Java – CICS TS V2.2

Performance test on zSeries

« Technical Report »

Contact

PSSC zSeries Benchmark Center Manager: Arnaud Tirmarche (a_tirmarche@fr.ibm.com)
PSSC Project Leader: Jocelyn Denis (jdenis@fr.ibm.com)

Preface

This Technical Report is the deliverable provided by the PSSC as result of the Java - CICS TS V2.2 performance test project for the final customer

Executive Summary

This performance test of Java-CICS TS V2.2 transactions (Cash Management and Card Management applications) was executed by IBM-PSSC in Montpellier (France) with the help of IBM Business Partner from July 1st to July 19th. It has the following characteristics:

1 - Final customer request :
- Find out how much resources (CPU, I/O, and memory) the 10 representative transactions will require in the defined zSeries high performance configuration.
- Find out what the end users response time will be.

2 - Final customer objectives :
- Simulate 200 concurrent active users, generating an approximate **transaction rate of 18 transactions per second**, and getting an **end users average response time lower than 1 second**.

3 – Performance test results:
- Capacity planning (required zSeries resources to run the customer transactions defined activity):
  - CPU load: approx. **165 MIPS** (approx. **9 MIPS per transactions/sec**) to execute 19 transactions/sec
  - I/O rate: approx. **120 I/O per second**
- High level of performance (end users average response time):
  - **Lower than 0.2 sec** when executing the 19 transactions/sec
- Good scalability:
  - For a transaction rate multiplied by 4 (more than 71 transactions/sec), the CPU load has been multiplied by 4.7 (780 MIPS), and the end users average response time has been multiplied by 2.5 (0.5 sec).
- Good stability:
  - A long run has been executed during **8 hours**, for a transaction rate of **19 transactions/sec**:
    - CPU load: approx 30% (approx. **245 MIPS = 12 MIPS per transaction/sec**) - Increase due to DB2 tables extends during the test
    - End users average response time **lower than 0.2 sec** (except for transaction GOC12C).

Table of Contents

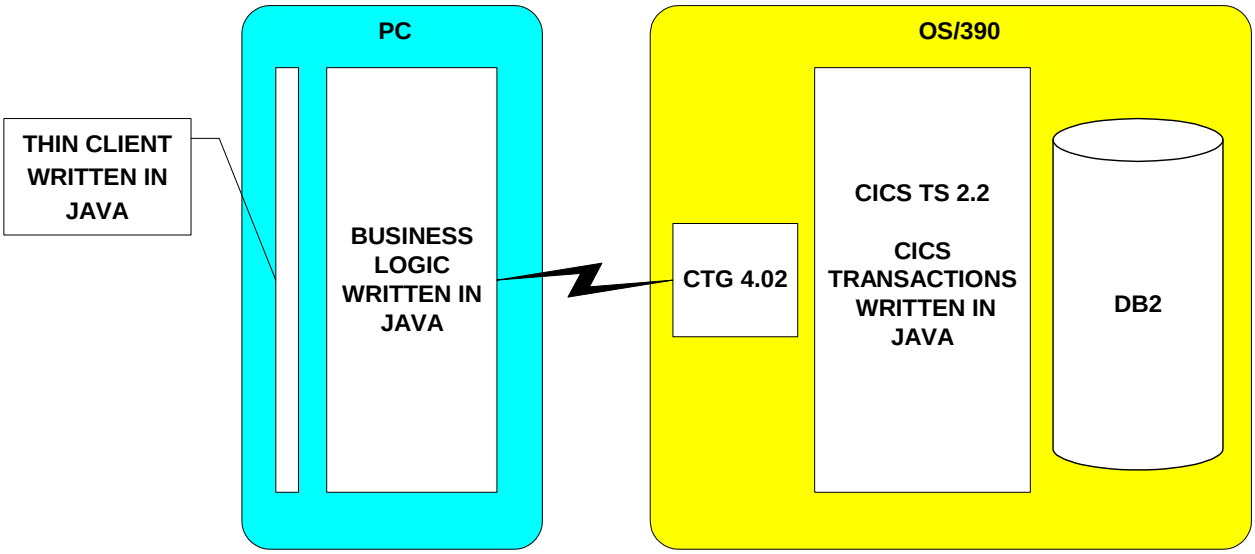
1. Project description

1.1 Customer request

There are two Software Development Projects in customer:
- Cash Management Software
- Back Office of the Card Management Software

The first one will have approx. 600 CICS transactions. The second one will have approx. 1500 CICS transactions. The IBM partner can estimate the number of CICS transactions because they spent eight months in the customer doing Business Process Reengineering.

When IBM partner signed the contract, the customer wanted IBM partner to develop software in Client/Server Architecture. The IBM partner thought there was a much better solution, using WebSphere and Thin Client or Web Browser on PC. At this time, WebSphere 3.5 on OS/390 was considered as not recommended for production environment. Before the end of the software development, WebSphere 4.1 for z/OS was released and ready for production environment. But the customer was prudent regarding this very new technology, and decided to have Client Server software using a proven architecture:

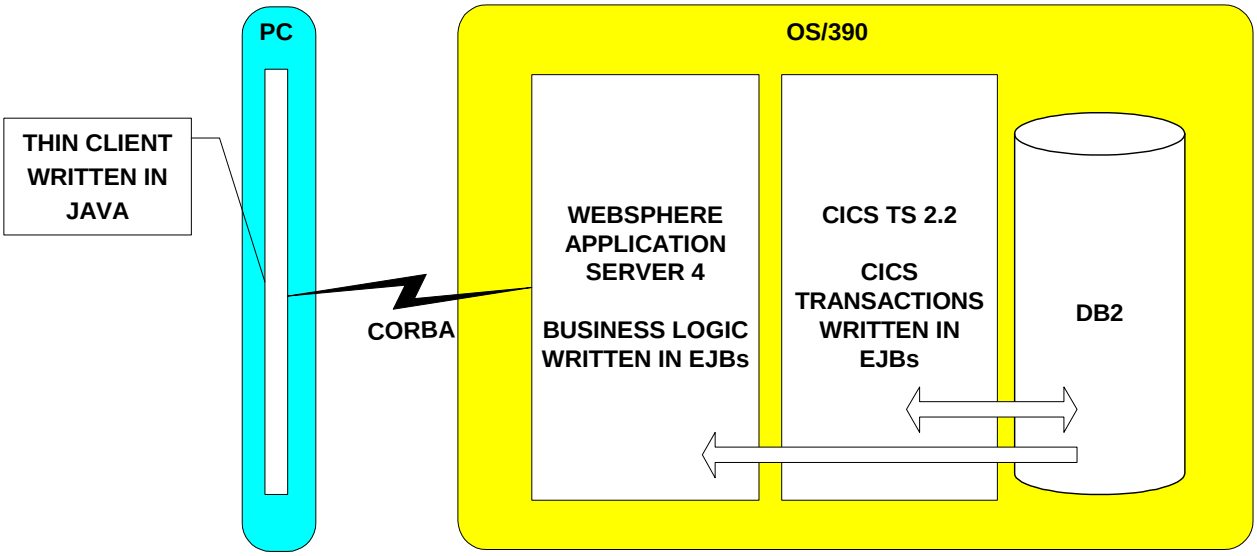


Client software has been written in Java and Server Software in COBOL. Development team was split into two groups; COBOL programmers and JAVA programmers. Every task has been divided to both of teams and development progress has depended on synchronization of this teams.
Many other difficulties have also been realized:
- problem determination was extremely difficult (debugging on two different platforms at the same time),
- when modifying program code in Java, COBOL program had to be modified and vise versa

- transferring data between client and server was possible only with mapping data and transforming all data types into strings and packing them into fixed length structure (extremely time consuming comparable to java-java communication which is done automatically using serializable objects)
- software design was inconsistent; object design (java) combined with COBOL design (procedural)
- etc.

After a short period of time, the IBM partner realized that so many transactions couldn't be produced in two different programming environments on time. Then IBM partner suggested having CICS transactions written in Java.
The customer IT staff considered that there was a big risk of using CICS TS V2.2 and CICS transactions written in Java because they were new products on the market. They were not confident in the stability and performance of the JVM inside the CICS. They requested proof that CICS TS V2.2 and the JVM would perform and be reliable enough in a highly loaded production environment.
Also they were interested in estimating the processor power and amount of memory required by the application.

Then the IBM partner tested the developed software on a MultiPrise 3000 mainframe, with good results in terms of response time, but with insufficient workload to be near the production environment of customer. Once the assumptions will be proved to be true, the software will be easily portable to WebSphere for z/OS. Then the future target architecture (not included in the benchmark) could be:



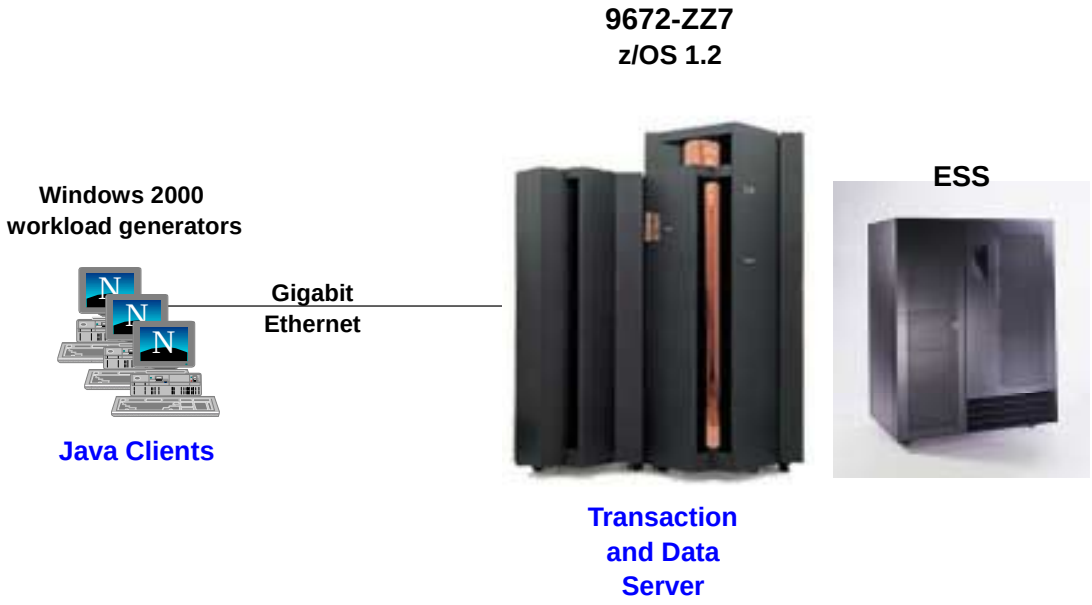
In conclusion, the benchmark request is to:

- Find out how many resources the 10 Java-CICS TS V2.2 representative transactions require in the defined zSeries configuration (z/OS 1.2 on zSeries 9672-ZZ7).
- Find out what the end users response time is.

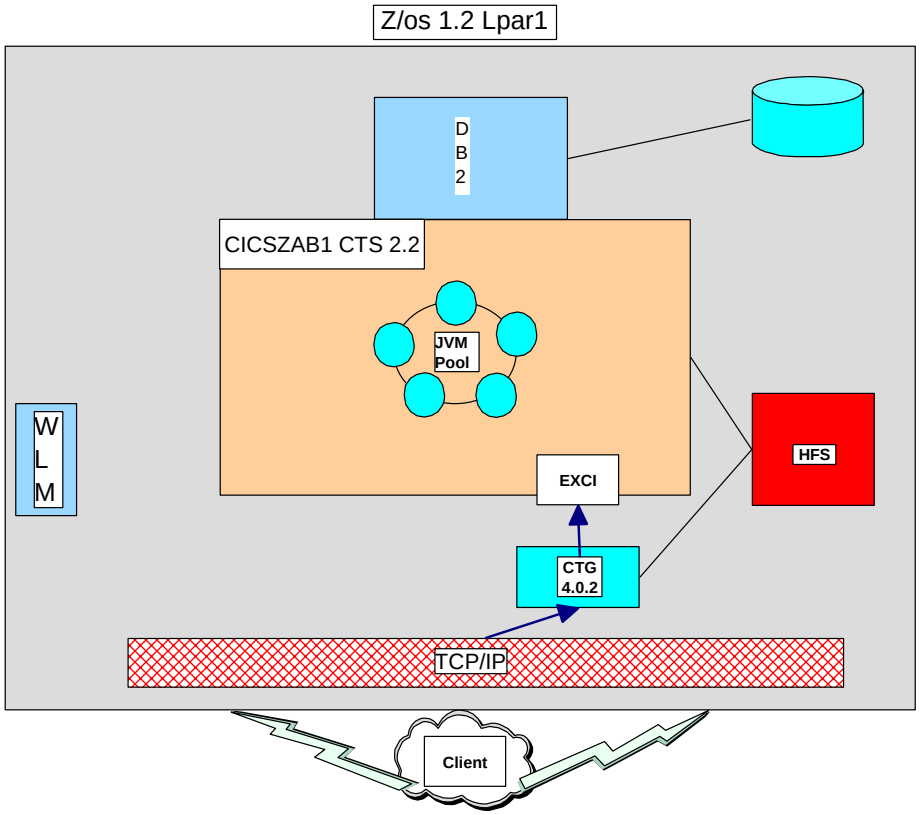
The hardware configuration used for this benchmark is probably more powerful than the one that will be implemented by the customer, but such an unconstraint environment is necessary to estimate the required resources (CPU, I/O, and memory) to run the application.

1.2 Technical architecture

The performance tests are executed with the following configuration:



The logical architecture implemented for the test (with CTG) is:



1.3 Objectives

The objectives of the online transactions performance test are:

- OLTP performance test:
  - Based on 10 representative CICS transactions:
    - GOC00C: application start.
    - GOC10C: cash desk monitoring select.
    - GOC11C: deposit box list.
    - GOC12C: bank branch summary select.
    - GOC14C: specification confirm.
    - GOC15C: specification confirm select.
    - GOC20C: local treasury state select.
    - GOC22C: cash desk status update.
    - GOC28C : cash specification insert and update.
    - GOC32C: insert deposit box cont.
  - Measure the end users average response time:
    - Simulate the activity of **200 concurrent active users.**
    - Generating a transaction rate of approx. **18 transactions/sec**.
    - The end users average response time objective is to be lower than **1 second**.

1.4 Confidentiality

As an IBM customer, the data provided by the customer are treated with respect of the confidentiality rules:

- Non-divulgation engagement.
- Confidential data will be physically erased at the end of the project.
- The customer provided media will be returned to the customer.

2. Project environment

2.1 PSSC resources

In this project, the PSSC is responsible to implement the hardware resources, the software resources, the skilled people resources, and to operate the defined performance tests.

2.1.1 Transaction and Database Server: IBM @server zSeries hardware

- One @server 9672-ZZ7 computer, running 1 dedicated Logical Partition:
  - 6 dedicated CP
  - 4 GB of memory (2 GB Central and 2 GB Expanded)
  - 8 ESCON channels
  - One Ethernet OSA card
  - One Gigabit Ethernet OSA card
- One Enterprise Storage Server DASD sub-system, 700 GB of storage (half capacity).

2.1.2 Workload generation hardware and workstations

- Three dedicated Windows 2000 Intel computers: Pentium IV 1.6 GHz - 512 MB RAM - 40 GB disks.
- Six dedicated Windows 2000 workstations.

2.1.3 Network

- Between the clients (Load Runner workload generators) and the network router, the network connections are dedicated Fast Ethernet ones.
- Between the network router and the CICS server, the network connection is a dedicated Gigabit Ethernet one.

2.1.4 Transaction and Database Server software

- z/OS V1R2.0
- z/OS V1R2.0 DFSMS
- z/OS V1R2.0 IBM Communication Server (TCP/IP)
- z/OS V1R2.0 RMF
- CICS TS V2.2 and CICS Performance Analyzer V1.2
- CTG 4.0.2
- JDK 1.3.1
- DB2 7.1

2.1.5 Workload generation software

- Windows 2000.
- Mercury Load Runner 7.5 workload generation software tool.

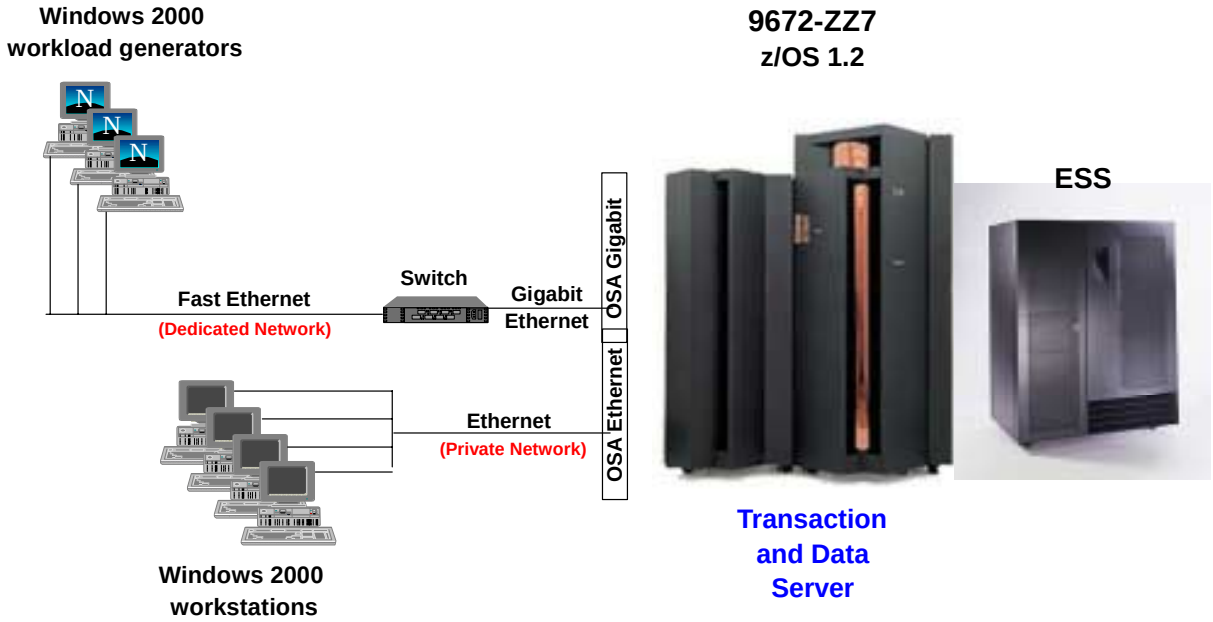
2.1.6 Performance measurement

- The PSSC is responsible to implement the standard IBM zSeries measurement products (SMF, RMF), and to collect the data.
- The PSSC is responsible to do, based on these figures, the required system tuning operations (z/OS, DFSMS, TCPIP, CICS TS, Java, DB2).

2.1.7 IBM PSSC people resources

- Project leader
- zSeries hardware installation support
- z/OS, DFSMS, and SMF/RMF support
- CICS support
- Java support
- DB2 support
- Load Runner support
- Network support

These skilled people are coming from the **PSSC zSeries Benchmark Center**.



2.2 IBM partner resources

The IBM partner team has provided with the following support for the benchmark:

- Provided the PSSC with **the 10 representative CICS transactions** to be tested, with the associated support for installation and maintenance.
- Provided the PSSC with Java skilled persons, who have developed the application, and participate to the workload generation scenarios development.

3. Performance test results

3.1 Scenario 1: performance test

3.1.1 Scenario description

This scenario consists in simulating the activity of **200 concurrent active users**, running the **10 representative transactions**, with a transaction rate of **19 transactions/sec**, during 22 minutes (6 to 7 minutes of ramp-up, and 15 to 16 minutes of stabilized activity).
A detailed description of the Load Runner programs used can be found in Appendix.

Based on previous tests, the optimum in terms of number of JVM was found to be **10** for the requested activity, each JVM memory being **56 MB**, and CICS MaxTasks being **60**: reasonable JVM wait, and reasonable memory usage were reached.

3.1.2 Scenario results

3.1.2.1 Load Runner metrics

Transactions:

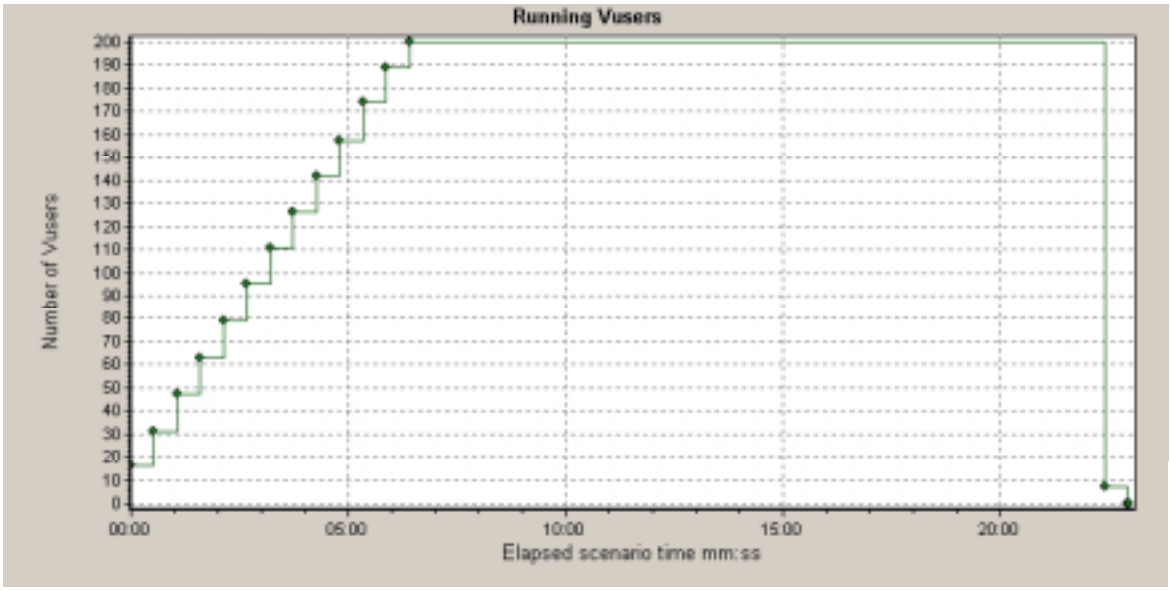
Total passed:	Total failed:	Total Stopped:
22,613	474	0

End users average response time:

Transaction Name	Mini	Average	Maxi	90 Pct	Pass	Fail	Stop
G22C	0.031	0.178	11.734	0.24	2,300	3	0
G28C	0.016	0.174	11.906	0.23	2,307	0	0
G32C	0.016	0.231	14.516	0.26	2,290	11	0
GOC00C	0.016	0.151	10.344	0.23	2,311	3	0
GOC10C	0.109	0.235	11.922	0.23	2,275	1	0
GOC11C	0.016	0.127	13.391	0.17	1,898	407	0
GOC12C	0.047	0.129	11.922	0.1	2,309	3	0
GOC14C	0.094	0.214	11.969	0.18	2,281	37	0
GOC15C	0.031	0.133	15.875	0.15	2,331	3	0
GOC20C	0.031	0.174	11.563	0.23	2,311	6	0

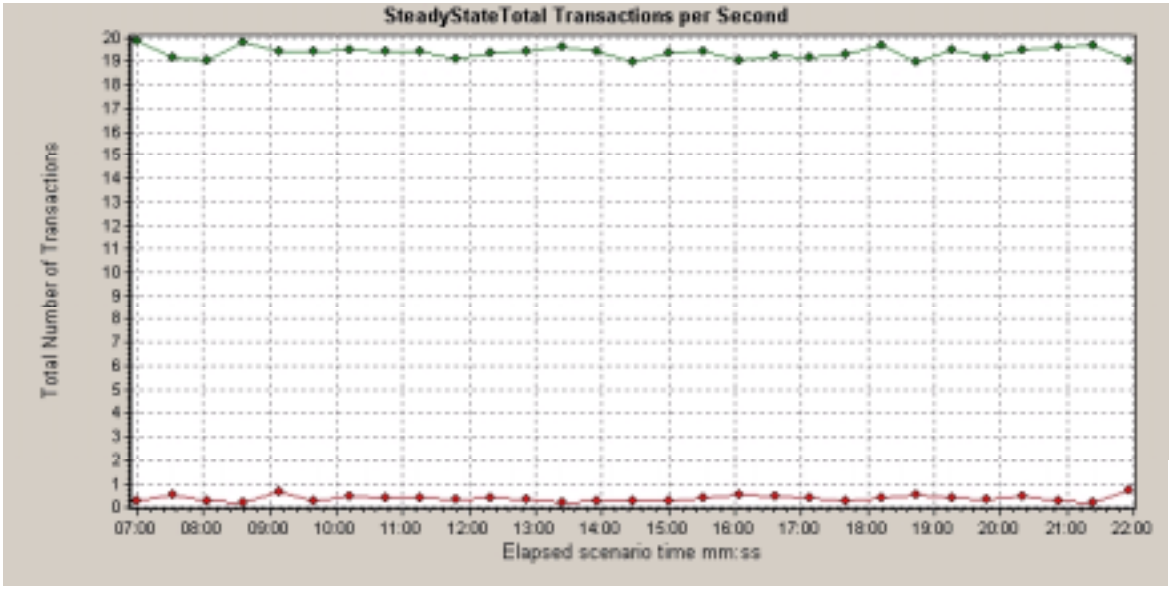
Note: failure is caused by business logic validation, locks, etc. (not caused by wrong system setup or bug in software).

Number of concurrent active users:



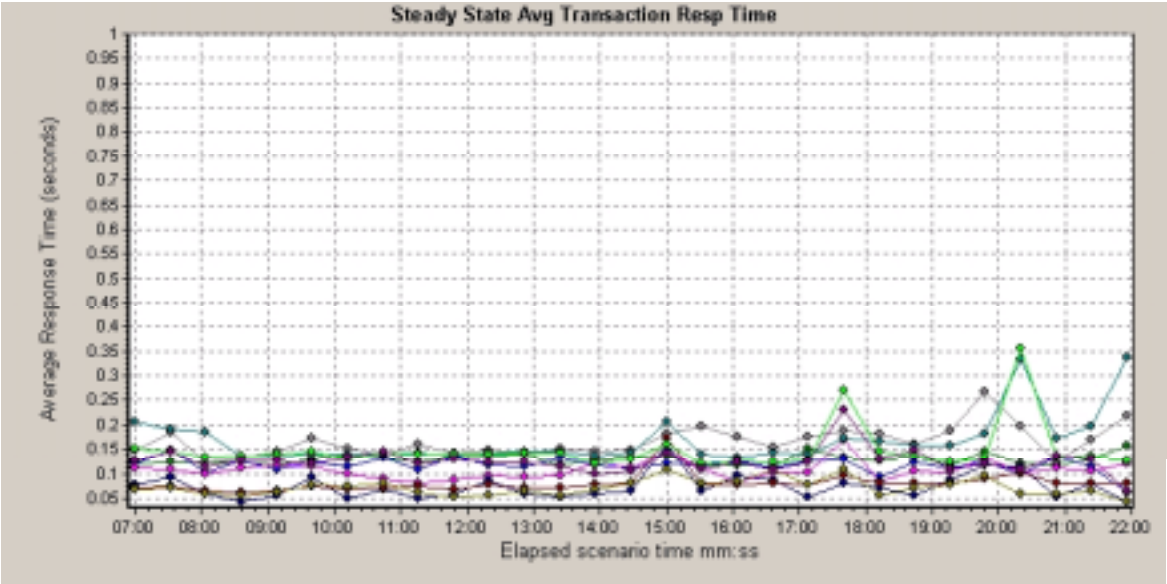
Color	Scale	Measurement	Graph's Minimum	Graph's Average	Graph's Maximum	Graph's Median	Graph's Std. Deviation
	1	Run	0	89.813	200	95	67.387

Total number of transactions per second (steady state):



Color	Scale	Measurement	Graph's Minimum	Average	Graph's Maximum	Graph's Median	Graph's Std. Deviation
	1	Fail	0.219	0.391	0.75	0.406	0.123
	1	Pass	18.938	19.35	19.844	19.406	0.237

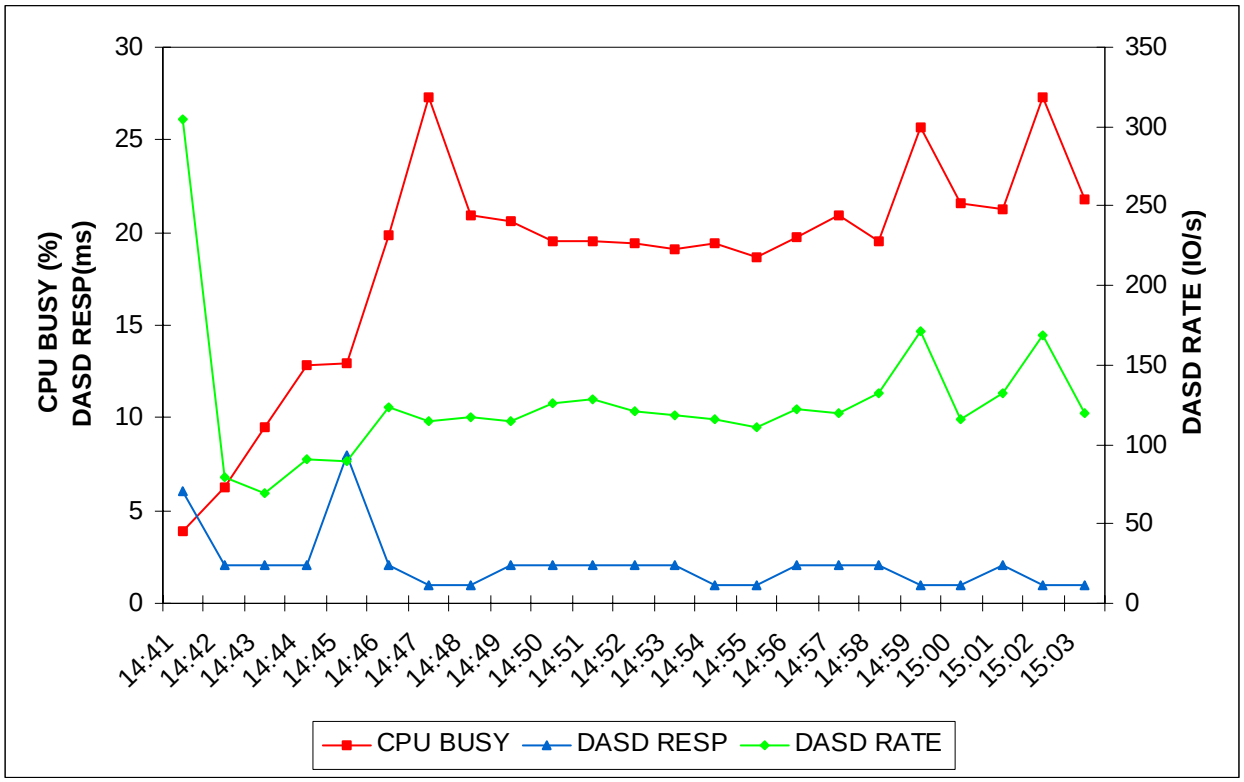
End users average transaction response time (steady state):



Color	Scale	Measurement	Minimum	Average	Maximum	Std. Deviation
	1	G22C	0.031	0.128	1.5	0.098
	1	G28C	0.016	0.119	1.547	0.09
	1	G32C	0.016	0.163	1.547	0.147
	1	GOC00C	0.016	0.107	5.469	0.16
	1	GOC10C	0.109	0.163	6.359	0.185
	1	GOC11C	0.016	0.074	2.203	0.126
	1	GOC12C	0.047	0.09	2.266	0.081
	1	GOC14C	0.094	0.148	6.391	0.264
	1	GOC15C	0.031	0.072	2.016	0.095
	1	GOC20C	0.031	0.127	5.781	0.165

3.1.2.2 zSeries metrics

zSeries CPU load and I/O activity:



zSeries memory usage:

- Main indicators for real memory:

```
  HIGH UIC  MIGR AGE
MIN   254   914083
MAX   254   916104
AVG   254.0 915092.8
```

```
                                                                    FRAME AND SLOT COUNTS
-------------------------------------------------------------------------------------------
                        CENTRAL STORAGE                          EXPANDED STORAGE
                    ---------------------------              ---------------------------
(161 SAMPLES)            MIN        MAX        AVG              MIN        MAX        AVG
AVAILABLE                164    196,175     55,302          460,110    494,529    478,304
SQA                   11,100     11,131     11,125                0          0          0
LPA                    5,221      5,223      5,221                0          0          0
CSA                    5,692      6,620      6,235                0         30          0
LSQA                  11,999     12,872     12,573              855      1,036        942
REGIONS+SWA          287,938    482,886    427,690           29,736     64,155     45,969
TOTAL FRAMES         528,381    528,381    528,381          524,288    524,288    524,288
                          FIXED FRAMES
                    ---------------------------
NUCLEUS               10,234     10,234     10,234
SQA                   10,431     10,462     10,456
LPA                       65         65         65
CSA                    1,194      1,215      1,199
LSQA                   9,334     10,223      9,949
REGIONS+SWA            2,498      4,317      3,279
BELOW 16 MEG             134        149        136
BETWEEN 16M-2G           N/A        N/A        N/A
TOTAL FRAMES          33,786     36,468     35,183
```

- Configuration for Virtual Memory:

```
          NUMBER OF SAMPLES     138
            STATIC STORAGE MAP
         AREA         ADDRESS   SIZE
       EPVT          17F00000  1665M
       ECSA           BB87000   196M
       EMLPA          BB86000     4K
       EFLPA                0     0K
       EPLPA          839A000  55.9M
       ESQA           37B7000  75.9M
       ENUC           1000000  39.7M
       ----- 16 MEG BOUNDARY ------
       NUCLEUS         FC1000   252K
       SQA             EB4000  1076K
       PLPA            CEE000  1816K
       FLPA                 0     0K
       MLPA                 0     0K
       CSA             B00000  1976K
       PRIVATE           1000  11.0M
       PSA                  0     4K
MAXIMUM POSSIBLE USER REGION - 11056K BELOW AND 1657M ABOVE
```

As shown in these reports, due to the system configuration, memory was not an issue during the benchmark. No paging occurred.

zSeries network load:

- Network load (Gigabit Ethernet OSA card):

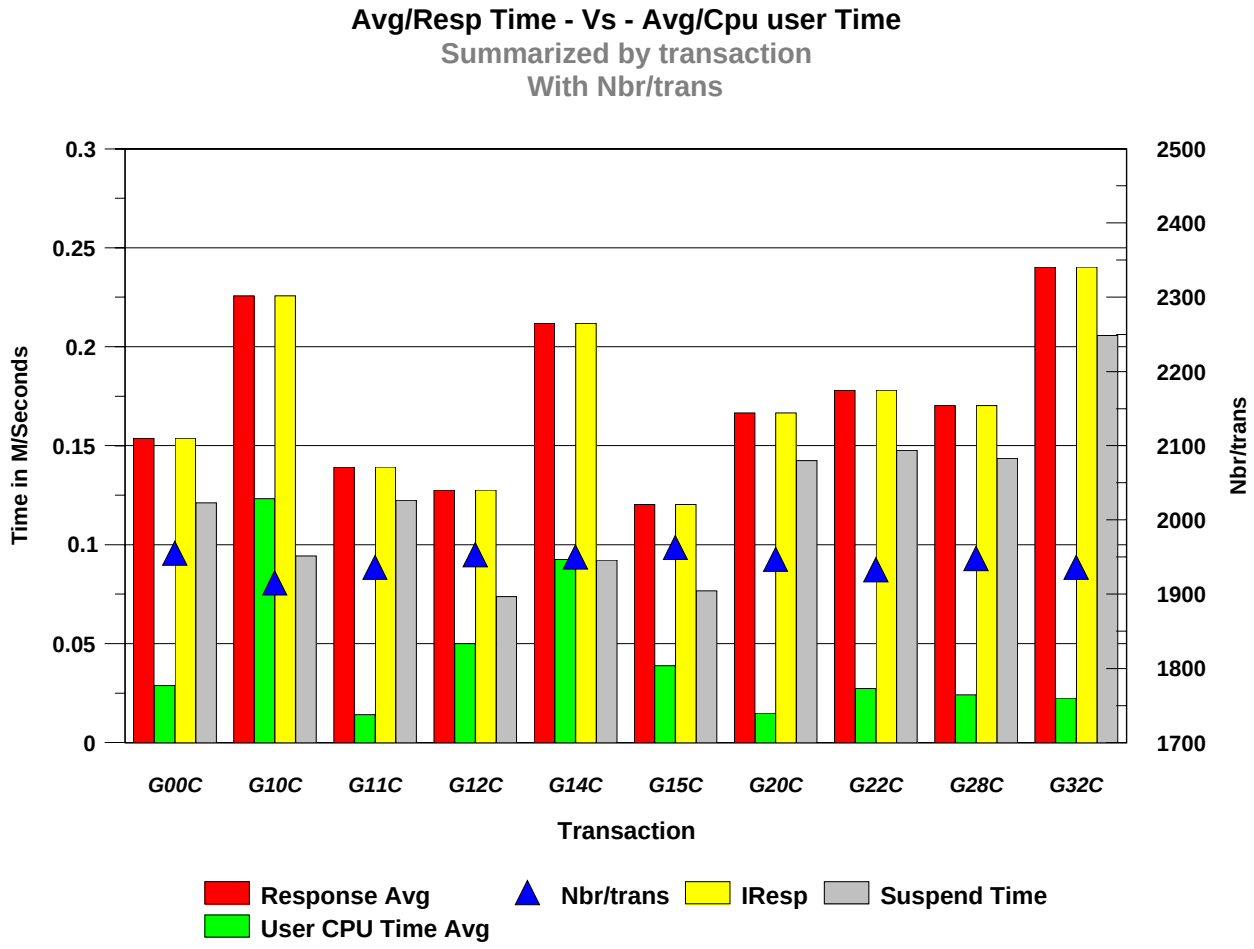
```
%Util: 3.96          (1)
%PCI Bus busy: 12.68 (2)
READ: 0.48 MB/s      (3)
WRITE: 0.49 MB/s     (4)

(1): Channel path utilization percentage
(2): Percentage of PCI bus cycles, the bus has been found busy for this channel in
relation to the theoretical limit
(3): Data transfer rates from the control unit to the channel
(4): Data transfer rates from the channel to the control unit
```

- Average transaction rate: 19.07 online transactions/sec

3.1.2.3 CICS and CTG metrics

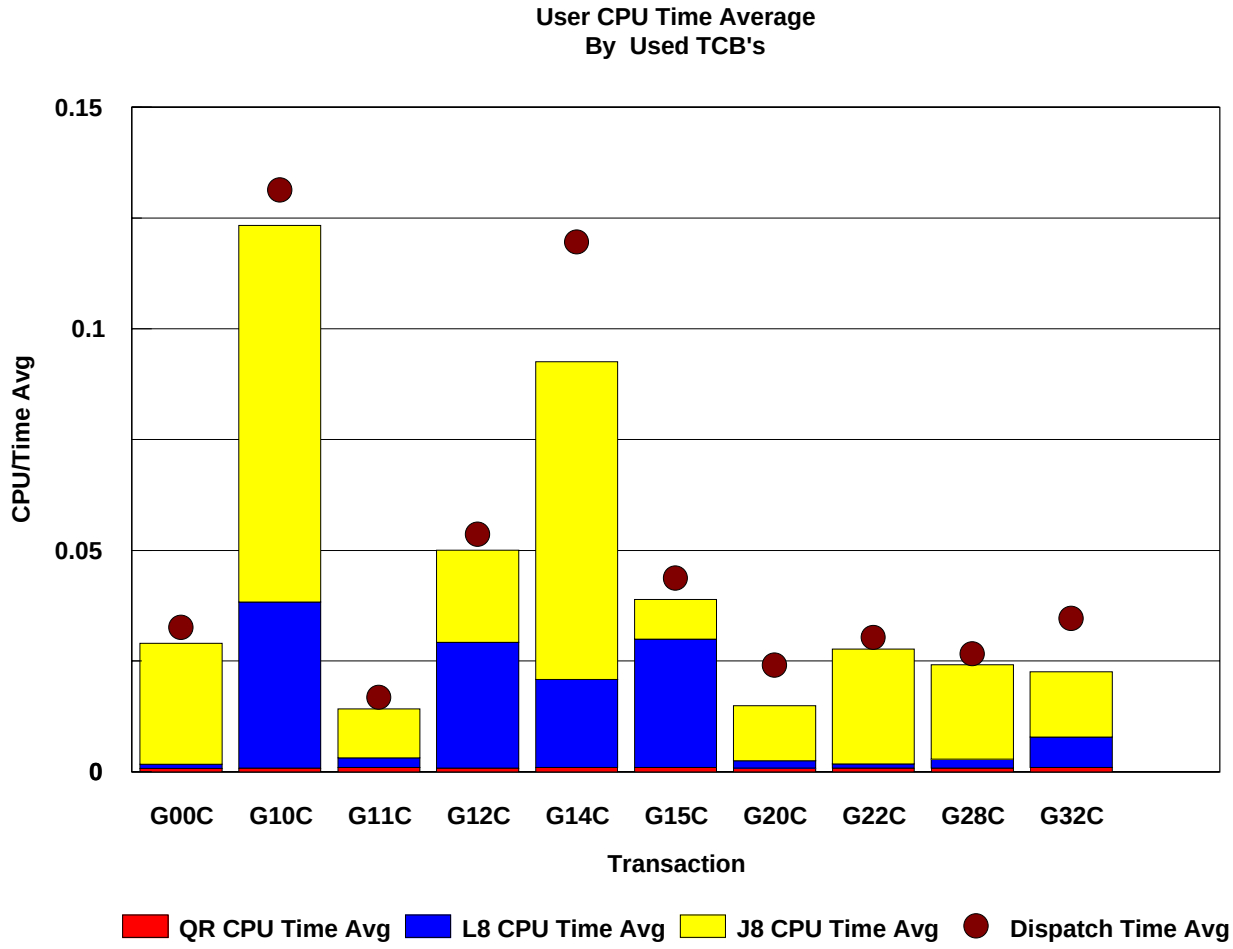
All CICS metrics come from the analyze of the SMF record using CICS Performance Analyzer V1.2. We have also used the online STAT provided by CICS and the End-of-Day statistics.

The following graphic represents the CICS average response time versus CICS average CPU time used, including the average Suspend time and Internal CICS response time, summarized by transaction name:



Remarks: the Avg Suspend Time is important and is described in Graphic 3 (Avg/Suspend Time Distribution). More investigation is required to reduce this Avg/Suspend Time.

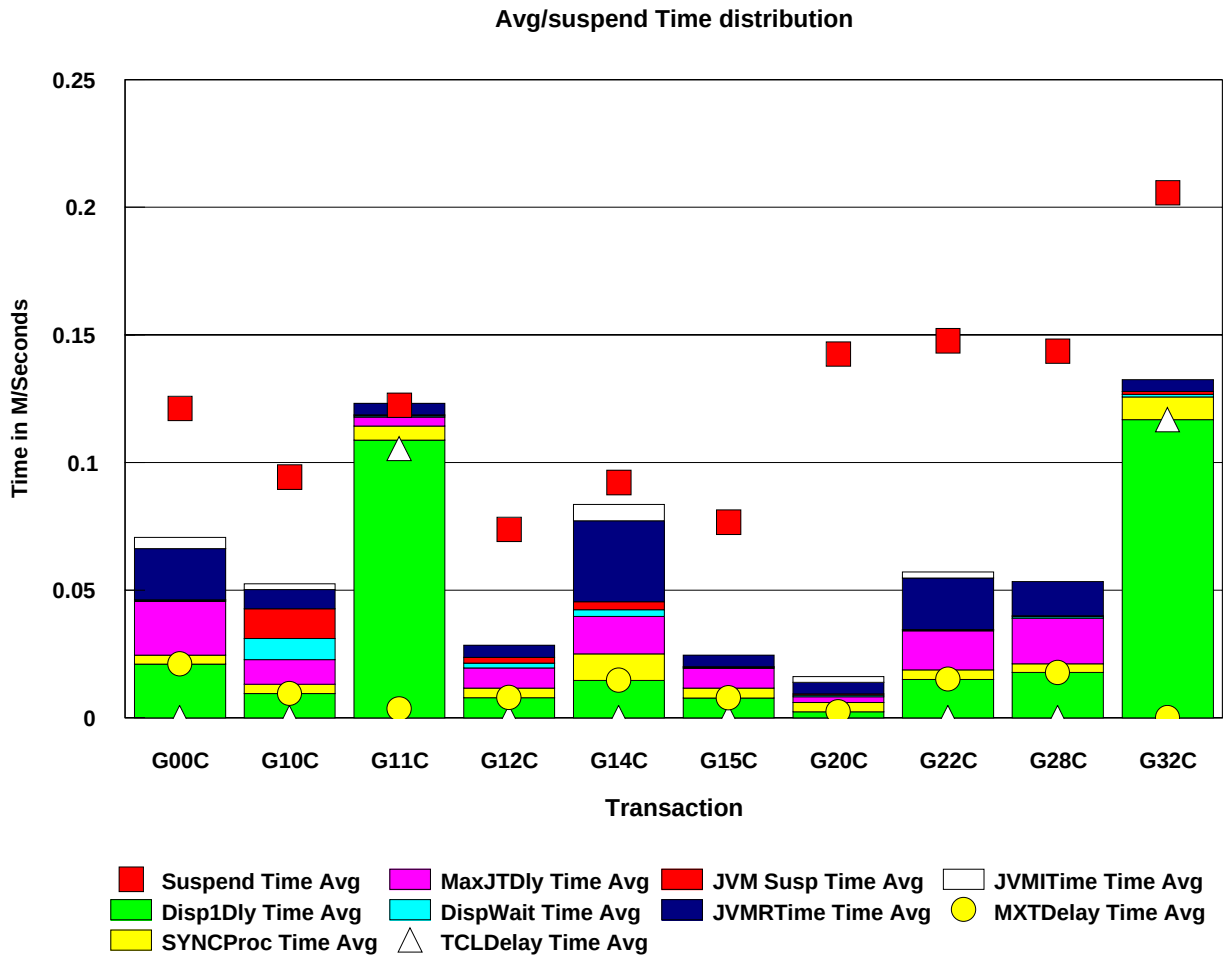
The following graphic represents the repartition in Average/Cpu time and Average/Dispatch Time for each used CICS TCB's during this benchmark (MS TCB & Key 8 TCB), summarized by transaction:



Remarks: for the Java application in CICS, the new OPEN TCB's J8 are used, and for this reason the principal CICS TCB QR remains free for classical CICS transactions (like COBOL, ASM, PL/1).

QR	=>	CICS TCB mode QR
L8	=>	CICS TCB mode L8 (Open TCB for DB2)
J8	=>	CICS TCB mode J8 (Open TCB for java , one TCB per JVM)
Avg/Dispatch Time	=>	Average elapsed time during which the user task was dispatched on each CICS TCB under which the task was executed.

The following graphic represents the repartition of the Average Suspend time summarized by transaction:



Remarks: Another type of wait was identified after the creation of this graphic and is not included in this chart; it is mainly caused by RMI suspend time (between 0,0009 to 0,03 sec); it is the total elapsed time the task was suspended by the CICS dispatcher while in the CICS Resource Manager Interface (RMI), a JC Wait time (for 0,002) , and RRMS Wait time (elapsed time in which the user task wait in doubt using resource recovery services for EXCI, it's the results of using the ECIrequest in extended mode). It should be necessary to further investigate.

Following is an extract of End-Of-Day statistics provided by DFHSTUP for JVMPOOL:

JVMPOOL STATISTICS

Total number of JVM program requests. . . . . :	9
Peak number of JVMs . . . . . :	10
Number of JVM program requests - Reuse specified. . . . . :	23072
Number of JVM program requests - JVM initialized. . . . . :	19
Number of JVM program requests - JVM mismatched . . . . . :	19
Number of JVM program requests - JVM terminated . . . . . :	9

Following is an extract of CICS ONLINE Statistics for Dispatcher TCB JVMpool:

TCB Pool . . . . . :	JVM
Current TCBs attached in this TCB Pool . .	10
Peak TCBs attached in this TCB Pool. . . . :	10
Peak TCBs attached in this TCB Pool. . . . :	10
Requests Delayed by Max TCB Pool Limit . . :	53
Total Max TCB Pool Limit delay time. . . . :	00:00:05.03215
Average Max TCB Pool Limit delay time. . . :	00:00:00.09494
Current TCBs in use in this TCB Pool . . . . . :	0
Peak TCBs in use in this TCB Pool. . . . . :	10
Times at Max TCB Pool Limit (MAXJVMTCBS) . . . :	61
Current Requests Delayed by Max TCB Pool Limit :	0
Peak Requests Delayed by Max TCB Pool Limit. . :	19
Total Delay time for current delayed . . . . . :	00:00:00.00000
Average Delay time for current delayed . . . . :	00:00:00.00000

Following is an extract of CICS Performance Analyzer WLM Services classes & Report classes for all user transactions classified:

V1R2M0 CICS Performance Analyzer
Workload Manager Activity Summary by Service Class

WKLD0001 Printed at 15:33:50 7/16/2002 Data from 14:29:25 7/16/2002 to 15:06:18 7/16/2002

Service Class	APPLID	Phase	Response Time Tasks	Average	Std Dev	90% Peak	Maximum
CICSCONV	A6PTZAT1	BTE	42	94.3597	335.307	524.223	1887.45
CICSDEF	A6PTZAT1	BTE	9	.1590	.4602	.7489	1.3857
CICSHI	A6PTZAT1	BTE	23072	.1712	.7568	1.1414	15.8654

V1R2M0 CICS Performance Analyzer
Workload Manager Activity Summary by Report Class

WKLD0001 Printed at 15:33:50 7/16/2002 Data from 14:29:25 7/16/2002 to 15:06:18 7/16/2002

Report Class	APPLID	Phase	Response Time Tasks	Average	Std Dev	90% Peak	Maximum
RCICCEMT	A6PTZAT1	BTE	7	295.378	367.502	766.517	1002.92
RCICS	A6PTZAT1	BTE	44	43.1111	284.513	407.857	1887.45
RCICZAT	A6PTZAT1	BTE	23072	.1712	.7568	1.1414	15.8654

Following is an extract of the WLM workload activity for the CICS transaction Gateway. This Workload total activity includes the start, the run, and the stop of each CTG and CICS:

REPORT BY: POLICY=POL1 REPORT CLASS=RCTG
DESCRIPTION =report CICS transaction Gateway

---SERVICE----		–SERVICE RATES--		PAGE-IN RATES		---STORAGE----	
IOC	9116	ABSRPTN	259685	SINGLE	0.0	AVG	110683
CPU	1412K	TRX SERV	259685	BLOCK	0.0	TOTAL	110684
MSO	356880K	TCB	171.8	SHARED	0.0	CENTRAL	110684
SRB	63346	SRB	7.7	HSP	0.0	EXPAND	0.11
TOT	358365K	RCT	0.0	HSP MISS	0.0		
/SEC	259685	IIT	0.0	EXP SNGL	0.0	SHARED	622.01
		HST	0.0	EXP BLK	0.0		
		APPL %	13.0	EXP SHR	0.0		

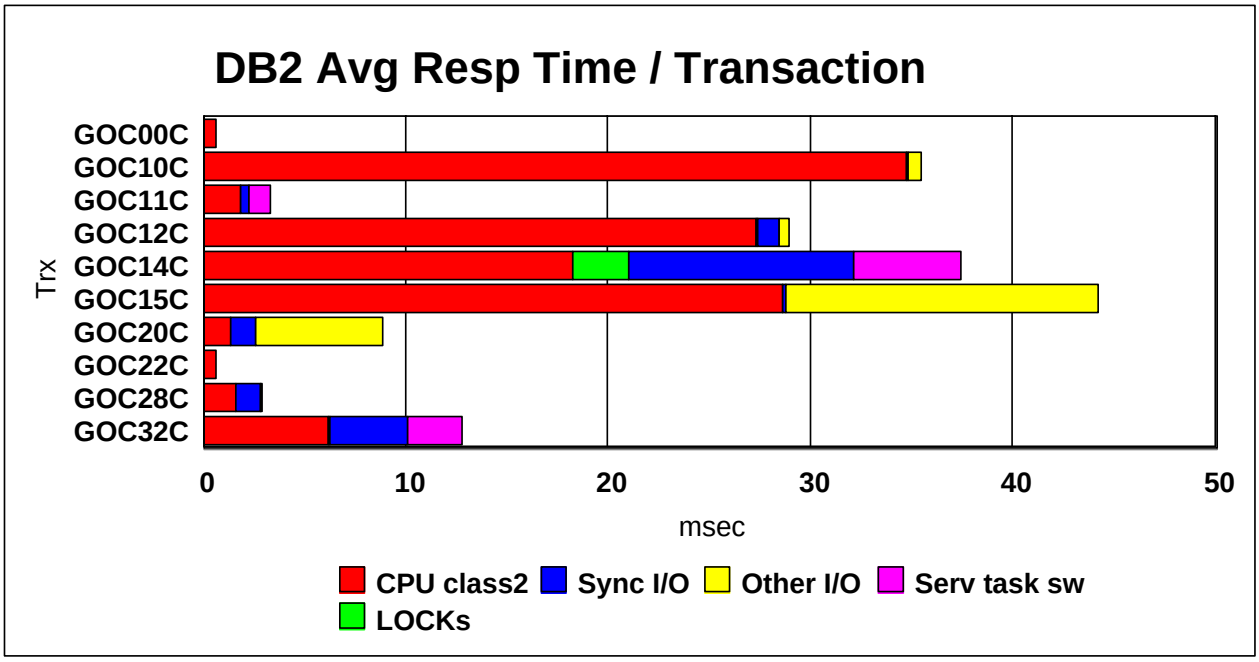
REPORT BY: POLICY=POL1 REPORT CLASS=RCICSZAT
DESCRIPTION =report CICS

---SERVICE----		-SERVICE RATES--		PAGE-IN RATES		---STORAGE----	
IOC	11776	ABSRPTN	1475892	SINGLE	0.0	AVG	99078.2
CPU	8870K	TRX SERV	1475892	BLOCK	0.0	TOTAL	99078.9
MSO	2028M	TCB	1079.4	SHARED	0.0	CENTRAL	99031.3
SRB	75365	SRB	9.2	HSP	0.0	EXPAND	47.56
TOT	2037M	RCT	0.0	HSP MISS	0.0		
/SEC	1476K	IIT	0.0	EXP SNGL	0.0	SHARED	33.00
		HST	0.0	EXP BLK	0.0		
		APPL %	78.9	EXP SHR	0.0		

3.1.2.4 DB2 metrics

The following graphic shows the response time distribution in DB2, dispatched as CPU, Lock Suspension, Synchronous I/O, Other Read I/O (Sequential I/O Waits), and Service Task Switch (commits, extents, open DS, …).
The response times are all below 50ms and the main part of the activity is CPU, except for 2 transactions:

- GOC14C which does updates (few lock contentions, sync I/O and service task switch)
- GOC15C which has a too high CPU consumption: indeed, the CPU consumption per DML (which is normally far below 1 ms) is 14 ms for GOC15C, so it means that the access path could be better tuned, and then could reduce the Other read I/O wait time as well.



3.1.3 Scenario conclusion

This scenario (200 concurrent active users, running the 10 representative transactions, with an average transaction rate of **19 transactions/sec**) allow us to:

- Demonstrate the performance level of the benchmark configuration (hardware and software) and the performance level of the Java transactions tested:
  - End users average response time being lower than **0.2 sec** (the objective was to reach an end users average response time lower than 1 sec).
    90% of the transactions had an end users average response time lower than **0.26 sec.**
- Evaluate the required resources to execute the defined scenario:
  - Average CPU load of approx. 20%
    (approx. **165 MIPS** = **9 MIPS per transaction/sec**, based on the 6 CPs LPAR processing capacity of 822 MIPS)
  - Average DASD I/O rate of approx. **120 I/O per sec**
  - Average DASD I/O response time of approx. **2 msec**
- Demonstrate the stability of the performance level, and the stability of the resources consumed.

3.2 Scenario 2: stress test

3.2.1 Scenario description

This scenario consists in simulating the activity of **200 concurrent active users**, running the **10 representative transactions**, with the **maximum transaction rate** allowed by the benchmark configuration, during 22 minutes (6 to 7 minutes of ramp-up, and 15 to 16 minutes of stabilized activity).
A detailed description of the Load Runner programs used can be found in Appendix.

Based on previous tests, the optimum in terms of number of JVM was found to be **20** for the requested activity, each JVM memory being **56 MB**, and CICS MaxTasks being **100**: reasonable JVM wait, and reasonable memory usage were reached.

3.2.2 Scenario results

3.2.2.1 Load Runner metrics

Transactions:

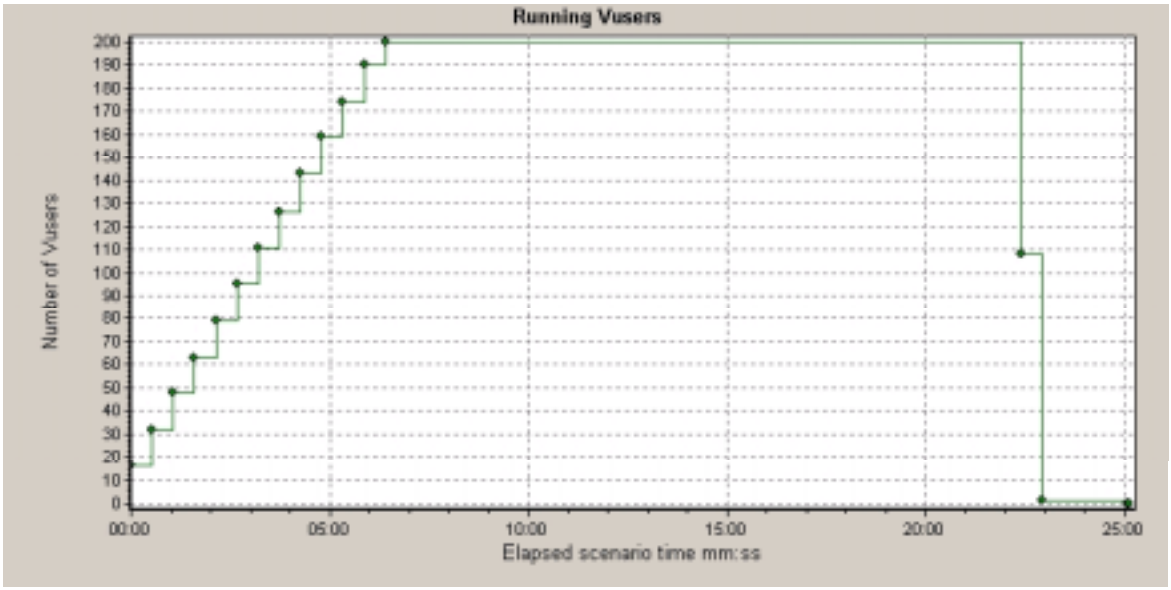
Total passed:	Total failed:	Total Stopped:
77,439	1,939	0

End users average response time:

Transaction Name	Mini	Average	Maxi	90 Pct	Pass	Fail	Stop
G22C	0.031	0.34	27.813	0.431	9,040	74	0
G28C	0.016	0.355	29.844	0.451	9,054	60	0
G32C	0.031	2.893	128.641	5.443	4,717	17	0
GOC00C	0.031	0.344	32.125	0.451	9,059	64	0
GOC10C	0.125	0.631	44.531	0.914	7,975	58	0
GOC11C	0.016	2.804	128.25	5.432	3,945	853	0
GOC12C	0.047	0.43	34.141	0.622	8,659	69	0
GOC14C	0.094	0.576	29.969	0.904	7,703	597	0
GOC15C	0.031	0.609	34.141	1.167	8,276	64	0
GOC20C	0.031	0.363	31.828	0.511	9,011	83	0

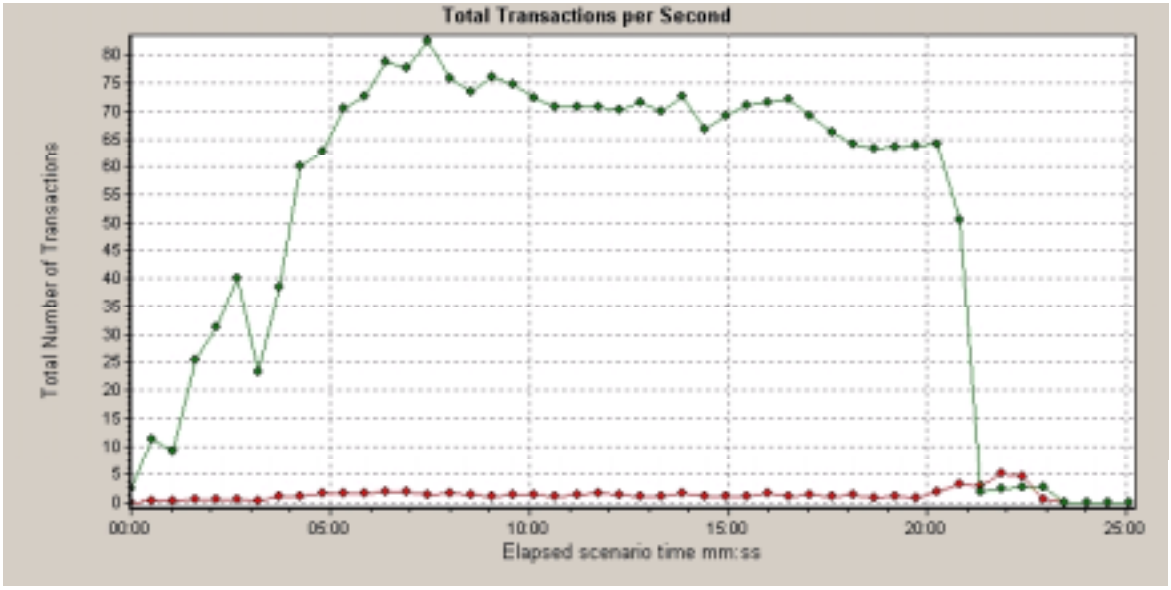
Note: failure is caused by business logic validation, locks, etc. (not caused by wrong system setup or bug in software).

Number of concurrent active users:



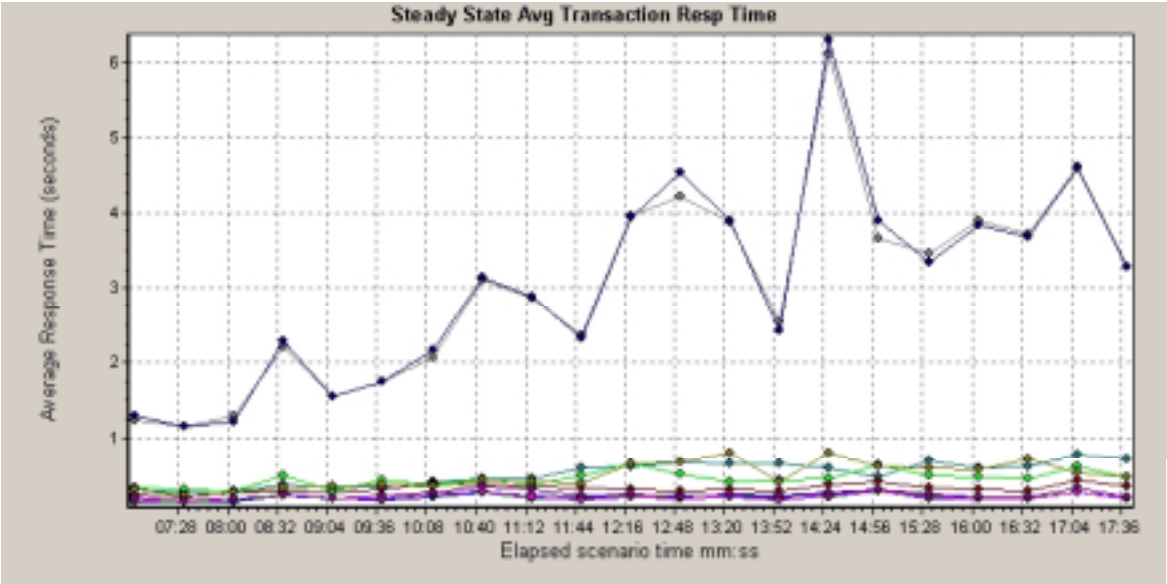
Color	Scale	Measurement	Graph's Minimum	Graph's Average	Graph's Maximum	Graph's Median	Graph's Std. Deviation
	1	Run	0	90.882	200	95	66.143

Total number of transactions per second:



Color	Scale	Measurement	Graph's Minimum	Average	Graph's Maximum	Graph's Median	Graph's Std. Deviation
	1	Fail	0	1.287	5.188	1.156	0.998
	1	Pass	0	51.386	82.5	66.281	28.843

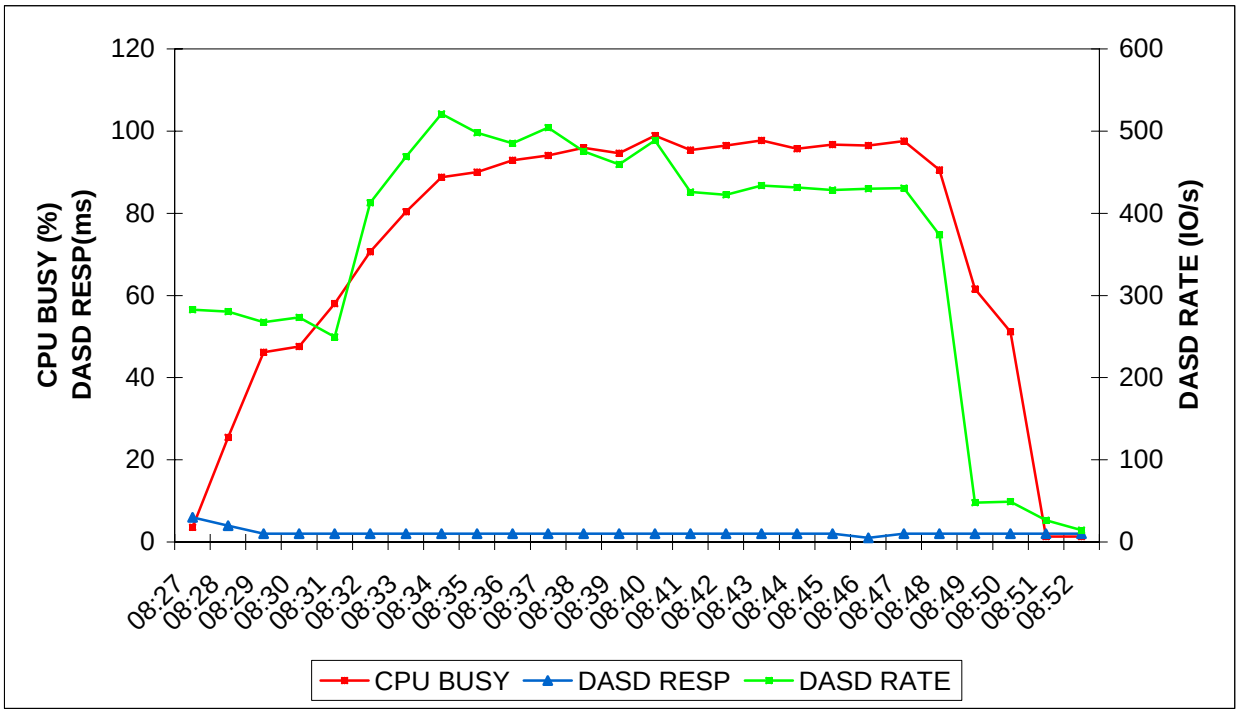
End users average transaction response time (steady state):



Color	Scale	Measurement	Minimum	Average	Maximum	Std. Deviation
	1	G22C	0.031	0.215	4.813	0.231
	1	G28C	0.016	0.222	5.547	0.269
	1	G32C	0.063	2.689	14.344	2.081
	1	GOC00C	0.031	0.217	5.172	0.232
	1	GOC10C	0.125	0.509	8.494	0.395
	1	GOC11C	0.031	2.704	14.203	2.104
	1	GOC12C	0.078	0.325	8.063	0.336
	1	GOC14C	0.094	0.448	10.625	0.583
	1	GOC15C	0.031	0.483	15.188	1.047
	1	GOC20C	0.031	0.249	8.594	0.371

3.2.2.2 zSeries metrics

zSeries CPU load and I/O activity:



zSeries memory usage:

- Main indicators for real memory:

```
   HIGH UIC  MIGR AGE
MIN   254    1008K
MAX   254    1010K
AVG   254.0  1008845
```

```
                                                        FRAME AND SLOT COUNTS
-------------------------------------------------------------------------------
                    CENTRAL STORAGE                    EXPANDED STORAGE
                ---------------------------    ---------------------------
(182 SAMPLES)         MIN        MAX        AVG         MIN        MAX        AVG
AVAILABLE             155    222,699     24,942     339,798    471,932    387,025
SQA                11,090     11,136     11,112           0          0          0
LPA                 5,226      5,226      5,226           0          0          0
CSA                 5,665      6,765      6,338           0          9          0
LSQA               11,095     12,053     11,700         935      1,650      1,447
REGIONS+SWA       262,371    483,760    458,828      52,333    183,836    137,292
TOTAL FRAMES      528,381    528,381    528,381     524,288    524,288    524,288
                      FIXED FRAMES
                ---------------------------
NUCLEUS            10,234     10,234     10,234
SQA                10,421     10,467     10,443
LPA                    65         65         65
CSA                 1,210      1,737      1,451
LSQA                8,675     10,047      9,777
REGIONS+SWA         2,569      4,121      3,319
BELOW 16 MEG          192        243        220
BETWEEN 16M-2G        N/A        N/A        N/A
TOTAL FRAMES       33,272     36,233     35,289
```

- Configuration for Virtual Memory:

```
   NUMBER OF SAMPLES     156
      STATIC STORAGE MAP
  AREA         ADDRESS   SIZE
EPVT          17F00000  1665M
ECSA           BB87000   196M
EMLPA          BB86000     4K
EFLPA                0     0K
EPLPA          839A000  55.9M
ESQA           37B7000  75.9M
ENUC           1000000  39.7M
----- 16 MEG BOUNDARY ------
NUCLEUS         FC1000   252K
SQA             EB4000  1076K
PLPA            CEE000  1816K
FLPA                 0     0K
MLPA                 0     0K
CSA             B00000  1976K
PRIVATE           1000  11.0M
PSA                  0     4K

MAXIMUM POSSIBLE USER REGION - 11056K BELOW AND 1657M ABOVE
```

As shown in these reports, due to the system configuration, memory was not an issue during the benchmark. No paging occurred.

zSeries Network load:

- Network load (Gigabit Ethernet OSA card):

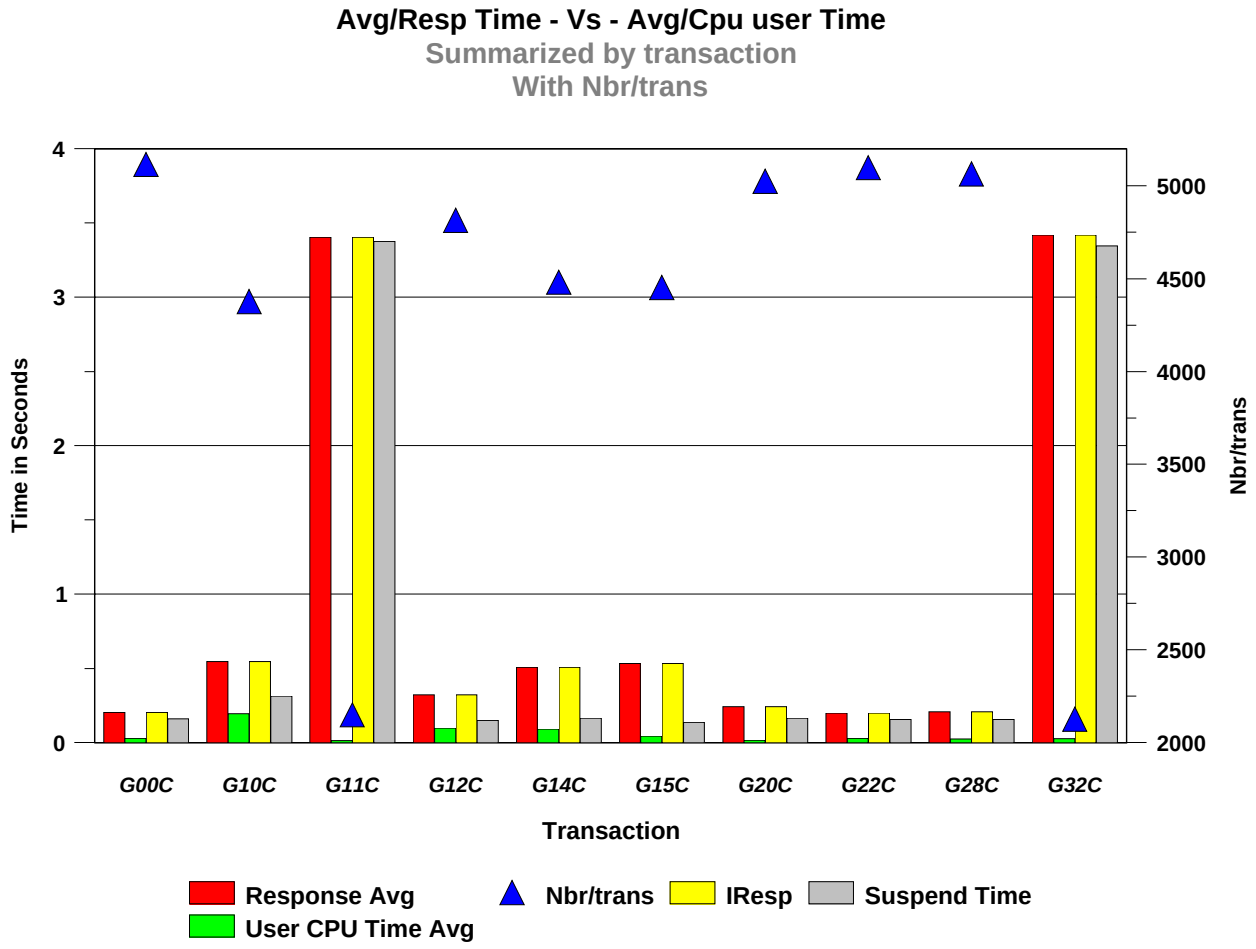
```
%Util: 8.04          (1)
%PCI Bus busy: 16.40 (2)
READ: 1.28 MB/s      (3)
WRITE: 1.29 MB/s     (4)

(1): Channel path utilization percentage
(2): Percentage of PCI bus cycles, the bus has been found busy for this channel in
relation to the theoretical limit
(3): Data transfer rates from the control unit to the channel
(4): Data transfer rates from the channel to the control unit
```

- Average transaction rate: 71.26 online transactions/sec

3.2.2.3 CICS and CTG metrics

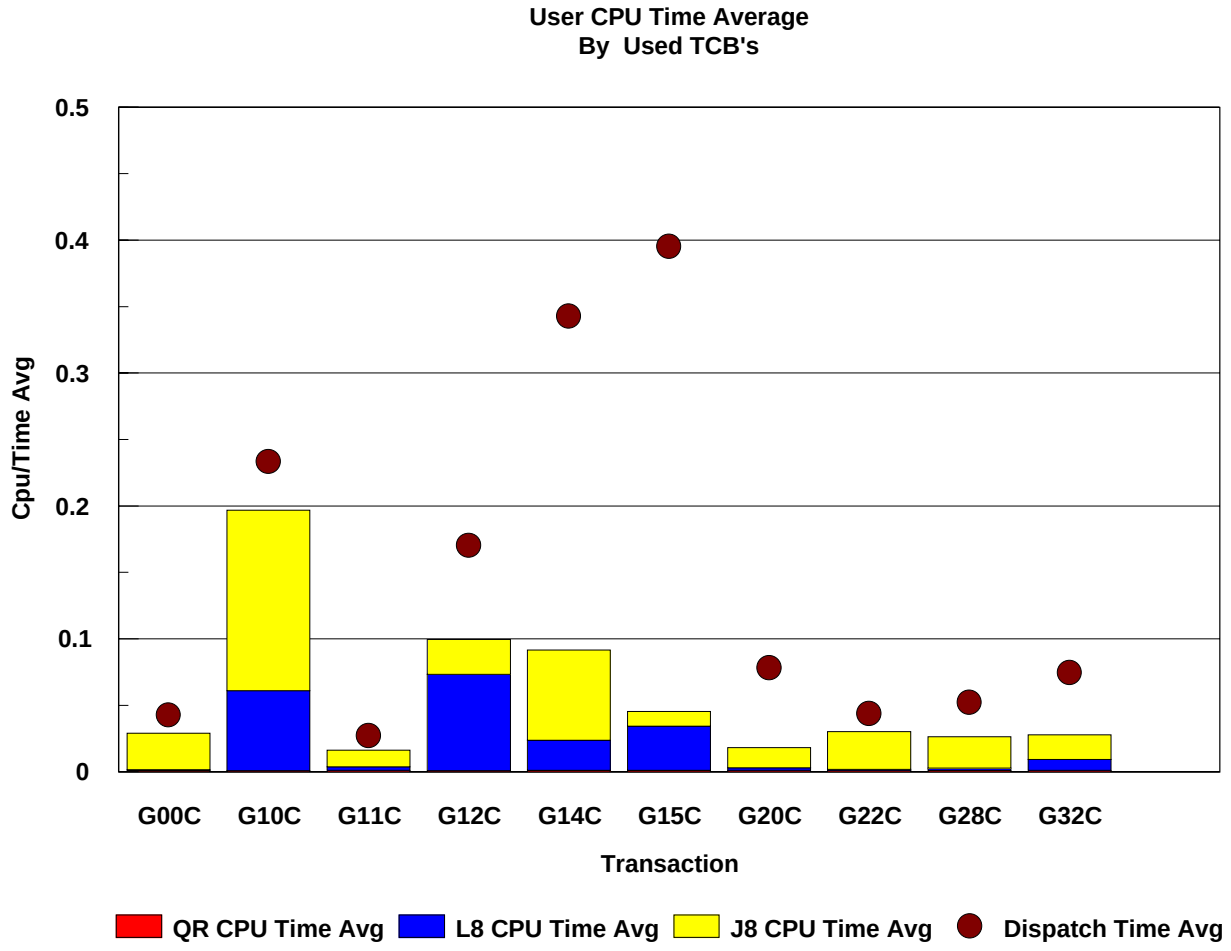
All CICS metrics come from the analyze of the SMF record using CICS Performance Analyzer V1.2. We have also used the online STAT provided by CICS and the End-of-Day statistics.

The following graphic represents the CICS average response time versus CICS average CPU time used, including the average Suspend time and Internal CICS response time, summarized by transaction name:



Remarks: the number of G11C and G32C transactions executed is lower than the others, because they are serialized in a CICS Transaction class with a MaxActive=1.

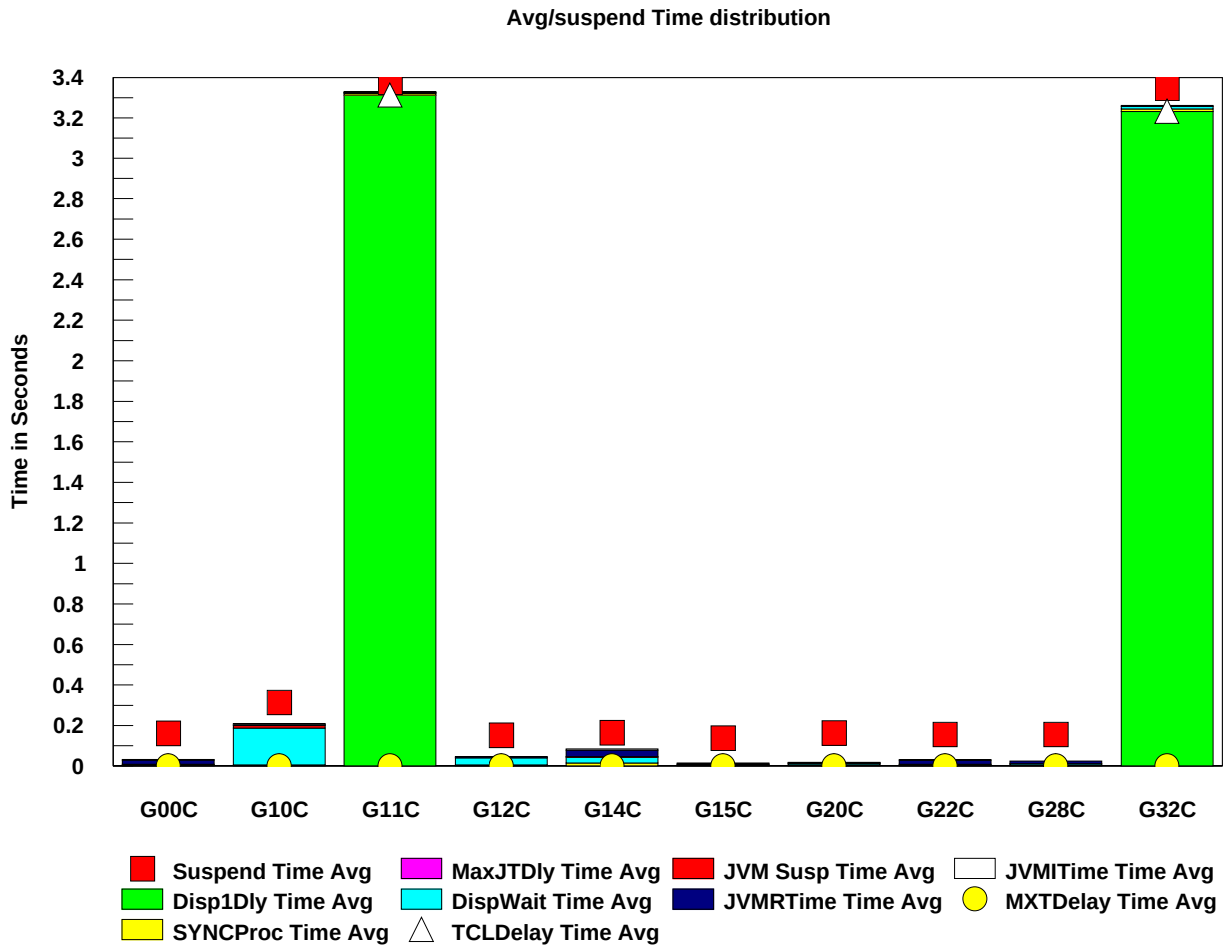
The following graphic represents the repartition in Average/Cpu time and Average/Dispatch Time for each used CICS TCB's during this benchmark (MS TCB & Key 8 TCB), summarized by transaction:



Remarks: for the Java application in CICS, the new OPEN TCB's J8 are used, and for this reason the principal CICS TCB QR remains free for classical CICS transactions (like COBOL, ASM, PL/1).

QR	=>	CICS TCB mode QR
L8	=>	CICS TCB mode L8 (Open TCB for DB2)
J8	=>	CICS TCB mode J8 (Open TCB for java , one TCB per JVM)
Avg/Dispatch Time	=>	Average elapsed time during which the user task was dispatched on each CICS TCB under which the task was executed.

The following graphic represents the repartition of the Average Suspend time summarized by transaction:



Remarks: Another type of wait was detected (same than the one in performance run).The First Dispatch Wait time for the G11C & G32C transaction is now very high for the CICS transaction class serialization with a MaxActive=1.

Following is an extract of the WLM workload activity for the CICS Transaction Gateway. This Workload total activity includes the start, the run, and the stop of each CTG and CICS:

REPORT BY: POLICY=POL1 REPORT CLASS=RCTG
DESCRIPTION =report CICS transaction Gateway

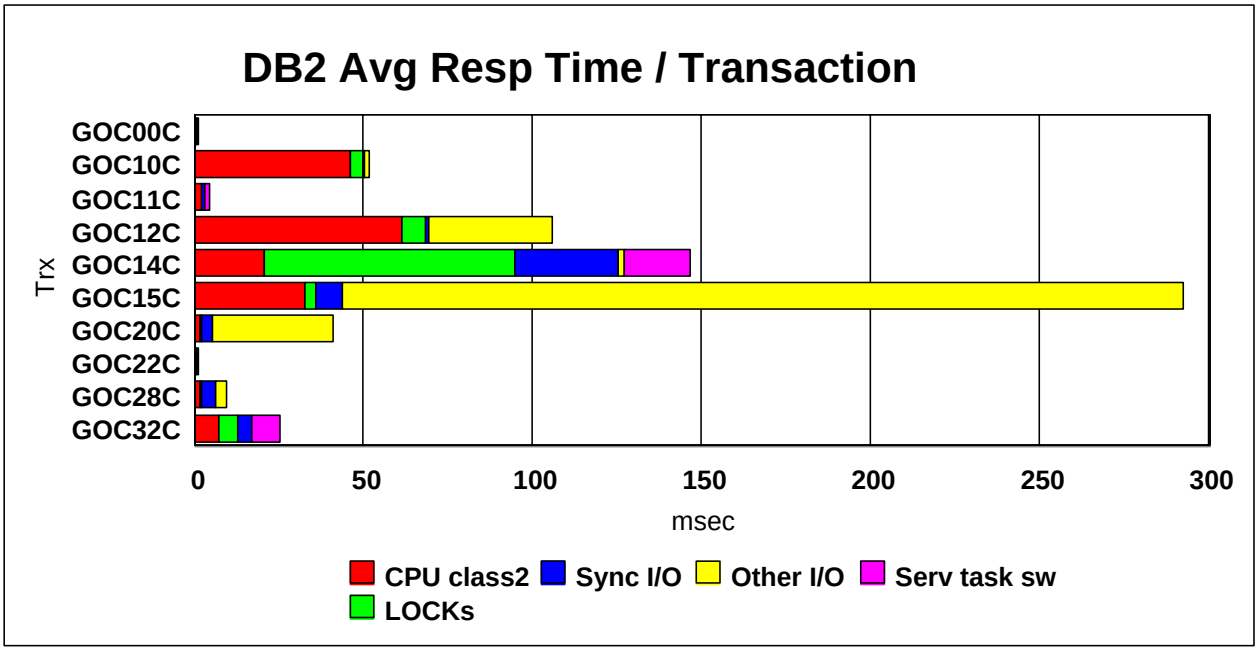
---SERVICE----		-SERVICE RATES--		PAGE-IN RATES		----STORAGE----	
IOC	31967	ABSRPTN	1227241	SINGLE	0.0	AVG	115070
CPU	7340K	TRX SERV	1227241	BLOCK	0.0	TOTAL	115071
MSO	1907M	TCB	893.2	SHARED	0.0	CENTRAL	112169
SRB	361771	SRB	44.0	HSP	0.0	EXPAND	2902.13
TOT	1914M	RCT	0.0	HSP MISS	0.0		
/SEC	1227K	IIT	0.0	EXP SNGL	4.5	SHARED	623.84
		HST	0.0	EXP BLK	0.0		
		APPL %	60.1	EXP SHR	0.0		

REPORT BY: POLICY=POL1 REPORT CLASS=RCICSZAT
DESCRIPTION =report CICS

---SERVICE----		--SERVICE RATES--		PAGE-IN RATES		---STORAGE----	
IOC	44745	ABSRPTN	12689K	SINGLE	0.0	AVG	212319
CPU	40123K	TRX SERV	12689K	BLOCK	0.0	TOTAL	212320
MSO	19754M	TCB	4882.5	SHARED	0.0	CENTRAL	211078
SRB	297115	SRB	36.2	HSP	0.0	EXPAND	1242.63
TOT	19795M	RCT	0.0	HSP MISS	0.0		
/SEC	12689K	IIT	0.0	EXP SNGL	0.0	SHARED	0.00
		HST	0.0	EXP BLK	0.3		
		APPL %	315.3	EXP SHR	0.0		

3.2.2.4 DB2 metrics

The following graphic shows that the response times increased from less than 50ms to 100-250ms in DB2, due to locking contentions and Other Read I/O Wait times.
As this run was a stress test with a high transaction rate, the locking is a normal situation, but the Other Read I/O indicate that the prefetch is not efficient enough, due to physical tablespace organization or bad access method. So, it may be a good idea to define explicit CLUSTER index for the tables, in order to have the data sorted in the expected sequence, and then reduce the Other Read I/O's.



3.2.3 Scenario conclusion

This scenario (200 concurrent active users, running the 10 representative transactions, with an average transaction rate of more than **71 transactions/sec**) allow us to:

- Demonstrate again the performance level of the benchmark configuration (hardware and software) and the performance level of the Java transactions tested:
  - End users average response time being lower than **0.5 sec** (the objective was to reach an end users average response time lower than 1 sec for a transaction rate of 18 transactions/sec).
  - Except for transactions GOC11C and G32C, that reached an end-users average response time of approx. 2.7 sec (due to application business logic, the 2 Transactions were serialized).
  - 90% of the transactions had an end users average response time lower than **0.9 sec** (except for transactions GOC11C and G32C).
- Evaluate the required resources to execute the defined scenario:
  - Average CPU load of approx. 95%
    (approx. **780 MIPS** = **11 MIPS per transaction/sec**, based on the 6 CPs LPAR processing capacity of 822 MIPS)
  - Average DASD I/O rate of approx. **450 I/O per sec**
  - Average DASD I/O response time of approx. **2 msec**
- Demonstrate the **scalability** of the overall solution:
  - For a transaction rate multiplied by 4, the CPU load has been multiplied by 4.7, and the end users average response time has been multiplied by 2.5.

3.3 Scenario 3: long run

3.3.1 Scenario description

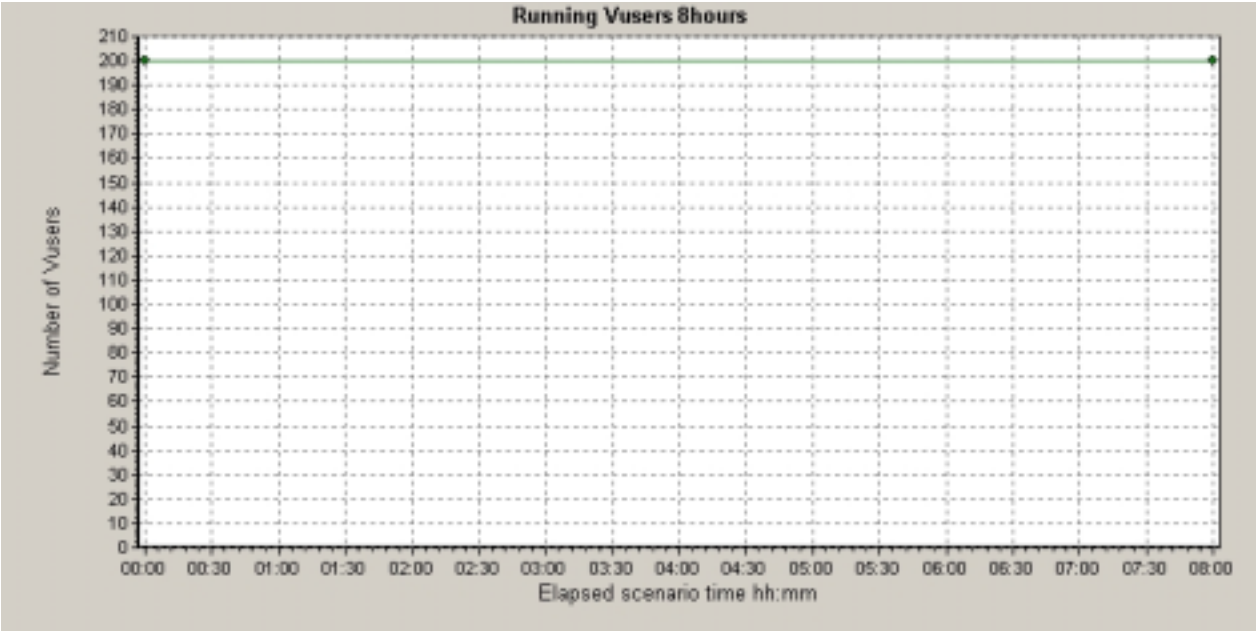
This scenario consists in simulating the activity of **200 concurrent active users**, running the **10 representative transactions**, with a transaction rate of **19 transactions/sec**, during 8 hours (6 to 7 minutes of ramp-up, and 473 to 474 minutes of stabilized activity).
A detailed description of the Load Runner programs used can be found in Appendix.

Based on previous tests, the optimum in terms of number of JVM was found to be **10** for the requested activity, each JVM memory being **56 MB**, and CICS MaxTasks being **60**: reasonable JVM wait, and reasonable memory usage were reached.

3.3.2 Scenario results

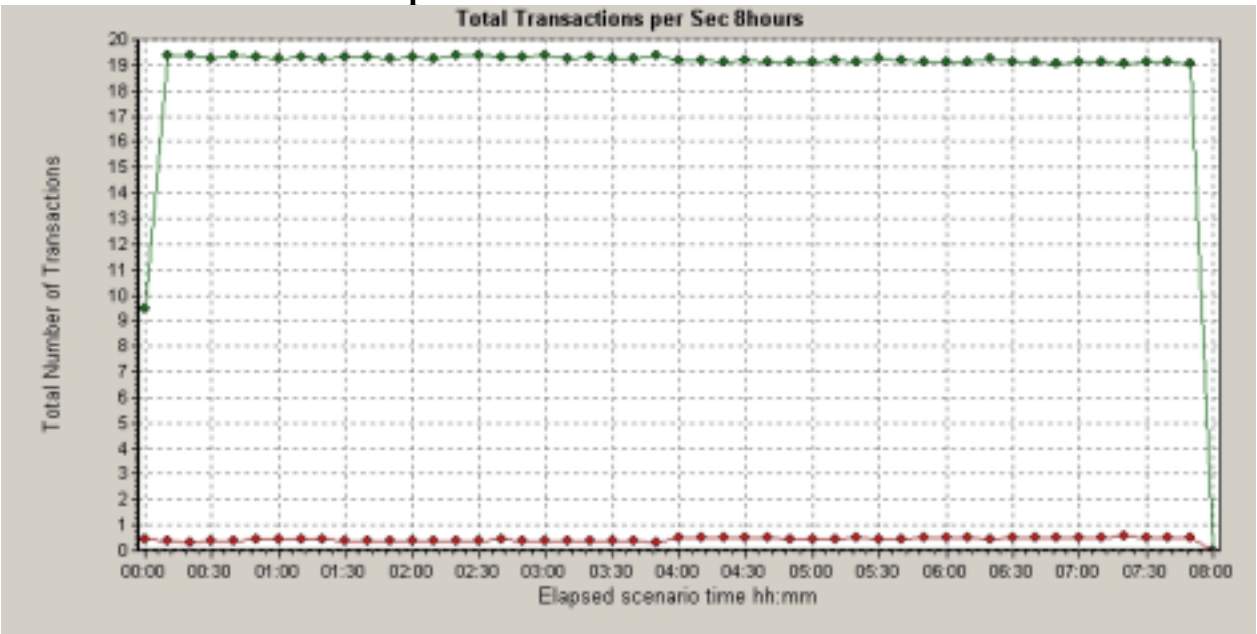
3.3.2.1 Load Runner metrics

Number of concurrent active users:



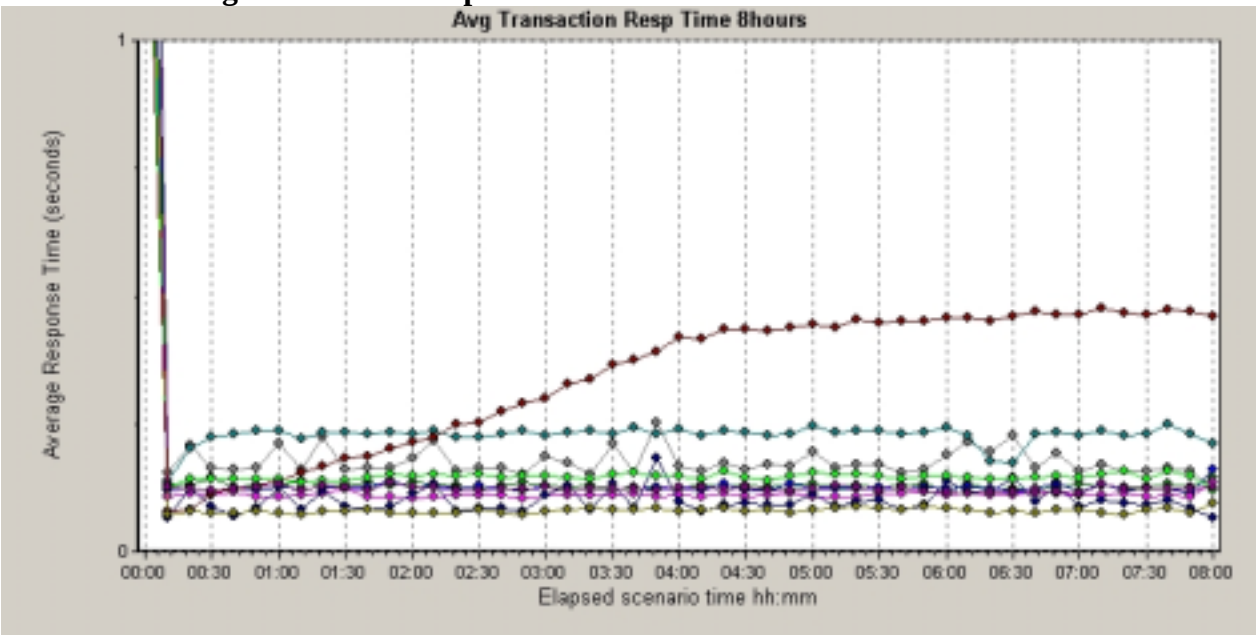
Color	Scale	Measurement	Graph's Minimum	Graph's Average	Graph's Maximum	Graph's Median	Graph's Std. Deviation
	1	Run	200	200	200	200	0

Total number of transactions per second:



Color	Scale	Measurement	Graph's Minimum	Average	Graph's Maximum	Graph's Median	Graph's Std. Deviation
	1	Fail	0	0.436	0.55	0.433	0.088
	1	Pass	0.03	19.014	19.397	19.222	3.02

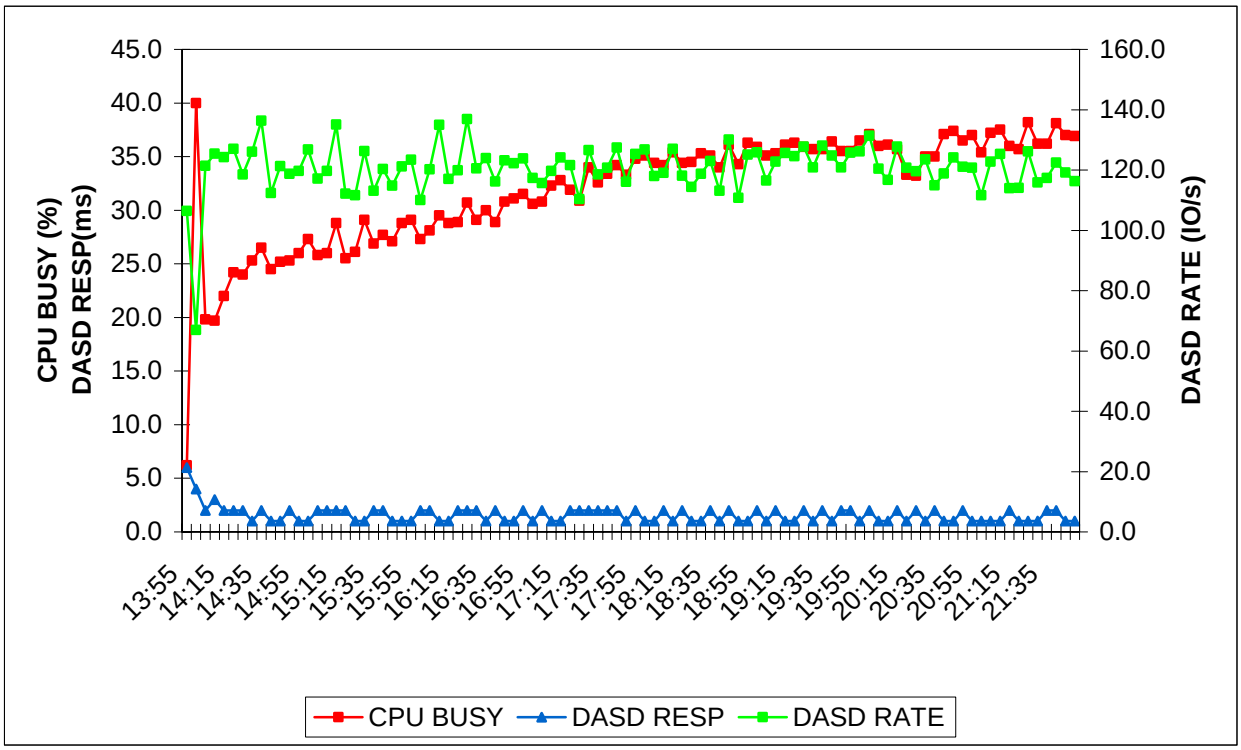
End users average transaction response time:



Color	Scale	Measurement	Minimum	Average	Maximum	Std. Deviation
	1	G22C	0.031	0.145	135.859	1.25
	1	G28C	0.016	0.143	135.844	1.39
	1	G32C	0.016	0.208	149.078	2.139
	1	GOC00C	0.016	0.127	135.859	1.311
	1	GOC10C	0.109	0.244	134.172	1.242
	1	GOC11C	0.016	0.134	149.203	2.129
	1	GOC12C	0.047	0.351	136.922	1.232
	1	GOC14C	0.094	0.163	137.625	1.282
	1	GOC15C	0.031	0.102	134.859	1.474
	1	GOC20C	0.016	0.145	137.766	1.595

3.3.2.2 zSeries metrics

zSeries CPU load and I/O activity:



zSeries memory usage:

- Main indicators for real memory:

```
   HIGH UIC  MIGR AGE
MIN   254    1795K
MAX   254    1838K
AVG   254.0  1816424
```

```
                                                        FRAME AND SLOT COUNTS
-------------------------------------------------------------------------------
                     CENTRAL STORAGE                    EXPANDED STORAGE
                   ---------------------------    ---------------------------
(2976 SAMPLES)          MIN        MAX        AVG        MIN        MAX        AVG
AVAILABLE                48    258,079     15,188    183,576    432,543    325,614
SQA                   6,855     11,190     11,169          0          0          0
LPA                   5,305      5,956      5,305          0          0          0
CSA                   7,435     11,907      8,556          0        155          0
LSQA                 10,005     11,535     11,216      1,284      1,779      1,550
REGIONS+SWA         226,179    482,261    466,707     91,171    340,372    198,119
TOTAL FRAMES        528,381    528,381    528,379    524,288    524,288    524,288
                        FIXED FRAMES
                   ---------------------------
NUCLEUS              10,234     10,234     10,234
SQA                   6,084     10,516     10,494
LPA                      65        130         65
CSA                   2,556      3,326      2,946
LSQA                  8,232     10,050      9,534
REGIONS+SWA           2,281      7,816      3,086
BELOW 16 MEG             91        190        146
BETWEEN 16M-2G          N/A        N/A        N/A
TOTAL FRAMES         34,426     37,728     36,361
```

- Configuration for Virtual Memory:

```
  NUMBER OF SAMPLES   2,880

      STATIC STORAGE MAP
  AREA         ADDRESS   SIZE
EPVT          17F00000  1665M
ECSA           BB87000   196M
EMLPA          BB86000     4K
EFLPA                0     0K
EPLPA          839A000  55.9M
ESQA           37B7000  75.9M
ENUC           1000000  39.7M
----- 16 MEG BOUNDARY ------
NUCLEUS         FC1000   252K
SQA             EB4000  1076K
PLPA            CEE000  1816K
FLPA                 0     0K
MLPA                 0     0K
CSA             B00000  1976K
PRIVATE           1000  11.0M
PSA                  0     4K

MAXIMUM POSSIBLE USER REGION - 11056K BELOW AND 1657M ABOVE
```

As shown in these reports, due to the system configuration, memory was not an issue during the benchmark. No paging occurred.
Nevertheless, in customer premises, the use of expanded storage in addition to the current 2GB of central storage is recommended until a migration to zOS is made.

zSeries Network load:

- Network load (Gigabit Ethernet OSA card):

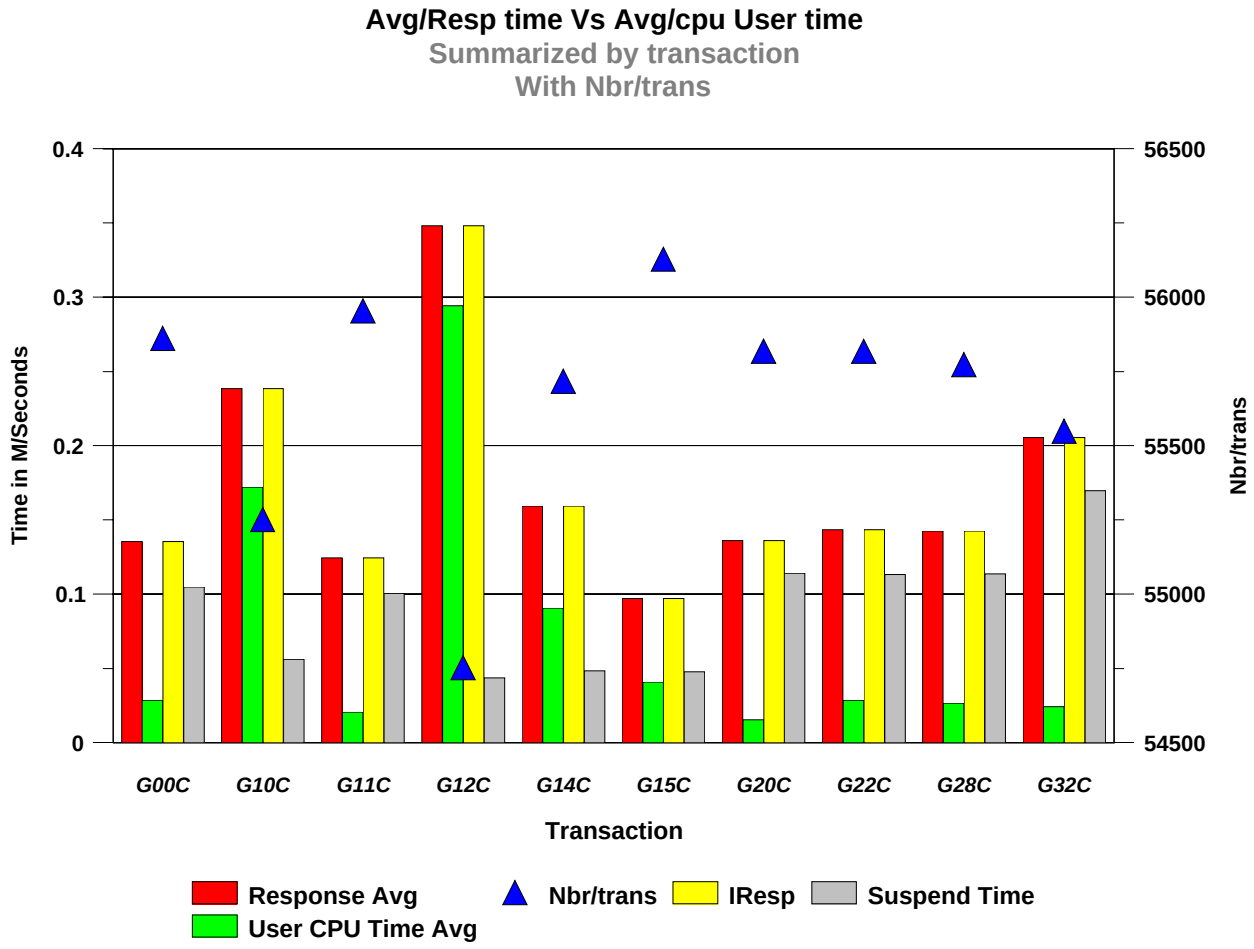
```
%Util: 3.99          (1)
%PCI Bus busy: 12.72 (2)
READ: 0.49 MB/s      (3)
WRITE: 0.50 MB/s     (4)

(1): Channel path utilization percentage
(2): Percentage of PCI bus cycles, the bus has been found busy for this channel in
relation to the theoretical limit
(3): Data transfer rates from the control unit to the channel
(4): Data transfer rates from the channel to the control unit
```

- Average transaction rate: 19.39 online transactions/sec

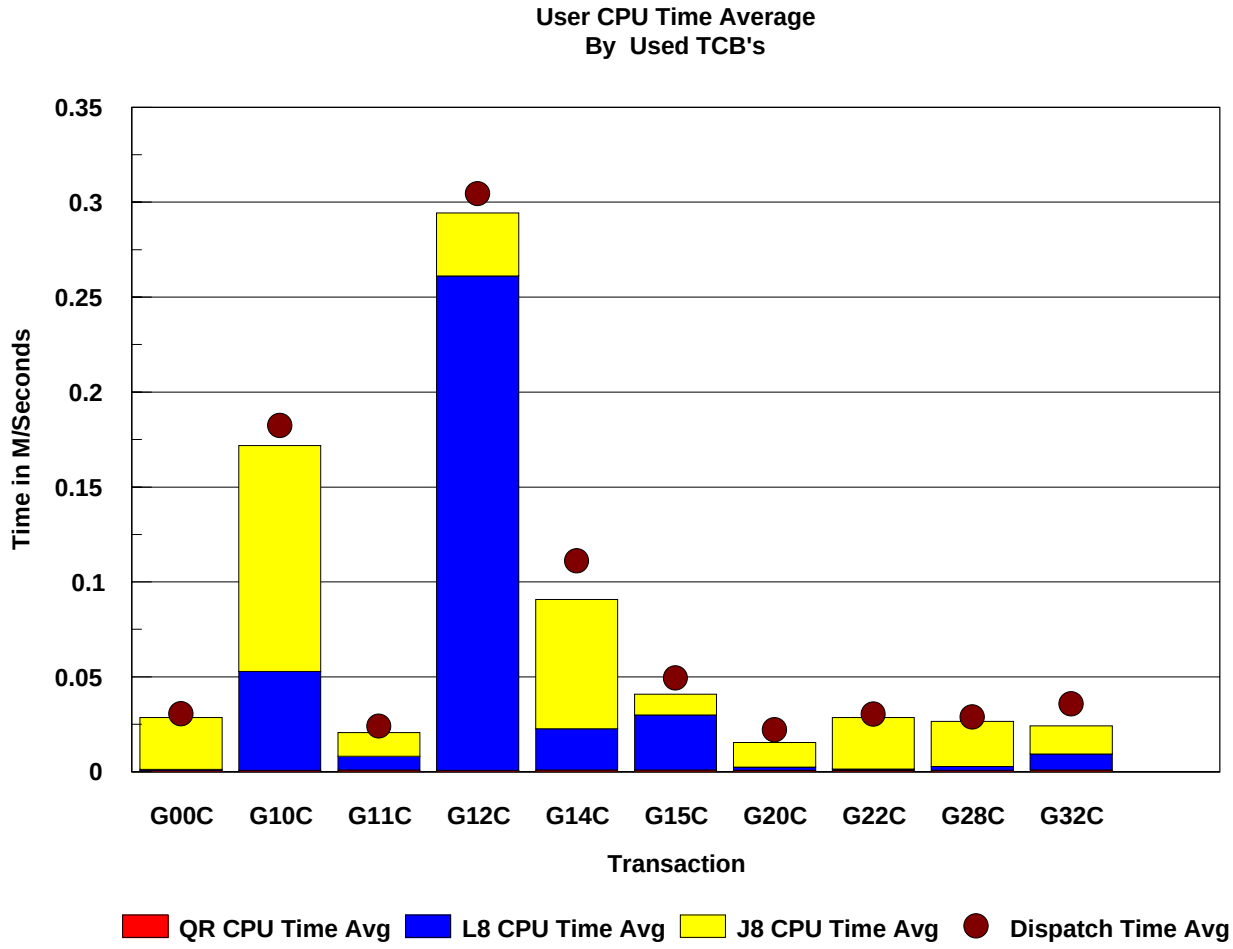
3.3.2.3 CICS and CTG metrics

All CICS metrics come from the analyze of the SMF record using CICS Performance Analyzer V1.2. We have also used the online STAT provided by CICS and the End-of-Day statistics.

The following graphic represents the CICS average response time versus CICS average CPU time used, including the average Suspend time and Internal CICS response time, summarized by transaction name:



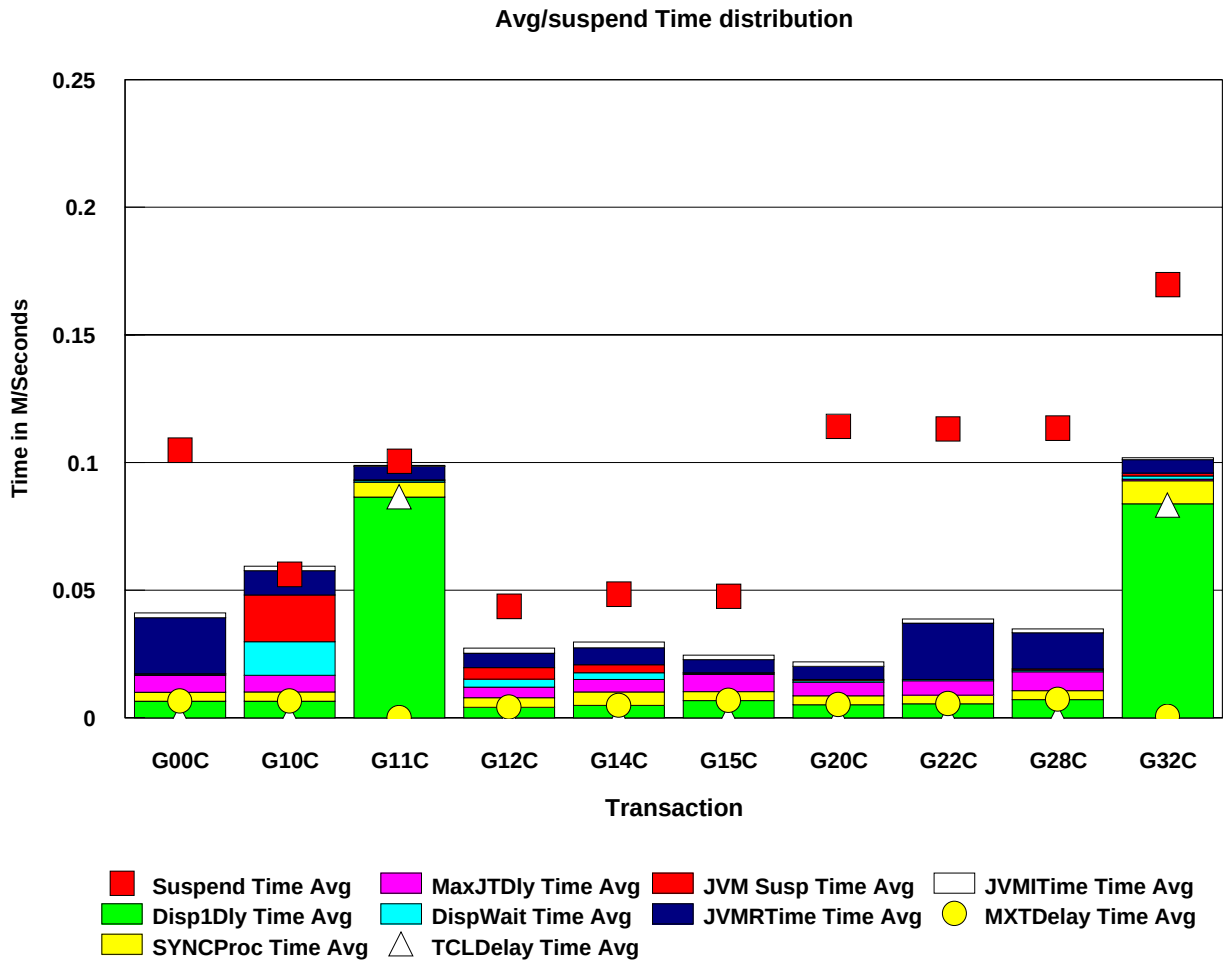
The following graphic represents the repartition in Average/Cpu time and Average/Dispatch Time for each used CICS TCB's during this benchmark (MS TCB & Key 8 TCB), summarized by transaction:



Remarks: for the Java application in CICS, the new OPEN TCB's J8 are used, and for this reason the principal CICS TCB QR remains free for classical CICS transactions (like COBOL, ASM, PL/1).

QR	=>	CICS TCB mode QR
L8	=>	CICS TCB mode L8 (Open TCB for DB2)
J8	=>	CICS TCB mode J8 (Open TCB for java , one TCB per JVM)
Avg/Dispatch Time	=>	Average elapsed time during which the user task was dispatched on each CICS TCB under which the task was executed.

The following graphic represents the repartition of the Average Suspend time summarized by transaction:



Remarks: Another type of wait was detected (same than the one in performance run).The First Dispatch Wait time for the G11C & G32C transaction is now very high for the CICS transaction class serialization with a MaxActive=1.

Following is an extract of the WLM workload activity for the CICS transaction Gateway. This Workload total activity includes the start, the run, and the stop of each CTG and CICS:

REPORT BY: POLICY=POL1 REPORT CLASS=RCTG
DESCRIPTION =report CICS transaction Gateway

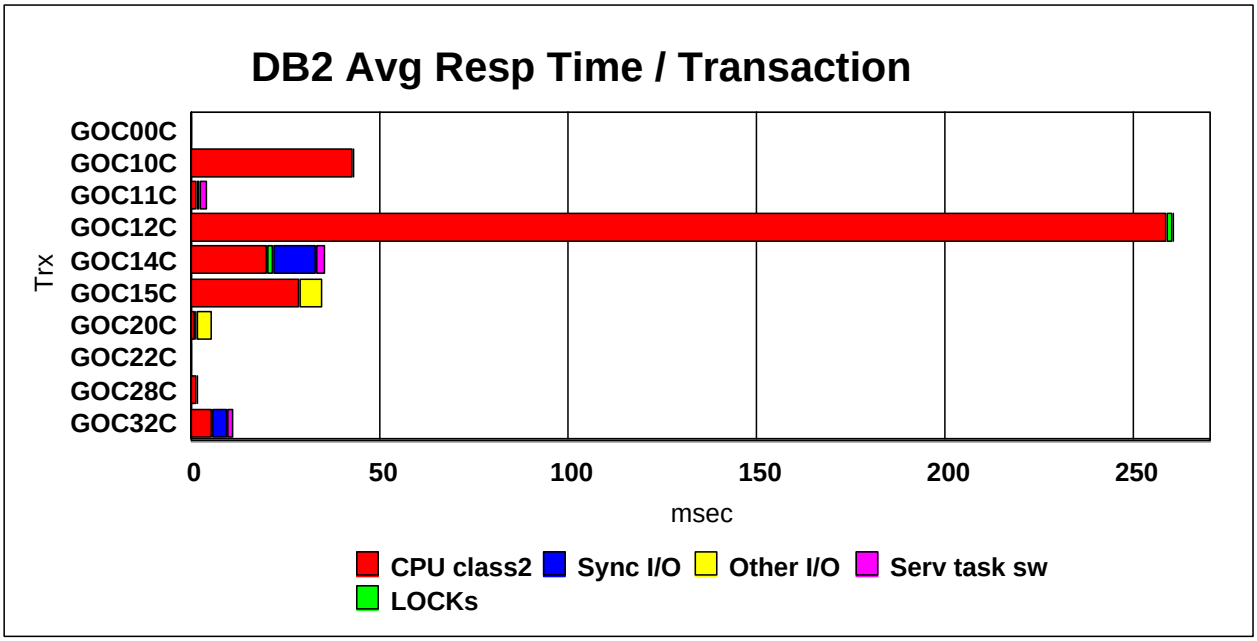
---SERVICE----		--SERVICE RATES--		PAGE-IN RATES		---STORAGE----	
IOC	223656	ABSRPTN	359765	SINGLE	0.0	AVG	145154
CPU	32277K	TRX SERV	359765	BLOCK	0.0	TOTAL	145154
MSO	10327M	TCB	3927.7	SHARED	0.0	CENTRAL	144653
SRB	1807K	SRB	219.9	HSP	0.0	EXPAND	501.75
TOT	10361M	RCT	0.0	HSP MISS	0.0		
/SEC	359765	IIT	0.0	EXP SNGL	0.0	SHARED	631.94
		HST	0.0	EXP BLK	0.0		
		APPL %	14.4	EXP SHR	0.0		

REPORT BY: POLICY=POL1 REPORT CLASS=RCICSZAT
DESCRIPTION =report CICS

---SERVICE----		--SERVICE RATES--		PAGE-IN RATES		----STORAGE----	
IOC	304347	ABSRPTN	3290955	SINGLE	0.0	AVG	18979
CPU	358447K	TRX SERV	3290955	BLOCK	0.0	TOTAL	118979
MSO	94419M	TCB	43618.5	SHARED	0.0	CENTRAL	118930
SRB	2062K	SRB	250.9	HSP	0.0	EXPAND	49.00
TOT	94779M	RCT	0.0	HSP MISS	0.0		
/SEC	3291K	IIT	0.1	EXP SNGL	0.0	SHARED	65.00
		HST	0.0	EXP BLK	0.0		
		APPL %	152.3	EXP SHR	0.0		

3.3.2.4 DB2 metrics

The following graphic shows that CPU consumption for transaction GOC12C increased a lot during this run.



Looking at the culprit SQL:

```
DECLARE DB2SQLJCURSOR9 CURSOR FOR
SELECT DBC.CUR_ID, CURR.CODE_CHAR, SUM (DBC.AMOUNT) AMOUNT
FROM CURRENCY CURR, DEPOSIT_BOX_CONT DBC, DEPOSIT_BOX D
WHERE
    D.ORG_UNI_ID = :d.org_uni_id
AND D.BANK_SIGN = :d.bank_sign
AND CURR.CUR_ID = DBC.CUR_ID
AND DBC.DEP_BOX_ID =D.DEP_BOX_ID
GROUP BY DBC.CUR_ID, CURR.CODE_CHAR
```

it turns out that the table DEPOSIT_BOX_CONT was empty and grew during the run, so that the access method (TableSpace Scan) was bad.
A RUNSTAT utility was executed when the table was filled with some rows, and it changed the TableSpace Scan to a matching index access that should solve this increasing response time problem.
The new access path is now DEPOSIT_BOX first, accessed with a TableSpace Scan as well.
So, an index on DEPOSIT_BOX columns (ORG_UNI_ID, BANK_SIGN) would make an index access with 2 matchcols.

3.3.3 Scenario conclusion

This scenario (200 concurrent active users, running the 10 representative transactions, with an average transaction rate of **19 transactions/sec**, running during **8 hours**) allow us to:

- Demonstrate again the performance level of the benchmark configuration (hardware and software) and the performance level of the Java transactions tested:
  - End users average response time being lower than **0.2 sec** (the objective was to reach an end users average response time lower than 1 sec).
  - Except for transaction GOC12C (DB2 CPU consumption increase due to table size increase during the test: Table Space Scan on DEPOSIT_BOX_CONT which was empty at start time of the run).
- Evaluate the required resources to execute the defined scenario:
  - Average CPU load of approx. 30%
    (approx. **245 MIPS = 12 MIPS per transaction/sec**, based on the 6 CPs LPAR processing capacity of 822 MIPS)
    Note: the CPU load increased slowly during the 8-hours run from 20% to 35% because of the DB2 tables increase and number of extends; this can be fixed by tuning the tables allocation parameters.
  - Average DASD I/O rate of approx. **120 I/O per sec**
  - Average DASD I/O response time of approx. **2 msec**
- Demonstrate the stability of the performance level, and the stability of the resources consumed, in a typical online daily activity of 8 hours.

3.4 Benchmark conclusion

This Java - CICS TS V2.2 benchmark demonstrate:

- The resources required to execute the defined workload:
  - 165 MIPS to execute up to 19 transactions/sec (approx. 9 MIPS per transaction/sec)
  - 780 MIPS to execute up to 70 transactions/sec (approx. 11 MIPS per transaction/sec)
- The performance level in terms of end users average response time for the defined workload:
  - 0.2 sec with a transaction rate of 19 transactions/sec
  - 0.5 sec with a transaction rate of 70 transactions/sec
  - In the benchmark unconstraint environment.
- The scalability of the configuration and the scalability of the application:
  - From 9 MIPS per transaction/sec up to 11 MIPS per transaction/sec (multiplied by 1.2)
  - When increasing the transaction rate from 19 to 70 transactions/sec (multiplied by 4)
  - With an end users average response time increasing from 0.2 sec to 0.5 sec (multiplied by 2.5).
- The stability of the performance level, and the stability of the required resources, in a typical online daily activity of 8 hours.

During the benchmark, we also demonstrated that the major tuning effort has to be done on the JVM area (number of JVM, JVM memory parameters…), and then on the DB2 area (application and databases design, tables, tables indexes, locking reduction to improve databases access concurrency…).

Else during this benchmark, the IBM partner had the opportunity to tune their transactions in terms of SELECT orders, to reach a high level of performance with reasonable resource consumption.

In these domains, more tuning can still be done at a detailed level, to win some more performance and to reduce IT resources consumption, but the major steps have already been covered during the benchmark, and we hope that this Technical Report (together with the final customer specialists visit in Montpellier) will allow the customer to benefit from this deep experience.

Based on the experience related in this Technical Report, and based on the additional skills acquired by the IBM partner,final customer have then the opportunity to start in the Java - CICS TS V2.2 domain with confidence.