deletes  Starts  Reads
_____
IYOT1.CICS22.DFHLOG        1     OK     YES  LOG_JG_20M        64,000   NO      0        NO    1,896     0       0
IYOT1.CICS22.DFHSHUNT      1     OK     YES  LOG_JG_20M        64,000   NO      0        NO        1     0       0

Logstreams - Requests
_____
                         Write                 Average   Buffer      Buffer      Force   Current   Peak    Retry
Logstream Name          Requests  Bytes Written  Bytes   Appends   Full Waits    Waits   Waiters  Waiters  Errors
_____
IYOT1.CICS22.DFHLOG     445,574   443,483,499    995    1,895,148       0       428,880     1       10       1
IYOT1.CICS22.DFHSHUNT         0             0      0            0       0             0      0        0       0
```

Figure 3   DFH0STAT for CICS Transaction Server V2.2: Part 2

DFHSTUP

This section describes the DFHSTUP program.

Use: Evaluate the CICS Interval or Daily Statistics

The CICS statistics utility program, DFHSTUP, is an offline utility that processes CICS statistics information contained in SMF Type 110 subtype 2 records. The CICS statistics domain records interval statistics when the system initialization parameter STATRCD=ON is specified. Other statistics record types (unsolicited, requested, and end-of-day) are written regardless of the setting of the STATRCD option. For additional information about DFHSTUP, refer to *CICS Transaction Server for z/OS: CICS Operations and Utilities Guide*, SC34-5991.

A separate version of DFHSTUP is provided with each release of CICS, so the correct version must be used to process the data.

In the sample DFHSTUP End of Day (EOD) report shown in Figure 4 on page 7, you will notice the same type of information provided by DFH0STAT is presented. However, in this case, it is for the entire execution of the CICS job, or for a 24-hour period if the CICS region was executing for multiple days.

In the sample report, notice that the log stream IYOT3.DFHLOG has been specified with a MAXBUFSIZE of 60000, is allocated to structure LOG_JG2_20M, has a retention period of 0, and that auto delete is set to NO.

The EOD report shows the overall summary for the day. To collect interval statistics, the interval can be set using the STATINT= system initialization parm, or it can be set through the CEMT master terminal command (in CICS TS V2.2).

```
CICS 5.3.0 Statistics Utility Program                    Report Date 05/31/2003   Report Time 09:10:01
End of Day Statistics Report
                        Collection Date-Time 05/31/2003-08:51:02  Last Reset 08:43:28  Applid IYOT3 Jobname IYOT3
_____
JOURNALNAME STATISTICS

  Journal  Journal                              Write    Bytes    Buffer
   Name     Type   Log Stream Name             Requests  Written  Flushes
  _____

  DFHLOG    MVS    IYOT3.DFHLOG                  N/A      N/A      N/A
  DFHSHUNT  MVS    IYOT3.DFHSHUNT                N/A      N/A      N/A
                                               _____
  *TOTALS*                                        0        0        0
_____
LOGSTREAM STATISTICS - RESOURCE
                        System                    Max Block DASD Retention  Auto    Delete
  Log Stream Name          Log   Structure Name    Length   Only  Period   Delete  Requests
  _____

  IYOT3.DFHLOG            Yes    LOG_JG2_20M        60000   No        0      No       45
  IYOT3.DFHSHUNT          Yes    LOG_JG_20M         64000   No        0      No        1
                                                                                   _____
  *TOTALS*                                                                            46
_____
LOGSTREAM STATISTICS - REQUEST
                       Write     Bytes    Buffer    Waits    Current       Peak      Total   Browse  Browse   Retry
  Log Stream Name Requests   Written   Appends  Buff Full Frce Wtrs   Frce Wtrs  Force Wts  Starts   Reads   Errors
  _____

  IYOT3.DFHLOG     90148  45213095    178839        0         0           3        554      46     111       0
  IYOT3.DFHSHUNT       0         0         0        0         0           0          0       0       0       0
                  _____
  *TOTALS*         90148  45213095    178839        0                              554      46     111       0
```

Figure 4   Sample DFHSTUP End of Day report

CICS Performance Analyzer

This section describes the CICS Performance Analyzer tool.

Use: Analysis of the SMF 110 Subtype 1 CICS performance records

CICS Performance Analyzer (PA) is a post processor for CICS SMF Type 110 records. It provides the ability to summarize by CICS monitoring fields, and display the API calls and resources used, along with CPU per transaction and response time.

In Release 2 of CICS Performance Analyzer, the product was enhanced to provide support for the System Logger SMF Type 88, and SMF Type 101 records produced by DB2®. The sample outputs from IXGRPT1 in Figure 5 on page 8 and from CICS PA in Figure 6 on page 8 show alternate methods for displaying the data.

```
                            BYT WRITTN  BYT WRITTN  BYT WRITTN                                          AVERAGE
                            BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------  BUFFER
LOGSTREAM NAME---- STRUCTURE NAME- IXGWRITES  STORAGE               INVOKED  TYPE1      TYPE2      TYPE3    SIZE

                            BYT DELETD  # DELETES   BYT DELETD  # DELETS  -------------EVENT--------------
                            INTERIM ST  W/O DASD    INTERIM ST    W/     OFF- DASD STRC NTRY STG  STG  RE-
                            W/O DASD    WRITE       W/DASD      WRITE    LOAD SHFT FULL FULL THLD FULL BLD
_____

IYOT1.CICS22.DFHLOG  LOG_JG2_5M   324226786  556595200    35060218  1138714 1072720     46589      19401         284
                                 302169464    1067720     30935658   103114   378   5    5 3470   0    0    0
```

Figure 5   Sample IXGRPT1 report

```
V1R2M0                                    CICS Performance Analyzer
                                       System Logger - Logstream Summary
                                 _____
LOGR0001 Printed at 11:15:59  5/28/2003     Data from 17:20:00:00  5/27/2003 to 17:20:00:00  5/27/2003
   Logger data


Logstream name         MVSID  Structure name   First interval start   Last interval stop      Total Interval
IYOT1.CICS22.DFHLOG    SYSD   LOG_JG2_5M       17:15:00.00  5/27/2003  17:20:00.00  5/27/2003     0000:05:00

            ----------- IXGWRITES ----------        --------------- DELETIONS ----------------
                                            Bytes      Count      Count      Bytes      Bytes
                                          Writn to     With      Without     After      Int Stor
                         Total    Average  Interim     DASD       DASD      Offload     w/o DASD
              Count      Bytes     Bytes   Storage     Write      Write     w. DASD      Write
            --------   --------  -------- --------   --------   --------   --------    --------
Total       1138714    324227K      285   556595K     103114    1067720    30936K      302169K
Rate(/Sec)     3795    1080756            1855317        343       3559    103119     1007232
Minimum     1138714    324227K            556595K     103114    1067720    30936K      302169K
Maximum     1138714    324227K            556595K     103114    1067720    30936K      302169K


            ----------------------------------------- EVENTS ------------------------------------------
                                   Demand                                                      Demand
                         Staging    DASD      Block     Staging    Entry     Struct            Init'd
              Offloads  Threshld   Shifts     Length     Full      Full      Full             Offloads
            --------   --------  --------   --------   --------   --------   --------          --------
Total           378         0        5                     0       3470        5                  0
Rate(/Sec)        1         0        0                     0         11        0                  0
Minimum         378         0        5        116           0       3470        5                  0
Maximum         378      3470        5       3308           0       3470        5                  0


            ---------------------- EVENTS ------------------------    -------------- DASD Writes ----------------
                                            Struct     Struct
                                           Rebuilds   Rebuilds                   Total
              Type1      Type2    Type3     Init'd     Complt'd    Count        Bytes      Average   Waits
            --------   --------  -------- --------   --------   --------     --------     --------  -------
Total       1072720     46589    19401        0          0        323        35060K          0      236
Rate(/Sec)     3575       155       64        0          0          1        116867                  0
Minimum     1072720     46589    19401        0          0        323        35060K               236
Maximum     1072720     46589    19401        0          0        323        35060K               236
```

Figure 6   CICS Performance Analyzer Logstream Summary

Figure 7 on page 9 is an example of the structure summary available in CICS Performance Analyzer.

```
V1R2M0                                      CICS Performance Analyzer
                                          System Logger - Structure Summary
                                        _____

LOGR0001 Printed at 11:15:59  5/28/2003      Data from 17:20:00:00  5/27/2003 to 17:20:00:00  5/27/2003
   Logger data


Structure name        MVSID      First interval start   Last interval stop                     Total Interval
LOG_JG2_5M            SYSD       17:15:00.00  5/27/2003  17:20:00.00  5/27/2003                     0000:05:00


            ----------- IXGWRITES ----------        --------------- DELETIONS -----------------
                                            Bytes     Count      Count       Bytes       Bytes
                                           Writn to   With      Without      After      Int Stor
                          Total    Average  Interim    DASD        DASD      Offload     w/o DASD
                Count     Bytes     Bytes   Storage    Write       Write     w. DASD      Write
              --------  --------  --------  --------  --------   --------    --------    --------
Total         1138714   324227K     285     556595K   103114    1067720      30936K     302169K
Rate(/Sec)       3795   1080756            1855317       343       3559      103119     1007232
Minimum       1138714   324227K             556595K   103114    1067720      30936K     302169K
Maximum       1138714   324227K             556595K   103114    1067720      30936K     302169K


            -------------------------------------------- EVENTS -----------------------------------------
                               Demand                                                          Demand
                      Staging    DASD      Block     Staging     Entry      Struct            Init'd
            Offloads  Threshld   Shifts    Length     Full        Full       Full            Offloads
            --------  --------  --------  --------   --------   --------    --------          --------
Total          378        0         5                   0         3470          5                0
Rate(/Sec)       1        0         0                   0           11          0                0
Minimum        378        0         5         116        0         3470          5                0
Maximum        378        0         5        3308        0         3470          5                0


            --------------------- EVENTS -------------------------       -------------- DASD Writes ----------------
                                            Struct     Struct
                                            Rebuilds   Rebuilds                    Total
              Type1     Type2     Type3     Init'd     Complt'd    Count      Bytes      Average      Waits
            --------  --------  --------  --------    --------   --------   --------    --------    --------
Total        1072720    46589     19401        0          0        323      35060K         0          236
Rate(/Sec)      3575       155        64        0          0          1      116867                     0
Minimum      1072720     46589     19401        0          0        323      35060K                   236
Maximum      1072720     46589     19401        0          0        323      35060K                   236
```

Figure 7   CICS Performance Analyzer Structure Summary

RMF reports

This section describes the RMF Coupling Facility Activity and Workload Activity reports.

RMF Coupling Facility Activity report

This section describes the RMF Coupling Facility Activity report.

Use: Analysis of the Coupling Facility (CF) structure service times

For CF-Structure log streams, the RMF Coupling Facility Activity report can provide important information about structure usage and service times. However, this report only provides information at the structure level; you must determine the number of log streams associated with the structure in question from either the SMF Type 88 data, or from the log stream definitions.

```
----------------------------------------------------------------------------------------------------
 COUPLING FACILITY NAME = CF103
 TOTAL SAMPLES(AVG) =   299  (MAX) =   299  (MIN) =   299
----------------------------------------------------------------------------------------------------
                                    COUPLING  FACILITY  USAGE  SUMMARY
----------------------------------------------------------------------------------------------------
 STRUCTURE SUMMARY
----------------------------------------------------------------------------------------------------
                                      % OF           % OF    AVG    LST/DIR  DATA      LOCK     DIR REC/
       STRUCTURE             ALLOC    CF        #    ALL     REQ/   ENTRIES  ELEMENTS  ENTRIES  DIR REC
 TYPE  NAME       STATUS CHG SIZE     STORAGE   REQ  REQ     SEC    TOT/CUR  TOT/CUR   TOT/CUR  XI'S

 LIST  DSNZPLEX_SCA  ACTIVE    32M    1.6       0    0.0     0.00   43K      86K       N/A      N/A
                                                                    128      819       N/A      N/A
       LOG_JG2_5M    ACTIVE     8M    0.4     1585K  100     5284.3 5725     11K       N/A      N/A
                                                                    1554     3143      N/A      N/A
 PROCESSOR SUMMARY
----------------------------------------------------------------------------------------------------
 COUPLING FACILITY      2064      MODEL 100     CFLEVEL  12
 AVERAGE CF UTILIZATION  (% BUSY)      20.2     LOGICAL PROCESSORS:   DEFINED   1    EFFECTIVE   1.0

                      C O U P L I N G   F A C I L I T Y   A C T I V I T Y

    z/OS V1R4            SYSPLEX WSCZPLEX          DATE 05/28/2003        INTERVAL 005.00.000
                        RPT VERSION V1R2 RMF      TIME 16.30.00          CYCLE 01.000 SECONDS
----------------------------------------------------------------------------------------------------
                                 COUPLING  FACILITY  STRUCTURE  ACTIVITY
----------------------------------------------------------------------------------------------------
 STRUCTURE NAME = LOG_JG2_5M       TYPE = LIST    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

 SYSD     1585K  SYNC  1458K 91.9  43.7   54.8       NO SCH  0   0.0   0.0     0.0     0.0
          5284   ASYNC 128K   8.1  323.5  406.6      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
```

Figure 8   RMF Coupling Facility Activity report

In the sample report shown in Figure 8, the structure size for LOG_JG2_5M has been altered to 8M, which is 0.4% of the total CF storage. The structure was defined with an INITSIZE of 5M, but altered to 8M due to an excessive number of structure full conditions while running the workload. This structure did 100% of the requests to the CF in the 5 minute interval shown. The average request rate was 5284 per second.

System Logger log streams are reported in the LIST section of the Usage Summary part of the report. Under the LST/DIR column, there are two lines per structure. The first line gives the total number of *entries* in the structure, and the second line shows the number currently in use. In the Data Elements column, the first line gives the total number of data *elements* in the structure, and the second line gives the number currently in use. The entries are in a common pool for the structure (the entries for all log streams come from the same pool). Dividing the number of data elements (11 K) by the number of entries (5725) gives an entry-to-element ratio of 1:2. The entry-to-element ratio is dictated by the log stream with the largest percentage of records in the structure. If the number of entries in use reaches 90% of the total number of entries available for the structure, System Logger will force an offload of *all* log streams in the structure.

Notice that the current usage (second row of data) indicates a 1:3 ratio (3143 in-use elements/1554 in-use entries), which is 2.022, rounded up to the next whole number. As previously noted, the number of entries (5725) and elements (11 K) in the structure are in a

1:2 ratio. With a MAXBUFSIZE set to 64000, the element size is 256, and consequently, because of the 1:2 ratio, this indicates an expected buffer size of 512 bytes. However, the current usage of 1:3 indicates at least one log stream is writing records equal to or greater than 750 bytes.

An important point to remember is that the number of data elements is divided equally among the connected log streams. So, if the number of data elements in the structure is 11 K, and there are three log streams connected to the structure, each log stream is allocated 3.66 K data elements.

The Coupling Facility Activity report also provides information about the request activity for the structure. From the SYSD system, there were 1585 K requests, giving an average of 5284 per second. Of the requests, 1458 K were synchronous with average service times of 43.7 microseconds.

The CF used in this example was a 2064, so response times for System Logger structures of less than 50 microseconds are considered good. If the CF resided in a 9672-R06, expected response times would be in the 60 to 100 microsecond range.

The average, or *mean*, represents the middle in the distribution of a set of individual measurements. The standard deviation measures the spread or variation of the individual measurements on either side of the average. Of all observations, 66% lie within plus or minus 1 standard deviation; 95% of all observations lie within plus or minus 2 standard deviations.

The report shows an average SYNC time of 43.7 microseconds with a standard deviation of 54.8 microseconds for the population of SYNC requests. In this case, 2 standard deviations to the negative side of the average would be less than zero. Consider 0 to +2 standard deviations on the positive side of the average defines the range where 95% of the requests will occur. Thus, 95% of all SYNC requests for this test would lie between 0 and 153.3 (2*54.8 +43.7) microseconds. If the standard deviation is large (for example, in the thousands of microseconds), it indicates that some portion of the CF configuration is non-responsive, causing large variability in individual measurements.

The most frequently seen reason for a non-responsive CF is the use of shared CF CPs, especially if Dynamic CF Dispatching (DYNDISP=YES) is enabled for the CF LPAR. Looking at the RMF CF Usage Summary report, whenever the number of logical processors defined is greater than the number of effective logical processors, the configuration might be experiencing performance issues. For production CICS regions, if you must use a CF with shared CPs, we strongly recommend DYNDISP=NO. The DYNDISP command is entered on the CF console.

Another point of caution, if the CF is actually an LPAR in the same CPC as the z/OS image, and the LPARs share CPs, *all* SYNC requests will be converted to ASYNC requests. Neither CICS nor the System Logger have control (or knowledge) of the change, and the reports will show the requests as SYNC, but the service times will be elongated.

You should also check for delayed requests. For example, if the number of NO SCH (that is, no subchannel), is greater than 10% of the total number of requests for the structure, you should consider adding more CF Link capacity, or reducing the load on the existing links.

Figure 9 on page 12 contains sample job control language (JCL) to produce the Coupling Facility Activity report.

```
//CFACT    JOB (????,????),'JIM GRAUEL',MSGLEVEL=(1,1),
// CLASS=A,MSGCLASS=O,NOTIFY=GRAUEL,REGION=0M
/*JOBPARM SYSAFF=SYSD
//**************************************************
//**      FORMAT THE CF ACTIVITY REPORT        ****
//**************************************************
//* PROVIDE THE SMF DATASET NAME                  *
//   SET    SMFIN1=GRAUEL.PERFTST.R222.RED1R.P1.MAY31.ELEVEN03
//**************************************************
//S1       EXEC PGM=ERBRMFPP,REGION=1000K
//MFPMSGDS DD SYSOUT=*
//SYSPRINT DD SYSOUT=*
//MFPINPUT DD  DSN=&SMFIN1,DISP=(OLD,KEEP)
//SYSIN DD *
 SYSOUT(O)
 SYSRPTS(CF)                      /*COUPLING FACILITY       */
//NULL  DD  *
```

Figure 9   Sample JCL to produce an RMF Coupling Facility Activity report

RMF Workload Activity reports

This section describes the RMF Workload Activity reports.

Use: Analysis of RMF data

RMF provides a wealth of information that is invaluable in the resolution of performance problems. This information can be used to measure the impact of log stream usage and definition changes.

When using the z/OS Workload Manager (WLM) in goal mode, it is possible to define a report class for each CICS region and the System Logger address space. Each CICS region and System Logger address space should be defined in a separate report class in order to understand the results of each change made during the log stream tuning process.

The same information can be gathered using a Report Performance Group if WLM is being used in compatibility mode. For guidance in defining a compatibility mode Report Performance Group, refer to the section titled "Defining Service Classes and Performance Goals" in *z/OS MVS Planning: Workload Management,* SA22-7602.

Figure 10 on page 14 contains a WLM Workload Activity report, which presents data collected for report classes RIYOT1 (a CICS region) and RLOGER (the System Logger address space). Report classes are defined using the WLM ISPF panels.

Sample JCL to produce the Workload Activity report is shown in Figure 11 on page 14.

The report interval is listed in the START and END times at the top of the report page. A word of caution: The minimum interval is defined by the INTVAL() parm in the SMFPRMxx member of SYS1.PARMLIB. In the samples collected, the interval was set to 5 minutes, as follows:

```
 INTVAL(05)                  /* SMF GLOBAL RECORDING INTERVAL */
```

It's important to ensure that the SMF 70 to 79 records are being collected, along with the CICS 110 records. The records that are to be collected are also defined in the SMFPRMxx member:

```
SUBSYS(STC,EXITS(IEFACTRT),INTERVAL(SMF,SYNC),
       TYPE(0,30,70:79,88,89,90,99,110,245))
SUBSYS(OMVS,NOEXITS,INTERVAL(SMF,SYNC),
       TYPE(0,30,70:79,88,89,90,99,110,245))
```

When the reports are formatted, it's possible to report a larger interval than was specified in the SMFPRMxx member by using the DINTV parm for the ERBRMPFF utility. However, remember that the smallest interval that can be reported is the value specified for INTVAL.

The following fields should be noted in the reports:

► **TRANSACTIONS**

- **MPL**: Multiprogramming level, the number of address spaces active in this service/report class during the interval The value should be one for the most accurate measurements for a given region.

► **DASD I/O**

- **SSCHRT**: Number of start subchannels (SSCH) per second in the reported interval.
- **RESP**: Average DASD response time (in milliseconds).

► **SERVICE RATES**

- **TCB**: CPU seconds accumulated in TCB mode during the collection interval.
- **SRB**: CPU seconds accumulated in SRB mode during the collection interval.
- **APPL%**: Percentage of a engine (CP) used during the collection interval.

► **PAGE-IN RATES**

- **SINGLE**: The average rate at which pages are read into main storage.

► **STORAGE**

- **AVG**: Average number of central and expanded storage frames allocated to ASIDs in the report class. A storage frame is 4096 bytes.

Refer to *z/OS V1R4.0 RMF Report Analysis,* SC33-7991, for more information about this report.

```
                              W O R K L O A D   A C T I V I T Y
                                                                                        PAGE   1
    z/OS V1R2                 SYSPLEX WSCZPLEX          START 01/01/2002-23.15.00 INTERVAL 000.04.59   MODE = GOAL
                              RPT VERSION V1R2 RMF      END   01/01/2002-23.19.59

                                        POLICY ACTIVATION DATE/TIME 11/21/2001 12.21.14
-------------------------------------------------------------------------------------------------- REPORT
CLASS(ES)

 REPORT BY: POLICY=WLMPOL                       REPORT CLASS=RIYOT1
                                                DESCRIPTION =CICS Report Class for Jim G.

 TRANSACTIONS     TRANS.-TIME HHH.MM.SS.TTT  --DASD I/O--  --SERVICE--   SERVICE RATES-  PAGE-IN RATES
 ----STORAGE----
 AVG       1.00  ACTUAL                   0  SSCHRT  11.4  IOC    343    ABSRPTN    631  SINGLE     0.0  AVG     8517.43
 MPL       1.00  EXECUTION                0  RESP     1.9  CPU   9348    TRX SERV   631  BLOCK      0.0  TOTAL   8517.43
 ENDED        0  QUEUED                   0  CONN     1.6  MSO 175902    TCB        0.8  SHARED     0.0  CENTRAL 8517.43
 END/S     0.00  R/S AFFINITY             0  DISC     0.1  SRB   3566    SRB        0.3  HSP        0.0  EXPAND     0.00
 #SWAPS       0  INELIGIBLE               0  Q+PEND   0.2  TOT 189159    RCT        0.0  HSP MISS   0.0
 EXCTD        0  CONVERSION               0  IOSQ     0.0  /SEC   631    IIT        0.0  EXP SNGL   0.0  SHARED     1.00
 AVG ENC   0.00  STD DEV                  0                             HST        0.0  EXP BLK    0.0
 REM ENC   0.00                                                         APPL %     0.4  EXP SHR    0.0
 MS ENC    0.00


 REPORT BY: POLICY=WLMPOL                       REPORT CLASS=RLOGER
                                                DESCRIPTION =Report for System Logger

 TRANSACTIONS     TRANS.-TIME HHH.MM.SS.TTT  --DASD I/O--  --SERVICE--   SERVICE RATES-  PAGE-IN RATES
 ----STORAGE----
 AVG       1.00  ACTUAL                   0  SSCHRT  12.3  IOC      1    ABSRPTN      8  SINGLE     0.0  AVG     5115.59
 MPL       1.00  EXECUTION                0  RESP     1.1  CPU     85    TRX SERV     8  BLOCK      0.0  TOTAL   5115.59
 ENDED        0  QUEUED                   0  CONN     0.8  MSO   1181    TCB        0.0  SHARED     0.0  CENTRAL 5115.59
 END/S     0.00  R/S AFFINITY             0  DISC     0.1  SRB   1085    SRB        0.1  HSP        0.0  EXPAND     0.00
 #SWAPS       0  INELIGIBLE               0  Q+PEND   0.2  TOT   2352    RCT        0.0  HSP MISS   0.0
 EXCTD        0  CONVERSION               0  IOSQ     0.0  /SEC     8    IIT        0.0  EXP SNGL   0.0  SHARED     0.00
 AVG ENC   0.00  STD DEV                  0                             HST        0.0  EXP BLK    0.0
```

Figure 10   Sample RMF Workload Activity report

```
//RCLASS   JOB (????,????),'JIM GRAUEL',MSGLEVEL=(1,1),
// CLASS=A,MSGCLASS=O,NOTIFY=GRAUEL,REGION=0M
/*JOBPARM SYSAFF=SYSD
//**************************************************
//**     FORMAT THE REPORT CLASSES           ****
//**************************************************
//* PROVIDE THE SMF DATASET NAME                    *
//   SET    SMFIN1=GRAUEL.PERFTST.R222.RED1R.P1.MAY31.ELEVEN03
//**************************************************
//RMFPP     EXEC PGM=ERBRMFPP,REGION=0M
//MFPMSGDS DD SYSOUT=*
//SYSPRINT DD SYSOUT=*
//MFPINPUT DD DSN=&SMFIN1,DISP=(OLD,KEEP)
//SYSIN DD *
 SYSOUT(O)
 SYSRPTS(WLMGL(RCLASS(RLOGER)))   /* REPORT CLASS FOR THE LOGGER*/
 SYSRPTS(WLMGL(RCLASS(RIYOT1)))   /* REPORT CLASS FOR IYOT1     */
 DINTV(0005)
 SUMMARY(INT)
 /*
```

Figure 11   Sample JCL to print the RMF Workload Activity report

SMF Type 88 records and IXGRPT1 program

This section describes the SMF Type 88 records and IXGRPT1 program.

Use: Analysis of log stream activity and performance

Each SMF 88 record reports System Logger activity for one log stream or structure. SMF record Type 88 has the following subtypes:

► **Subtype 1**: Reports log stream usage. The records contain the information required for log stream or structure diagnosis and performance analysis.

► **Subtype 11**: Records CF structure alter activity. Subtype 11 records are produced when the System Logger changes the entry-to-element ratio for CF-Structure log streams.

Note the IXGRPT1 program used to format and report on the SMF Type 88 data does *not* have the ability to summarize at an interval larger than the interval used for data collection (the INTVAL value specified in the current SMFPRMxx).

SMF 88 subtype 11 records are also recorded if the size of the structure is altered using a SETXCF alter command, for example:

```
SETXCF START,ALTER,STRNM=LOG_JG2_5M,SIZE=10240
```

An SMF 88 subtype 11 record is written at the completion of the SMF interval, and when the application disconnects from the log stream.

SMF Type 88 records are mapped using the IXGSMF88 macro found in SYS1.MACLIB.

The IXGRPT1 program is supplied as PL/I source in SYS1.SAMPLIB. In addition, member IXGRPT1J is supplied as self-documenting JCL to link-edit and execute a precompiled version of the utility, IXGRPT1L, which is also supplied in SYS1.SAMPLIB.

The utility reads the SMF Type 88 records from a sequential data set and formats the data by log stream within the SMF interval.

Figure 12 provides a sample IXGRPT1 output that includes a DASD-only and a CF-Structure log stream.

```
                         BYT WRITTN  BYT WRITTN  BYT WRITTN                                        AVERAGE
                         BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------   BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE            INVOKED  TYPE1      TYPE2     TYPE3   SIZE
                         BYT DELETD  # DELETES   BYT DELETD  # DELETS  --------------EVENT--------------
                         INTERIM ST  W/O DASD    INTERIM ST    W/     OFF- DASD STRC NTRY STG  STG   RE-
                         W/O DASD    WRITE       W/DASD        WRITE  LOAD SHFT FULL FULL THLD FULL BLD
----------------------------------------------------------------------------------------------------------
12/11/02  3:05:00 AM  (SMF INTERVAL TIMESTAMP 'B8A8D68F92F00000'X)
Example 1 - DASD-only Log Stream
 IYOT1.CICS22.DFHLOG  *DASDONLY*  11240854   32935936         0     8041      0        0          0    1397
                                  30801920       7520   1368064      334      4   0    0    0    4   0   0
Example 2 - Coupling Facility Log Stream
 IYOT2.CICS22.DFHLOG  LOG_JG2_20M 34224020   49171968    425972    77860  77860        0          0     439
                                  35502061      80766    391092      872      7   0    0  121    0   0   0
```

Figure 12   Sample IXGRPT1 output

Notice the time stamp. This is related to the SMF interval, but unlike most other reports based on SMF data, the time stamp reflects the *end* of the interval, *not* the start.

The report layout takes some getting used to. Each "line" of the report is actually two lines long. Similarly, the column headings are two lines long. For example, taking the far-right

column in the report, the heading says "AVERAGE BUFFER SIZE ------- RE- BLD". This means in the far-right column, the "first" line (the one that contains the log stream name) contains the average buffer size for this log stream, and the "second" line contains the number of times the structure was rebuilt. Although this is a bit confusing at first, you will soon get used to it.

IXGRPT1 field summary and IXGSMF88 cross reference

The following information is intended to provide simple definitions for the report fields and cross reference the report columns to the actual SMF Type 88 data fields. Remember that this report only represents information for one system. If you have a log stream which has connectors on multiple systems (forward recovery log, for example), you must view the reports from all systems to get the total view of the activity against the log stream.

Note: The time stamps in the reports are given in Greenwich mean time, not local time.

The following fields should be noted:

► **Logstream name** (SMF88LSN)

Log stream name.

► **Structure Name** (SMF88STN)

Name of the structure used for this log stream. For a DASD-only log stream, this field will show *DASDONLY*.

► **BYT Written by users IXGWRITES** (SMF88LWB)

This is the total number of bytes passed to System Logger on IXGWRITE requests over the SMF interval. It does not contain any "hot air"[1] assumptions or rounding to the interim storage medium's storage boundary.

The computation is the same for CF-Structure log streams and DASD-only log streams.

► **BYT Written to interim storage** (SMF88SWB)

The amount of storage used in the "interim" storage portion of the log stream.

This is the total number of bytes passed to System Logger on IXGWRITE requests (from SMF88LWB), plus the "hot air" used to fill out the storage medium increments.

For CF-Structure log streams, this value includes the amount of element space, rounded up to the element boundary (either 256 or 512 bytes), for each IXGWRITE. For CF-Structure log streams that are duplexed to staging data sets, this field still represents the amount of space in the CF that is used; space usage for the staging data sets is not reported.

For DASD-only log streams, this value includes the amount of staging data set CI space (rounded up to the CI boundary, 4096 bytes, for each IXGWRITE) required to hold the log data written to the log stream, plus System Logger's staging data set block management control information.

A point of interest: If the same data was written to a DASD-only log stream and a CF-Structure log stream, SMF88SWB will normally be larger for the DASD-only because of the larger storage increments for DASD-only log streams (4096 bytes versus 256 or 512 bytes).

---

[1] "Hot air" is the difference between the increments of space in the interim storage (for example, 4096-byte CIs in the staging data sets) and the amount of data passed on the IXGWRITE request. For example, if 1096 bytes of data were passed on IXGWRITE for a DASD-only log stream, the "hot air," or empty space, would be 3000 bytes.

► **BYT Written to DASD** (SMF88LDB)

The number of bytes written to the offload data set or sets. This includes System Logger control information (40 bytes per log block, at the time of writing). This data is *not* rounded to a CI boundary.

► **# Writes invoked** (SMF88LWI)

The number of IXGWRITEs issued by the connectors on this system.

► **# Writes Completed**

  – **TYPE1** (SMF88SC1)

  The number of IXGWRITEs that completed normally while the log stream had not yet reached its HIGHOFFLOAD threshold and less than 90% of elements were in use. Prior to APAR OA03055, this field is only used for CF-Structure log streams. With the application of OA03055, this field will also contain data for DASD-only log streams.

  – **TYPE2** (SMF88SC2)

  The number of IXGWRITEs that completed when the log stream was at or above the HIGHOFFLOAD threshold value, but less than 90% of elements for the log stream were in use. Prior to APAR OA03055, this field is only completed for CF-Structure log streams. With the application of OA03055, this field will also contain information for DASD-only log streams.

  – **TYPE3** (SMF88SC3), applies only to CF-Structure log streams

  The number of writes completed when 90% or more of the elements for the log stream are in use. This field is not applicable to DASD-only log streams.

  This field should be zero for best performance.

► **Average Buffer size**

Average amount of data passed on each IXGWRITE request during the interval.

This information is not included in the SMF record; it is calculated by the IXGPRT1 program by dividing SMF88LWB by SMF88LWI.

► **BYT Deleted interim ST w/o DASD** (SMF88SIB)

This is the number of bytes deleted from the interim storage, including the "hot air," as a result of IXGDELET requests, where the data had not already been moved to an offload data set. For active log streams, this value should be close to SMF88SWB, meaning most data is being deleted before incurring the overhead of being moved to offload data sets.

For CF-Structure log streams, this value includes the amount of element space, rounded up to the element boundary for each IXGWRITE. The data is deleted from the CF structure without moving the data to an offload data set.

For DASD-only log streams, this value includes the amount of staging data set space, rounded up to the CI boundary for each IXGWRITE. The data is deleted from the log stream staging data set without moving the data to an offload data set.

A non-zero value in this field reflects log stream tail trimming is being initiated by the log stream connector.

► **# Deletes w/o DASD write** (SMF88SII)

The number of times a deletion from storage was performed where the data had not been moved to an offload data set.

This is another reflection of the log stream tail trimming activity.

► **BYT Deleted interim ST w/DASD** (SMF88SAB)

During the offload process, System Logger first physically deletes any logically deleted log blocks. If this is not sufficient to reduce the log stream to the LOWOFFLOAD value, it then moves log blocks, starting with the oldest, to offload data sets until the LOWOFFLOAD threshold is reached.

This field contains the number of bytes that were deleted from interim storage during this interval after the data had been moved to offload data sets.

For CF-Structure log streams, this value includes the amount of "hot air" in each element.

For DASD-only log streams, this value includes the amount of "hot air" in the staging data set CIs. It's worth noting that System Logger does not read the staging data set when moving log blocks to the offload data set; it reads the local buffer (data space) copy and writes it to the offload data set. However, the SMF88SAB value represents the amount of space used in the staging data set because this resource is of limited size.

For both CF-Structure and DASD-only active log streams that do tail trimming with the expectation the data will be held within interim storage, this field should be zero for best performance.

► **# Deletes w/write** (SMF88SAI)

The number of times a deletion from storage was performed where the data had been moved to an offload data set.

This is another reflection of insufficient log stream tail trimming activity. For optimum performance, this field should be zero.

► **EVENT**

– **OFFLOAD** (SMF88EO)

The number of offload *processes* that ran in the interval.

– **DASD SHFT** (SMF88EDS)

The number of times an additional offload data set is allocated during the interval.

Monitor this field to see if the number of offload data sets being allocated is consistent with your requirements for this log stream. For critical log streams, this field should be zero during the prime processing period for best performance and availability.

– **STRC FULL** (SMF88ESF) (CF-Structure log streams only)

Number of times a structure full condition was reached.

Following the application of OA03055, the System Logger will increment SMF88ESF when a log stream exceeds its allowable limit in the structure. The structure might not be completely full.

Prior to APAR OA03055, SMF88ESF only reflects the number of times the structure was completely full, rather than being log stream based. The change provides a better correlation of SMF88ESF and Return Code 8 Reason Code 866 errors.

This field should be zero for best performance and availability.

– **NTRY FULL** (SMF88EFS) (CF-Structure log streams only)

Number of times an offload was initiated because the number of entries in use in the structure reached 90%. Note that when this condition occurs, an offload is initiated for all log streams in the structure.

This field should be zero for best performance.

– **STG THLD** (SMF88ETT)

Number of times an IXGWRITE was issued, and the staging data set had reached or exceeded its HIGHOFFLOAD threshold.

For a CF-Structure log stream, this value should be zero. If the value is greater than zero, it indicates that the offload is being triggered because the staging data set reached its HIGHOFFLOAD threshold before the CF-Structure log stream. The size of the staging data set should be increased so that it has the capacity to hold the same or a greater amount of data compared to the CF-Structure log stream storage.

For a DASD-only log stream, the value should be equal to or slightly higher than the number of offloads in the interval. A number greater than the number of offloads in the interval indicates a number of IXGWRITE calls were issued to the log stream while it was at or above the HIGHOFFLOAD threshold. A very large number indicates a possible slow down in the offload process.

– **STG FULL** (SMF88ETF)

The number of times the staging data set filled in the interval. This field applies to both CF-Structure and DASD-only log streams.

This field should be zero for best performance and availability. A non-zero value in this field commonly indicates one of the following:

- There is a problem with the offload process, meaning the System Logger could not move the data to the offload data set quickly enough.
- The staging data is too small for the volume of data being written to this log stream.
- The HIGHOFFLOAD threshold is too high, meaning the space between HIGHOFFLOAD and the data set being full is too small.

– **RE-BLD** (SMF88ERI) (CF-Structure log streams only)

The number of times during the interval that the structure was rebuilt.

SMF Type 88 observations and possible actions

Consider the following:

▶ For DFHLOG and DFHSHUNT, the number of "BYT DELETD INTERIM ST W/O DASD" should be very close to the "BYT WRITTN TO INTERIM STORAGE". A non-zero value in "BYT DELETD INTERIM ST W/DASD" indicates that data is being offloaded and then deleted, costing additional processing and I/O. The BYT WRITTN to DASD should be very low.

Factors to be considered include:

– There are long running CICS tasks, preventing log tail deletion.

This is not simply CICS transactions with high response times, but rather long running tasks that use recoverable resources and do not issue a sync point. In addition, the tasks are causing log records to be written during each activity keypoint interval.

If message DFHLG0743 (`Tail of log stream log_stream_name deleted at block id X'nnnn'`) is not being issued for DFHLOG with each activity keypoint (message DFHRM0205), a long running task is preventing tail trimming. However, it is not unusual to see infrequent DFHLG0743 messages for DFHSHUNT. Units of work might exist on DFHSHUNT for extended periods of time. Examples are conversational tasks that have updated a recoverable resource and mirror tasks awaiting a forget flow from the connected region.

In CICS Transaction Server, the forget flow (part of 2-phase commit) is carried with the next task attach to flow across the MRO link. This improves performance on most links, but if the usage is low, the log records for the mirror might reside on DFHSHUNT for an

extended period of time. A DFHLG0743 message being issued for DFHSHUNT indicates that units of work, which had been inactive, have completed. CICS APARs PQ22563 and PQ14796 reduce the amount of data recorded on DFHSHUNT. In addition, APARs PQ56315 (1.3) and PQ56341 (2.2) should be applied to address the "forget flow" problem.

– AKPFREQ is set too high.

Use the DFHRM0205 messages to determine how often an activity keypoint is being taken.

– Problems related to interim storage:

• CF-Structure log stream structure size is too small, or staging data set cannot hold as much data as the log stream structure storage.
• Incorrect entry-to-element ratio. This can happen when log streams with very different block sizes or very different write rates are in the same structure. Dynamic changes to the entry-to-element ratio will not be made more frequently than once every 30 minutes.
• DASD-only log streams. The size of the staging data set might be insufficient.

– HIGHOFFLOAD.

Should be set no higher than 85%.

– LOWOFFLOAD:

• For DFHLOG and DFHSHUNT, set in the range of 40 - 60%.
• For user journals, set to 0%. All data should be offloaded each time an offload is initiated.

► # WRITES COMPLETED (note that prior to the installation of the PTF for APAR OA03055, this is only provided for CF-Structure log streams):

– TYPE1: Normal, this number should be high.
– TYPE2: Normal, but the log stream reached or exceeded the HIGHOFFLOAD value.
– TYPE3: Writes issued when 90% of the elements for the log stream are in use:
  • Look for HIGHOFFLOAD set to 90% or higher.
  • Tail trimming is not happening or is not being effective (see item Ê on page 19).
  • CICS is filling the space above the HIGHOFFLOAD point faster than the log stream is being offloaded.

► EVENTS:

– NTRY FULL indicates the number of times *all* log streams in the structure were offloaded due to reaching 90% of the structure entries in use.

This could be the result of the entry-to-element ratio being too large (AVGBUFSIZE is specified larger than necessary), or a poorly behaving application that is causing many small records to be written to a log stream that normally contains large records.

– OFFLOADs are good if they are triggered by the HIGHOFFLOAD value. However, offloads are bad if they are triggered by an NTRY FULL condition. In addition, for a CF-Structure log stream, offloads should be triggered because the CF-Structure log stream reached its HIGHOFFLOAD value, *not* because the staging data set reached its HIGHOFFLOAD point (see STG THLD on page 21).

– DASD Shifts indicates the number of times a new offload data set is allocated:

• For DFHLOG and DFHSHUNT, this number should be very small. A large value here indicates too much data is being offloaded (see item Ê on page 19).
• Verify that the log stream has a sufficiently large LS_SIZE value specified.
• For user journals, each offload data set should be capable of holding multiple offloads of the log stream.

- If the offload data set size has not been specified in the log stream definition (LS_SIZE) or in the SMS data class, the size will be determined by the value specified in the ALLOCxx member of SYS1.PARMLIB, which defaults to 2 tracks.

– STRC Full indicates the number of times a structure full condition was reached; this should always be 0.

– STG THLD is the number of times the HIGHOFFLOAD percentage was reached for the staging data set.

   This is good for DASD-only log streams but should not happen for CF-Structure log streams. If non-zero values are seen here for a CF-Structure log stream, the staging data set size needs to be increased so that it will hold at least as much data as the log stream in the structure. A value here for a CF-Structure log stream also indicates that the CF is either volatile or failure dependent, or that DUPLEXMODE(UNCOND) has been specified for the log stream.

– Rebuild indicates the number of structure rebuilds in the interval. If this happens on a regular basis, it needs investigation.

Performance analysis

This section describes considerations about performance for CICS log streams.

Sample workload and methodology for the test cases

Throughout the following discussion, there are references to the sample workload. The tests were run on an IBM 2064-216 with IBM ESS (2105) configured as 3390-3 DASD devices and a stand-alone CF (2064-100) and two Integrated Coupling Facilities. The z/OS LPAR had four shared CPs.

The workload consists of up to three CICS regions (CICS Transaction Server Versions 2.2 and 1.3) executing under z/OS R1.4. For the test runs, a transaction (JORM) was started through terminal input. The application reads and writes seven records from/to a VSAM file (FILEA), issues a sync point, and then repeats the process six times. It then issues an EXEC CICS Start for 10 transactions (JORA-J). Each of these transactions reads and writes seven records from/to a unique VSAM (FILEA-J), issues a sync point, and repeats the process six times. It will then issue an EXEC CICS START for itself. This will repeat up to the number of repetitions passed from JORM (x'7500'). The 10 tasks run in parallel for 30 minutes, and the region is then canceled. In cases where three regions are used, the process is the same in each region.

In some cases, the only log streams for each CICS region are DFHLOG and DFHSHUNT. The log streams are varied between a CF-Structure log stream and DASD-only. No data is ever written to DFHSHUNT during the tests. In other cases, a forward recovery log is added to each region.

Data collected includes the SMF Type 88 records, SMF Type 70 to 79 records, and the SMF Type 110 (CICS) records. The data was combined in the following charts by sample. Prior to running each test, the System Logger address space was canceled (FORCE IXGLOGR,ARM) and restarted (S IXGLOGRS). The SMF data sets were switched prior to each sample run. The system was left to idle for 10 minutes in order to establish the "idle" level of the storage used by the System Logger address space. The tests are run for 30 minutes; a 5 minute interval is then selected for the display from 15 minutes into the run.

Evaluation of the SMF Type 88 data allows us to fine-tune the log stream activity. But to carry the tuning to the next level, it's necessary to look at the system resources required to sustain

the performance. The RMF Workload Activity reports (taken from the SMF 70 to 79 records) provide the CPU used, average DASD I/O rates and response times, TCB and SRB times, and the number of storage frames consumed.

To ease analysis, a separate WLM reporting class has been defined for each CICS region and the System Logger address space. The data is formatted using the RMF Post Processor (ERBRMFPP).

Refer to Figure 10 on page 14 for a further explanation of Workload Activity reports, report classes, and sample JCL to produce the reports.

Combining the information found in the Workload Activity reports with the information from the SMF Type 88 records helps us to understand the costs associated with a given log stream configuration. CICS Performance Analyzer was used to summarize transaction data from the CICS 110 records.

The data presented in each sample begins with the SMF Type 88 data followed by a combination (handcrafted to reduce the presentation space) of data from RMF and the CICS 110 records. The RMF information is taken from the Workload Activity reports for the CICS and System Logger address spaces. The CICS-specific information is taken from summarized CICS 110 records using the CICS Performance Analyzer.

```
Type CICS  DASD APPL%  # Tran/ Average  Average  Average Average  Logr Logr  DASD  APPL% Storage Idle    Net
     TCB   I/O          Second CPU/Task Resp     JC Wait File WT TCB  SRB   I/O         Frames  Frames  Frames
DASD 0.4   8.3  0.2     .197   0.005    50.1582  50.0519 .0988   0.1  0     2.3   0.1   2571.42 1025.06 1546.36
```

Figure 13   Sample of the handcrafted reports

Using the example in Figure 13, the data on the left is for the CICS address space. The TCB, DASD I/O, and APPL% are taken from the RMF report class data.

The number of transactions per second is taken from the transaction report class.

Average CPU per task, Average response time, Average JC Wait, and Average File Wait times are taken from the CICS Performance Analyzer (PA) reports.

The data on the right starting with LOGR TCB is for the System Logger address space (IXGLOGR), taken from the RMF report class reports.

The meaning of each heading is:

Type   The log stream type.

    DASD   Indicates that this is a DASD-only log stream.

    CF64   Indicates that this is for a CF-Structure log stream defined with MAXBUFSIZE(64000).

    CF32   Indicates that this is for a CF-Structure log stream defined with MAXBUFSIZE(32000).

CICS TCB   This field reports the CICS TCB time in seconds for the interval, taken from the RMF report class information for the CICS address space.

DASD I/O   This field gives the average number of I/Os per second initiated by the region during the interval. The data is obtained from the RMF report class for the CICS region. Because the workload is based on VSAM files, this number provides a good indication of transaction throughput.

APPL%   This field reports the percent of a CP used by the region in the report interval.

# Tran/Second	The number of transactions per second is found in the RMF transaction report class and in the output of DFH0STAT.
Average CPU/Task	The average CPU per transaction is taken from the CICS PA reports.
Average Resp.	The average transaction response time is provided in the CICS PA reports and the RMF transaction report class.
Average JC Wait	The average journal control wait time is found in the CICS PA reports and provides the average time tasks spent waiting for journal requests to complete.
Average File WT	This field reports the average time spent waiting for VSAM file requests to complete.
LOGR TCB	This field reports the TCB time consumed in the System Logger address space (IXGLOGR). The information is taken from the RMF report class for the System Logger address space.
LOGR SRB	This field provides the SRB CPU time used by the System Logger address space during the interval. The information is found in the RMF report class for the System Logger address space.
DASD I/O	This field gives the average number of I/Os per second initiated by the System Logger during the interval. The data is obtained from the RMF report class for the System Logger address space. This number is a good indicator of the amount of staging and offload data set activity.
APPL%	This field reports the percent of a CP used by the System Logger address space in the report interval.
Storage Frames	This field reports the average storage frames allocated to the System Logger address space during the interval. A storage frame is a 4 K page. The data is taken from the RMF report class for the System Logger.
Idle Frames	The number of pages used when System Logger is idle, that is, prior to starting the CICS region or regions.
Net Frames	For System Logger, the increase due to CICS logging activity in the interval (represented by Net Frames) is calculated by subtracting the number of idle frames from the average storage frames.

Initially, it was difficult to define the System Logger steady state until OPERLOG was varied off. With OPERLOG running, the amount of storage required for System Logger would grow as more console messages were placed in the OPERLOG log stream. To stop OPERLOG use, so only syslog is used, issue the following z/OS command:

```
v operlog,hardcpy,off
```

Log stream characteristics and placement

The placement and size of the log stream or streams you want to tune must be understood. For example, are the staging and offload data sets allocated to DASD pools that provide the best performance? This not only includes the actual device access time, but activity that could make the device unavailable, for example, a reserve from another system or high activity data sets that could result in the request being queued.

Is there possible contention between log streams? For example, when there are multiple log streams in a given structure, do they have similar characteristics, that is, the volume and block size of the data written?

If the log streams in question reside in a Coupling Facility, determine the largest average buffer size for log streams in the structure. This value is used to determine the

entry-to-element ratio. The element size is determined by the MAXBUFSIZE value specified when the structure is defined in the System Logger policy (using IXCMIAPU). When the value is equal to or less than 65276, the element size is 256. If the value specified is greater than 65276, the element size is 512.

For example, if the element size is 256, and the IXGWRITE is for 734 bytes, three elements are required.

The entry-to-element ratio is dynamically adjusted based on a current snapshot of the usage of all log streams connected to the structure. The snapshot will be taken about every 30 minutes and the ratio adjusted if necessary.

If there are three log streams in the structure, with average buffer sizes of 1800, 1450, and 734, System Logger will use the current real time average (how many records of each size) to define the entry-to-element ratio. Assuming an element size of 256 bytes, if most of the records are in the 1800-byte range, the average would be 1:8. This means that we expect to use eight elements for each entry.

Consider the case where the application writes many short records, say 200 bytes, to the log stream, resulting in an average size of 734. In this situation, one entry is still used for each write, but only one of the eight assumed elements is used. The net effect is more entries are used than predicted, leading to an entry full condition before HIGHOFFLOAD can be reached. Offload will be triggered, but there is no extra room to write additional records until some space is recovered. In this situation, the NTRY FULL count will be equal to or greater than the number of offloads.

The important point to remember is that a log stream performs best when it is contained in a structure where all log streams have similar characteristics (that is, average buffer size and amount of data written). Typically, the log streams for TORs, AORs, and FORs are different and should not be mixed. Even among AORs, two AORs might have completely different log stream characteristics. In addition, DFHLOG and DFHSHUNT should not be placed in the same structure.

Prior to z/OS R1.2, CF writes greater than 4 KB are written asynchronously, which can increase response time. However, z/OS introduced a new heuristic algorithm for deciding whether a given CF request should be sent synchronously or asynchronously. This new algorithm is discussed in "z/OS Performance: New Heuristic Algorithm for Managing CF Request Conversion," Flash10159, available at:

    ftp://ftp.software.ibm.com/software/mktsupport/techdocs/heuristic3.pdf

The important point is that the System Logger knows whether a given IXGWRITE request is processed synchronously or asynchronously. For synchronous requests, it does not respond to the IXGWRITE requestor until the command completes; for asynchronous writes, it responds with a return code indicating the write was asynchronous and the result will be POSTed later.

Coupling Facility (CF) hardware

It's very important to understand if the log streams will be in a structure that resides in a stand-alone CF or in a CF that resides in the same CPC as one or more of the connected systems.

A stand-alone CF is considered volatile when it no longer has a battery backup, that is, in the event of a power loss, the data could not be preserved. If the CF is in the same physical CPC as the z/OS LPAR, it is considered to be failure dependent, or in the same failure domain. In either case, System Logger will (if STG_DUPLEX(YES) is specified for that log stream)

duplex the data to the staging data set rather than its data space. This ensures data integrity, but increases the cost and response time for log writes.

A symptom of this condition is when the CICS address space is initializing and messages are received indicating a staging data set is being formatted.

How log defer interval (LGDFINT) can affect performance

The LGDFINT CICS SIT parameter defines the log defer interval. The interval is the length of time, specified in milliseconds, the CICS log manager is to wait before invoking the System Logger for a forced write request. The value chosen can have an impact on transaction response time. Prior to CICS Transaction Server V2.2, the default is 30 milliseconds. With CICS Transaction Server V2.2, the default has been changed to 5 milliseconds.

To understand the effects of different LGDFINT values, we examine three samples using a DASD-only log stream; however, the results are equally applicable to CF-Structure log streams. The log stream was defined with a HIGHOFFLOAD of 80, LOWOFFLOAD of 50, and a STG_SIZE of 2518. AKPFREQ was set at 1000.

The first sample is shown in Figure 14. LGDFINT was set at 30 milliseconds, which was the default value in CICS Transaction Server V1.3.

```
                                  BYT WRITTN  BYT WRITTN  BYT WRITTN                                       AVERAGE
                                  BY USERS    TO INTERIM    TO DASD    #WRITES  ---# WRITES COMPLETED------  BUFFER
-LOGSTREAM NAME- STRUCTURE NAME--- IXGWRITES    STORAGE               INVOKED  TYPE1     TYPE2     TYPE3    SIZE
                                  BYT DELETD  # DELETES   BYT DELETD  # DELETS --------------EVENT---------------
                                  INTERIM ST  W/O DASD    INTERIM ST     W/    OFF- DASD STRC NTRY STG  STG  RE-
                                  W/O DASD    WRITE       W/DASD       WRITE   LOAD SHFT FULL FULL THLD FULL BLD
-------------------------------------------------------------------------------------------------------------------
(1) AKPFREQ 1000 LGDFINT 30 STG_SIZE 2518 HIGHOFFLOAD 80 LOWOFFLOAD 50
12/11/02  2:20:00 AM  (SMF INTERVAL TIMESTAMP 'B8A8CC80A7400000'X)
 IYOT1.CICS22.DFHLOG *DASDONLY*   3032896     9834496          0      2401        0         0          0      1263
                                 7585792     1852             0         0        1  0      0    0    1    0    0


Type CICS  DASD APPL%  # Tran/ Average   Average   Average Average Logr Logr  DASD   APPL% Storage  Idle     Net
     TCB   I/O          Second CPU/Task  Resp      JC Wait File WT  TCB  SRB   I/O          Frames   Frames   Frames
DASD 1.6   37.8 0.6     1.57   .0044     11.11     11.042  0.0608  0.1  0.1   8.4    0.1   2949.35  859.09   2090.26
                                                                   Logstream Writes/sec: 10.42
```

Figure 14   Log Defer sample 1

The STG_SIZE of 2518 equates to roughly 10.3 MB (2518 CIs at 4096 bytes per CI). There are no user journals in this test.

At first glance, the SMF Type 88 data indicates the log stream is working fine; no data is being offloaded to the offload data set (BYT DELETD INTERIM ST W/DASD is zero). There was a single staging threshold and one offload.

The problem is not the log stream function, but rather the fact very little work is being processed in the CICS region.

From the RMF data, observe that the CICS TCB time is very low (1.6 seconds), with a response time of 11.11 seconds. Also note that CICS is only issuing an average of 37.8 DASD I/O per second. This DASD I/O has nothing to do with System Logger, but instead the CICS VSAM file accesses, which are a direct reflection of the number of file control requests in the region. Finally, notice that the CICS region is only using 0.6% of an engine during this interval.

In the System Logger address space, the CPU, SRB, and APPL% values are very low, indicating a minimum of work is being passed to System Logger. The average number of DASD I/O is 8.4 per second, and the average net storage frames is 2090.26. The number of storage frames is an important value to watch, especially in an LPAR that is storage constrained. In this sample, the System Logger storage is 8.56 MB (2090.26 * 4096).

The average response time per transaction is 11.11 seconds with an average CPU per task of 0.0044 seconds. Note, of the 11.11 seconds response time, an average of 11.042 seconds is spent waiting for journal I/O to complete.

DFH0STAT reported the region was issuing 10.42 log writes per second.

In the second sample, shown in Figure 15, LGDFINT was reduced to 5 and all other parms remained the same. This change brought the transaction throughput up from 1.57 per second in sample 1 to 4 per second in sample 2, and the average response time dropped from 11.11 to 3.0288 seconds.

```
                              BYT WRITTN  BYT WRITTN  BYT WRITTN                                          AVERAGE
                              BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------    BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE               INVOKED  TYPE1     TYPE2     TYPE3     SIZE
                              BYT DELETD  # DELETES   BYT DELETD  # DELETS  ---------------EVENT---------------
                              INTERIM ST  W/O DASD    INTERIM ST    W/    OFF- DASD STRC NTRY STG  STG   RE-
                              W/O DASD      WRITE       W/DASD     WRITE   LOAD SHFT FULL FULL THLD FULL  BLD
------------------------------------------------------------------------------------------------------------
(2) AKPFREQ 1000 LGDFINT 5 STG_SIZE 2518 HIGHOFFLOAD 80 LOWOFFLOAD 50
12/11/02  3:05:00 AM  (SMF INTERVAL TIMESTAMP 'B8A8D68F92F00000'X)
 IYOT1.CICS22.DFHLOG  *DASDONLY*  11240854   32935936         0      8041      0       0           0      1397
                              30801920        7520   1368064       334      4   0    0    0    4    0    0


Type CICS  DASD APPL% # Tran/ Average   Average  Average Average Logr Logr DASD APPL% Storage  Idle     Net
     TCB   I/O         Second CPU/Task  Resp     JC Wait File WT  TCB  SRB  I/O        Frames   Frames   Frames
DASD 5.4   140.4 2.2    4     .0039     3.0288    2.9631  0.059   0.1  0.2  27.3  0.1  2850.76  859.08  1991.68
                                                                         Logstream Writes/sec:  29.85
```

Figure 15   Log Defer sample 2

When evaluating CICS transaction performance issues, it's important to remember that there are multiple parts in the transaction response time. The two main categories are execution and suspend time. Execution time is when the application, CICS management modules, exit programs, and so forth are processing as part of the transaction. In the workload used, the suspend time consists of waiting for VSAM file I/O and log/journal writes. Wait for redispatch is the portion of suspend time after the ECB is posted until the task is redispatched.

Reducing the value for LGDFINT from 30 to 5 means that each force write will only be held for 5 milliseconds (ms), rather than 30, before the IXGWRITE is issued. This reduces the opportunity for additional records to be added to the buffer, but it also reduces the time a task must wait for the log write to complete.

Due to the speed of the processor and short application path length, it's not unusual to have all 10 tasks waiting for the buffer to be written. This is true with LGDFINT set to 5, and with it set to 30, every task will be waiting a great portion of the 30 ms for each log write request. Although not shown in the reports, DFH0STAT shows 10 tasks waiting each time a sample is taken. This is reported in the number of Force Waits and the current number of waiters.

As previously noted, wait for redispatch is time from the point the ECB is posted to when the task is redispatched. Because all of the tasks are waiting for most of the buffers, when the log write completes, there will be a variety of wait for redispatch times as the tasks are resumed.

In samples 1 and 2, DFHLOG is a DASD-only log stream. The staging data set I/O response time averages about 1.2 ms. Setting LGDFINT to 30 essentially elongates the log write time to 31.2 ms, or reduces it to 6.2 ms when LGDFINT is set to 5.

The TCB time reflects the fact that the region is processing more work, along with the fact that it is now averaging 140.4 DASD I/O per second for VSAM file activity. The CPU per transaction is now 0.0039 seconds. Notice that the JC Wait time has dropped to 2.9631 seconds. This clearly indicates transaction throughput was being throttled by having the log defer interval (LGDFINT) set to 30.

The System Logger DASD I/O increased from 8.4 to 27.3 per second. The average System Logger storage for the CICS log stream decreased slightly to 8.16 MB (1991.68 * 4096).

However, this region is capable of much better throughput. As outlined earlier, the workload consists of a very small assembler application that will generate 49 log writes per task (one for each read/update and one for each sync point). The data indicates that journal wait time is still the largest component of the response time, accounting for 2.9631 seconds of the 3.0288 second response time.

Due to the workload and the journal wait time in sample 2, we decided to go against the recommendation and remove the log defer interval by setting it to 00. Figure 16 shows some interesting results.

```
                          BYT WRITTN  BYT WRITTN  BYT WRITTN                                    AVERAGE
                          BY USERS    TO INTERIM   TO DASD     #WRITES   ---# WRITES COMPLETED------  BUFFER
-LOGSTREAM NAME- STRUCTURE NAME--- IXGWRITES   STORAGE                 INVOKED   TYPE1     TYPE2     TYPE3    SIZE
                          BYT DELETD  # DELETES   BYT DELETD  # DELETS  --------------EVENT--------------
                          INTERIM ST  W/O DASD    INTERIM ST     W/     OFF- DASD STRC NTRY STG  STG  RE-
                          W/O DASD    WRITE       W/DASD       WRITE    LOAD SHFT FULL FULL THLD FULL  BLD
--------------------------------------------------------------------------------------------------------------
(3) AKPFREQ 1000 LGDFINT 00 STG_SIZE 2518 HIGHOFFLOAD 80 LOWOFFLOAD 50
12/11/02  3:50:00 AM  (SMF INTERVAL TIMESTAMP 'B8A8E09E7EA00000'X) )
 IYOT1.CICS22.DFHLOG *DASDONLY*   184275302  699846656          0  170861       0       0        0     1078
                                 701902848    171363           0       0      87   0    0   0 156  0    0



Type CICS  DASD APPL%  # Tran/ Average   Average  Average Average  Logr Logr  DASD  APPL% Storage  Idle    Net
     TCB   I/O         Second CPU/Task Resp       JC Wait File WT   TCB  SRB   I/O         Frames  Frames  Frames
DASD 85.3  2277 35.2   56.83   .0038   0.1845    0.1224  0.0556   0.2  3.9   572.8 1.8 3092.09 859.08 2233.01
                                                                           Logstream Writes/sec:  569.75
```

Figure 16   Log Defer sample 3

Setting LGDFINT to 0 means that CICS will issue the IXGWRITE immediately when the log data is appended to the buffer.

The number of IXGWRITES has increased from 8041 in sample 2 to 170861 in sample 3, an increase of over 21 fold. It's also interesting to note there are still multiple records per buffer (the average buffer size is 1078), indicating tasks are adding data to the buffers in the time it takes for each IXGWRITE to complete.

With the effective reduction in IXGWRITE time, there's a significant increase in transaction throughput. The transaction rate has increased from 4 transactions per second in sample 2 to 56.83 transactions per second.

The average response time dropped to 0.1845, with a large increase in the DASD I/O for both CICS and System Logger. The increase in CICS TCB time from 5.4 to 85.3 CPU seconds reflects a large increase in the amount of work being processed in the CICS region. Note that there is still an average of 0.1224 seconds per task spent waiting for log writes to complete.

The increase in DASD I/O in the System Logger address space clearly reflects the fact CICS is issuing 569.75 log writes per second. From the RMF data (although not shown in Figure 16 on page 27), the DASD requests in the System Logger address space are averaging 1.3 milliseconds.

The average System Logger storage frames for the CICS log streams has grown from 8.56 MB in the first sample to 9.15 MB in the third sample. The .59 MB increase is a small cost increase for the increase in throughput from 4 to 56.83.

Rerun log defer interval samples

Reviewing the numbers from the first three samples, it seems that there was still a delay in the I/O for the staging data sets. The problem seemed to be related to the fact the LPAR was not busy enough leading to a condition called "elongation of I/O." This is a condition where LPAR management has taken the CPs from the LPAR and there is a slight delay in processing the I/O completion.

Although this might seem a bit unlikely, consider the results of samples 4 and 5. Sample 6 shows very little improvement over sample 3, due to the much greater activity recorded in sample 3 with LGDFINT set to 0.

To overcome this problem a soaker job was started in the LPAR, which was constantly issuing I/O, and thus keeping the LPAR active.

The results presented in Figure 17 show the improvements of having the LPAR busier. Compared to sample 2, the transaction throughput increased from 4 transactions per second to 11.76, with the average response time dropping from 3.0288 seconds to 0.850 seconds. The CPU time per transaction also dropped from .0039 to .0030 seconds, and the journal wait time was reduced from 2.9631 to 0.773.

```
                             BYT WRITTN  BYT WRITTN  BYT WRITTN                                    AVERAGE
                             BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------  BUFFER
 -LOGSTREAM NAME- STRUCTURE NAME--- IXGWRITES  STORAGE              INVOKED  TYPE1     TYPE2     TYPE3   SIZE
                             BYT DELETD  # DELETES   BYT DELETD  # DELETS  --------------EVENT--------------
                             INTERIM ST  W/O DASD    INTERIM ST     W/    OFF- DASD STRC NTRY STG  STG   RE-
                             W/O DASD    WRITE       W/DASD       WRITE   LOAD SHFT FULL FULL THLD FULL  BLD
 ----------------------------------------------------------------------------------------------------------
 (4) AKPFREQ 1000 LGDFINT 5 STG_SIZE 2518 HIGHOFFLOAD 80 LOWOFFLOAD 50
 06/02/03  5:25:00 PM  (SMF INTERVAL TIMESTAMP 'B9831971E9300000'X)
  IYOT1.CICS22.DFHLOG *DASDONLY*  39246795   94347264         0    23033    0    0    0    0        1703
                             93634560     22859     1134592     277    12    0    0    0   12    0    0


 Type CICS  DASD APPL%  # Tran/ Average   Average  Average Average  Logr Logr DASD APPL% Storage  Idle    Net
      TCB   I/O         Second CPU/Task   Resp     JC Wait File WT  TCB  SRB  I/O        Frames  Frames  Frames
 DASD 14.1  493.8 5.7   11.76   .0030      .850     0.773  0.077    0.0  0.5  77.4  0.2  14123.2 12918.0 1205.2
```

Figure 17   Sample 4: Rerun of sample 2 with LGDFINT set to 5

The System Logger storage frames for the CICS log stream also dropped to 1205.2 (4.9 MB).

However, on the negative side, 1134592 bytes of data were offloaded prior to it being deleted, indicating, due to the increased transaction rate, the staging data set is reaching its HIGHOFFLOAD threshold faster.

Figure 18 on page 29 presents the results of rerunning sample 3 (LGDFINT set to 00), with the soaker job in the LPAR.

```
                            BYT WRITTN  BYT WRITTN  BYT WRITTN                                              AVERAGE
                            BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------       BUFFER
-LOGSTREAM NAME- STRUCTURE NAME--- IXGWRITES  STORAGE                    INVOKED  TYPE1      TYPE2      TYPE3      SIZE
                            BYT DELETD  # DELETES   BYT DELETD  # DELETS  --------------EVENT--------------
                            INTERIM ST  W/O DASD    INTERIM ST    W/     OFF- DASD STRC NTRY STG  STG     RE-
                            W/O DASD    WRITE       W/DASD      WRITE    LOAD SHFT FULL FULL THLD FULL   BLD
---------------------------------------------------------------------------------------------------------------
(5) AKPFREQ 1000 LGDFINT 00 STG_SIZE 2518 HIGHOFFLOAD 80 LOWOFFLOAD 50
06/02/03  5:45:00 PM  (SMF INTERVAL TIMESTAMP 'B9831DEA51F00000'X)
 IYOT1.CICS22.DFHLOG *DASDONLY* 196402515  848240640         0    207090       0        0          0         948
                               755175424     184369   94830592     23152     104    0    0   0 146   0     0


Type CICS  DASD APPL% # Tran/ Average   Average  Average Average Logr Logr  DASD  APPL% Storage  Idle      Net
     TCB  I/O         Second CPU/Task  Resp     JC Wait File WT  TCB  SRB   I/O         Frames   Frames   Frames
DASD 71.1 2410  30.1  57.38  .0030     0.1770   0.1116  0.0600   0.1  4.6   693.4  2.0  14118.6  12918.0  1200.6
```

Figure 18   Sample 5: Rerun of sample 3 with LGDFINT set to 0

The results are interesting because the transaction rate is only up slightly from sample 3 (Figure 16 on page 27), the CPU per transaction is down to .0030 seconds, the response time is also reduced to .1770, and the TCB time is also reduced. The other interesting point is the amount of System Logger storage dropped from 2233.01 to 1200 page frames. The reduction is over 1032 frames, which equates to over 4.2 MB storage in the System Logger space.

However, the negative fact of having a large amount of data deleted after it offloaded remains. Assuming the CISIZE of the offload data set is the default of 4096, the offloading of 94830592 bytes of data resulted in 23152 I/O requests, only to have the data deleted later by CICS. We know that the application is extremely efficient, with sync points being issued on a regular basis. This leads to the conclusion that the size of the staging data set should be increased.

In the next sample, shown in Figure 19, STG_SIZE was increased to 3500.

```
                            BYT WRITTN  BYT WRITTN  BYT WRITTN                                              AVERAGE
                            BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------       BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE                     INVOKED  TYPE1      TYPE2      TYPE3      SIZE
                            BYT DELETD  # DELETES   BYT DELETD  # DELETS  --------------EVENT--------------
                            INTERIM ST  W/O DASD    INTERIM ST    W/     OFF- DASD STRC NTRY STG  STG     RE-
                            W/O DASD    WRITE       W/DASD      WRITE    LOAD SHFT FULL FULL THLD FULL   BLD
---------------------------------------------------------------------------------------------------------------
(6) AKPFREQ 1000 LGDFINT 00 STG_SIZE 3500 HIGHOFFLOAD 80 LOWOFFLOAD 50
06/02/03  6:35:00 PM  (SMF INTERVAL TIMESTAMP 'B983291757D00000'X)
 IYOT1.CICS22.DFHLOG *DASDONLY* 196979915  846868480         0    206755       0        0          0         952
                               854085632     208517         0         0      76    0    0   0 148   0     0


Type CICS  DASD APPL% # Tran/ Average   Average  Average Average Logr Logr  DASD  APPL% Storage  Idle      Net
     TCB  I/O         Second CPU/Task  Resp     JC Wait File WT  TCB  SRB   I/O         Frames   Frames   Frames
DASD 71.5 2418  30.2  57.57  .0030     0.1770   0.1110  0.0600   0.1  4.6   692.2  2.0  14603.3  12918.0  1685.3
```

Figure 19   Sample 6: Rerun of sample 3 with STG_SIZE set to 3500

The log stream is now operating in a more cost-effective manner. Notice that the data is no longer being deleted after it is offloaded. There is an increase in the System Logger storage associated with the log stream; however, this is more than offset by saving the 23 K I/O requests every 5 minutes.

Conclusion: Reducing the log defer interval can have a very positive impact on transaction throughput and log stream activity. However, the size of the log stream is key to retaining the backout and recovery data within interim storage.

These tests were run on a IBM @server zSeries® 900 Model 2xx (Turbo) processor. The use of LGDFINT=0 on a slower processor might not show the same improvements.

> **Tip:** We recommend that you start with a LGDFINT value of 5 milliseconds for all releases. A value of 00 might be valid in some cases based on the speed of the processor and the overall CICS workload.
>
> The log defer interval value (LGDFINT) can be changed using CEMT I SYS. The evaluation of the results is much easier if the changes coincide with an SMF interval.

How activity keypoint frequency (AKPFREQ) can affect performance

The activity keypoint frequency (AKPFREQ) and log defer interval (LGDFINT) can have a significant impact on log stream performance.

Although activity keypoint frequency was used in CICS/ESA® Version 4.1, it has taken on a more significant role in CICS Transaction Server. In CICS/ESA V4.1, AKPFREQ was a measure of the physical I/O issued to the system journal (DFHJ01). In CICS Transaction Server, it is a measure of the number of appends to the log stream buffer for DFHLOG. In CICS/ESA V4.1, it was common practice to specify AKPFREQ to provide an activity keypoint every 15 minutes. In CICS Transaction Server, the value should be set to ensure an activity keypoint is taken more frequently.

It's important to remember that tail trimming is initiated for DFHLOG and DFHSHUNT during the CICS activity keypoint process. The value specified for AKPFREQ defines the frequency an activity keypoint has taken, thus the frequency of tail trimming, which has a major impact on the data residency time. Trimming the tail of the log removes records that are no longer required for backout and recovery, thus making space available in interim storage.

With a larger AKPFREQ value, the length of time between activity keypoints increases, thus additional data is retained in interim storage. Either data will be offloaded, costing I/O and overhead to the offload data sets, or the size of interim storage will need to be increased, leading to additional storage requirements in the System Logger data space.

In a very busy region, it is not uncommon to observe an activity keypoint being taken every minute.

To understand the effects of AKPFREQ, we examine three samples, all using a CF-Structure log stream; the results are equally applicable to DASD-only log streams. Example 7 is shown in Figure 20 on page 31. In this sample, AKPFREQ was set to 7500, LGDFINT 00, and the log stream was connected to a 10 MB structure with HIGHOFFLOAD set to 80 and LOWOFFLOAD set to 50.

```
                          BYT WRITTN  BYT WRITTN  BYT WRITTN                                   AVERAGE
                          BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------  BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE                INVOKED  TYPE1    TYPE2      TYPE3    SIZE
                          BYT DELETD  # DELETES   BYT DELETD  # DELETS  --------------EVENT--------------
                          INTERIM ST  W/O DASD    INTERIM ST    W/     OFF- DASD STRC NTRY STG  STG  RE-
                          W/O DASD    WRITE       W/DASD      WRITE   LOAD SHFT FULL FULL THLD FULL  BLD
_____
(7) AKPFREQ 7500 LGDFINT 00 10M structure HIGHOFFLOAD 80 LOWOFFLOAD 50
08/25/03  1:55:00 PM  (SMF INTERVAL TIMESTAMP 'B9EC872C55500000'X)
IYOT1.CICS22.DFHLOG LOG_JG2_10M   332803618  571351296   137814804  1169531  951494    174502     43536         284
                          278135862   1004842   121753484   401533    324  18   0  ****   0   0    0

Type CICS  DASD  APPL% # Tran/ Average   Average  Average Average Logr Logr  DASD  APPL% Storage  Idle    Net
     TCB   I/O         Second CPU/Task  Resp     JC Wait File WT  TCB  SRB   I/O         Frames  Frames  Frames
CF64 179.0 3550  65.04 65.04  .0063     0.1183   0.0068  0.1032   0.5  12.2  101.1  6.0  14906.1  8933    5973.1
                                                                        Logstream Writes/sec: 2981.83
```

Figure 20   Sample 7 with AKPFREQ set to 7500

Examination of the SMF Type 88 data reveals 121753484 bytes were deleted after having been moved to the offload data set. If the offload data set CISIZE is 4096, this equates to 29,725 I/Os to move the data, which is deleted soon thereafter. The number of NTRY FULL conditions was so large that it does not fit in the report, symbolized by ****. Both conditions suggest either that the data residency time is too long or that the log stream size is too small. If this log stream was DASD-only, the symptom would probably be STG FULL.

In the information recorded in the RMF report classes and the CICS SMF 110 records, CICS is processing 65.04 transactions per second with an average CPU of .0063 seconds. The CICS region is using 65.04% of an engine at 179 CPU seconds in the 5 minute interval.

The number of log stream writes per second as reported by DFH0STAT is 2981.83.

The RMF information collected for the System Logger in this period shows a very low TCB time (which is normal) and 12.2 CPU seconds of SRB time, using about 6.0% of an engine. The average number of storage frames is 5973.1 (24.465 MB).

Example 8 (Figure 21 on page 32) shows the effect of reducing AKPFREQ to 4000. Examination of the SMF Type 88 data for this sample shows that the data is no longer being deleted after having been moved to the offload data set, with no negative events. The number of offload processes has decreased to 302, indicating better use of interim storage. This is all the result of trimming the tail of log stream more often, thus making the offload process more efficient, that is, reaching the LOWOFFLOAD point without the need to move data to the offload data sets.

```
                              BYT WRITTN  BYT WRITTN  BYT WRITTN                                       AVERAGE
                              BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------  BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE              INVOKED  TYPE1     TYPE2     TYPE3     SIZE
                              BYT DELETD  # DELETES  BYT DELETD  # DELETS  ---------------EVENT---------------
                              INTERIM ST  W/O DASD   INTERIM ST     W/      OFF- DASD STRC NTRY STG  STG   RE-
                              W/O DASD    WRITE      W/DASD        WRITE    LOAD SHFT FULL FULL THLD FULL  BLD
_____
(8) AKPFREQ 4000 LGDFINT 00 20M structure HIGHOFFLOAD 80 LOWOFFLOAD 50
08/25/03  2:05:00 PM  (SMF INTERVAL TIMESTAMP 'B9EC896889B00000'X)
IYOT1.CICS22.DFHLOG LOG_JG2_10M 330844117  567308032        0  1159263  1147390    11873        0       285
                               330323643    1157378         0        0      302   0   0   0   0   0   0

Type CICS  DASD  APPL% # Tran/ Average   Average  Average Average Logr Logr  DASD  APPL% Storage  Idle   Net
     TCB   I/O         Second  CPU/Task  Resp     JC Wait File WT TCB  SRB   I/O         Frames   Frames Frames
CF64 173.7 3529  66.1  73      0.0053    0.1191   0.0069  0.1035  0.2  15.3  90.1  5.3   15286    8933   6353
                                                                       Logstream Writes/sec: 3370.32
```

Figure 21   Example 8 with AKPFREQ set to 4000

The RMF information and the CICS SMF 110 records show the transaction rate is now 73 per second, and the average CPU time remains constant. The CICS region used 66.1% of an engine with 173.7 TCB seconds CPU seconds in the 5 minute interval.

The number of log stream writes per second as reported by DFH0STAT is 3370.32.

The System Logger TCB, SRB, and APPL% remained constant for the period.

Example 9 (Figure 22) shows the results of further reducing AKPFREQ to 2500. The SMF Type 88 data shows the results of trimming the tail even more frequently. The log stream continues to run very well. The number of offloads is down, the number of TYPE2 writes is down, and there are no negative events recorded.

```
                              BYT WRITTN  BYT WRITTN  BYT WRITTN                                       AVERAGE
                              BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------  BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE              INVOKED  TYPE1     TYPE2     TYPE3     SIZE
                              BYT DELETD  # DELETES  BYT DELETD  # DELETS  ---------------EVENT---------------
                              INTERIM ST  W/O DASD   INTERIM ST     W/      OFF- DASD STRC NTRY STG  STG   RE-
                              W/O DASD    WRITE      W/DASD        WRITE    LOAD SHFT FULL FULL THLD FULL  BLD
_____
(9) AKPFREQ 2500 LGDFINT 00 20M structure HIGHOFFLOAD 80 LOWOFFLOAD 50
08/25/03  2:15:00 PM  (SMF INTERVAL TIMESTAMP 'B9EC8BA4BE100000'X)
IYOT1.CICS22.DFHLOG LOG_JG2_5M 332796138  570890752        0  1167419  1155716    11702        0       285
                               332663638    1166945         0        0      243   0   0   0   0   0   0

Type CICS  DASD  APPL% # Tran/ Average   Average  Average Average Logr Logr  DASD  APPL% Storage  Idle   Net
     TCB   I/O         Second  CPU/Task  Resp     JC Wait File WT TCB  SRB   I/O         Frames   Frames Frames
CF64 178.9 3545  66.2  77.38   0.0053    0.1185   0.0069  0.1030  0.2  15.2  92.1  5.2   14844.5  8933   5911.5
                                                                       Logstream Writes/sec: 3524.03
```

Figure 22   Example 9 with AKPFREQ set to 2500

The RMF information and the CICS SMF 110 records show that the transaction rate is now 77.38 per second. The CICS region used 66.2% of an engine, at 178.9 CPU seconds of TCB time in the 5 minute interval.

The number of log stream writes per second as reported by DFH0STAT is 3524.03.

The System Logger TCB, SRB, and APPL% remained constant for the period.

Conclusion: Reducing the AKPFREQ value caused CICS to trim the tail of the log more frequently. Trimming the tail more often makes it possible for the offload process to reach the LOWOFFLOAD point without the overhead of issuing I/O to move data to the offload data sets.

The number of offload processes run in each interval decreased as AKPFREQ was decreased, indicating better use of interim storage.

The negative side of reducing the AKPFREQ value is that it increases the number of CSKP transactions (activity keypoint) that run in the interval.

There is a balance point between the number of activity key points taken, the log defer interval, and the size of the log stream interim storage. Each should be considered when evaluating log stream performance.

Not every region will respond in this manner, so each region should be evaluated individually. The number of activity keypoints should also be monitored using the DFHRM0205 messages. The intent is not to cause an excessive number of CSKP tasks to run, but to optimize the log stream operation. In a busy region, a target of one activity keypoint every 40 to 60 seconds is good.

Also, each DFHRM0205 message should be accompanied by a DFHLG0743 message indicating the log is being trimmed. If log tail trimming is not happening consistently, the transactions active during the interval should be investigated to understand why sync points are not being issued.

> **Tip:** When changing the AKPFREQ (activity keypoint frequency) value using CEMT I SYS, it is best to make the changes coincide with an SMF interval, making it easier to evaluate the results of the change.

Multiple DASD-only log streams considerations

The next step is the examination of multiple regions running the workload within the same z/OS image. Figure 19 on page 29 provides a single region baseline. Examples 10 and 11 (Figure 23 on page 34 and Figure 24 on page 34) assist in the analysis of a three region scenario. Each CICS region is executing the same workload that is used in the preceding samples.

In sample 10, the three regions have AKPFREQ set to 1000, LGDFINT is 5, HIGHOFFLOAD 80, LOWOFFLOAD 50, and STG_SIZE is 3500.

For the samples where three regions were part of the test, the results are provided for only the IYOT1 region, to conserve space. The results in each region portray the same story.

The data presented in Figure 23 on page 34 is the result of setting STG_SIZE to 3500 for three identical regions (IYOT1, IYOT2, and IYOT3). Each region was executing the same workload. LGDFINT was set to 0 in the sample shown in Figure 19 on page 29, while in this case, it was set to 5, which explains the reduction in workload for this region.

```
                                  BYT WRITTN  BYT WRITTN  BYT WRITTN                                         AVERAGE
                                  BY USERS   TO INTERIM   TO DASD   #WRITES  ---# WRITES COMPLETED------    BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE                INVOKED  TYPE1     TYPE2     TYPE3       SIZE
                                  BYT DELETD # DELETES  BYT DELETD  # DELETS ---------------EVENT---------------
                                  INTERIM ST W/O DASD   INTERIM ST    W/     OFF- DASD STRC NTRY STG  STG    RE-
                                  W/O DASD   WRITE      W/DASD       WRITE   LOAD SHFT FULL FULL THLD FULL   BLD
-----------------------------------------------------------------------------------------------------------------
(10) 3 DASD-only regions AKPFREQ 1000 LGDFINT 5 STG_SIZE 3500 HIGHOFFLOAD 80 LOWOFFLOAD 50
06/03/03  1:15:00 AM  (SMF INTERVAL TIMESTAMP 'B983827F86D00000'X)
 IYOT1.CICS22.DFHLOG *DASDONLY*   42072697   107237376          0   26179        0        0          0      1607
                                  103055360  25157      1253376     306      9    0    0    0   12    0     0

Type CICS  DASD APPL% # Tran/ Average  Average  Average Average Logr Logr  DASD  APPL% Storage  Idle     Net
     TCB   I/O        Second CPU/Task Resp      JC Wait File WT TCB  SRB   I/O         Frames   Frames   Frames
DASD 15.4  528.8 6.3  12.60   .0029    0.8111    0.7204 0.0846  0.0  1.8   268.6  0.8  18097.5  12949.2  5148.3
```

Figure 23   Three DASD-only CICS regions with LGDFINT set to 5

After the data was collected with LGDFINT set to 5, the value was set to 00 in each of the regions. Figure 24 presents the results.

```
                                  BYT WRITTN  BYT WRITTN  BYT WRITTN                                         AVERAGE
                                  BY USERS   TO INTERIM   TO DASD   #WRITES  ---# WRITES COMPLETED------    BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE                INVOKED  TYPE1     TYPE2     TYPE3       SIZE
                                  BYT DELETD # DELETES  BYT DELETD  # DELETS ---------------EVENT---------------
                                  INTERIM ST W/O DASD   INTERIM ST    W/     OFF- DASD STRC NTRY STG  STG    RE-
                                  W/O DASD   WRITE      W/DASD       WRITE   LOAD SHFT FULL FULL THLD FULL   BLD
-----------------------------------------------------------------------------------------------------------------
(11) 3 DASD-only regions AKPFREQ 1000 LGDFINT 00 STG_SIZE 3500 HIGHOFFLOAD 80 LOWOFFLOAD 50
06/03/03  1:25:00 AM  (SMF INTERVAL TIMESTAMP 'B98384BBBB300000'X)
 IYOT1.CICS22.DFHLOG *DASDONLY*   158089404  772702208         0   188648       0        0          0       838
                                  728211456  177786     39841792    9727     66    0    0    0  122    0     0


Type CICS  DASD APPL% # Tran/ Average  Average  Average Average Logr Logr  DASD  APPL% Storage  Idle     Net
     TCB   I/O        Second CPU/Task Resp      JC Wait File WT TCB  SRB   I/O         Frames   Frames   Frames
DASD 64.2  1925 27.6  45.83   .0033    0.2289    0.1199 0.0996  0.2  12.9  1910   5.6  18197.5  12949.2  5248.3
```

Figure 24   Three DASD-only CICS regions with LGDFINT set to 0

The results are a little surprising. Because the regions are running in the same configuration used in sample 6 (Figure 19 on page 29), the expectation was the results for this test would be the same. However, notice that the transaction throughput dropped to 45.83 per second.

A quick look at the Workload Activity report for the CICS region in the interval (Figure 25 on page 35) reveals that this is not a System Logger issue, but a DASD contention problem. The DASD response time has gone from 1.3 milliseconds to 1.9 with the average Q+Pend time being 0.9 milliseconds.

```
 z/OS V1R4              SYSPLEX WSCZPLEX          START 06/02/2003-21.20.00 INTERVAL 000.05.00   MODE = GOAL
                        RPT VERSION V1R2 RMF        END   06/02/2003-21.25.00

                                        POLICY ACTIVATION DATE/TIME 05/07/2003 14.14.58

REPORT BY: POLICY=WLMPOL                        REPORT CLASS=RIYOT122
                                        DESCRIPTION =


TRANSACTIONS    TRANS.-TIME HHH.MM.SS.TTT  --DASD I/O-  --SERVICE---  -SERVICE RATES-  PAGE-IN RATES  ----STORAGE----
AVG    1.00  ACTUAL               0  SSCHRT 1925  IOC    57810  ABSRPTN  50594  SINGLE     0.0  AVG    7628.01
MPL    1.00  EXECUTION            0  RESP     1.9  CPU   840470  TRX SERV 50594  BLOCK      0.0  TOTAL  7627.96
ENDED     0  QUEUED               0  CONN     1.0  MSO    14086K  TCB       64.2  SHARED     0.0  CENTRAL 7627.96
END/S  0.00  R/S AFFINITY         0  DISC     0.0  SRB   193389  SRB       14.8  HSP        0.0  EXPAND    0.00
#SWAPS    0  INELIGIBLE           0  Q+PEND   0.9  TOT   15178K  RCT        0.0  HSP MISS   0.0
EXCTD     0  CONVERSION           0  IOSQ     0.0  /SEC   50594  IIT        3.8  EXP SNGL   0.0  SHARED    1.00
AVG ENC 0.00 STD DEV              0                             HST        0.0  EXP BLK    0.0
REM ENC 0.00                                                    APPL %    27.6  EXP SHR    0.0
MS ENC  0.00
```

Figure 25   RMF Workload Activity report

As it turns out, the DASD contention was the result of the CICS VSAM file placement, such that files required for each region were on the same packs.

The numbers reported for the System Logger address space reflect the fact it's handling log streams from two additional regions. Even considering that the System Logger is handling requests from three regions, the amount of DASD I/O is up significantly. The amount of data being offloaded is consistent across each of the regions. The increase in logging activity and the offload activity is driving the DASD I/O rate to 1053 per second. There are no staging full or other error indicators lists under the EVENTS columns. Dividing the number in the STG THLD column by the number of offloads in the interval helps you understand the number of times an I/O to the staging data set was issued while at or above the HIGHOFFLOAD point. With the addition of OA03055, this condition will be reflected in the TYPE2 writes information in the same way it is reported for CF-Structure log streams.

A large ratio of STG THLD to OFFLOADS (or TYPE2 writes) indicates CICS is issuing many IXGWRITEs during the period of time that the log stream is at or above the HIGHOFFLOAD threshold value. This could indicate possible delays in the offload process. The SMF data for the System Logger address space indicates that there are also delays in its DASD I/O time (Q+PEND + DISC = 0.3 milliseconds).

Remember, staging data sets are deallocated when CICS disconnects from the log stream, so information about data set placement must be obtained while CICS is connected.

Another important factor is the increase in the storage associated with the System Logger address space. The average number of storage frames in use for all three regions (DFHLOG and DFHSHUNT for three regions) is 5248.3 frames (20.5 MB). The number of System Logger storage frames in use is slightly more than three times the number reported in sample 6 (Figure 19 on page 29), which is to be expected.

Note: These tests were run on a z900 turbo processor. The use of LGDFINT=0 on a slower processor could create enough overhead to offset any gains in throughput.

Conclusions: When working with DASD-only log streams, the settings for AKPFREQ, LGDFINT, and the size specified for STG_SIZE determine the performance characteristics of CICS system log streams (DFHLOG and DFHSHUNT). The greater the number of CICS

regions, the more System Logger activity. The number of storage frames used by the System Logger is equal to the data duplexed in the data space for all active log streams.

For user log streams (user and auto journals, and forward recovery log streams), the STG_SIZE and LS_SIZE for offload data sets and data set placement determine the performance characteristics of the log stream.

Although not specified as part of the log stream definitions, the physical data set placement might have an impact on performance. Data set placement is influenced by the SMS data classes and Automatic Class Selection routines (ACS).

CF-Structure log streams

The next several examples use the same methodologies to evaluate the performance of CF-Structure log streams, noting additional factors to be considered.

In looking at the CF-Structure log streams, the first sample is a single DFHLOG log stream that resides in a 10 MB structure. Notice the structure name indicates that it is a 5 MB structure (LOG_JG2_5M); the structure was defined with INITSIZE 5 MB and SIZE as 10 MB.

With the changes implemented as part of CF Level 12, structures to be used by System Logger should be allocated with larger amounts of storage. For additional information, refer to *z/OS MVS Setting Up a Sysplex,* SA22-7625.

Observations indicate a guideline would be to add 4 MB plus 20% of the original pre-CF Level 12 structure to estimate the new size. At older CF Levels, the tests were run with 5 MB, and with CF Level 12, 9 MB was needed. Note that this rule is not hard and fast, but rather a way to estimate changes that might be required when migrating to CF Level 12.

For most of the CF tests, AKPFREQ was set at 1000, and the LGDFINT value was varied.

In the first test, AKPFREQ was set to 1000, LGDFINT set to 5, and the structure has been altered, with a SETXCF command, to 10 MB:

```
SETXCF START,ALTER,STRNM=log_jg2_5m,size=10240
```

The results are shown in Figure 26.

```
                            BYT WRITTN  BYT WRITTN  BYT WRITTN                                      AVERAGE
                            BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------ BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE            INVOKED  TYPE1      TYPE2     TYPE3   SIZE
                            BYT DELETD  # DELETES  BYT DELETD  # DELETS  --------------EVENT--------------
                            INTERIM ST  W/O DASD   INTERIM ST     W/     OFF- DASD STRC NTRY STG  STG  RE-
                            W/O DASD    WRITE      W/DASD       WRITE   LOAD SHFT FULL FULL THLD FULL  BLD
---------------------------------------------------------------------------------------------------------
(12) AKPFREQ 1000 LGDFINT 5 10M HIGHOFFLOAD 80 LOWOFFLOAD 50
06/03/03  4:15:00 AM  (SMF INTERVAL TIMESTAMP 'B983AABB35900000'X)
 IYOT1.CICS22.DFHLOG LOG_JG2_5M   9077199   11419904         0    13621 13559      62        0       666
                                 7302499      10952         0        0    3    0    0    0    0    0    0


Type CICS  DASD APPL%  # Tran/ Average   Average  Average Average  Logr Logr  DASD APPL% Storage  Idle     Net
     TCB   I/O          Second CPU/Task  Resp     JC Wait File WT  TCB  SRB   I/O        Frames  Frames  Frames
CF64 4.0  108.7 1.5     2.50   .0041     3.9085   3.8194  .0820    0.0  0.1   0.0   0.1  14776.0 12957.0 1819.0
                                                                                        44.90 Log writes per second
```

Figure 26   Coupling Facility 10 MB structure: Log Defer 5

In the sample presented in Figure 26, there is a single region with the DFHLOG log stream defined to a 10 MB structure. MAXBUFSIZE is set at 64000 (meaning the element size will be

256 bytes). DFHLOG is the only log stream in the LOG_JG2_5M structure. DFHSHUNT was defined as a DASD-only log stream. It should be noted there was no activity to DFHSHUNT during the sample.

AKPFREQ is set to 1000, LGDFINT to 5, HIGHOFFLOAD to 80, and LOWOFFLOAD to 50. See Example 1.

Example 1   IXCMIAPU output

```
STRUCTURE NAME(LOG_JG2_5M) LOGSNUM(1)
          MAXBUFSIZE(64000) AVGBUFSIZE(400)
          EFFECTIVE AVERAGE BUFFER SIZE(400)


          LOGSTREAM NAME                      CONNECTION
          --------------                      ----------
          IYOT1.CICS22.DFHLOG                 NO

          LOGSTREAMS CURRENTLY DEFINED TO THIS STRUCTURE(1)

LOGSTREAM NAME(IYOT1.CICS22.DFHLOG) STRUCTNAME(LOG_JG2_5M) LS_DATACLAS(MVSLOGR)
          LS_MGMTCLAS() LS_STORCLAS() HLQ(GRAUEL) MODEL(NO) LS_SIZE(5000)
          STG_MGMTCLAS() STG_STORCLAS() STG_DATACLAS() STG_SIZE(0)
          LOWOFFLOAD(50) HIGHOFFLOAD(80) STG_DUPLEX(YES) DUPLEXMODE(COND)
          RMNAME() DESCRIPTION() RETPD(0) AUTODELETE(NO) OFFLOADRECALL(NO)
          DASDONLY(NO) DIAG(YES) LOGGERDUPLEX(COND) EHLQ(NO_EHLQ)
```

The SMF Type 88 data shown in Figure 26 on page 36 indicates that the log stream is working well, but the RMF data reports the CICS TCB time is very low (4.0 seconds in the 5 minute SMF interval) and the transaction throughput is only 2.5 per second.

The response time is 3.9085 seconds, of which 3.8194 seconds is JC Wait time, indicating the transactions are waiting for log activity.

Figure 27 contains a portion of log stream information collected during the test, using DFH0STAT. There have been 41,942 IXGWRITE requests in the interval, but there were 94,881 Force Waits, indicating multiple CICS log records are being written in a single IXGWRITE call. Now, this is good from a buffering standpoint, but will tend to have a negative impact on response time.

It's also interesting to note the Peak Waiters value was 11. Refer to "Sample workload and methodology for the test cases" on page 21 for a discussion of the workload. There are a maximum of 10 user tasks active at any given time, so to have a depth of 11 waiting for a log write to complete would indicate that there was at least one point in time when all user tasks were waiting for the log write to complete. The 11th task is the CSKP task performing an activity keypoint.

```
Logstreams - Requests
_____

                      Write                  Average   Buffer     Buffer     Force    Current   Peak     Retry
Logstream Name        Requests  Bytes Written  Bytes   Appends    Full Waits  Waits    Waiters  Waiters  Errors
_____

IYOT1.CICS22.DFHLOG    41,942    27,643,879     659   114,823           0    94,881        7       11       0
IYOT1.CICS22.DFHSHUNT       0             0       0         0           0         0        0        0       0
```

Figure 27   DFH0STAT log stream requests

The length of time CICS will wait before issuing the IXGWRITE call is defined by the LGDFINT parm. Based on this information, LGDFINT should be reduced to improve throughput.

For sample 13, shown in Figure 28, LGDFINT was again reduced to 00.

```
                          BYT WRITTN  BYT WRITTN  BYT WRITTN                                            AVERAGE
                           BY USERS  TO INTERIM    TO DASD    #WRITES  ---# WRITES COMPLETED------  BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES   STORAGE                 INVOKED   TYPE1     TYPE2     TYPE3     SIZE
                          BYT DELETD  # DELETES  BYT DELETD  # DELETS  --------------EVENT--------------
                          INTERIM ST  W/O DASD   INTERIM ST     W/     OFF- DASD STRC NTRY STG  STG   RE-
                          W/O DASD     WRITE      W/DASD       WRITE   LOAD SHFT FULL FULL THLD FULL  BLD
--------------------------------------------------------------------------------------------------------
(13) AKPFREQ 1000 LGDFINT 0 10M HIGHOFFLOAD 80 LOWOFFLOAD 50
06/03/03  4:25:00 AM  (SMF INTERVAL TIMESTAMP 'B983ACF769F00000'X)
 IYOT1.CICS22.DFHLOG LOG_JG2_5M    411127971  699720448         0  1419492  1405697     13795         0       289
                          410859909    1418563             0        0      250  0   0    0    0   0    0


Type CICS  DASD APPL% # Tran/ Average   Average  Average Average  Logr Logr  DASD  APPL% Storage  Idle    Net
     TCB   I/O         Second CPU/Task  Resp       JC Wait File WT TCB  SRB   I/O        Frames   Frames  Frames
CF64 206.8 4378 75.8  104.25  .0061    .0990     .0062   .0812    0.2  19.1  0.0   6.4  16555.8  12957.0 3598.8
.
                                                                              1,522.94 Log writes per second
```

Figure 28   Coupling Facility 10 MB structure: Log Defer 0

Once again the 5 MB structure was expanded to 10 MB prior to the test.

Comparing sample 13 (Figure 28) with sample 12 (Figure 26 on page 36), it's easy to see the improvement. Transactions are now running at a rate of 104.25 transactions per second, with an average response time of .0990 seconds. JC Wait time has been reduced to .0062 seconds.

In the SMF Type 88 data, we see 1419412 IXGWRITE calls were issued in the 5 minute SMF interval, of which 1405697 were TYPE1, and only 13795 were TYPE2. The TYPE2 requests are defined as IXGWRITE requests that completed when the log stream is at or above the HIGHOFFLOAD threshold value. Refer to "IXGRPT1 field summary and IXGSMF88 cross reference" on page 16.

It's also interesting to note that there were 250 offload processes run during the 5 minute interval. Because no data is being physically moved to the offload data sets (BYT DELETED INTERIM ST W/DASD is 0), the offload process executes very quickly.

The System Logger storage used for this log stream is 3598.8 frames (14.7 MB).

On the CICS side, the TCB time is 206.8 CPU seconds out of a possible 300 for the interval. RMF reports the APPL% as 75.8% of an engine for the interval. Clearly there is room for additional growth. The CPU per transaction has increased significantly from sample 12 due to the use of synchronous writes to the CF.

DFH0STAT reports CICS was issuing 1,522.94 log writes per second during the sampled interval.

Figure 29 on page 39 presents the Coupling Facility Activity report showing the data for structure LOG_JG2_5M in the sample 13 time frame.

```
                    C O U P L I N G   F A C I L I T Y   A C T I V I T Y

      z/OS V1R4           SYSPLEX WSCZPLEX           DATE 06/03/2003          INTERVAL 005.00.000
                          RPT VERSION V1R2 RMF       TIME 00.20.00           CYCLE 01.000 SECONDS
  -----------------------------------------------------------------------------------------------------------
   COUPLING FACILITY NAME = CF103
  -----------------------------------------------------------------------------------------------------------
                            COUPLING  FACILITY  STRUCTURE  ACTIVITY
  -----------------------------------------------------------------------------------------------------------
   STRUCTURE NAME = LOG_JG2_5M     TYPE = LIST   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

   SYSD     1417K   SYNC   1328K  93.8   42.8    55.2     NO SCH   0    0.0    0.0      0.0     0.0
            4723    ASYNC   88K    6.2  422.4   480.4     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
  -----------------------------------------------------------------------------------------------------------
   PROCESSOR SUMMARY
  -----------------------------------------------------------------------------------------------------------
   COUPLING FACILITY   2064   MODEL 100   CFLEVEL 12
   AVERAGE CF UTILIZATION  (% BUSY)        18.4    LOGICAL PROCESSORS:   DEFINED   1    EFFECTIVE   1.0


                             COUPLING  FACILITY  USAGE  SUMMARY
  -----------------------------------------------------------------------------------------------------------
   STRUCTURE SUMMARY
  -----------------------------------------------------------------------------------------------------------
                                    % OF              % OF   AVG   LST/DIR  DATA      LOCK     DIR REC/
        STRUCTURE             ALLOC  CF           #   ALL    REQ/  ENTRIES  ELEMENTS  ENTRIES  DIR REC
   TYPE NAME       STATUS CHG SIZE   STORAGE  REQ    REQ    SEC   TOT/CUR  TOT/CUR   TOT/CUR  XI'S
   LIST DSNZPLEX_SCA ACTIVE    32M    1.6       0    0.0    0.00   43K      86K       N/A      N/A
                                                                   128      819       N/A      N/A

        LOG_JG2_5M  ACTIVE    10M    0.5    1417K   100   4722.9  8278     16K       N/A      N/A
                                                                  4040      8055      N/A      N/A
```

Figure 29   Coupling Facility Activity report for the sample 13 period

CF103 is a stand-alone CF and is performing very well. Notice that the structure (LOG_JG2_5M) for the CICS log streams has issued 100% of the requests to this CF during this interval.

Of the requests, 93.8% are synchronous, which means from the time the IXGWRITE request is issued in the CICS logger, traveling through System Logger, through XES, to the CF, and back to CICS is considered as one very long instruction. The process is run under the CICS QR TCB, so the time is all considered CPU time against the transaction.

There are a number of reasons why the request might be converted to an asynchronous request. When this happens, wait/post logic is invoked, possibly reducing the CPU time attributed to the task.

Multiple CF-Structure log streams considerations

The next three samples were run with three CICS regions (IYOT1, IYOT2, and IYOT3) running the identical workload in the same LPAR. IYOT1 and IYOT2 are CICS Transaction Server Version2.2 and IYOT3 is CICS Transaction Server Version 1.3. Only information for IYOT1 is presented.

The DFHLOG log streams for each region were connected to a 20 MB structure, which was altered to 30 MB for the third sample. The DFHSHUNT log streams for IYOT1 and IYOT2 are DASD-only, and DFHSHUNT for IYOT3, through a system programmer oversight, was assigned to a separate 20 MB structure. In all cases, DFHSHUNT has no activity.

```
                             BYT WRITTN  BYT WRITTN  BYT WRITTN                                       AVERAGE
                             BY USERS    TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------   BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE               INVOKED  TYPE1     TYPE2      TYPE3    SIZE
                             BYT DELETD  # DELETES  BYT DELETD  # DELETS  --------------EVENT--------------
                             INTERIM ST  W/O DASD   INTERIM ST     W/     OFF- DASD STRC NTRY STG  STG    RE-
                             W/O DASD    WRITE      W/DASD       WRITE    LOAD SHFT FULL FULL THLD FULL   BLD
-------------------------------------------------------------------------------------------------------------
(14) 3 CICS regions using CF-Structure log streams  AKPFREQ 1000 LGDFINT 0 20M HIGHOFFLOAD 80 LOWOFFLOAD 50
06/03/03  2:35:00 PM  (SMF INTERVAL TIMESTAMP 'B984354FE4D00000'X)
 IYOT1.CICS22.DFHLOG LOG_JG2_20M 194600278  330642432          0   669792 665555    4237            0    290
                             194353465    668978           0        0   142   0   0    0   0    0    0


Type CICS  DASD APPL% # Tran/ Average   Average  Average Average Logr Logr DASD  APPL% Storage  Idle    Net
     TCB   I/O        Second CPU/Task Resp      JC Wait File WT TCB  SRB  I/O         Frames   Frames  Frames
CF64 110.4 2074 40.9   49.37   .0066    .2640     .0163   .2336  0.4 32.3 0.0   10.9  22971.6  12961.0 10010.6

                                                                       1,560.54 Log writes per second
```

Figure 30   Three CF-Structure log stream regions

Analysis of the data presented in Figure 30 reveals the same problem observed in the tests with three DASD-only regions. Throughput is again throttled by the placement of the VSAM files for the CICS regions. The transaction rate is 49.37 with an average response time of .2640. The JC Wait time has dropped to 0.0163, but the average File Wait time shows 0.2336 seconds, clearly the cause of the poor response time. The RMF data reports the DASD I/O response time for IYOT1 is averaging 3.4 milliseconds with Q+PEND time at 2.1 milliseconds.

The log streams are functioning very well with no numbers in the "negative" columns, such as BYT DELETD W/DASD, TYPE3 writes, or negative events. It's interesting to note that DFH0STAT reports an average of 1,560.54 log writes per second in this region for the interval.

The RMF numbers shown for the System Logger represent activity for all three regions. APPL% is up to 10% with the SRB time at 32.3 seconds. The storage frames include storage for all six log streams (three DFHLOG and three DFHSHUNT), but because there is no activity for any of the DFHSHUNT log streams, consider the 10010.6 (41 MB) frames to be for the three DFHLOG log streams. It now should be apparent how storage use for System Logger can grow very quickly as the number of connected log streams increases in the LPAR.

In the next portion of this run, AKPFREQ was set to 7500. See Figure 31 on page 41.

```
                              BYT WRITTN  BYT WRITTN  BYT WRITTN                                     AVERAGE
                              BY USERS   TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------  BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE                INVOKED  TYPE1     TYPE2      TYPE3   SIZE
                              BYT DELETD  # DELETES  BYT DELETD  # DELETS  --------------EVENT--------------
                              INTERIM ST  W/O DASD   INTERIM ST    W/    OFF- DASD STRC NTRY STG  STG   RE-
                              W/O DASD    WRITE      W/DASD       WRITE   LOAD SHFT FULL FULL THLD FULL  BLD
----------------------------------------------------------------------------------------------------------------
(15) AKPFREQ 7500 LGDFINT 0 20M HIGHOFFLOAD 80 LOWOFFLOAD 50
06/03/03  2:45:00 PM  (SMF INTERVAL TIMESTAMP 'B984378C19300000'X)
 IYOT1.CICS22.DFHLOG LOG_JG2_20M  198408124   336027904    91873967    677378 611535     64869        975       292
                              166333443     567866    80944767    273230   291  11    0   0   0   0    0


Type CICS  DASD APPL%  # Tran/ Average  Average  Average Average  Logr Logr  DASD  APPL% Storage  Idle    Net
     TCB   I/O          Second CPU/Task  Resp     JC Wait File WT  TCB  SRB   I/O         Frames   Frames  Frames
CF64 114.0 2130 42.7    50.71   .0066    .2613     .0136   .2352   5.9  35.8  0.0   13.9  26125.5  12961.0 13164.5

                                                                           1,761.04 Log writes per second
```

Figure 31   Three CF-Structure log streams AKPFREQ set to 7500

During the first segment of the test when AKPFREQ was set to 1000, activity keypoints were being taken about twice per second. To reduce the number of activity keypoints, AKPFREQ was raised to 7500. The result was the number of activity keypoints were reduced to one every three seconds.

Figure 31 shows the results. Transaction throughput remains about the same as sample 14 in Figure 30 on page 40. The file response time is still high due to the DASD I/O contention. DASD response time is still averaging 3.4 milliseconds.

CICS APPL% increased 2%, and TCB time has increased from 110.4 to 114 seconds. Response time and CPU times are basically flat. The log writes per second is up 1761.04 per second.

The down side of raising the AKPFREQ value is the data remains in interim storage longer due to the delay in trimming the tail. The result is that data is once again being offloaded (BYT DELETD INTERIM ST W/DASD was 80944767 bytes in the interval) and the System Logger storage usage for CICS log streams is up to 13164.5 frames (53.9 MB).

Also notice that there were 11 DASD shifts (allocation for offload data sets) in the interval. DASD shifts are a direct reflection of data being offloaded. If you stop the offloading of data, the DASD shifts will go away.

For the third segment of this test, the structure was altered to 30 MB through a SETXCF command:

```
SETXCF START,ALTER,STRNM=log_jg2_5m,size=30720
```

In sample 16 (Figure 32 on page 42), things are once again looking good from the System Logger side. Data is no longer being offloaded, there are no DASD shifts, and the number of offloads has dropped significantly.

```
                               BYT WRITTN  BYT WRITTN  BYT WRITTN                                    AVERAGE
                               BY USERS   TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------  BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES   STORAGE              INVOKED  TYPE1     TYPE2      TYPE3    SIZE
                               BYT DELETD  # DELETES  BYT DELETD  # DELETS  --------------EVENT--------------
                               INTERIM ST  W/O DASD   INTERIM ST    W/     OFF- DASD STRC NTRY STG  STG   RE-
                               W/O DASD    WRITE      W/DASD       WRITE   LOAD SHFT FULL FULL THLD FULL  BLD
----------------------------------------------------------------------------------------------------------
(16) 3 CICS regions using CF-Structure log streams AKPFREQ 7500 LGDFINT 0 30M HIGHOFFLOAD 80 LOWOFFLOAD 50
06/03/03  2:55:00 PM  (SMF INTERVAL TIMESTAMP 'B98439C84D900000'X)
 IYOT1.CICS22.DFHLOG LOG_JG2_20M  187306756   319705600          0   650661  647672       2990        0     287
                                186124374     646446           0        0      91   0   0   0    0    0   0


Type CICS  DASD APPL%  # Tran/ Average   Average  Average Average  Logr Logr  DASD  APPL% Storage   Idle     Net
     TCB  I/O         Second  CPU/Task  Resp      JC Wait File WT   TCB  SRB   I/O         Frames  Frames   Frames
CF64 106.1 2003 39.2   47.70    .0066    .2827     .0173   .2513    0.3  29.6  0.0    9.9  29859.9 12961.0  16898.9

                                                                         1,870.57 Log writes per second
```

Figure 32   Three CF-Structure log streams 30 MB with AKPFREQ set to 7500

One again, the down side is an increase in the System Logger storage used for the CICS log streams. The average number of frames is now 16898.9 (69.2 MB). With the increase in structure size, there will be more data in interim storage before the HIGHOFFLOAD point is reached, thus accounting for the large increase in storage frames.

Conclusion: Use of the CF for log streams makes it possible to obtain very high transaction throughput, but outside bottlenecks, in this case DASD I/O contention, can negate a large portion of the benefits.

Tuning log streams is actually more than simply giving the structure more storage, but also addressing all the issues that cause interim storage data residency time to be elongated.

Non-system log streams

User journals, forward recovery logs, and auto journals present a slightly different performance tuning opportunity. The data is always offloaded so the main considerations are interim storage residency time and how quickly the data can be offloaded.

The next samples show the effects of having a forward recovery log for one VSAM file in the region.

In sample 17 (Figure 33 on page 43), there is a single CICS region, with DFHLOG allocated in a CF structure (LOG_JG2_20M). The region also has a forward recovery log (&applid..&SYSID..DFHJ03) for one of the files (FILEC). The user journal was defined with STG_SIZE of 25000 4 K CIs, that is, it can hold 102.4 MB of data.

```
                              BYT WRITTN  BYT WRITTN  BYT WRITTN                                    AVERAGE
                              BY USERS   TO INTERIM   TO DASD    #WRITES  ---# WRITES COMPLETED------ BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES  STORAGE               INVOKED  TYPE1    TYPE2     TYPE3  SIZE
                              BYT DELETD  # DELETES  BYT DELETD  # DELETS  --------------EVENT--------------
                              INTERIM ST  W/O DASD   INTERIM ST    W/     OFF- DASD STRC NTRY STG  STG  RE-
                              W/O DASD    WRITE      W/DASD       WRITE   LOAD SHFT FULL FULL THLD FULL  BLD
----------------------------------------------------------------------------------------------------------
(17) AKPFREQ 1000 LGDFINT 0 20M HIGHOFFLOAD 80 LOWOFFLOAD 50 --- Forward Recovery Log STG_SIZE 25000
06/03/03  9:45:00 PM  (SMF INTERVAL TIMESTAMP 'B984956CB0F00000'X)
 IYOT1.CICS22.DFHLOG LOG_JG2_20M  418554024   720508928          0  1480321 1475486      4836        0      282
                               421389900     1490611          0        0      90   0   0  0   0    0    0

 IYOT1.IY01.DFHJ03     *DASDONLY*    30473250   101478400   29913580    24775      0        0        0     1230
                                0           0   88190976    21531       2   0   0  0 268    0    0


Type CICS  DASD APPL%  # Tran/ Average   Average  Average Average  Logr Logr  DASD  APPL% Storage  Idle   Net
     TCB   I/O        Second  CPU/Task  Resp      JC Wait File WT   TCB  SRB   I/O        Frames  Frames Frames
CF64 215.3 4433 79.7   105.53   .0062    .1007     .0060   .0842    0.1  17.6  83.4   6.0  31757.0 12760  18997

                                                                            3619.24 logwrites/second
```

Figure 33   User journal: STG_SIZE 25000, single region

This sample provides you a sample report of what can happen from a resource impact point of view when adding a user journal. System Logger is issuing DASD I/O at the rate of 83.4 per second. DFHLOG is connected to a CF. No data is being offloaded from DFHLOG, and it does not have a staging data set. Therefore in this case, all DASD I/O is associated with the user journal.

The greatest impact on the system is the 18997 page frames (77 MB) allocated by System Logger to contain the data for both CICS log streams. To be technically accurate, there are three log streams; however, DFHSHUNT is connected, but no data is being written. Notice the excellent transaction rate (105.53 per second) with a response time of 0.1007 seconds.

DFHLOG is running very well, and in fact, DFHJ03 is also working very well.

There are some subtle differences between the log streams. For the user journal, because all data is offloaded, there should not be any information in the BYT DELETD INTERIM ST W/O DASD column. Likewise, there should not be any data reported under #DELETES W/O DASD WRITE.

Under the BYT DELETD INTERIM ST W/DASD column is the amount of data offloaded in the interval. The number of STG THLD (the number of writes when the log stream was at or above the HIGHOFFLOAD threshold point) is more than 130 times the number of offloads.

Because the log stream is so large, CICS is able to issue 134 write requests in the time it take System Logger to perform the offload. Using the average buffer size shown (1230) means that CICS is writing 164820 bytes while the offload process is taking place.

To determine if that is a problem, we must determine how much space is available above the HIGHOFFLOAD threshold. STG_SIZE was specified as 25000 4 K CIs, so using a HIGHOFFLOAD of 80%, the threshold point is at 20000 CIs. The storage above the threshold point is 5000 CIs or 20.48 MB. Because this is a DASD-only log stream, and the average buffer size is less the 4 K, there is room for 5000 writes before filling the staging data set. Remember each buffer is rounded to a 4 K CI boundary.

Using the sizing formula, a different size was used as the STG_SIZE value for the allocation of the forward recovery log, 2549, that is, approximately 10.4 MB of space; for more details,

refer to *Systems Programmer's Guide to: z/OS System Logger*, SG24-6898. The results of this configuration are shown in Figure 34.

```
                               BYT WRITTN   BYT WRITTN   BYT WRITTN                                    AVERAGE
                                BY USERS    TO INTERIM    TO DASD      #WRITES   ---# WRITES COMPLETED------  BUFFER
-LOGSTREAM NAME- STRUCTURE NAME-- IXGWRITES   STORAGE                  INVOKED   TYPE1      TYPE2     TYPE3    SIZE
                               BYT DELETD   # DELETES    BYT DELETD   # DELETS   ---------------EVENT---------------
                               INTERIM ST   W/O DASD     INTERIM ST      W/     OFF- DASD STRC NTRY STG  STG   RE-
                               W/O DASD     WRITE        W/DASD        WRITE    LOAD SHFT FULL FULL THLD FULL  BLD
-----------------------------------------------------------------------------------------------------------------
(18) AKPFREQ 1000 LGDFINT 0 20M HIGHOFFLOAD 80 LOWOFFLOAD 50 --- Forward Recovery Log STG_SIZE 2549
06/03/03 10:20:00 PM  (SMF INTERVAL TIMESTAMP 'B9849D3F68400000'X)
 IYOT1.CICS22.DFHLOG LOG_JG2_20M  445609444   767130624          0   1576179 1571029    5150        0       282
                                  444776164    1573177           0         0       95 0    0    0    0   0    0

 IYOT1.IY01.DFHJ03    *DASDONLY*   29163300    97116160    31417260     23710        0       0        0      1230
                                         0           0    93859840     22915       10 0    0    0  290   0    0


Type CICS  DASD APPL%  # Tran/ Average  Average   Average Average  Logr Logr  DASD  APPL% Storage  Idle     Net
     TCB   I/O          Second CPU/Task Resp      JC Wait File WT   TCB  SRB   I/O        Frames   Frames   Frames
CF64 225.8 4720 83.3   112.36  .0061    .0921      .0061   .0751    0.1  19.1  79.9   6.5 22579.9  12044.2  10535.7

                                                                                    4533.17 logwrites/second
```

Figure 34   User journal: STG_SIZE 2549, single region

Analysis of the data shows an offload is happening every 30 seconds, rather than every minute as expected. This is not a problem, but the mystery must be solved.

The assumption was that the JORC transaction would be executed at a rate of 6 per minute. However, the data from sample 13 (Figure 28 on page 38), which is the base region without a forward recovery log, shows the rate was actually 11 transactions per second.

Plugging the transaction rate of 11 per second into the formula gives:

> 11 transactions * 6 UOW/transaction = 66 UOW per second
> Each UOW writes 1230 rounded to a 4096 byte boundary or 1 CI per UOW = 66 CI/sec.
> Therefore, 66 CI/second * 60 = 3960 CI/minute.
> With HIGHOFFLOAD set to 80, STG_SIZE = 3960 * 1.18 = 4672.8, round to 4673.

This accounts for the discrepancy and validates the formula.

Comparing sample 18 with sample 17 (Figure 33 on page 43), there is a decrease in the number of page frames required by System Logger (10535.7 versus 8997). This is a reduction of 34.7 MB storage required for a *single* log stream.

Sample 17 illustrates a common error, where the size of a user journal is defined to contain the same amount of data as a journal in a R410 system. The nature of user journals is that they are never accessed by the online CICS region, so all data will be offloaded, that is, logtail management does not apply to user journals. As such, there is little benefit in keeping the data in the interim storage; it's better to reduce the size (STG_SIZE), so offloads happen on a regular basis, and let the data reside in the offload data sets.

LS_SIZE should be specified to be large enough to contain the offloaded data during the prime processing period of the day.

Conclusion: The STG_SIZE value for user journals should be calculated and specified to provide a data residency time of one minute or less, thus reducing the amount of storage required by System Logger to back the data in the data space.

Summary

Tuning CICS log streams for performance involves understanding the CICS region, the operating environment, and the workload. External factors, such as DASD response time and CF access time, make the task more challenging, but not insurmountable.

Data residency time is key to controlling the amount of storage required by the System Logger for duplexing the data to the data space. Data residency time is also key to establishing an even movement of data from non-system logs to the offload data sets.

It is important to understand the tools that are available and how they can be used for log stream analysis.

As the environment changes, such as an increase in workload, a new level of the Coupling Facility Control Code, new applications, a new release of the operating system, or a new release of CICS, ensure that the log streams sizes are still valid.

The team that wrote this Redpaper

This Redpaper was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center.

Jim Grauel is an IBM Senior Technical Staff Member. He joined IBM in 1965 as a Hardware Customer Engineer in St. Louis, Missouri. In 1972, he transferred to software as a Program Support Representative, supporting OS R21.7 and CICS (Release 2.3). In 1980, he transferred to the CICS support center in San Jose, California, and later to Raleigh in 1984, and has been working as Level 2 Support in the CICS Support Center.

Frank Kyne is a Senior IT Specialist at the International Technical Support Organization, Poughkeepsie Center. He writes extensively and teaches IBM classes worldwide on all areas of Parallel Sysplex®. Before joining the ITSO five years ago, Frank worked in IBM Global Services in Ireland as an MVS™ Systems Programmer.

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:
This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

47

Send us your comments in one of the following ways:
- ► Use the online **Contact us** review redbook form found at:
  ibm.com/redbooks
- ► Send your comments in an Internet note to:
  redbook@us.ibm.com
- ► Mail your comments to:
  IBM Corporation, International Technical Support Organization
  Dept. HYJ Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400 U.S.A.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

CICS®	@server™	Parallel Sysplex®
CICS/ESA®	@server™	Redbooks(logo) ™
DB2®	IBM®	RMF™
DFSORT™	ibm.com®	z/OS®
@server™	MVS™	zSeries®

Other company, product, and service names may be trademarks or service marks of others.