" Sysplex/CICSplex/DB2 Data Sharing:
Assessing the Costs and Analyzing New Measurement Data"

Ellen Friedman, SRM Associates

Abstract

The S/390 CMOS technology coupled with Parallel Sysplex/ CICSplex and DB2 data sharing provides a scalable, flexible growth path for large CICS/DB2 applications. Workload balancing can be achieved across a multiple CPC sysplex environment with the advent of CICSplex and DB2 data sharing since there is no affinity between data requests and the processor executing the transaction.

The IBM benchmarks for DB2 have shown that in a 100% data sharing environment the throughput loss is less than 17% with an additional fixed cost of 2-3% for Sysplex management. The benchmarks which we conducted in a Sysplex/CICSplex/DB2 data sharing environment yielded less than 12% overall overhead, but more importantly we can demonstrate that only those workloads with a high degree of data sharing were impacted.

An overview of the Parallel Sysplex architecture particularly as it relates to DB2 data sharing is presented to familiarize the reader with the issues and measurements. We review the metrics reported by the new RMF and DB2 Performance Monitor reports that are used in to analyze the costs of data sharing and the performance impact to system workloads. The factors contributing to performance overhead, how to measure this overhead and how to analyze the new data metrics and reports which are part of the new sysplex data sharing environment are detailed in the paper. Benchmark results from a sysplex/CICSplex DB2 data sharing environment operating on an RX3 CMOS platform are reviewed in detail.

Introduction

The S/390 CMOS microprocessor technology offers better price/performance ratios than bipolar technology and provides good scalability characteristics. Parallel sysplex preserves a "single system image" view for the end user and provides easier management than multiple systems. Data sharing performance is quite good for workloads with high read-write inter-system interest and the overhead for additional systems is relatively flat. Parallel sysplex also provides for flexible growth, as one can add additional processors and capacity in a non-disruptive fashion. The larger capacity of the sysplex enables increased concurrency of transactions accessing the same DB2 database. With DB2 datasharing, work can be balanced across all processors in the sysplex since there is no affinity between data and the processor. DB2 data sharing is the key to to balancing large CICS/DB2 applications across multiple CPCs offering higher availability to workloads. In the event of a failure, a workload can be processed on one of the other systems in the sysplex environment and still be able to access all required data.

There are still some drawbacks with the new CMOS technology; most specifically, the processor speeds currently available. The Bipolar technology in IBM and Amdahl is in the 60 MIPS range and 120 MIPS on HDS as compared with 24 currently on CMOS. Successful migration to a CMOS sysplex environment assumes that the application is not CPU bound and can be split into multiple regions to take advantage of a multi-processor and a multiple CPC environment.

The IBM benchmarks for DB2 have shown that in a 100% data sharing environment the throughput loss is less than 17% with an additional fixed cost of 2-3% for sysplex management. The benchmarks which we conducted in a sysplex/CICSplex/DB2 data sharing environment yielded less than 12% overall overhead, but more importantly we can demonstrate that only those workloads with a high degree of data sharing were impacted. The inquiry workloads were only impacted to the degree that less overall capacity was available, but response time and transaction CPU time for inquiry workloads were not impacted. Additionally, we executed benchmarks at varying transaction rates and demonstrated the scalability of the CMOS processor complexes. Specific details of the workloads and the results are provided in the discussion of the case study.

This paper is divided into two parts. The first half provides an overview of the parallel sysplex architecture particularly as it relates to DB2 data sharing. This section contains a discussion of the new metrics and RMF and DB2 Performance Monitor reports that are used to analyze the performance costs of data sharing and its impact on system workloads. In the second section, a case study is presented to illustrate the use of these reports and to provide some additional insights into the data sharing impact for specific types of work. The case study includes a discussion of benchmark results from a sysplex/CICSplex data sharing environment operating on an RX3 CMOS platform.

Sysplex/CICSplex DB2 Data Sharing Overview

The performance costs arising from the implementation of parallel sysplex and data sharing are dependent upon the hardware configuration, workload characteristics and the amount of data sharing required by the applications. These factors must be considered and evaluated in order to determine the performance impact of this new technology on the application and system environment.

The actual data sharing cost is most dependent on the coupling facility access rates for locking, caching and local buffer pool invalidation. The access rate to the coupling facility is a function of both the degree to which the workload requires access to shared data and the frequency of that access. Some applications may not require any access to shared data while others may have a high data sharing requirement. This section reviews the factors that contribute to performance overhead, how to measure these costs and analyze the new performance metrics and reports for a sysplex, data sharing environment.

Sysplex/CICSplex

In order to understand where the overhead functions come from, we need to review the architecture of sysplex and CICSplex and some of the functional capabilities and performance concerns. In addition, we need to review the coupling facility structure and performance impacts of data sharing and lock contention.

Parallel Sysplex: A sysplex is defined as a set of MVS systems communicating and cooperating with each other through applicable hardware and software to process the work. The sysplex increases the overall processing capacity available to applications by allowing up to 32 central processing complexes, “CPCs”, to communicate with each other in the sysplex.

The primary hardware component facilitating the communication is the cross system coupling facility (XCF) and the coupling facility channels providing high bandwidth fiber optic links between CPCs. In a sysplex, CPCs can communicate using both CTC (channel-to-channel) or using the coupling facility. The coupling facility is a microprocessing unit or a CPU, that provides high speed access to shared data across applications and subsystems running on different MVS systems.

MVS/ESA Version 5 provides coupling facility support. This support includes the workload manager, sysplex failure management, coupling facility resource management and automatic restart manager and system logger that allow the installation to set policies for handling specific events in the sysplex environment. The workload manager defines performance goals for workloads. The coupling facility resource management (CFRM) couple data set defines how MVS will manage coupling facility resources. The sysplex failure management (SFM) couple data set defines how MVS will manage system and signaling connectivity failures and PR/SM reconfiguration actions. The automatic restart management (ARM) couple data set defines how to handle an outage including how to automatically restart batch jobs and started tasks. In addition, the system logger (LOGR) couple data set defines, updates and deletes structure or log stream definitions.

Data Sharing: Data sharing allows concurrent subsystems to access and change the same data while maintaining data integrity. Support exists for shared IMS and DB2 databases.

Prior to DB2 V4, only distributed data supported read and write access between multiple DB2 subsystems. With DB2 V4, data sharing allows read-write access to DB2 data from different DB2s residing in multiple CPCs in a parallel sysplex. DB2 data sharing in a parallel sysplex provides much better performance than distributed data.

The coupling facility is the vehicle allowing data base managers and MVS components to share this data using the buffer pools and lock structures stored in it. But it is the high performance of the coupling facility that makes the sysplex environment viable. The performance impact of data sharing in terms of resources and response time impact are reviewed as part of the case study. The percentage of applications that use shared databases is an significant element in determining the performance costs of running a parallel sysplex.

The coupling facility itself has its storage dynamically partitioned into structures: cache structure, list structure and lock structure. The cache structure contains the global buffer pools and is used as a high-speed buffer for storing shared data with common read/write access. The cache structure is used to store changed and optionally unchanged pages. The changed data can support a quick refresh of local buffers after a cross invalidation. DB2 data sharing creates a shared communication area (SCA) in the list structure that is used to interchange information between sharing DB2 subsystems. The lock structure enables rapid shared and exclusive locking for serialization of shared resources maintained by the DB2 lock manager.

CICSplex Systems Management: CICSplex System Manager/ESA (CICSplex SM) provides a real-time single-system image system manager for CICS/ESA and allows an entire CICSplex to be managed as a single CICS system. Workload management within CICSplex and MVS optimizes service by automatically routing transactions to the most appropriate CICS address space in the CICSplex. It provides workload balancing and separation and handles transaction affinities. In its simplest form, a CICSplex operates within a single MVS image and uses CICS MRO facilities. In a parallel sysplex, a CICSplex can be configured across all the MVS systems using the enhanced MRO facilities

Many of the same reasons and benefits for employing CICS MRO can be extended to the sysplex environment. These benefits including splitting an application to run in multiple AORs to better utilize the processor complex and CPCs. Workload separation to provide greater availability of multiple AORs to reduce or minimize the impact of a failure. Additionally workloads can be separated based on transaction and workload affinities such as requiring access to the same programs/data for specific departments or to separate work based on resource usage and response time requirements.

The CICS Dynamic Transaction Routing program was extended in Version 4 of CICS to interact with the MVS/ESA workload manager to achieve optimum workload balancing for CICS transactions. The transaction routing is accomplished via MRO using standard cross memory services if the region resides on the same MVS system. If the region resides on a different MVS system in the parallel sysplex environment, the XCF is utilized to connect the systems which provides a significant performance advantage over the previously required VTAM LU6.2 link. The overhead for CICSplex is approximately 2% as compared with closer to 10% overhead for VTAM/ISC communication in earlier releases of CICS.

DB2 Data Sharing/ Data Sharing Groups and Group Buffer Pools: For transactions within a CPC to access DB2 data, there must be one or more DB2 instances defined to that CPC. The set of DB2 instances that access a specific set of shared data is called a “group”. Each individual DB2 instance is called a “member” of the group and only members of a group can share data in read-write mode. A data sharing group can have up to 32 members.

A data sharing member is a DB2 subsystem containing subsystem id, IRLM, address spaces, BSDS files, active and archive logs, ZPARM definitions etc. A data sharing member shares DB2 resources with other members of the group. All members of the group share the DB2 catalog and directory, all databases, table spaces, tables and indexes, plans and packages. In addition, all members share the group buffer pools and share the same lock structure and recovery information. As a result of sharing, a plan bound on one member can be executed on another and recovery of a table space can be executed on one member resulting in all members having the table space recovered.

Group Buffer Pools are cache structures in the coupling facility that are extensions of the local buffer pools in the individual members. Group buffer pools are used when two or members are sharing the same pageset or partition and at least one of the members has write interest. In this case the pageset or partition becomes “GBP dependent” due to inter-DB2 read-write interest”.

There must be one GBP for each local buffer pool containing table spaces to be shared. Group buffer pools are used in a data sharing environment to:

- store updated pages
- register the pages stored in the local buffer pools of members in a “Page directory”.
- to optimally clean unchanged pages in the group buffer pool.

The following table is taken from “DB2 for MVS/ESA, Data Sharing Performance Topics”, [Ref. 3], which indicates the conditions causing GBP dependence and when the mode of a page set physical lock changes. For Read-only tables, there is no group buffer pool dependence nor any page validation checking and therefore no data sharing overhead. Similarly, a Read-Write Request from one DB2 will not have any GBP dependence or page validation checking if there is no interest from other members.

Interest of Requesting DB2	**Interest of Other Members**	**Page Set Physical lock mode**	**Group Buffer Pool Dependence**	**Check Page Validity by Requester**
READ-ONLY	NONE, Read only	Shared	NO	NO
READ-ONLY	Read-Write	Intent Share (IS)	YES	YES
READ-WRITE	NONE	Exclusive (X)	NO	NO
READ-WRITE	Read-Only	Share with Intent Exclusive (SIX)	YES	NO/YES
READ-WRITE	Read-Write	Intent Exclusive (IX)	YES	YES

Table1. Group Buffer Pool Dependence and Physical Lock Modes

Page set physical locks are used to track inter-DB2 read-write interest of different members for a page set or partition. Each member acquires one page set physical lock for each open table space or partitioned index.

Switching of a page set or partition between Read-Only and Read-Write interest affects the mode of page set physical locks and the GBP dependency. The switching is caused by Read-Only switching and by updates to the page set or partition. Updated pages are stored in the GBP to improve transaction response time at commit and to allow quick refresh of these pages without having to resort to DASD I/O. The “page directory’ in the GBP registers the pages stored in the local buffer pools and provides a facility for managing the consistency of data pages with the local buffer pools of all members.

Lock Contention: Contention occurs when different data sharing group members need to use the same resource at the same time. There are three types of contention:

- False contention
- XES contention
- Real contention

Real contention occurs when the lock requests are incompatible. False contention occurs when the hashing algorithm for the lock table provide the same hash value for two different resources.

Since the MVS XES component is only aware of share and exclusive lock, an IRLM contention exit must be invoked that can examine the lock at a finer level of granularity to determine whether the lock modes are truly incompatible. The MVS XES manager will give control to the IRLM contention exit associated with the global lock manager if it cannot resolve the lock. If the IRLM contention exit determines that the contention is not real i.e., two compatible locks, the contention is deemed XES contention.

The data displayed in the table below shows a sample CF activity report for the lock structure. It indicates the total number of requests and rate per second, average delay in microseconds and number of requests experiencing contention for one of the system. The report is repeated for each system.

STRUCTURE NAME = DSNDB2B_LOCK1 TYPE = LOCK

SYSTEM NAME	TOTAL AVG/SEC		REQ	% OF ALL	Service Time Average (MIC)	STD_DEV	REQUEST	CONTENTIONS
SYS1	30179	SYNC	30K	50.20%	150.4	29.1	# REQ	29K
	251.5	ASYNC	0	0.00%	0	0	# REQ DELAYED	517
		CHNGD	0	0.00%	INCLUDED IN ASYNC		-CONT	515
							-FALSE CONT	199

Table 2. RMF Coupling Facility Activity Report- Lock Structure

The data displayed in Table 2 is for a DB2 system. It should be noted that DB2 does not make asynchronous requests to the coupling facility and therefore this value is always 0. The average service time of the coupling facility is recorded in microseconds. The column labeled "REQ" equals the number of synchronous requests without contention. The #REQ delayed is a subset of this, which should be small and indicates the number of lock requests which could not be immediately processed due to some delay. The field DELAYED FALSE CONT indicates the number of requests which were delayed due to false contention. It should be noted that although MVS XES detects false contentions, the RMF XES definition of false contention is not the same as the DB2 definition reported by DB2PM. DB2 can have more information concerning the type of contention, particularly if the XCF Global Lock manager needs to be invoked. Therefore DB2PM can report a higher number of contentions than the XCF Activity report. In this example, the percentage of false contentions is so small that there is no reason to be concerned with the difference in reporting.

DB2 Data Sharing and Sysplex Performance Reports

DB2 data sharing allows databases to be shared by multiple DB2 subsystems across multiple processor complexes with full read/write capability. The costs of data sharing depend upon the degree of sharing between systems. Coupling Facility accesses are made to access shared data in the group buffer pool and to resolve locks, some of which are due to real contention, while other accesses are required to resolve XES contention and false contention. There are specific RMF and DB2 PM reports that provide data on coupling facility access rates, lock service times, and group buffer pool accesses. The RMF reports include the Coupling Facility (CF) report and the XCF Activity report. DB2PM provides information on locking and data sharing locking as part of the statistics reports. Group buffer pool total accesses are also provided by the Statistics reports based on IFCID 230 record.

The total access rate to the Coupling Facility can be obtained for each system by totaling the Number of Lock Requests Per Second and Number of Group Buffer Pool Accesses per Second for each DB2 subsystem. The RMF XCF Activity Report details the coupling facility access rate for each Group Buffer Pool and for each XES Lock Structure. Figure 1a below shows the cache structures summarized for T1CF1LP for the case study and their allocation sizes. Figure 1b below shows the SCA and lock structures summarized for T1CF2LP for the case study and their allocation sizes.

The total coupling facility access rates are relatively low for this application and the CF utilization was < 5%.

COUPLING FACILITY NAME = T1CF1LP							
TOTAL SAMPLES(AVG) = 119 (MAX) = 119 (MIN) = 119							
COUPLING FACILITY USAGE SUMMARY							
STRUCTURE SUMMARY							
TYPE	STRUCTURE NAME	STATUS CHG	ALLOC SIZE	% OF CF STORAGE	# REQ	% OF ALL REQ	AVG REQ/SEC
LIST	IXCXCF	ACTIVE	4M	0.40%	924	3.40%	7.7
CACHE	DSNDB2U1_GBP0	ACTIVE	2M	0.20%	13	0.00%	0.11
	DSNDB2U1_GBP1	ACTIVE	73M	7.80%	12910	47.10%	107.58
	DSNDB2U1_GBP2	ACTIVE	73M	7.80%	7603	27.70%	63.36
	DSNDB2I1_GBP0	ACTIVE	2M	0.20%	22	0.10%	0.18
	DSNDB2I1_GBP1	ACTIVE	73M	7.80%	1588	5.80%	13.23
	DSNDB2I1_GBP2	ACTIVE	73M	7.80%	2669	9.70%	22.24
	DSNDB2I1_GBP4	ACTIVE	34M	3.70%	11	0.00%	0.09
	DSNDB2I1_GBP5	ACTIVE	34M	3.70%	12	0.00%	0.1
	DSNDB2I1_GBP6	ACTIVE	147M	15.60%	560	2.00%	4.67

Figure 1a. Group Buffer Pool Allocations defined in Coupling Facility T1CF1LP

C O U P L I N G F A C I L I T Y A C T I V I T Y							
MVS/ESA	SYSPLEX CPLEX1	DATE	5/29/96	INTERVAL 002.00.000			
SP5.2.2	RPT VERSION 5.2.0	TIME	14.04.00	CYCLE 01.000 SECONDS			
COUPLING FACILITY NAME = T1CF2LP							
TOTAL SAMPLES(AVG) = 119 (MAX) = 119 (MIN) = 119							
COUPLING FACILITY USAGE SUMMARY							
STRUCTURE SUMMARY							
TYPE	STRUCTURE NAME	STATUS CHG	ALLOC SIZE	% OF CF STORAGE	# REQ	% OF ALL REQ	AVG REQ/SEC
LIST	DSNDB2U1_SCA	ACTIVE	4M	0.40%	48	0.20%	0.4
	DSNDB2I1_SCA	ACTIVE	4M	0.40%	44	0.10%	0.37
LOCK	DSNDB2U1_LOCK1	ACTIVE	8M	0.90%	18514	62.30%	154.28
	DSNDB2I1_LOCK1	ACTIVE	8M	0.90%	11096	37.40%	92.47
	STRUCTURE	TOTALS	24M	2.60%	29702	100%	247.52

Figure 1b. Lock Structure Allocations defined in Coupling Facility T1CF2LP

Both the RMF XCF Activity report and DB2PM reports provide the same access rates to the group buffer pools for each of the systems in the sysplex.

The data displayed in Figure 2 below shows the total number of reads, writes and castout operations for Group Buffer Pool 6 for the DB2I1 subsystem on SYS1. The information is reported as totals for a 10 minute period. The statistics reports in DB2PM are obtained for summarized periods by providing start and end times.

There are 7 group buffer pools defined for this DB2 subsystem and the report is therefore repeated for each group buffer pool.

The total number of accesses to the coupling facility for requests due to ***DB2I1 for GBP6***=

Total Number of Reads+Total Number Writes+ Total Pages Castout+ Other Requests= 429+430+1+1231=2091.

The access rate to the coupling facility for GBP6 in DB2I1 can be obtained by dividing the total number of accesses by the duration of report=2091/10=209 accesses per minute.

The total number of accesses to the coupling for ALL requests due to DB2I1=
SUM Number of Accesses for (GBP1+GBP2+GBP3+GBP4+GBP5+GBP6+GBP7).

Read Operations	Reads Due To Cross Invalidation	= 12%	Reads Due To Page Not Found in Buffer Pool	=	88%
	Data Returned	= 53	Data Returned	=	4
	Data Not Returned - R/W Interest	= 0	Data Not Returned - R/W Interest	=	347
	Data Not Returned - No R/W Interest	= 0	Data Not Returned - No R/W Interest	=	25
	Reads Done by Prefetch	= 0%	**Total Number of Reads**	=	429
	Data Returned	= 0	**Read Hit Percentage**	=	**13%**
	Data Not Returned - R/W Interest	= 0	Percent Not Returned - R/W Interest	=	81%
	Data Not Returned - No R/W Interest	= 0	Percent Not Returned - No R/W Interest	=	6%
	Other Requests	= 1231	Reads Failed - No Storage	=	0
Write Operations	Changed Pages - Sync	= 430	Castout Engine Not Available	=	0
	Clean Pages - Sync	= 0	Write Engine Not Available	=	0
	Changed Pages - Async	= 0	Write Engine Failed - No Storage	=	0
	Clean Pages - Async	= 0	**Pages Castout from GBP to DASD**	=	1

Figure 2. DB2PM report- Subset of Statistics Report for Group Buffer Pool 6 for SYS1

The data displayed in Figure 3 below shows the XCF Activity report for DB2I1 for GBP6. The figure represents a subset of the XCF Activity report and shows the coupling facility requests for GBP6 for DB2I1 and is summarized over the RMF interval. The DB2PM data was summarized over a 10 minute period whereas the RMF XCF Activity report is for a 2 minute subset of this 10 minute period. Unfortunately, there is no way to obtain summarized RMF XCF Activity reports for a specified duration as with other RMF reports. DB2PM and RMF both indicate approximately 3 accesses/second to the coupling facility for DB2I1 due to GBP6.

SYSTEM NAME	TOTAL AVG/SEC		#REQ	% OF ALL
SYS1	316	SYNC	316	56.40%
	2.63	ASYNC	0	0.00%
		CHNGD	0	0.00%
SYS2	244	SYNC	244	43.60%
	2.03	ASYNC	0	0.00%
		CHNGD	0	0.00%
TOTAL	560	SYNC	560	100%
	4.67	ASYNC	0	0.00%

Figure 3: A subset of the RMF XCF Activity Report for DB2I1, GBP6.

These reports illustrate the metrics used in determining CF access rates. CF access rates will be large if there is a high degree of data sharing. If datasharing is low than the delays experienced will not be of significant concern.

Case Study

The application defined for the case study was a CICS/DB2 application utilizing sysplex, CICSplex and DB2 data sharing. The application was written using prototype code which was modified slightly to execute in this environment. The original code was developed to execute in a single region and then modified to run in multiple CICS regions to take advantage of a sysplex environment with multiple CPCs.

Benchmarks were conducted to demonstrate CPU, application and architecture scalability in the sysplex environment. Transactions were executed at varying rates to demonstrate scalability in a 2 CPC environment.

The benchmark facility consisted of the following hardware configuration:

- **2 9672-RX3 (10 processors) each with 2 GB Central Storage and 2 GB Expanded Storage.**
- **RAMAC-2 RAID Storage with 2 GB Cache backed by 180 GB emulating 64 3390-3 devices.**
- **RAMAC-1 RAID Storage with 1 GB Cache backed by 45 GB emulating 16 3390-3 devices.**
- **2 Coupling Facilities each with 2 High Speed Links**.

The hardware configuration was logically partitioned to run 2 Coupling Facilities, two 9672-R53 processors and the TPNS driver as follows:

- 1 9672 -RX3 Processor which was logically partitioned to run as 3 images (2 CFs, 1LPAR for TPNS):
  - ⇒ 2 Coupling Facilities T1CF1LP and T1CF2LP each with 800 MB Central Storage each with 3 processors.
  - ⇒ 1 LPAR to execute TPNS with 4 processors and 400 MB of Central Storage and 2GB of Expanded Storage.

- 1 9672-RX3 Processor to run as:
  - ⇒ 2 LPARs each with 5 engines or logically 2 9672-R53 processors.
  - ⇒ Workloads would be balanced across the two systems utilizing sysplex and CICSPlex.

The DASD configuration was spread as follows:

- DB2 Databases were spread uniformly across 32 of the 64 RAMAC-2 Class devices.
- MVS work packs, DB2 Logs and CICS Journals were placed on the RAMAC-1 Class devices.

The application was response time critical, required 24x7 availability and minimal to no down time. The advantage of parallel sysplex and DB2 data sharing was high availability and the ability to run a fully operational and responsive system in a multiple CPC environment. The application architecture, transaction processing and routing can be best described in Figure 4 below.

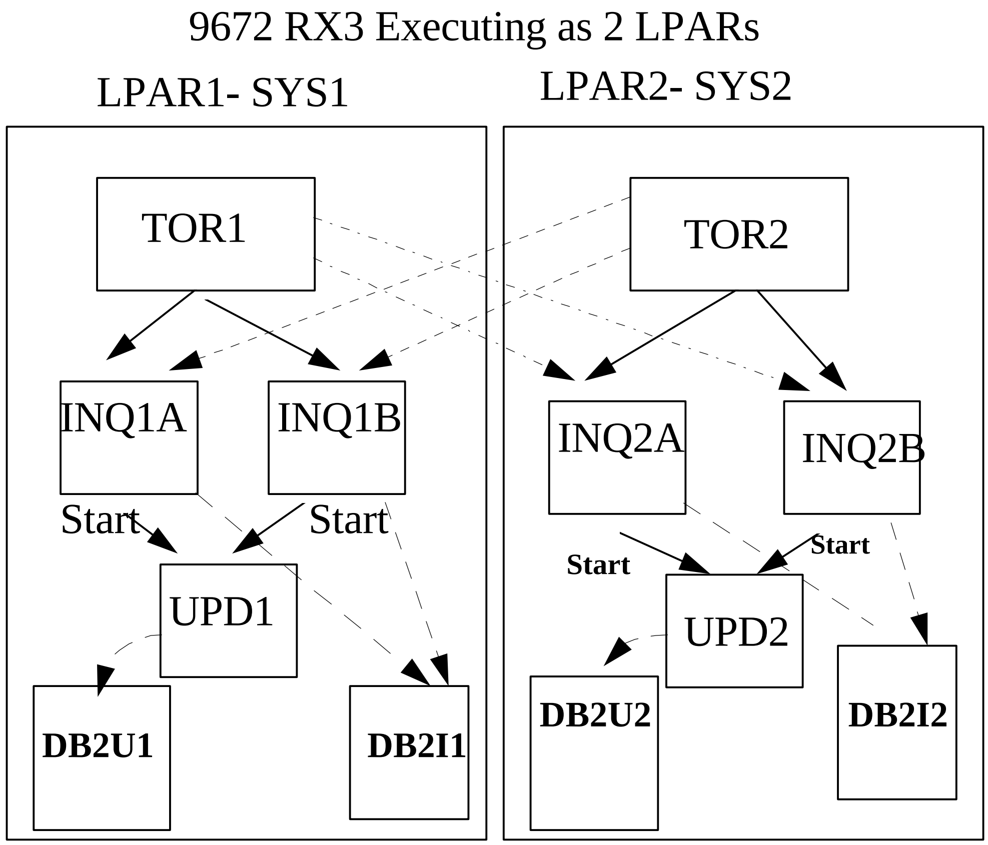

Figure 4. CICS Configuration in Parallel Sysplex/CICSplex Case Study

The CICS regions include: 2 primary AORs, 1 secondary AOR and 1 TOR per complex and 2 DB2 subsystems per CPC. DB2I supported the primary Inquiry-based AOR and DB2U supported the secondary Update work. When executing as 2 1x5 LPARs there are a total of 4 DB2 subsystems DB2I1 and DB2U1 on LPAR1 and DB2I2 and DB2U2 on LPAR2.

To evaluate the impact of data sharing, there were two runs executed, one with data sharing and the other without. Additionally, a second set of runs was executed to evaluate the impact of sysplex. All work executed at an arrival rate of approximately 27 transactions/second, since some performance problems due to insufficient I/O pathing were experienced at higher arrival rates. The non-data sharing and non-sysplex runs were executed on a 1X10 RX3 CMOS processor complex whereas the data sharing and sysplex runs were executed on the RX3 running in LPAR mode as 2-1X5 CMOS processor complexes.

Evaluating the impact of data sharing, sysplex and CICSplex overhead could be best achieved utilizing the same exact hardware environment. Unfortunately, the test team was unable to schedule additional performance tests with the system configured as a stand-alone 1X5 processor. There was limited test time with multiple goal sets including contingency and availability evaluations which were time consuming. Fortunately, the speed differences between an RX3 and R53 processor complexes have been documented by IBM and the scalability of these processors has also been documented in terms of ITR and ETRs for the 9672-CMOS processor class. We were therefore able to compare CPU timings from different processor configurations by normalizing the data. The processor speeds are detailed in Table 3 below.

Processor Model	9672-RX3	9672-R53
#CECs	1	2
# Processors per CEC	10	5
Central Storage per CEC	2GB	1 GB
Expanded storage per CEC	2 GB	1 GB
MIPS per Engine	16.5	19.4
MIPS per CEC	165.2	96.9

Table 3. Comparison between 9672 RX3 and R53 Systems for Case Study

In order to affect any comparison in terms of data sharing or sysplex cost for these test cases, we first needed to normalize the CPU measurements to the same processor class. CPU timings measured on an RX3 will be longer than those taken on an R53 and therefore all overhead figures must be adjusted and normalized to account for the speed differences. Additionally, we learned based on discussions with IBM Poughkeepsie that the LPAR management overhead reported by RMF represents a trivial portion of the total LPAR management cost. In fact, for a CMOS RX3 executing as 2 LPARs each with 5 processors, the overhead for LPAR can be as high as 10%. We therefore also accounted for this overhead when comparing CPU timings for the sysplex and data sharing benchmark runs.

Case Study Workload Analysis

All of the Inquiry transactions and the Simple Update UPDA transaction could execute in any one of the four Inquiry AORs. TPNS was used to drive the work to the TORs and CICSplex was used to balance the transaction workload across the four regions on the 2 CPCs. The secondary transactions were initiated via a CICS START with Protect from the primary Inquiry or Simple Update transaction. A mirror transaction was invoked in the secondary Update AOR which started either the secondary UPDB transaction or the INSERT transaction. There was no CICSplex or workload balancing involved in the secondary AORs processing the Update UPDB and INSERT transactions.

In the fully configured production system, there will be approximately 70 Inquiry transactions/second, and 8 Update transactions/second. There is also one UPDB transaction per Simple Update generated and one Secondary Insert per Inquiry generated. The total workload requirement for the production system is expected to be approximately 156 transactions/second. The CPU requirements to support this work is in the 200 MIPS range.

The application was designed to consist of primary units of work which spawned secondary units of work. This was done to satisfy very stringent response time requirements for the primary units of work. The secondary workload which was created asynchronously had minimal to no respone time requirements.

The following tables detail the profile of CICS transactions which was fairly similar across all regions. The tables represent average response times and resource usage across both systems. The transaction volumes are for one of the four Inquiry regions and the secondary AOR on SYS1.

Table 4a shows the average response times, CICS application CPU time and DB2 Class 2 CPU time, DB2 I/O time and DB2 Class 2 Elapsed time for the primary inquiry, primary update and secondary update and insert transactions.

Transaction	Volume	MEAN Response Time (sec)	CICS CPUTM	DB2 ELAPSED Time (sec)	DB2 TCB Time (sec)	DB2 I/O TIME (sec)
Inquiry type1 simple	19936	0.0674	0.0299	0.0051	0.0034	0.0010
Inquiry type 1 complex	1143	0.5248	0.1092	0.2752	0.0710	0.1973
Inquiry type 2 simple	24459	0.0631	0.0270	0.0046	0.0031	0.0010
Inquiry type 2 complex	395	0.3427	0.0691	0.1924	0.0409	0.1438
Simple Update UPDA	1587	0.4685	0.0679	0.2991	0.0642	0.1775
Secondary Update UPDB	764	0.0390	0.0050	0.0210	0.0080	0.0075
Secondary Insert	23754	0.0340	0.0040	0.1000	0.0030	0.0400

Table 4a. CICS/DB2 transaction profile for INQ1 and UPD1 on SYS1- profile is similar for all regions

The data presented in Table 4b. below gives an indication of the CRUD profile in terms of total logical I/Os (Select/Insert/Update/Delete), Selects and Insert activity, fetches, getpages and physical I/Os. The Simple Type1 and Simple Type 2 Inquiries represent the bulk of the transaction load and have minimal I/O processing requirements and trivial response times. The Complex Inquiry transactions are much more CPU and I/O intensive but still possess a service level requirement of .5 seconds for average response time. The secondary transactions do not have any service level requirements for response time but do have a high availability requirement.

The Secondary Update transaction UPDB uses more DB2 CPU than the Insert transaction and does approximately 3 fetches per update transaction. One INSERT transaction is generated for each Inquiry transaction and performs approximately 1.2 INSERTS per Inquiry.

Transaction	Volume	Logical I/O COUNT	GET PAGES	SYNCH I/O	PRE-FETCH	FETCH	#Inserts	#Selects
Inquiry type1 simple	19936	1.33	3.36	0.05	0.180	0.002		1.3
Inquiry type 1 complex	1143	9.49	80.48	11.74	4.066	20.600		9.5
Inquiry type 2 simple	24459	1.03	3.10	0.05	0.177	0.002		1.0
Inquiry type 2 complex	395	7.32	58.93	6.37	3.815	5.400	0.0	7.1
Simple Update UPD1	1587	11.47	106.10	9.05	0.444	5.800	7.4	4.2
Secondary Update UPDB	764	0.00	8.05	0.75	0.280	3.200	0.0	0.0
Secondary Insert INS1	23754	1.16	5.60	0.00	0.000	0.000	1.2	0.0

Table 4b. DB2 CRUD profile for transactions executing in INQ1 and UPD1 on SYS1- profile is similar for all regions

Case Study Benchmark Scenarios

The following benchmarks were executed in order to evaluate the impact of DB2 data sharing and Parallel sysplex. We were not able to perform any specialized tests in which we could isolate the impact of CICSplex. We have detailed the analysis and performance measurement data used in analyzing the performance impact of DB2 data sharing and Parallel sysplex for this environment.

By comparing the B1 and B2 scenarios we could evaluate the impact of data sharing for this application. The difference in CPU utilization between the corresponding runs could then be attributed to DB2 Data Sharing overhead.

B1 vs. B2 DB2 Data Sharing Overhead Scenarios:

B1: Executed all Inquiry workloads that had corresponding maintenance/update work associated with it. The transactions included Query Type2 Simple and Complex, Query Type3 Simple and Complex, Type2 and Type 3 Primary Updates and all corresponding secondary Insert and Update transactions. These secondary transactions were auto-initiated via a CICS start once the primary unit of work was completed.

B2: Executed Queries that had corresponding maintenance/update work associated with it. The transactions included Query Type2 Simple and Complex, Query Type3 Simple and Complex, Type2 and Type 3 Primary Updates and all corresponding secondary Insert and Update transactions. This work executed the same script as B1 except that it ran on a single RX3 CMOS processor complex with no data sharing or sysplex management. The same number of TORs and AORs as B1 were utilized for processing the load. However, there was only one DB2 subsystem to support the primary query and update transactions and one DB2 to support the secondary update transactions.

The purpose of B3 and B4 were to evaluate the impact of sysplex overhead. For these scenarios, only a subset of the transaction code was executed in order to eliminate the need to call DB2. Essentially the code that was processed eliminated all calls to DB2. As a result, there would be no necessity for data sharing in these runs. Therefore, the only difference between these runs is the sysplex overhead for the B3 run as compared with no sysplex overhead in the B4 run that executed on an RX3 processor complex.

B3 vs. B4 DB2 Sysplex Overhead Scenarios:

B3: The full transaction mix executing only a subset of the transaction code was exeucted on 2 1x5 processor complexes with the same CICS configuration and DB2 configuration as B1.

B4: This run is the same workload as B3 except that it is executed on a single CPC, an RX3 with 10 CPUs. The same number of TORs and AORs as B3 were utilized for processing the load.

Evaluation of Sysplex Overhead

In order to compare the CPU utilizations for the sysplex vs. non-sysplex executions, we adjusted the CPU utilization for the non-data sharing execution on the RX3 to what it would have used on a single image 9672-R53 (1x5 CMOS CPC that was not LPAR'd).

The utilization of the non-data sharing work would therefore be 15% less utilized on a 9672- R53. Table 3 above shows that there is a 15% difference in the MIPS capacity per engine for an 9672-RX3 as compared with a 9672-R53. The results in Table 5 below indicate that there is an approximate 11% increase in CPU requirements for B3 over B4. However since there is between 5-10% overhead for running a CMOS box in LPAR mode, the net cost of sysplex for this configured system is probably in the 2% range that IBM has indicated in their documentation.

	B3	B4	% Difference	SYSPLEX Overhead
Normalized volume	15234	14456		
CPU Busy (%)	20.35	17.43		
ITR	74859.951	82914.713		
Capacity Delta*			9.71%	
Coupling Efficiency (%)*			90.29%	
Increased * CPU Time			10.76%	
B3 vs B4				**1-2%**

Table 5. Differences in measurements for B3 vs. B4
**** NOTE: Differences include LPAR management overhead, actual difference is only 1-2% cost for sysplex.***

1. ITR is calculated by dividing ETR (transactions) by CPU % Busy
2. Coupling Efficiency is calculated as (non-CF scenario ITR)/(CF scenario ITR)
3. Capacity Delta is computed as 1-Coupling Efficiency
4. Increased CPU Time is computed as (non-CF scenario ITR - CF scenario ITR) / (CF scenario ITR)

Evaluation of Data Sharing Overhead

In order to determine the overall impact of Data Sharing on the application, we compared the Internal throughput Rates (ITRs) and computed percent increase in CPU from the non-data sharing run to the data sharing. In order to affect the comparison, the measured CPU utilizations were adjusted to account for both the speed differences between the machines and the LPAR management overhead. The CPU timings for B1and B2 were normalized to the speed of a single image R53. Therefore, the RX3 figures were adjusted to be 15% less busy on a faster 1X5 engine speed and the 1X5 figures were adjusted to be 10% less busy to minimize the impact of LPAR management. The adjusted CPU utilizations are detailed in Table 6 below. Additionally, since the two runs did not execute for the same amount of time, the volumes were normalized for the same interval length. The volumes used in this analysis are based on transaction load through the TOR.

	B1	B2	
Normalized Volume	32814	31615	
CPU Busy (%)	17.05	14.72	
ITR	192457.48	214775.82	
Capacity Delta			10.39%
Coupling Efficiency (%)			89.61%
Increased CPU Time			11.60%

Table 6. Comparison of Runs B1 vs. B2 Data Sharing Overhead

The data displayed in Table 6 above represents the overall impact of data sharing. The next step is to identify where the overhead occurs based on additional performance measurement data.

Table 7 details the overhead by CICS region and DB2 address space. Most of the overhead occurs in the secondary AOR regions UPD1 and UPD2 which contains the transactions performing most of the update activity. The inquiry activity is much less impacted by DB2 data sharing, since DB2 CPU is a much smaller percentage of the total work performed by queries. The secondary units of work workloads have a trivial CICS CPUtime since there is minimal application code. These transactions are therefore much more impacted by the increase to the DB2 component of CPU in the data sharing environment.

Workload	Adjusted B1 CPU/tran	Adjusted B2 CPU/tran	B2 vs. B1 % Difference
DB2I1IRLM	0.0004	0.0004	
DB2I1MSTR	0.0010	0.0009	
DB2I1DBM1	0.0014	0.0012	
DB2I1 Total	0.0028	0.0024	15.00%
TOR1	0.0097	0.0084	11.11%
AOR: INQ1A	0.0382	0.0332	6.69%
DB2U1IRLM	0.0002	0.0002	
DB2U1MSTR	0.0058	0.0050	
DB2U1DBM1	0.0011	0.0010	
DB2U1 Total	0.0071	0.0062	15.00%
AOR: UPD1	0.0140	0.0122	20.36%
AOR: INQ1B	0.0392	0.0341	8.93%

Table7. Data Sharing Overhead Comparisons by CICS region and DB2 Subsystem

The data displayed in Table 8 details the specific components of CPU time which increased with data sharing. While the average increase for the Inquiry AORs was in the 7-9% range, a decomposition of CPU time indicates that there was only a 5% increase to CICS application CPU for queries and a 4% increase to the DB2 component for the inquiries. The 1.4% delta increase from 4.6% to 6% is the increase to MVS overhead as a result of data sharing. Similarly the Primary Update transactions experienced an approximate 6% increase in CICS application CPU time and a 32% increase in the DB2 component of CPU time due to data sharing. The overall CICS+DB2 component increase was approximately 19% and there was a delta increase of an additional 1.3% for MVS overhead for the Primary Updates.

We were unable to produce this same detailed analysis for the secondary Update regions because detailed DB2 performance measurements were not collected. It would have been interesting to review the detailed transaction performance data since there was as a 20% increase in CPU processing requirement for the CICS region due to data sharing.

The secondary update workloads UPDA and UPDB are the primary users of the group buffer pools in the coupling faciltiy since update transactions will have read/write intent. It is therefore not surprising that the resulting increases in CPU time are more significant for the update transactions than the inquiry work.

	CICS Application CPU time	DB2 TCB time	CICS+DB2 CPU Time	CPU time with CICS O/H	CPU time with MVS O/H
Inquiry- Type 2 Simple	4.90%	4.07%	4.60%	4.55%	6.01%
Inquiry -Type2 Complex	5.02%	22.65%	11.48%	11.43%	12.79%
Inquiry- Type3 Simple	4.98%	8.57%	5.18%	5.13%	6.58%
Inquiry -Type3 Complex	9.36%	-27.17%	-3.68%	-3.74%	-2.16%
Primary Update	5.85%	32.33%	18.71%	18.66%	19.90%

Table 8. Comparison between B1 and B2: % Increase in CPU time for Data sharing over non-datasharing.

In summary, this analysis confirms that data sharing overhead is highly dependent upon the workload requirements. Inquiry based work will not require access to the coupling facility so long as the pages requested are not marked with read-write intent. Therefore if there is no corresponding update transactions accessing the same pages, the inquiry transactions will not be impacted.

In this case study, the Inquiry transactions were only directly impacted by the Primary Update transactions that executed in the same regions as the Inquiry work. These transactions updated many of the same tables that the Inquiry transactions accessed, however the transaction rate for the Primary Update transactions was on the order of 3-4 transactions/second as compared with 69-70 transactions/second. The Secondary Update transactions UPD2, which also had a transaction rate equivalent to 69-70 transactions/second accessed different tables than the Inquiry transactions. Therefore although it had extremely high data sharing requirements and experienced approximately a 20% in DB2 Data sharing processing overhead, the Inquiry transactions were not impacted by it.

Buffer Pool Analysis

A series of benchmark scenarios were executed to evaluate the scalability the configuration in terms of CPU requirements, response times and buffer pool requirements. The series of runs executed utilized the same mix of work but executed at varying transaction loads. These runs were defined as follows:

A1: Executed at approximately 80 Transactions per second included all Type1-Type 3 Inquiries
including Simple and Complex, secondary updates,
and simple updates for Type 2 and Type 3 transactions.
A2: Executed the same workload as A1 at 40 transactions per second
A3: Executed the same workload as A1 at 27 transactions per second
A4: Executed the same workload as A1 at 20 transactions per second
A5: Executed the same workload as A1 at 70 transactions per second
A6: Executed the same workload as A1 at 13 transactions per second

For the purposes of this paper and to illustrate the buffer pool analysis, we have restricted the reporting of results to the A1 and A3 executions.

The mapping of tables and indexes to local buffer pools was the same as that defined for Group Buffer Pools in the Coupling Facility. The data displayed in Table 9 below represents buffer pool statistics for the A3 run executing 27 transactions/second in a data sharing environment and includes the Buffer Pool Hit Percentages and ratios of getpages/synchronous I/O counts for the primary transactions per buffer pool. The number of commits is equivalent to the number of transactions for this sample period from A3. Therefore summary statistics such as Getpages/transaction can be obtained. The BPHIT% represents the percentage of time that the required page was found in the buffer pool. The percentage of time that a physical I/O is required = 1-BPHIT%.

A key inquiry index was defined to BP6 and the corresponding key Inquiry Table was mapped to BP7. The simple inquiry transactions will always require access to the key Inquiry Index. The BPHIT% for the Key table is 0 for all inquiry transactions. The BP HIT % for Buffer Pools 4 and 5 containing the bulk of the tables and their corresponding indices is in the 80-90% range for inquiry transactions.

TRANID	BPOOL	COMMIT	GET PAGE	GET PAGE/ Tran	SYNCH I/O	GET PAGE/ SYNCHIO	BP HIT%	PRE FETCH	PRE FETCH PAGES
Secondary UPD2	1	748	361	0.483	261	1.38	28%	0	
	2	748	2310	3.088	296	7.8	87%	0	
Primary UPDB	1	737	8788	11.924	1741	5.05	80%	0	
	2	737	48198	65.398	3172	15.19	93%	0	
	4	737	883	1.198	210	4.2	75%	3	
	5	737	9209	12.495	67	137.45	99%	3	
	6	737	8246	11.189	76	108.5	83%	272	1289
	7	737	2745	3.725	1217	2.26	9%	51	1272
Average Type 3 Inquiry	1	466	125	0.268	61	2.05	0%	59	1208
	2	466	587	1.260	134	4.38	0%	49	1061
	4	466	303	0.650	76	3.99	75%	0	0
	5	466	505	1.084	63	8.02	88%	3	0
	6	466	1652	3.545	62	26.65	12%	197	1390
	7	466	120	0.258	42	2.86	0%	86	1942
Average Type 2 Inquiry	1	11616	1854	0.160	534	3.47	0%	224	5275
	2	11616	7382	0.636	963	7.67	38%	229	3640
	4	11616	1086	0.093	66	16.45	94%	0	0
	5	11616	1949	0.168	29	67.21	99%	7	0
	6	11616	36958	3.182	547	67.56	75%	2104	8661
	7	11616	613	0.053	167	3.67	0%	336	7598
Average Type1 Inquiry	1	9389	3342	0.356	1257	2.66	0%	985	20041
	2	9389	16424	1.749	3400	4.83	49%	412	4964
	4	9389	10115	1.077	1786	5.66	82%	0	0
	5	9389	15020	1.600	2605	5.77	83%	7	6
	6	9389	38684	4.120	554	69.83	83%	1967	5960
	7	9389	2693	0.287	555	4.85	0%	1127	22345

Table 9. A3 Results: Buffer Pool Statistics for Primary Updates, Secondary Update 2 and all Inquiries averaged.

The data displayed in Table 10 compares the BPHIT% and ratios of Getpages/synchronous I/Os at two arrival rates A3=27 Transactions/second and A1=83 Transactions/second. The results indicate no difference in BPHIT% as a function of increased volume. The buffer pools were sized based on table size requirements and the fact that the buffer pool hit percentage is invariant to volume suggests that one only need consider table space requirements in sizing the buffer pools.

TRANID	BPOOL	A3 GETPAGE/ SYNCHIO	A3 BPHIT%	A1 GETPAGE/ SYNCHIO	A1 BPHIT%
Secondary UPD2	1	1.38	28%	1.35	25%
	2	7.8	87%	7.89	88%
Primary UPDB	1	5.05	80%	4.6	78%
	2	15.19	93%	14.56	93%
	4	4.2	75%	4.02	74%
	5	137.45	99%	149.06	99%
	6	108.5	83%	159.5	87%
	7	2.26	9%	2.26	0%
Average Type 3 Inquiry	1	2.05	0%	2.45	0%
	2	4.38	0%	5.06	0%
	4	3.99	75%	5.29	81%
	5	8.02	88%	13.23	92%
	6	26.65	12%	26.66	11%
	7	2.86	0%	3.52	0%
Average Type 2 Inquiry	1	3.47	0%	2.37	0%
	2	7.67	38%	6.63	20%
	4	16.45	94%	16.36	94%
	5	67.21	99%	94.6	99%
	6	67.56	75%	73.32	76%
	7	3.67	0%	3.04	0%
Average Type1 Inquiry	1	2.66	0%	2.61	0%
	2	4.83	49%	4.89	51%
	4	5.66	82%	5.65	82%
	5	5.77	83%	5.87	83%
	6	69.83	83%	90.6	85%
	7	4.85	0%	5.23	0%

Table 10. Comparison between BP Hit % for A1 vs. A3

Group Buffer Pool Analysis

The following tables show the amount of activity to each of the group buffer pools in the coupling facility and the resulting delays for A1 and A3. Group buffer pool data is not captured by transaction and plan name. It is only captured for the entire DB2 Subsystem. Therefore, some assumptions must be made based on application knowledge as to which transactions require access to the GBP and coupling facility.

The data displayed in Tables 11 shows the transaction rates per second on SYS1 and SYS2 respectively for scenario A3 executing 27 primary transactions per second.

Tranid	SYS1 Rate/second	SYS2 Rate/second
Inquiry Transaction Type 3	.27	.23
Inquiry Transaction Type 2	6.9	6.8
Inquiry Transaction Type 1	6.0	5.9
Primary Update UPDA	.48	.47
Secondary Update: INS2	13.2	13.1
Secondary Update UPDB	.42	.43

Table 11. Transaction rates per second for SYS1and SYS2 for execution A3.

The data indicates that the systems are fairly well balanced between SYS1 and SYS2 although the SYS1 does seem to have a slightly higher transaction rate. The data displayed in Table 12. shows the Group Buffer pool statistics for the A3 execution. Requests to the group buffer pool are made if the page requested is marked with read-write intent. The group buffer pools correspond to the DB2 buffer pools in main processor storage.

The group buffer pool request rate per second can be used in determining the probability that a particular transaction type will need to access the GBP. The transaction rate for Secondary transaction INSERT=13/second for SYS1. Therefore, a typical INSERT transaction will require 43/13.2=3.3 accesses/tran to the GBP. Similarly there are approximately 13 inquiry transactions/second, and therefore there is approximately 1.1 accesses/ran to GBP2. Additionally the inquiries will access the key Inquiry Index and key Inquiry Table buffer pools approximately .4 accesses/tran. The total average delay for GBP accesses for inquiry transactions on SYS1 is fairly trivial and can be computed as follows:

=1.1*225.6+.46*294.5+.46*286.4=515.4 microseconds
=0.00052 seconds

The Read Hit % indicates the percentage of time that the data requested was found in the group buffer pool. If the data page is not found in the group buffer pool, an I/O must be performed to bring the page back from DASD. Therefore if the Read Hit % for an index = 60%, it implies that 40% of the time an external I/O will be required to read the index page. The delay for accessing the Group Buffer Pool in this case equals the average delay in GBP or 242 microseconds + the I/O time or approximately 20 milliseconds for conventional 3390 DASD.

The # Requests/Second can be compared with the transaction arrival rates to arrive at a likelihood that a transaction requires access to the Coupling Facility or Group Buffer Pool. Group Buffer Pool 2 in DB2 is accessed at least once per Inquiry and the key Inquiry Index and the key Inquiry Table is accessed approximately 30% of the time for an inquiry transaction.

A3		#REQ/SEC	AVG SERVICE Time (MICROSEC)	STD_DEV	READ HIT %	ASYNCH CAST OUT RATE
SYS1	DB2D GBP1	51.34	220.6	74.3	54%	14.43
SYS2	DB2D GBP1	43.83	217.3	68.9		
SYS1	DB2D GBP2	23.92	242.7	71.8	60%	4.06
SYS2	DB2D GBP2	23.63	235.2	65.8		
SYS1	DB2T GBP1	8.65	262.6	164.2	39%	0.00
SYS2	DB2T GBP1	6.88	246.1	139.1		
SYS1	DB2T GBP2	14.81	225.6	101.9	33%	0.00
SYS2	DB2T GBP2	13.77	218.1	87		
SYS1	DB2T GBP6	4.67	294.5	138.6	13%	0.00
SYS2	DB2T GBP6	5	281.3	131.9		
SYS1	DB2T GBP7	5.75	286.4	187.7	24%	0.00
SYS2	DB2T GBP7	4.57	291	186.8		

Table 12. Group Buffer Pool Statistics for A3

The data displayed in Table 13 below shows the transaction rates/second for SYS1 and SYS2. In addition, the data displayed in Table 14 shows the Group Buffer Pool Statistics for the A1 run at 84 transactions/second. The Group Buffer Pool Read Hit Percentages were approximately the same as at A3=N/3 except that the hit rate for the Key Inquiry Index was better at the higher arrival rate.

Tranid	SYS1 Rate/second	SYS2 Rate/second
Inquiry Transaction Type 3	.72	.78
Inquiry Transaction Type 2	21.1	21.2
Inquiry Transaction Type 1	18.6	18.1
Primary Update UPDA	1.44	1.5
Secondary Update: INS2	40.8	40.4
Secondary Update UPDB	1.3	1.34

Table 13. Transaction rates per second for SYS1 and SYS2 for execution A1 at 84 transactions/second.

The group buffer pool access rates for A1 is consistent with A3. There is approximately 1 access to the GBP2 for each query transaction. There are approximately 4 accesses to GBP1 per secondary Update transaction and approximately 35-40% of the time a query will access GBP6 or GBP7.

A1		#REQ/SEC	AVG SERVICE (MICROSEC)	STD_DEV	READ HIT %
SYS1	DB2D GBP1	171.3	228.3	83.4	55%
SYS2	DB2D GBP1	151.4	227.7	79.1	
SYS1	DB2D GBP2	79.8	254.1	77	68%
SYS2	DB2D GBP2	79.1	252.2	71.9	
SYS1	DB2T GBP1	22.9	273.5	175	42%
SYS2	DB2T GBP1	22.5	268.7	160	
SYS1	DB2T GBP2	41.92	235.4	111.2	37%
SYS2	DB2T GBP2	41.27	228.8	94.3	
SYS1	DB2T GBP6	12.4	296.3	131.7	31%
SYS2	DB2T GBP6	16.07	268.6	140.7	
SYS1	DB2T GBP7	14.64	313.6	215.2	29%
SYS2	DB2T GBP7	14.53	304.8	206	

Table14. Group Buffer Pool Statistics for A1

Summary

The benchmarks which were executed to evaluate parallel sysplex and DB2 data sharing for a high availability, response time sensitive CICS/DB2 application confirmed that the application and processor architecture is scalable and could satisfy all business requirements for availability and service. We showed that the cost of DB2 data sharing are indeed highly dependent upon the workload requirements. The cost is dependent upon the access rates to the coupling facility for locking, caching and local buffer pool invalidation which is a function of the degree to which the workload requires access to shared data. Inquiry-only tables will have no group buffer pool dependence nor any need for page validation checking and therefore no accesses to Coupling Facility will be required. Additionally, if there is Read/Write Request from one DB2, but no Write Interest from any other DB2 in the sysplex, Coupling Facility access will not be required. The case study illustrated that the only workloads which experienced an increase in CPU requirements due to DB2 data sharing were the update transactions. The case study also demonstrated that the inquiry workloads experienced minimal impact since the update transactions that required access to same the tables as the inquiry work were very limited in volume and the secondary update transactions did not access those same tables.

In summary, the new CMOS technology exploiting parallel sysplex, and DB2 data sharing offers an excellent scalable solution for response time critical CICS/DB2 applications with better price performance than bipolar technology.

References

1. System/390 MVS Parallel Sysplex Performance, IBM SG24-4356-01
2. MVS/ESA SP Version 5, Sysplex Migration Guide Version 5.1.0, Version 5.2.0, IBM SG24-451-00
3. DB2 for MVS/ESA Version 4, Data Sharing Performance Topics, IBM SG24-4611-00