White Paper
September 2009

Expecting the unexpected: How to manage high peak workloads and maintain your service level agreements

By Paul Johnson, CICS System Management

Contents

2 *Executive summary*
2 *Introduction*
3 *Workload management*
6 *Establishing a dynamic workload management environment*
9 *Operational characteristics*
10 *CICSPlex System Manager sysplex-optimized workload management*
15 *Summary*
15 *For more information*

Executive summary

Establishing IBM CICS® environments that can cope with unexpected fluctuations in workloads might seem to be a difficult task. However, such an environment can be achieved by employing the dynamic workload management capabilities of IBM CICSPlex® System Manager and automation products such as IBM Tivoli® NetView®.

This paper concentrates on the use of the dynamic workload management and operational capabilities of CICSPlex System Manager, along with automation products for implementing systems that can provide highly available applications capable of coping with both predictable and unpredictable demand.

Introduction

When CICS was originally introduced, transaction processing needs were significantly different than they are today. Previously, these needs were addressed by a single CICS system on a single CPU, started cold each morning and shut down each evening so that the CPU could run overnight batch. At that time, networks were in their infancy, consisting of hundreds of terminals connected by IBM System Network Architecture (SNA). Applications were simple BMS map set applications running back-office workloads.

As the evolution of CICS progressed and the demands of business increased, the limitations of the single address space began to be reached due to increased numbers of terminals; exhaustion of dynamic storage areas (DSAs); increased demands for access to VSAM, IBM IMS™, and IBM DB2® data; and increasing sophistication as applications no longer resided only in CICS but also had components in IBM WebSphere® Application Server and IBM WebSphere MQ. The hardware changed as well, providing the ability to dynamically dispatch work over multiple processors.

Highlights

A highly diverse set of workloads exploit CICS, ranging from traditional applications to Web-facing workloads, Web services, and the latest Atom capabilities in CICS Transaction Server for z/OS V4.1.

Demands on the workload also changed with 24x7 operations and strict service level agreements (SLAs) requiring highly available, customer-facing applications. The parallel sysplex and the CICSPlex as we know it had been born. Efficiently managing and dynamically exploiting multiple processors and the many address spaces that resulted from this change gave rise to new technologies such as CICSPlex System Manager single system image management and dynamic workload management capabilities.

Today, a highly diverse set of workloads exploit CICS, ranging from traditional applications to Web-facing workloads, Web services, and the latest Atom capabilities in CICS Transaction Server for z/OS® V4.1. CICS provides all the capabilities to unlock your existing data and applications using service-oriented architecture (SOA). Event-based architecture can be exploited to further unlock existing assets. Multiprocessors can be leveraged through the exploitation of open transaction environment (OTE). Connectivity with TCP/IP becomes closer as more CICS transports are enabled.

Customer-facing applications across the Internet commonly demand 24x7 availability, and customer expectations mean that businesses must be constantly connected to ensure customer retention. This paper concentrates on the use of the latest dynamic workload management and operational capabilities of CICSPlex System Manager, along with automation products for implementing systems that provide highly available applications, capable of coping with both predictable and unpredictable demand.

Workload management

The term "workload management" is used in many ways—to refer to network balancing, IBM zSeries® System Resource Manager, and IBM Workload Manager for z/OS and CICS Transaction Server.

Network balancing

Requests across a TCP/IP or SNA network are directed to CICS residing on IBM zSeries. Requests across the various boxes in the network are balanced and dynamically routed to optimize traffic in the network.

The request then arrives at the sysplex boundary where the session traffic is balanced using capabilities such as z/OS Sysplex Distributor, virtual IP address (VIPA), DNS, port sharing for TCP/IP, and IBM VTAM® generic resource sharing for SNA. These technologies work in cooperation with IBM Workload Manager for z/OS and balance sessions with the listener layer of CICS systems in the sysplex.

IBM zSeries System Resource Manager

zSeries System Resource Manager dynamically manages processor storage, I/O priority, and CPU cycles for address spaces running on z/OS based upon goal-based policy. This policy is specified in terms of an active service policy, which defines service classes by describing the performance objectives of part of the workload.

Goals can be defined by:

- *Response time – typically transaction response time–including average response time and percentile response time.*
- *Velocity – how fast work should be run, typically used for address space startup (for example, CICS initialization).*
- *Discretionary – work with no goals.*

Goals are associated with workloads in various subsystems through classification rules.

Highlights

The CICSPlex System Manager component of CICS Transaction Server for z/OS provides administration and runtime capabilities to dynamically distribute workload requests

Workload Manager for z/OS and CICS Transaction Server

CICS initializes under a z/OS velocity goal. When active, it switches to z/OS Workload Manager performance block mode and a performance block is then allocated to each active task. CICS interacts with Workload Manager for z/OS to inform it of transaction attach, dispatch, and ultimately task termination.

CICS also provides exit points for identifying the system on which to execute a given workload request for various types of workload (for example, transaction routing, dynamic starts, and program links). These exits are typically exploited in the listening (router) layer. Exit points are also provided to reject workload requests and for asynchronous requests (such as STARTs) in the regions that receive the workload to execute (target regions).

Among many other management capabilities, the CICSPlex System Manager component of CICS Transaction Server for z/OS provides administration and runtime capabilities to dynamically distribute workload requests, utilizing these exit points. These capabilities fall into three main areas:

- *Workload balancing – Workload balancing consists of choosing, from a set of candidate regions, the best region to process this given request based on a balancing algorithm (queue or goal).*
- *Workload separation – Workload separation–identifying different sets of candidate regions for a given request, based on administration policy– is typically used in associating a set of candidate regions with a geographical location or an application or set of applications.*
- *Affinity management – Affinity management ensures that affinity rules are not violated in dynamic routing environments. Identifying affinities can be achieved through IBM CICS Interdependency Analyzer. When defined to CICSPlex System Manager, CICS Interdependency Analyzer will ensure that affinity rules are not violated.*

Various other factors such as target system health, type of connectivity between the router and target, abend history, and system events are taken into account when decisions are made about routing. In essence, a weight is calculated for each candidate region utilizing this data along with current load and the region. The region with the lowest weight is chosen (subject to affinities). CICS then routes the request to that region.

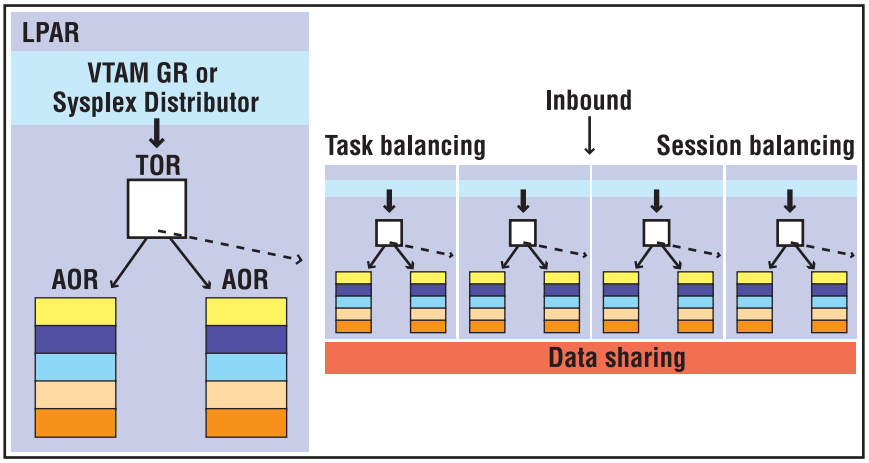
Two types of balancing algorithms are provided:

- *Queue, which takes into account the above factors to decide the appropriate region to route to. This algorithm optimizes throughput.*
- *Goal, which takes the same factors into consideration, but also takes into account the response time goal objective specified in the zSeries System Resource Manager.*

More information about routing can be found at the CICS Information Center[1] and in Xephon *CICS Update*.[2]

Establishing a dynamic workload management environment

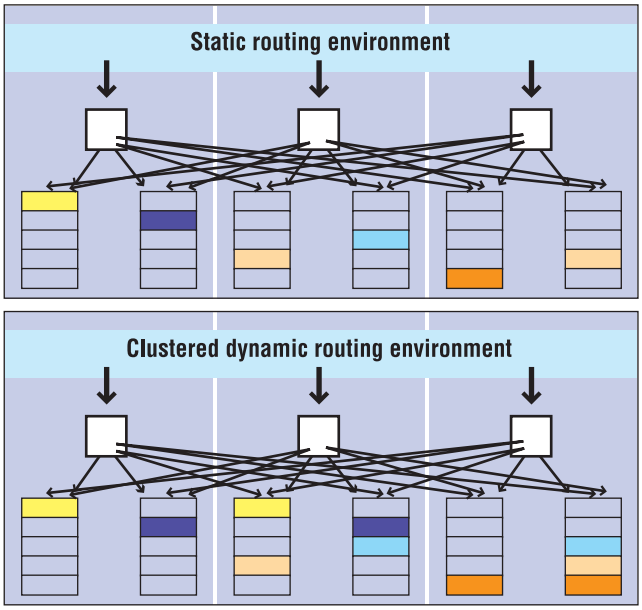
Figure 1 illustrates the classic sysplex heterogeneous setup. Sessions are balanced across the available set of listener regions on each logical partition (LPAR) through the appropriate technology for SNA or TCP/IP. Each listener region can accept any request and can route those requests to any CICS application-owning region (AOR) in the sysplex. These AORs can run any of the available applications (represented by colored bands). Data is accessed using appropriate data sharing technology, such as VSAM record level sharing (RLS) or DB2 data sharing.



The classic sysplex model

This type of configuration eliminates single point of failure at the address space and LPAR level, dynamically redistributing requests to balance the workload across the set of available AORs. While the availability of an individual system might not be 100 percent, this configuration gives the impression of 100 percent application availability and can cope with unforeseen demands on capacity, maximizing the exploitation of a multiprocessor configuration with high communication bandwidth.

Figure 2 shows a more realistic environment. Applications were originally statically routed to a given AOR (application partitioning). As the application availability or resource consumption demands dictated, these applications were analyzed, the AORs were cloned, and dynamic routing was employed.



A more realistic sysplex model

The general steps for moving into this environment are:

- *Select an application to enable.*
- *If this application is not already statically routed to an AOR, create an AOR for this application and statically route to the AOR. At this point, any problems with disassociation with the terminal-owning region (TOR) will be uncovered.*
- *Clone the AOR and dynamically route to the set of AORs. (Placement of the AOR depends on availability requirements.) You now have some balancing and failover ability at the AOR level.*
- *Clone the listener region to give you failover at this layer and enhanced session balancing from the communications layer.*

Highlights

Automation products such as Tivoli NetView allow you to prepare for planned and unplanned outages and cope with universal or reduced capacity demands by leveraging base and integration capabilities.

By leveraging this sysplex environment, you can split out given applications with little impact on existing applications.

Operational characteristics

In reality, an environment is not static. Automation products such as Tivoli NetView allow you to prepare for planned and unplanned outages and cope with universal or reduced capacity demands by leveraging base and integration capabilities.

Operational switchover from LPAR1 to LPAR2 can be achieved at the session level by switching routing tables in the communications layer. Existing sessions are bound until closure to LPAR1, while new sessions are bound to LPAR2. This technique can be used for switching over to a different physical box, because LPAR1 might be required for other processing overnight. LPAR1 might also be used as a regular switch to a set of disaster failure systems to ensure that a switch could indeed occur in the event of a catastrophic failure.

Application or region maintenance can be achieved by using CICSPlex System Manager Workload Manager "quiesce and activate" capability to remove the region from the candidate list. Existing threads then run to completion and new threads are distributed elsewhere. When quiesced, maintenance can be applied without the end user ever seeing an unavailable application. The region can then be activated back into the workload, and the change rippled across the AORs in the same manner.

Highlights

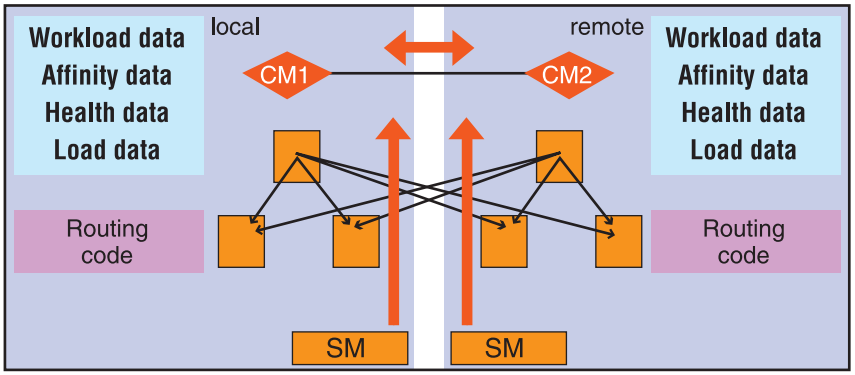
Even though dynamic workload management can balance work across regions, ultimately all systems will be filled to capacity. To accommodate peak loads, a common practice is to over-configure the workload manager. For example, your candidate target regions could be AOR1-10, with only AOR1-5 being employed normally. When the CICSPlex System Manager Real Time Analysis (RTA) component detects that AOR1-5 can no longer cope, AOR6-10 can be employed. This mode can be in either a hot or cold standby. Hot standby minimizes reaction time and is defined as the state when the AORs are initialized but quiesced. Activation is achieved simply by activating the region. Cold standby is achieved by starting the AOR. In this case, only the active systems are consuming resources.

Activation and starting a region can be achieved with CICSPlex System Manager API programs running in an automation product. A similar mechanism can be employed for "quiesce and shutdown" when the additional AORs are no longer needed. Other schemes employ Tivoli NetView to ensure that a minimum number of AORs are available on a given LPAR. Many schemes can be implemented with CICSPlex System Manager APIs, perfectly fitting the solution to the customer's needs.

CICSPlex System Manager provides management facilities—including workload management capabilities—that are not restricted by the sysplex boundary.

CICSPlex System Manager sysplex-optimized workload management

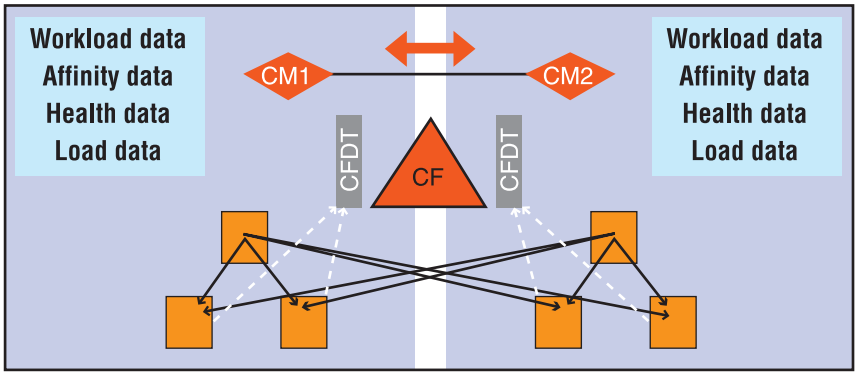
CICSPlex System Manager provides management facilities that are not restricted by the sysplex boundary. The same is true for its workload management capabilities. Some aspects of the classic CICSPlex System Manager solution are illustrated in Figure 3. CICSPlex System Manager management code runs in CICS address spaces, referred to as CICS Managing Address Spaces (CMAS) and illustrated as CM1 and CM2. The CMASs communicate together to provide a Single System Image (SSI) for all tasks supported by CICSPlex System Manager. Management agents reside in the CICS regions running the application workload. CICSPlex System Manager routing code also resides in the routing regions accessing data maintained by the CMAS in data spaces. Each component of CICSPlex System Manager has its own data space.



Workload management capabilities of CICSPlex System Manager

Workload data pertaining to routing policy, active affinities, system health data, and load data is among the data maintained in the workload manager data space. Data about target regions is collected by agents and transmitted among the CMASs so that agents in the routers can reference this information when making a routing decision. Information about targets on other LPARs is updated by CMAS-to-CMAS communication. The time to communicate this information introduces latency into the process, which in some types of routing (particularly asynchronous routing requests such as STARTs) can reduce the efficiency of the overall workload management solution.

Although this mechanism has proven itself over many years of customer use, the introduction of ever-faster processors and the wider adoption of sysplex coupling facilities by customers has enabled a more efficient mechanism to be employed for managing state data in a sysplex environment. This new facility in CICS Transaction Server for z/OS V4.1 provides sysplex-optimized workload management, outlined in Figure 4.

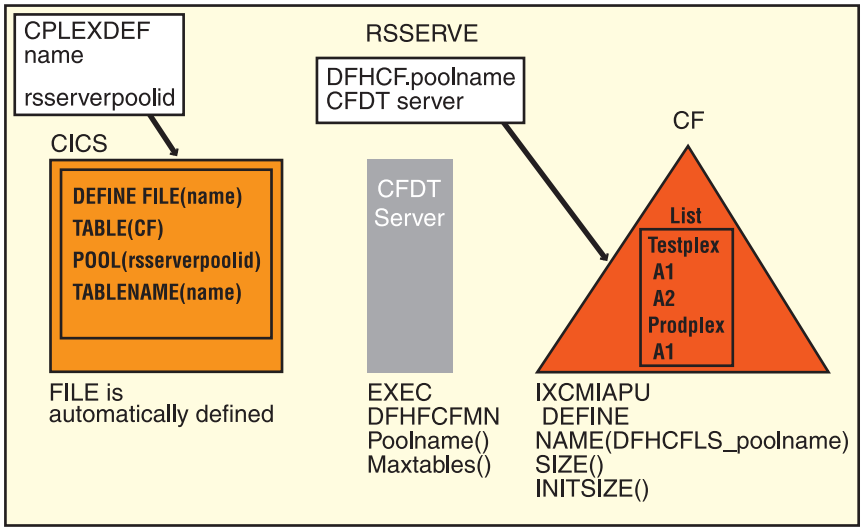


Sysplex-optimized workload management capabilities of CICSPlex System Manager

The solution has several key features:

- *Leverages a coupling facility data table (CFDT) server for maintaining load and state data. This CFDT server can be either existing or dedicated. This server is defined and managed in a standard fashion, as shown in Figure 5.*
- *Records state data by a new CICS domain. RS domain in target regions records data directly into the corresponding record in the CFDT.*
- *Routes regions reference data cached in the workload manager data space from the CFDT records. Updating from the CFDT server is based upon an aging algorithm.*
- *Controls frequency of access to the CFDT by introducing banding schemes and upper and lower bounds when the region is at low utilization and close to maxtask.*

Highlights

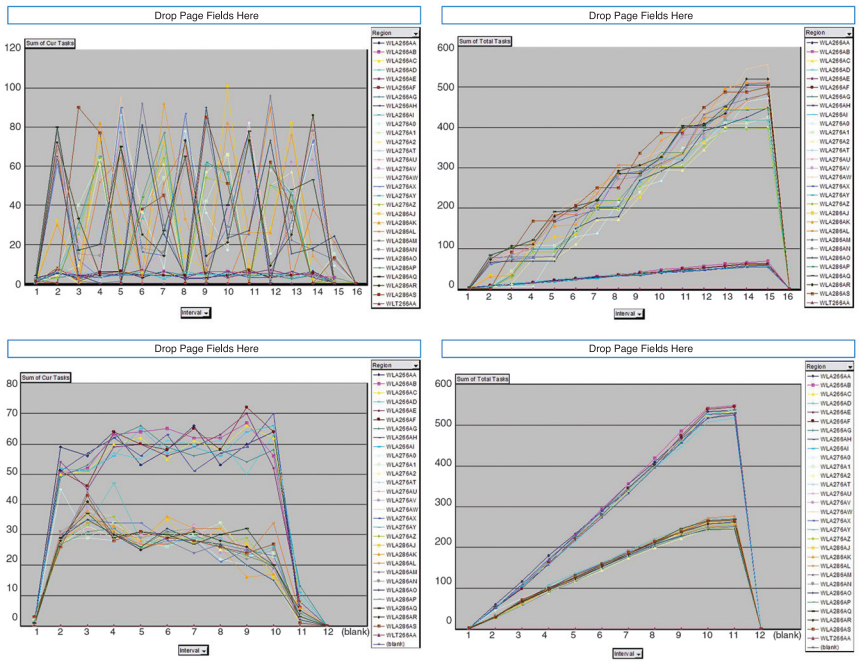


CICSPlex System Manager workload management and the coupling facility

All of this activity is customizable, and it coexists with existing workload management schemes for CICSPlex System Manager. Furthermore, if the coupling facility (CF) becomes unavailable for any reason, the CICSPlex System Manager workload manager will seamlessly fall back to its classic mode until the CF availability is reestablished. The user controls whether or not this new scheme is employed. The amount of CF storage used is minimal; each target region occupies approximately 40 bytes of storage.

Initial testing for distributed START requests has shown a more balanced distribution of workload on the newest processors, with a reduced overall execution time for the same workload.

While specific tests in a controlled lab environment should not be extrapolated to a customer's constantly varying workload, initial testing for distributed START requests has shown a more balanced distribution of workload on the newest processors, with a reduced overall execution time for the same workload, as shown in Figure 6.



Sysplex-optimized workloads enabled by CICSPlex System Manager

As well as introducing this new CICSPlex System Manager workload management capability, improved information has also been provided by introducing dynamic routing statistics and selection factor data to better understand the execution of your dynamic routing environment.

- *Dynamic routing statistics provide information about the total number of routing requests received by request type, such as route selects, terminates, and abends.*
- *Selection factor data provides you with a snapshot of the various factors that are used as input to the routing decision.*

All of this data is available online using the Web user interface.

Highlights

The sophisticated workload management capabilities provided by CICS Transaction Server for z/OS can optimize throughput and enable increased availability.

Summary

The sophisticated workload management capabilities provided by CICS Transaction Server for z/OS, in combination with automation products, can provide systems that enable increased availability and optimize throughput to the desired criteria. Application workloads can be managed without application change, minimizing the time to exploitation of these facilities. With the latest release of CICS Transaction Server for z/OS, sysplex-optimized workload management facilities further facilitate throughput and smoother workload distribution, enabling you to successfully establish systems that can cope with changing needs.

For more information

To learn more about how IBM can help your organization manage high peak workloads, or to upgrade to IBM CICS Transaction Server for z/OS V4.1, please contact your IBM marketing representative or IBM Business Partner, or visit: **ibm.com**/cics

IBM Corporation
IBM Systems and Technology Group
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
September 2009

[1] "IBM CICS Information Center. https://publib.boulder.ibm.com/infocenter/cicsts/v4r1/index.jsp?topic=/com.ibm.cics.ts.sampleplugin.doc/overview.html

[2] *CICS Update*, Xephon Inc., Issues 204-208, 223. www.xephonusa.com

Recyclable, please recycle

ZSW03131-USEN-00