Redbooks
ibm.com/redbooks

IBM CICS Performance Series: CICS TS V5.3 Benchmark on IBM z13

John Burgess

Ian Mitchell

Catherine Moxey

z Systems

IBM

Redpaper

Introduction

This IBM® Redpaper™ publication documents the performance characteristic benchmarks of two high-transaction rate IBM CICS® Transaction Server for IBM z/OS® workloads:

- The first workload executes on a single IBM z13™ logical partition (LPAR) with 18 central processors (CPs) up to a rate of 174,000 CICS transactions per second.
- The second workload executes on a single z13 LPAR with 26 CPs up to a rate of 227,000 CICS transactions per second (TPS).

The motivation for the benchmark is to demonstrate some of the benefits of CICS Transaction Server (CICS TS) V5.3, compared with previous releases in terms of:

- Improved performance
- Enhanced interaction with IBM z/OS Workload Manager (WLM) for transaction CPU tracking and IBM Mobile Workload Pricing (MWP)
- The ability to demonstrate the stability and scalability of CICS handling high-transaction rates in excess of those tested previously

In particular, CICS TS V5.3 has optimized transaction dispatching for HTTP requests. In this release, CICS is able to directly dispatch an application transaction for each request that it receives, without first needing to run a CWXN (web attach) transaction.

> **Note:** The CWXN transaction is still used if your application uses an analyzer program or cannot take advantage of Application Transparent Transport Layer Security (AT-TLS) for its Secure Sockets Layer (SSL) needs.

Eliminating CWXN reduces CPU usage and in addition, reduces the amount of direct-access storage device (DASD) space used to store standard System Management Facility (SMF) records when CICS Performance monitoring is active.

Together with z/OS V2.1 and later WLM enhancements, transactions running in CICS TS V5.3 can be classified into service and report classes that directly measure and report their CPU consumption. CICS has been enhanced to continuously measure CPU consumption of every transaction, without the need to activate CICS Performance class monitoring. These measurements are aggregated by z/OS WLM to provide enhanced reporting in SMF records and IBM Resource Measurement Facility™ (RMF™) reports. The CPU measurements for workloads classified in this way make it simpler and more efficient to collect the data required for Mobile Workload Pricing.

These enhancements and optimizations make it possible to drive CICS TS to even higher transaction rates than were previously possible, allowing applications that need to run at high volumes of requests and with increasing workloads to be hosted in CICS with confidence.

The primary audience for this paper includes technical specialists who need to make architectural, capacity, and performance-related decisions regarding new or existing application workloads on the IBM z Systems™ platform. While this benchmark is deeply technical in nature, the results have been laid out in a straightforward manner that should also appeal to technical executives who are making platform choices, especially those where extremely high levels of throughput and linear scalability are important consideration factors.

The CICS TS application

For our testing, the CICS workload consisted of an application providing two distinct services accessing common data. One service is accessed via terminals using the LU2 protocol, and the other using TCP/IP clients with a CICS HTTP WEB interface. The majority of transactions issued an **EXEC CICS STARTBROWSE** command to locate a starting position in a recoverable CICS maintained shared data table and then issued five **EXEC CICS READNEXT** commands. The CICS Maintained Data Table (CMT) was backed by a Virtual Storage Access Method (VSAM) data set. Approximately 10% of the LU2 client transactions issued an **EXEC CICS READ UPDATE**, followed by an **EXEC CICS REWRITE**, updating the VSAM file record. A smaller proportion of the TCP/IP clients issued updates. All updates needed to be function shipped to the file-owning region (FOR) by the use of a multiregion operation (MRO) connection where data integrity and recoverability were maintained.

As shown in Figure 1 on page 3, the LU2 transactions were sent using IBM Enterprise Extender into a set of six terminal-owning regions (TORs). All of these transactions were then routed to a set of 10 application-owning regions (AORs) using IBM CICSPlex® SM Sysplex Optimized routing over MRO XM.

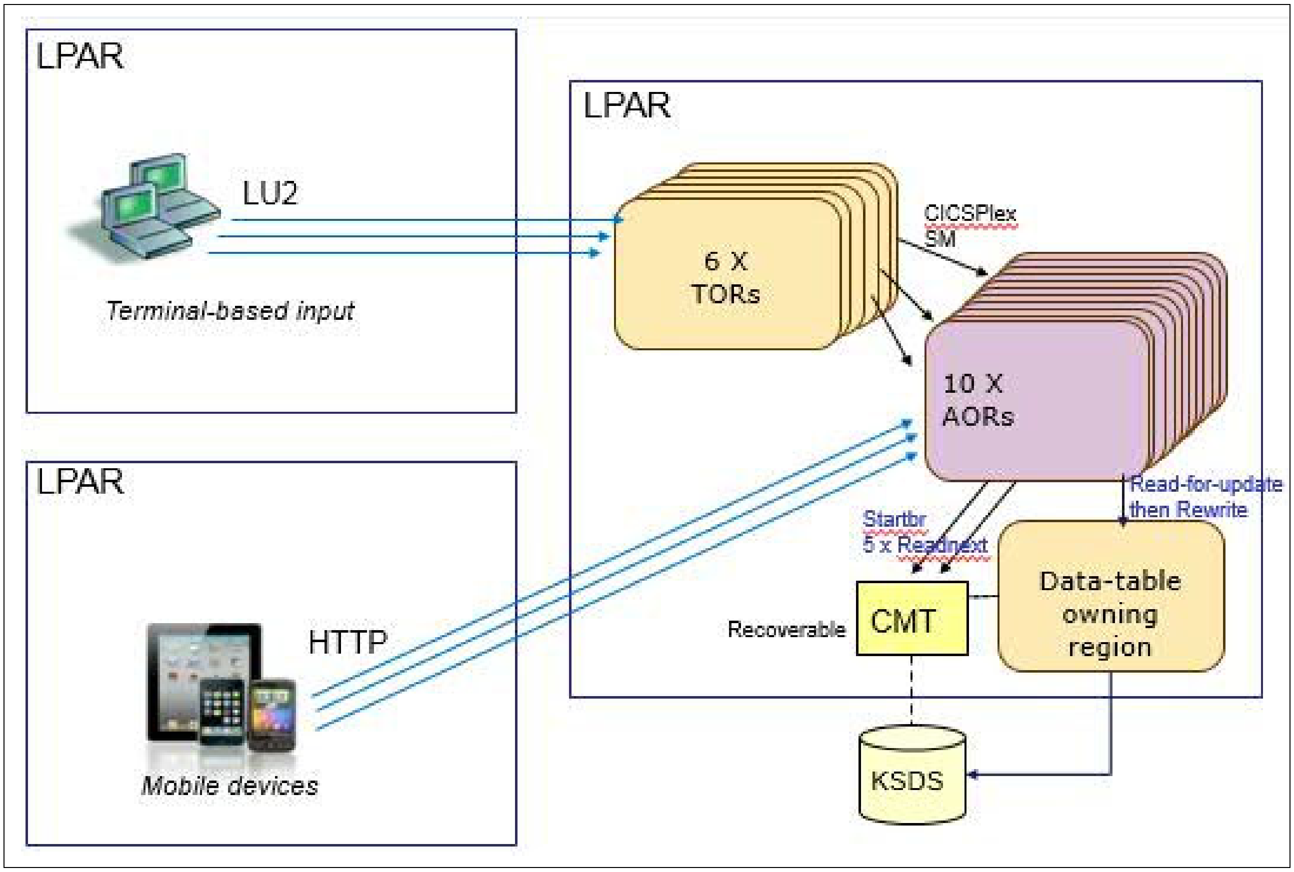

Figure 1 CICS configuration for benchmark testing of CICS TS V5.2 and CICS TS V5.3

You can also see in Figure 1 that the TCP/IP clients entered the system directly into the AORs via TCP/IP port-sharing, maintaining persistent connections.

IBM Workload Simulator for z/OS (WSim) was used on two other LPARs to simulate the clients.

Retail catalog shopping scenario

The CICS workload used for this benchmark might represent the back-end processing for a number of applications. We describe in this paper how it might correspond to a retail application scenario, involving shopping from an online catalog.

A retailer provides kiosks in-store for customers to browse for items they want to reserve. Also provided is a mobile application on which customers can find and purchase items. Each request results in a browse through, on average, five items of a particular category of goods (such as training shoes or T-shirts) before a customer moves on to look at a different category of items, initiating another request. Customers typically browse through many items before deciding upon a purchase.

During a sales promotion, the retailer needs to be able to handle a very high volume of requests being made at the same time, and throughout the day. The retailer wants to be able to account for CPU usage by the requests that come from the mobile application, but without the requirement to run with CICS performance-class monitoring enabled.

Figure 2 depicts this retail catalog shopping scenario.

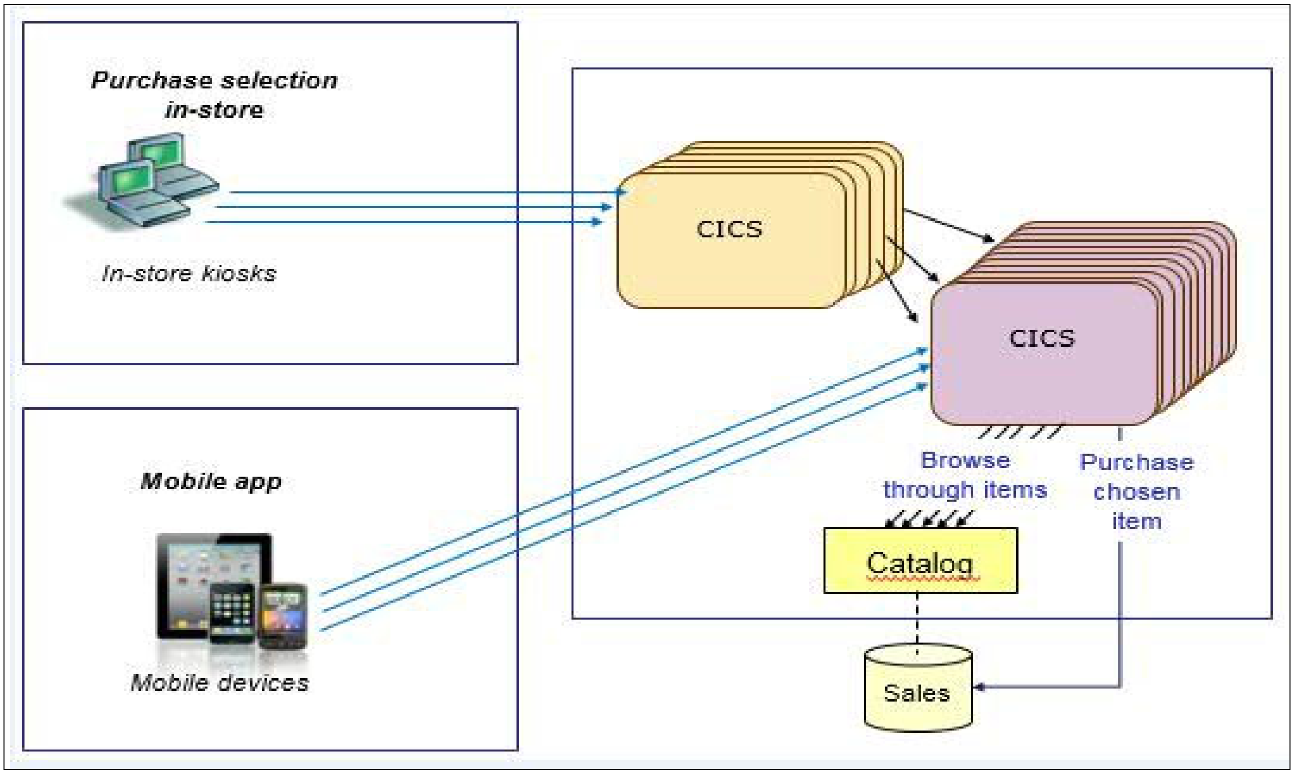

Figure 2 Retail shopping scenario

The retail catalog is held in a recoverable CICS maintained shared data table, backed by a Virtual Storage Access Method (VSAM) key sequenced data set (KSDS) file. Browsing the catalog results in direct access to the shared data table, whereas purchase requests map to updates that are also applied to the backing VSAM data set.

Requests made at the kiosks are processed by LU2 transactions, which are sent using the Enterprise Extender protocol into six terminal-owning regions, and then routed across 10 application-owning regions using CICSPlex SM Sysplex optimized routing.

Requests made from the mobile application are processed by using HTTP web requests directly into the application-owning regions using TCP/IP shared ports.

Hardware and software configuration

The following section provides the hardware and software configurations that are necessary for the retail catalog shopping scenario.

Hardware

The following hardware was used for this benchmark:

- IBM z13 (type 2964) model NE1
- LPAR with 18 dedicated processors (CPs) for the system under test
- An alternate configuration with 26 dedicated CPs was also used for some measurements
- Internal coupling facility with interval control program (ICP) links
- DS8870 (2424-961) DASD

Software

The following software was used for this benchmark:

- CICS Transaction Server V5.3 and CICS Transaction Server V5.2
- z/OS V2R1 plus WLM and RMF Authorized Program Analysis Reports (APARs) OA47042, PI47638, and OA48466
- WSim
- RMF Monitor I and III

Performance monitoring tools

For our benchmark sampling period, we used the tools described in this section.

RMF Monitor I

RMF Monitor I was used to record system resource usage, including CPU, direct access storage device (DASD), and storage, and was also used with the z/OS WLM configuration to record the CPU, transaction rates, and response times for CICS service classes and report classes. SMF records 70-79 are written on an interval basis and can be post-processed by using the RMF postprocessor (program ERBRMFPP).

A new function in CICS TS V5.3 passes the CPU times used by transactions to z/OS WLM. This enables CPU time collection for groups of transactions based on WLM classifications. Also introduced are two new CICS classifications based on TCP/IP service name and transaction class. Within a classification, a value of *MOBILE* can be specified for a reporting attribute, which enables WLM to accumulate all of the CPU time for this classification and report it directly as Mobile CPU time for consideration for discounted mobile pricing. These values are also displayed in RMF postprocessor reports.

z/OS WLM APAR OA47042 and RMF APAR OA48466 are needed to use this new function, and also, if z/OSMF is used, APAR PI47638.

RMF Monitor III

RMF Monitor III was used to record the coupling facility activity for the Logger and the Coupling Facility Data table used by CICSPlex SM Sysplex Optimized Routing. RMF Monitor III can also be used on an interactive basis and the data can also be written to VSAM data sets for later review via the interactive RMF TSO application.

CICS TS statistics

CICS statistics were used to monitor and report CICS resource usage, including CPU, storage, file accesses, number of request transactions routed, number of requests function shipped, and so on.

CICS interval statistics have most of the counters reset at the start of each interval so that any resource consumption reported only relates to the measurement period being observed. Interval statistics can be activated using CEMT SET STATISTICS. When you set this, the first interval can be adjusted to a shorter time so that all the intervals are synchronized to the end-of-day time (**STATEOD**) parameter. For example, in Example 1, if you were to use CEMT to set and an interval of 15 minutes at 10 minutes past the hour, the first interval would expire in 5 minutes so that all future intervals line up on 15-minute wall clock boundaries. The values in this first report can also be associated with a much longer period, depending on the time of the last reset.

Another alternative to using interval statistics would be to use CEMT to reset the counters and then at the end of the measurement period, use CEMT to record all the statistics. Resetting the statistics requires a change of state from ON to OFF or from OFF to ON. To ensure this happens, the commands in Example 1 serve as an example of resetting the statistics in one of the AORs in this configuration.

Example 1 Collecting CICS statistics

```
F CICSA001,CEMT SET STAT OFF RESET
F CICSA001,CEMT SET STAT ON RESET

        Measurement period is between the RESET and the RECORD

F CICSA001,CEMT PERFORM STAT ALL RECORD
```

Regardless of whether the statistics are ON or OFF, when PERFORM RECORD is issued, a statistics record is written.

CICS statistics are written as SMF 110 subtype 2 records and can be post-processed by using the CICS statistics utility program, DFHSTUP, or by using CICS Performance Analyzer (CICS PA).

Benchmark methodology

This section describes the approach used in running and measuring the benchmark.

Workload Simulator, running on two separate LPARs, was used to simulate the clients that drive the workload. For the 18 CP measurements, we had 14700 LU2 3270 simulated clients and 6000 web clients. For the 26 CP measurements, we increased the web clients to 9000.

The LU2 traffic flowed over Enterprise Extender, which allows SNA data to flow over TCP/IP. The HTTP traffic flowed over IBM HiperSockets™.

Simulated client think times were selected to produce a specific overall throughput rate in the CICS address spaces. The system was allowed to settle into a steady state, and then performance measurement data was collected for a 5-minute interval. RMF I, RMF III, and CICS statistics were collected for this interval, and then the simulated think time was reduced to create a higher throughput rate. The process was then repeated. Data for five separate measurement intervals was collected.

Throughput was measured in terms of business transactions per second (BTPS). For example, consider a transaction entered a TOR, was routed to the AOR, then function shipped a file request to the FOR. In terms of CICS transactions, this would involve one transaction in the TOR, another in the AOR, and a third CSMI (mirror) transaction in the FOR.

We counted this as *one* business transaction, and, in terms of RMF service and report class, this would be counting only the Begin to End (BTE) transactions, not the executed transactions. That is, in this case, the three individual transactions are counted as one business transaction.

We also did not count the separate CWXN transaction in CICS TS V5.2 that is used in processing the HTTP requests entering the AORs. We only counted the CWBA equivalent transaction that carried out the business processing. In CICS TS V5.3, the need to run this CWXN transaction has, in many circumstances, been eliminated, as was the case in this benchmark. Excluding the counting of the CWXN transaction in CICS TS V5.2 gave a direct comparison in terms of business transaction rates with CICS TS V5.3.

z/OS Workload Manager classifications

This section describes the z/OS classifications that were used for this benchmark and explains how you can use these classifications to collect CPU time for separate types of workload, such as work initiated from mobile devices.

From a z/OS WLM point of view, using CICS TS V5.3 for this benchmark, we were able to classify all of the transactions arriving on a specific TCP/IP service installed in the AORs. In this workload, all of the AORs had a TCP/IP service that we called *HTTP1* enabled. Any transaction arriving over this TCP/IP service was classified into a service class called *CICSJB* and reported in a report class called *SCALHTTP*. All of these transactions arriving on this TCP/IP service also had their CPU times attributed to Mobile CPU time. This facility has enabled the collection of CPU times for transaction types to be accumulated without the need to turn on CICS Performance class monitoring. This accumulation includes any CPU times for transactions in other regions running on behalf of these arriving on this TCP/IP service, such as a function ship or transaction route.

Transactions arriving in the TORs were also classified into service class CICSJB and reported in a report class called SCAL3270. These were not classified as mobile in this workload.

Figure 3 and Figure 4 show the CICS subsystem classification definition used for this workload. Figure 4 shows the view when using PF11 to page right from the view shown in Figure 3, including the new reporting attribute, MOBILE.

```
Command ===> ____________________________________________ Scroll ===> CSR

Subsystem Type . : CICS         Fold qualifier names?   Y  (Y or N)
Description  . . . CICS Transaction classification

Action codes:   A=After     C=Copy        M=Move      I=Insert rule
                B=Before    D=Delete row  R=Repeat    IS=Insert Sub-rule
                                                         <=== More
        -------Qualifier--------         Storage   Reporting  Manage Region
Action    Type       Name       Start      Critical  Attribute  Using Goals Of

 ____    1 CT       HTTP1      ___        NO        MOBILE     N/A
 ____    1 SI       IYCUT0*    ___        NO        NONE       N/A
 ____    1 SI       IYCUA0*    ___        NO        NONE       N/A
 ____    1 SI       IYCUZC07   ___        NO        NONE       N/A
```

Figure 3 CICS WLM classification

```
Command ===> ____________________________________________ Scroll ===> CSR

Subsystem Type . : CICS         Fold qualifier names?   Y  (Y or N)
Description  . . . CICS Transaction classification

Action codes:   A=After     C=Copy        M=Move      I=Insert rule
                B=Before    D=Delete row  R=Repeat    IS=Insert Sub-rule
                                                         <=== More
        -------Qualifier--------         Storage   Reporting  Manage Region
Action    Type       Name       Start      Critical  Attribute  Using Goals Of

 ____    1 CT       HTTP1      ___        NO        MOBILE     N/A
 ____    1 SI       IYCUT0*    ___        NO        NONE       N/A
 ____    1 SI       IYCUA0*    ___        NO        NONE       N/A
 ____    1 SI       IYCUZC07   ___        NO        NONE       N/A
```

Figure 4 WLM CICS classification (PF11)

Before CICS TS V5.3, using this new z/OS WLM facility, the CICS subsystem service class CICSJB would not show any CPU time related to CICS. Now this CICS subsystem service class shows the transaction CPU APPL% related to CICS transactions running in it. It also separates out the CPU time associated with those transactions that were classified as MOBILE.

The two report classes, SCALHTTP and SCAL3270, will also have the transaction CPU times recorded, but only the SCALHTTP will show any Mobile CPU time. CPU times captured in this method will be similar in magnitude to the times that would be collected with CICS Performance class monitoring.

Figure 5 shows the Job Entry Subsystem (JES) classification rules for the CICS address spaces.

```
Subsystem Type . : JES        Fold qualifier names?   Y  (Y or N)
Description  . . . ________________________________

Action codes:   A=After     C=Copy        M=Move     I=Insert rule
                B=Before    D=Delete row  R=Repeat   IS=Insert Sub-rule
                                                             More ===>
         -------Qualifier--------                 -------Class--------
Action   Type      Name     Start                 Service     Report
                                        DEFAULTS: BATCH       BATCH
 ____     1 TN      CICST0*  ___                  JESTORS     CICSTORS
 ____     1 TN      CICSA0*  ___                  CICSBTCH    CICSAORS
 ____     1 TN      CICS2A07 ___                  CICSBTCH    CICSFORS
```

Figure 5   JES subsystem classification rules

When CICS initially starts, it is classified either as a JES or STC subsystem type. In our case, the CICS regions were started by submitting job control language (JCL), so they were JES classified. These CICS regions are managed to the service classes specified in the JES subsystem classifications until transactions start running in them. In our case, we specified that these CICS regions should be *Transaction Managed* so while transactions are running, the regions will be managed by the service classes specified in the associated CICS subsystem classification rules, rather than the JES subsystem.

In the JES subsystem classification, we have associated the TOR, AOR, and FOR report classes with the same corresponding report classes that we used for the CICS regions in the CICS subsystem classification.

This provides the opportunity to report CICS transaction counts and response times, plus the CICS region CPU usage, and now, with CICS TS V5.3, the CPU times used by transactions are also reported.

Note: An important factor when using service and reporting classes for reporting transactions through RMF is that CICS transactions only get reported in the regions in which they were originally initiated, classified, and then ended.

For example, in the TOR/AOR/FOR environment suppose that 5000 transactions arrived in the TOR, these will be classified accordingly and if they are all routed to the AOR they will carry that classification with them. The same is true for any transaction running on behalf of these in the FOR such as the CSMI (mirror) transaction. When looking at an RMF report, the only transaction count that will appear will be the 5000 in the TOR. This count will contain the number of transactions ended, which is the total number of transactions that were initiated in the region or set of regions represented by this report class or service class.

RMF only reports the count of the initial transactions; it is not accumulating the count of any transaction that was started as a result of a transaction route or function ship. In this example, where all the transactions were routed to an AOR, there would be an ENDED count of 5000 and an Executed (EXCTD) count of 5000 in the RMF report for the TOR. Ended transactions represent the BTE counts.

EXCTD count is a count of transactions that have been routed to another CICS region.

Also shown in a CICS report class in an RMF report are the Actual and Execution Response times. The Actual Response time refers to the overall response time of transactions, and Execution Response time refers to the time spent executing in the AORs and FORs.

If you look at an RMF report in the TOR/AOR/FOR configuration where all transactions arrived in the TOR, the only report class with a transaction count would be the TOR because all of the transactions running in the AOR and FOR are in the Execution phase running on behalf of those arriving in the TOR. The report class would show the total address space APPL% CPU time, but it would now also show the actual transaction APPL% CPU time.

> **Note:** This new CPU time for transactions follows the same rules as BTE counted transactions. That is, CPU times recorded for those transactions in execution states in the AORs and FORs will be accumulated back to the TOR where they originated.

TOR transaction APPL% can therefore be greater than the TOR address space APPL%. Pure AORs and FORs will have little transaction APPL% compared to z/OS region APPL%.

Figure 6 shows an example of a report class that has both a JES and a CICS subsystem classification associated with it.

```
REPORT BY: POLICY=POLICY                                REPORT CLASS=CICSTORS
                                                        DESCRIPTION =Report for WLM TORs - Johnb

-TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---   SERVICE TIME  ---APPL %---
AVG       2.00  ACTUAL                13  SSCHRT   0.0  IOC          0  CPU    5.755  CP      10.16
MPL       2.00  EXECUTION             13  RESP     0.0  CPU     458149  SRB    1.344  AAPCP    0.00
ENDED    74001  QUEUED                 0  CONN     0.0  MSO          0  RCT    0.000  IIPCP    0.00
END/S  1059.11  R/S AFFIN              0  DISC     0.0  SRB     107001  IIT    0.000
#SWAPS       0  INELIGIBLE             0  Q+PEND   0.0  TOT     565150  HST    0.000  AAP       N/A
EXCTD    67577  CONVERSION             0  IOSQ     0.0  /SEC      8088  AAP      N/A  IIP      0.00
AVG ENC   0.00  STD DEV               20                                IIP    0.000
REM ENC   0.00                                          ABSRPTN   4044
MS ENC    0.00                                          TRX SERV  4044

TRANSACTIONS APPL% :   TOTAL :  CP   37.24  AAP/IIP ON CP   0.00   AAP/IIP   0.00
                       MOBILE :  CP    0.00  AAP/IIP ON CP   0.00   AAP/IIP   0.00
```

Figure 6 Example of a TOR report class

As shown in Figure 6, we see both transaction rates from the CICS classification and the total address CPU APPL % from the JES classification. Notice that the TRANSACTION APPL% of 37.24 is greater than the address space APPL % of 10.16. This is because 67577 transactions out of the 74001 were routed to other regions, and the CPU from those EXECUTED transactions has been accumulated back to the TORs’ Transaction APPL%.

In this example, no Mobile CPU time has been recorded for this report class.

Figure 7 shows the flow of transactions in this workload through a TOR to an AOR using transaction routing, and in some cases, on to an AOR region using function shipping. This illustrates the meaning of BTE and EXEcution in z/OS WLM terms.

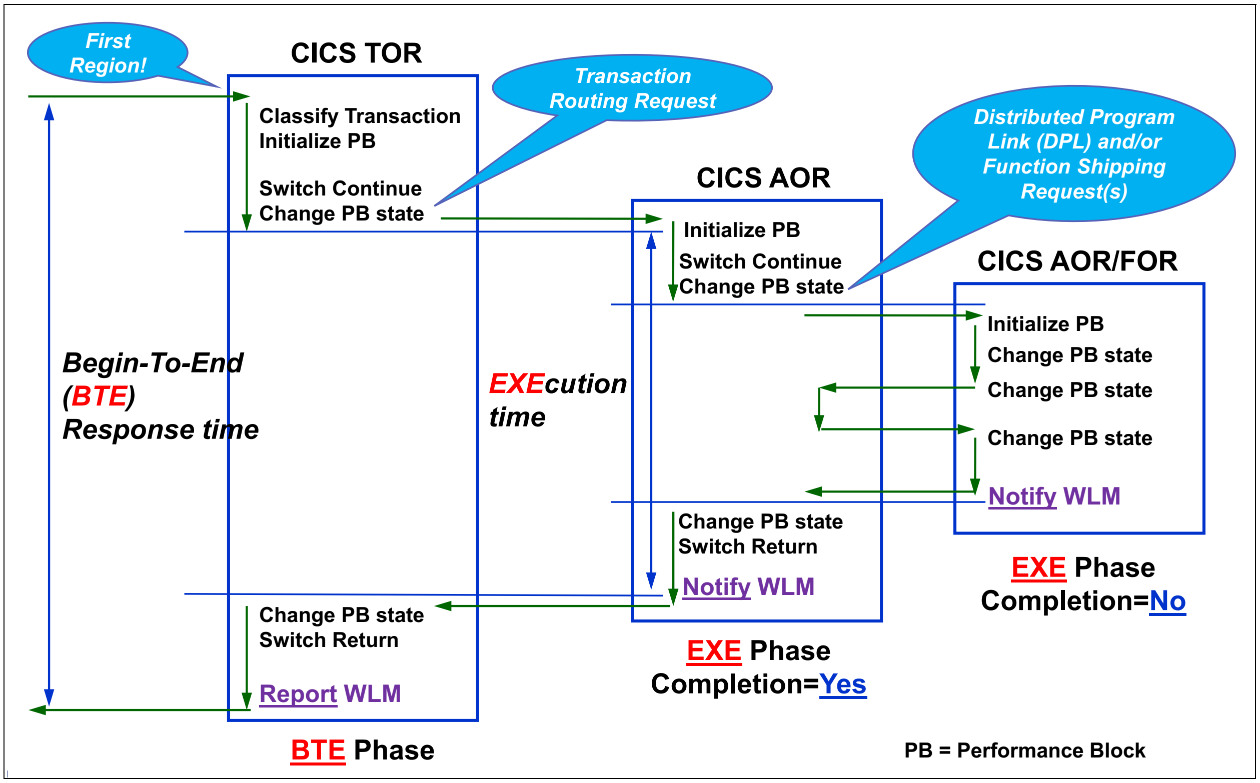

Figure 7 CICS and z/OS transaction tracking

Figure 8 shows the relationship between the CICS regions and the service classes that transactions were being managed by, in addition to the report classes associated with them. The TORs were classified with an attribution of BOTH. This means that they are managed to the performance goal of the JES service class, but will nevertheless classify and track all transactions in the correct CICS specified service class so that z/OS WLM can still manage transactions to their response time goals.

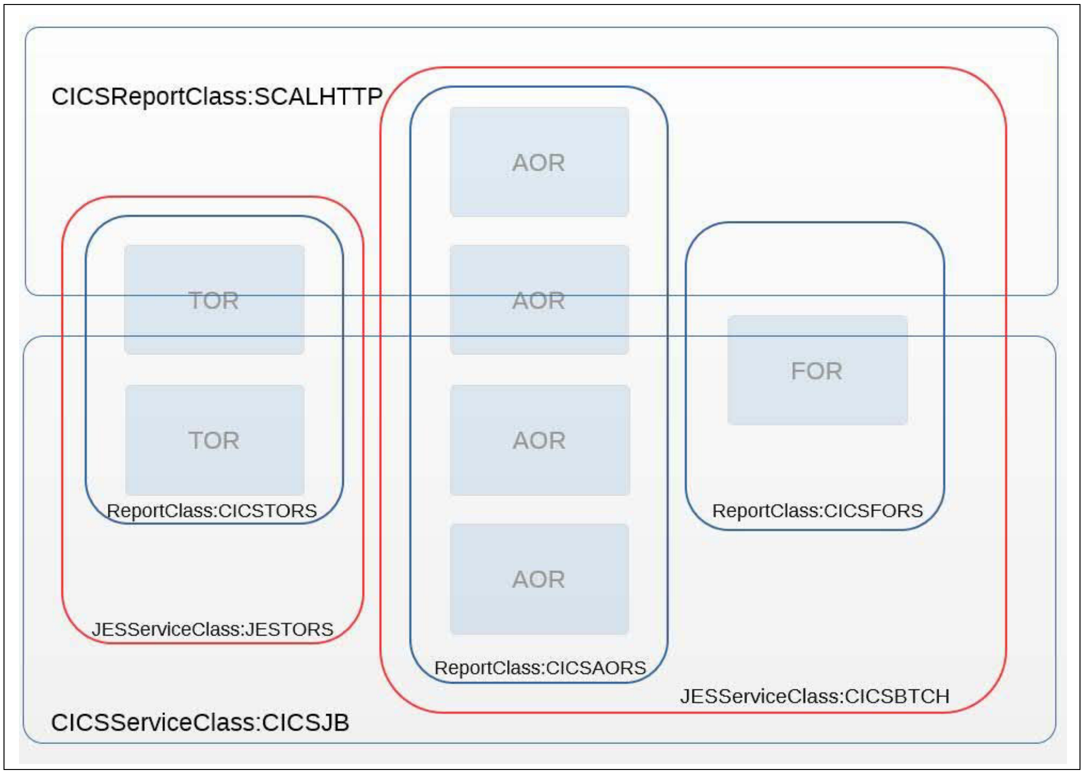

Figure 8 Service classes and report classes

Terms used in the results

The following list gives a description of the terms used in the following tables of results:

- ETR: The transaction rate per second based on Begin to End (BTE) transactions only and excluding any CWXN transaction that we did not consider as one of the business transactions.
- TOR CPU%: The percent of 1 CP that is being used by the CICS regions to which all the LU2 clients were attached. This is as reported by RMF Monitor I under the ---APPL %--- heading for that report class. This region might also be known as a *routing region*.
- AOR CPU%: The percent of 1 CP that is being used by the CICS regions that all the LU2 transactions were routed to and that the TCP/IP clients were connected to. This is as reported by RMF Monitor I under the ---APPL %--- heading for that report class.
- FOR CPU%: The percent of 1 CP that is being used by the CICS region owning the Shared Data table to which all file updates were function shipped to. This is as reported by RMF Monitor I under the ---APPL %--- heading for that report class.

- Total CICS CPU%: The percent of 1 CP that is being used by the combination of all the CICS address spaces. This is as reported by RMF Monitor I under the ---APPL %--- heading for that Workload report that contains all the service classes.
- LPAR Busy %: The LPAR busy as taken from the RMF CPU Activity reports.
- CPs used: Total CPU used by the CICS address spaces to run the workload.

Results

This section contains the results extracted from the RMF reports.

Table 1 shows key performance data extracted from the RMF I reports. It shows a range of five different transaction rates running on the 18 CP LPAR, starting at 38 K TPS, up to a maximum in this configuration of 174 K TPS.

Table 1 CICS V5.3 results with 18 CPs

ETR	TOR CPU%	AOR CPU%	FOR CPU%	Total CICS CPU%
38667.05	34.99	310.84	14.90	360.73
57796.81	57.15	439.70	22.38	519.22
67206.91	63.80	508.98	25.82	598.59
110560.90	120.99	793.78	42.41	957.18
174624.30	210.45	1232.60	65.34	1499.40

Figure 9 shows only the CPU used by the CICS address space. CPU usage grows linearly as the transaction rate increases, demonstrating good scalability.

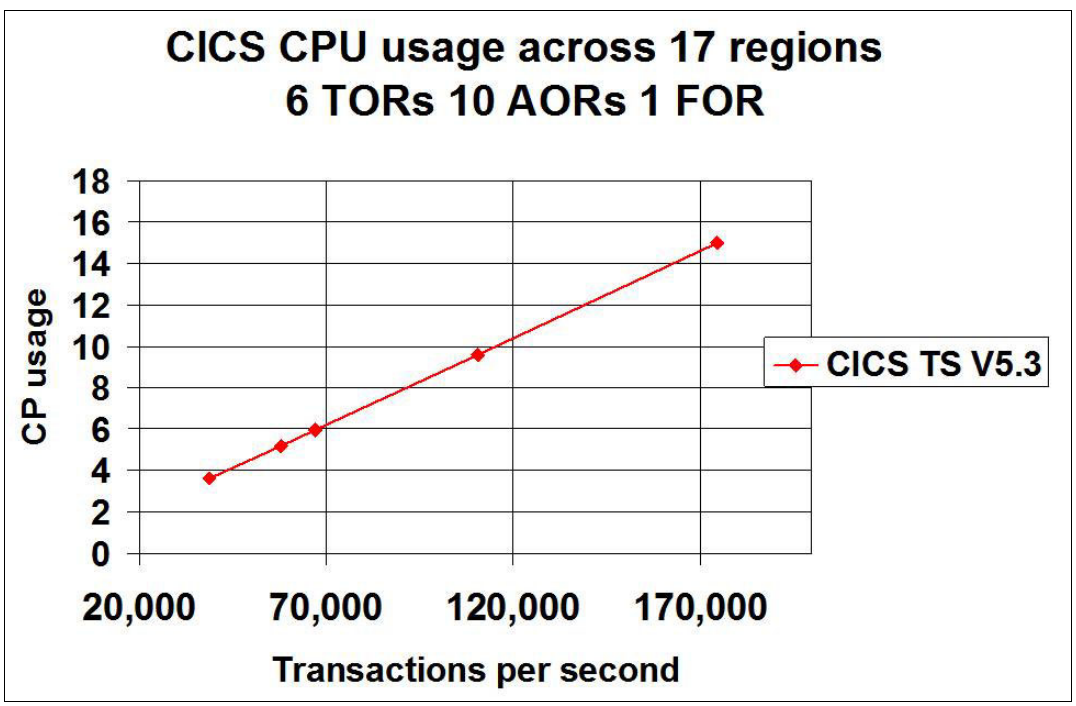

Figure 9 CICS CPU versus ETR for the 18 CP LPAR

More detail from the higher ETR

This section highlights more of the performance data from the upper measurement viewpoint.

Figure 10 is the RMF Workload Five-Minute Interval report extracted from RMF. It represents the 17 CICS regions and shows a total of 174624 TPS. Note also that 52545856 transactions ended in total and of those, 6996693 were executed (routed from TORs to AORs). The difference between the ended and the executed represent the transactions that arrived directly into the AORs via TCP/IP.

```
REPORT BY: POLICY=POLICY     WORKLOAD=CICSCPU
                                    All cics batch region cpu

-TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---   SERVICE TIME  ---APPL %---
AVG      17.00  ACTUAL                 1  SSCHRT 16368  IOC      2464K  CPU 4021.378  CP    1499.4
MPL      17.00  EXECUTION              0  RESP     0.0  CPU    241887K  SRB  473.494  AAPCP   0.00
ENDED 52545856  QUEUED                 0  CONN     0.0  MSO          0  RCT    0.000  IIPCP   0.00
END/S 174624.3  R/S AFFIN              0  DISC     0.0  SRB     28481K  IIT   16.978
#SWAPS       0  INELIGIBLE             0  Q+PEND   0.0  TOT    272832K  HST    0.000  AAP      N/A
EXCTD  6996693  CONVERSION             0  IOSQ     0.0  /SEC    906696  AAP      N/A  IIP     0.00
AVG ENC   0.00  STD DEV                4                                IIP    0.000
REM ENC   0.00                                          ABSRPTN    53K
MS ENC    0.00                                          TRX SERV   53K

TRANSACTIONS APPL% :   TOTAL :   CP 1493.8   AAP/IIP ON CP   0.00   AAP/IIP   0.00
                       MOBILE :  CP 648.53   AAP/IIP ON CP   0.00   AAP/IIP   0.00
```

Figure 10 RMF Workload Five-Minute Interval report

The new capability for capturing CICS transaction CPU time is shown in Figure 10, where the total time is 1493.8% of 1 CP. Out of that, 648.53% of a CP has been identified as Mobile CPU usage.

Figure 11 shows the report class for the transactions arriving on the new TCP/IP service (Connection type) classification, named HTTP1. It shows a transaction rate of 151342 TPS, all of which ran in the regions in which they arrived (EXCTD 0). All of these transactions were classified as MOBILE and as shown, used 648.53% of 1 CP (6.48 CPs). This ties in with the ended – executed value in Figure 10 on page 14.

```
REPORT BY: POLICY=POLICY                                 REPORT CLASS=SCALHTTP
                                                         DESCRIPTION =

-TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT
AVG       0.00  ACTUAL                1
MPL       0.00  EXECUTION             0
ENDED 45540056  QUEUED                0
END/S 151342.1  R/S AFFIN             0
#SWAPS       0  INELIGIBLE            0
EXCTD        0  CONVERSION            0
AVG ENC   0.00  STD DEV               1
REM ENC   0.00
MS ENC    0.00

TRANSACTIONS APPL% :   TOTAL :   CP 648.53   AAP/IIP ON CP   0.00   AAP/IIP   0.00
                       MOBILE :  CP 648.53   AAP/IIP ON CP   0.00   AAP/IIP   0.00
```

Figure 11 CICS TCP/IP service report class

In total, there were 151 K HTTP TPS (52545856 – 6996693 / 300) = 151 K TPS (shown in Figure 12 on page 17) and 23 K LU2 transactions that were routed from the TORs to the AORs.

Data summary at maximum throughput

Table 2 shows some of the key data points for the 18 CP measurement at maximum throughput.

Table 2 LPAR data summary for 18 CPs

Measurement	**Value**
Total TPS	174,624
HTTP TPS	151,342
LU2 TPS	23,251
CICSPlex SM routed TPS	23,251
Shared data table reads per second	831,835
Reads for update per second	8,257
Rewrites per second	8,257
Physical coupling facility log writes per second	10,878

CICS V5.3 comparison with CICS V5.2

This section contains a comparison of CICS TS V5.3 with CICS TS V5.2 running the same workload and using the same configuration. The main advantage that CICS TS V5.3 displays in this benchmark from a purely performance point of view is that it can eliminate the need to run the CWXN transaction, although there are other performance benefits in CICS TS V5.3 that also contribute to the differences with CICS TS V5.2.

One of the performance improvements that we were able to show in this benchmark is the more efficient TCP/IP processing within CICS TS V5.3. HTTP requests that arrive in CICS (and that do not require the use of an analyzer program and are not using CICS SSL) are optimized in CICS TS V5.3. The transaction that runs the target application or pipeline (the 'CWBA' or 'CPIH' transaction) can be attached directly by the socket listener task ('CSOL') without the need to attach the 'CWXN' transaction first. Eliminating CWXN reduces CPU usage and also reduces the DASD space required to store SMF records when CICS Performance monitoring is active.

CICS TS V5.3 has a new interaction with z/OS WLM in which it records transaction CPU usage for later post processing with RMF and, which enables a simpler method for the identification of the CPU time eligible for mobile workload pricing. Transaction CPU times can now be recorded without the need to turn on CICS Performance monitoring.

Also in CICS TS V5.3, there is an improved algorithm for selecting MRO sessions to be used for transaction routing and function shipping.

There is no transaction CPU reporting using z/OS WLM in CICS TS V5.2. This is a feature of CICS TS V5.3.

CICS TS V5.2 18 CP results summary

Table 3 shows some of the key data points for the 18 CP measurement when using CICS TS V5.2.

Table 3 Table of CICS TS V5.2 results

ETR	TOR CPU%	AOR CPU%	FOR CPU%	Total CICS CPU%
35207.41	31.06	311.57	13.09	355.73
39058.23	35.79	342.59	14.60	392.98
57805.34	58.31	488.47	21.75	568.53
68290.13	66.50	571.52	25.24	663.26
132781.79	141.53	1071.40	48.51	1261.40

As shown in Figure 12 and in the results in Figure 10 on page 14, CICS TS V5.3 reaches a higher throughput rate and uses less CPU than CICS TS V5.2.

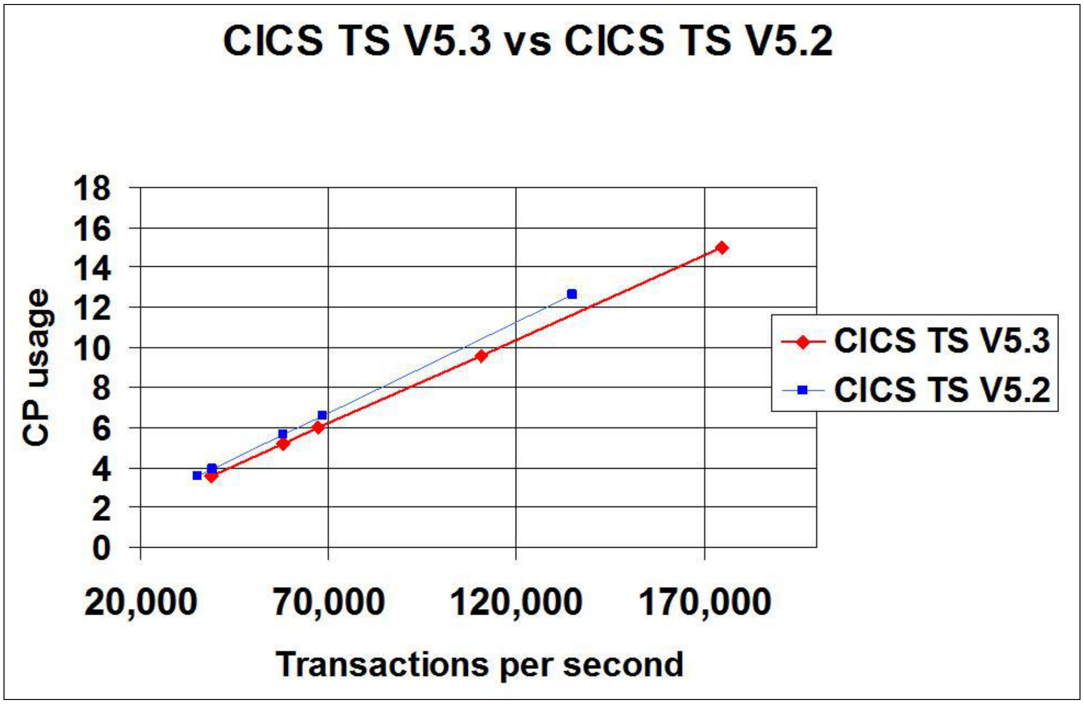

Figure 12 Comparison of TPS and CPU between CICS TS V5.2 and CICS TS V5.3

The main reason for the lower throughput rate for CICS TS V5.2, as shown in Figure 12, is because the AORs became constrained by the QR TCB in each of the AORs at a lower transaction rate. A slight increase in response times also caused more concurrent transactions in the system at any one time, leading to an increase in MXT conditions, compared to CICS TS V5.3. Higher throughput in CICS TS V5.2 can be achieved by increasing the number of AORs, but that did not form part of this exercise.

CICS TS with performance class monitoring

CICS TS V5.3 now provides the ability to identify and collect transaction CPU times using WLM classifications for possible inclusion for mobile pricing discounts. Before CICS TS V5.3, to identify and accumulate CPU times for certain transaction types, performance class monitoring would need to be turned on. This has a CPU cost associated with it, and additional DASD would be needed to store the SMF data.

This section compares the same CICS TS V5.3 workload running with and without performance class monitoring.

Figure 13 shows a comparison of ETR and CPU for CICS TS V5.3, with and without performance class monitoring. With performance class monitoring active, the highest ETR achieved is reduced from 174 K to about 135 K TPS. The main reason for this is the increased use of the QR TCB becoming a constraint at the higher transaction rate.

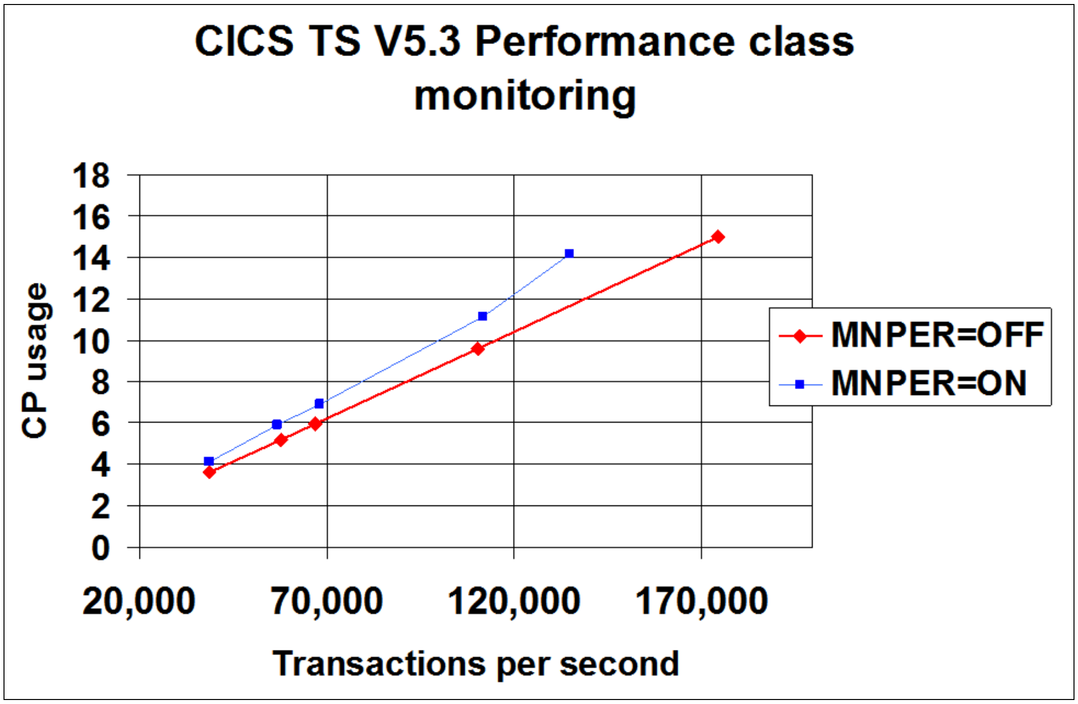

Figure 13 CICS with and without performance class monitoring

Also in Figure 13, the non-linear line shows the increase in CPU per transaction at the higher rate. When CICS regions become constrained, like any software system, it takes more CPU to process the higher numbers of active, concurrent transactions that are competing for resources. Also, the LPAR was already running at 97% before monitoring was turned on, so not only was the QR constraint causing a drop in throughput, but the LPAR was also constrained by the number of CPs available.

There can be other benefits to running CICS with performance class monitoring enabled, because the more detailed data that it provides can be used to understand performance characteristics of the system and diagnose any unexpected behavior in your application. While the results in this section do show a reduced throughput with performance class monitoring active, it also demonstrates that a considerably high transaction rate is still achievable with CICS performance class monitoring enabled. The results show that even when monitoring is enabled, the transaction rate we achieved is high (135,000 tps).

If performance monitoring is only being used to measure the CPU time used by transactions that are eligible for mobile pricing, the new facility in CICS TS V5.3 for accumulating transaction CPU time using z/OS WLM classifications can be used instead, giving benefits in terms of CPU savings.

Figure 14 shows an example of CPU attribution using z/OS WLM as an alternative to using performance class monitoring. This figure is an extract from RMF of the report by z/OS WLM Policy that shows the total MOBILE CPU for this system across all address spaces.

```
REPORT BY: POLICY=POLICY
                  Policy for plex 3

-TRANSACTIONS-  TRANS-TIME HHH.MM.SS.TTT  --DASD I/O--  ---SERVICE---   SERVICE TIME  ---APPL %---
AVG       91.43  ACTUAL                 1  SSCHRT 16399  IOC       2469K  CPU 4029.814  CP     1669.7
MPL       91.43  EXECUTION              0  RESP     0.0  CPU     242395K  SRB  977.529  AAPCP    0.00
ENDED 52545857  QUEUED                 0  CONN     0.0  MSO           0  RCT    0.004  IIPCP    0.01
END/S 174624.3  R/S AFFIN              0  DISC     0.0  SRB      58799K  IIT   17.040
#SWAPS       59  INELIGIBLE             0  Q+PEND   0.0  TOT     303662K  HST    0.000  AAP       N/A
EXCTD  6996693  CONVERSION             0  IOSQ     0.0  /SEC      1009K  AAP      N/A  IIP      0.00
AVG ENC   2.00  STD DEV                4                                 IIP    0.000
REM ENC   0.00                                          ABSRPTN    11K
MS ENC    0.00                                          TRX SERV   11K

TRANSACTIONS APPL% :    TOTAL :  CP 1664.1   AAP/IIP ON CP   0.01   AAP/IIP   0.00
                       MOBILE :  CP 648.53   AAP/IIP ON CP   0.00   AAP/IIP   0.00
```

Figure 14 Mobile CPU at the Policy level

In this particular case, only the transactions classified as arriving over a named CICS TCP/IP service are considered MOBILE. Figure 14 shows 6.48 CPs that are considered as Mobile CPU usage. It is this CPU service, expressed in millions of service units (MSUs) that is used as input into the Mobile Workload Pricing Tool (MWRT).

Results of running 26 CPs in CICS TS V5.3

The main focus for this paper has been on measurements in an 18 CP environment. However, we also had access to an additional 8 CPs for a limited time, and we used this opportunity to scale the workload up to an even higher limit.

This section contains the results from running the same applications on an LPAR with 26 CPs. To achieve higher throughput and to capitalize on the additional CPs, another five AORs were added to the configuration, producing 15 AORs total, and the number of TCP/IP clients was increased from 6000 to 9000. This adjusted the proportion of HTTP transactions compared to LU2 transactions.

Table 4 shows extracts from the RMF reports from the 26 CP benchmark. Only data from the highest ETR measurement point has been shown where a transaction rate of 227 K per second was reached. This demonstrates the considerable qualities that CICS TS has in terms of throughput scalability.

Table 4 LPAR data summary for 26 CPs

Measurement	Value
Total TPS	227,884
HTTP TPS	196,597
LU2 TPS	22,564
LPAR busy	96.13%
TOR Appl%	237.19%
AOR Appl%	1811.50%
FOR Appl%	77.4%
Shared data table reads per second	1,095,805
Reads for update per second	8,685
Rewrites per second	8,685
Physical CF log writes per second	11,960

Summary

This performance benchmark demonstrates how CICS TS for z/OS can scale to handle workloads with notably high transaction rates, while retaining the reliability and efficiency demanded by enterprise business.

Optimizations within CICS TS V5.3 are shown to provide a marked improvement in performance for this benchmark. Simpler and more efficient ways to track mobile transaction CPU times for use with mobile workload pricing have been demonstrated in this paper.

The inherent capabilities of CICS, combined with these operational efficiencies, enable you to open your CICS solutions to new systems of engagement. This can allow you to provide new services, either internally within your organization or externally to clients, to gain further value from your IBM z Systems investments.

Authors

This book is one in a series focused on CICS performance. It is written by members of the IBM Hursley CICS development community. The subject matter in this series is based on feedback from CICS customers.

John Burgess is a Senior Software Engineer working in the IBM Hursley CICS performance team. He has over 20 years' experience with CICS and specializes in the performance of large systems.

Ian Mitchell is an IBM Distinguished Engineer and CTO for z Systems Middleware Runtimes. He has 27 years' experience as a technical leader in IBM Hursley and has been in the forefront of creating and introducing many of CICS' important innovations including Workload Management, Sysplex Support, Business Transaction Services, Java, Web Services, Event Processing and the IBM CICS Explorer®. In this role, Ian has worked with many of the technical leaders across all the mainframe technology components. Ian also has close relationships with many of the largest clients of IBM running critical systems using CICS and the mainframe.

Catherine Moxey is an IBM Senior Technical Staff Member in CICS technical strategy and planning, based at IBM Hursley in the UK. Catherine has over 30 years of experience as a software engineer, 26 of those with IBM. Her areas of focus include CICS Performance and Optimization, event processing support, and analytics.

Thanks to the following people for their contributions to this project:

- David Harris, Andy Bates, Fraser Bohm, and Ian Burnett for contributing ideas and support for the benchmark
- Martin Cocks for leading the performance improvement work in CICS TS V5.3
- Chris Baker for his numerous improvements to CICS Performance Monitoring
- Irene Stahl and Horst Sinram for providing z/OS WLM support for CPU time classification
- Peter Mailand for the RMF processor enhancements

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

For more information about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  https://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new IBM Redbooks® publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

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 grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

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 jurisdictions 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 websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

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

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

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 actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate 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. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

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

CICS®	IBM z Systems™	Resource Measurement Facility™
CICS Explorer®	IBM z13™	RMF™
CICSPlex®	Redbooks®	z Systems™
HiperSockets™	Redpaper™	z/OS®
IBM®	Redbooks (logo) ®	z13™

The following terms are trademarks of other companies:

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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

REDP-5320-00

ISBN 0738454958

Printed in U.S.A.