1/1999 - Enclaves: Managing business transactions

Enclaves are a new method of managing business transactions that gives you more effective control over non-traditional workloads. Although it is not necessary to understand the implementation of enclaves, it is valuable to understand why they were invented and which products are using them. If you use one of these products, you need to manage enclaves as part of your performance policy. To get the most out of your S/390 investment, you should know what enclaves are and how to use them effectively.

Introduction
OS/390 Workload Manager (WLM) uses a goal-oriented approach to managing workloads running in the sysplex. You define performance goals - response times, for example - for each type of work. Such goals apply to interactive workloads whose transactions, as reported to OS/390, closely match the end-user view of transactions.

Enclaves provide a way to account for the resources used by a business transaction, even when those resources are consumed by multiple work units, even across multiple address spaces. Since goals are known and resource consumption can be monitored, WLM can control the transaction directly rather than indirectly through the associated address spaces. Understanding how WLM manages enclaves will help you interpret feedback from RMF, both in the Monitor I Workload Activity report and the new *RMF Monitor III Enclave report* introduced in OS/390 Version 2 Release 7.

What is an enclave?
An enclave is a construct introduced in MVS/ESA V5R2 and extended in OS/390 V1R3 *that represents a business transaction, or a unit of work*. An enclave is an anchor for accumulating the resources consumed by the transaction, regardless of where it is executing. The intent is to have the enclave map closely to the end user's view of a transaction. The transaction represented by an enclave consists of pieces that can span many server address spaces and simultaneously share these address spaces with other transactions.

One of the most novel aspects of enclave transactions is that OS/390 can manage them independently, even though several may be executing concurrently in the same address space.

Why are enclaves needed?
Before enclaves existed, WLM could manage transactions only by adjusting the resources of the transactions' server address spaces. Even if interactive transactions (such as CICS or IMS transactions) are known to WLM, resources are distributed indirectly by allocating resources to the address spaces serving the transactions. For example, each address space is assigned one dispatching priority and one I/O priority for all work running there. These priorities are determined by WLM, based on the work with the most stringent goals.

This method of transaction management works well to the extent that transactions with different goals or importance can be segregated into different address spaces. If transactions with multiple goals and importances are using the same server address spaces concurrently, WLM is limited in how precisely it can allocate resources to the work that needs them.

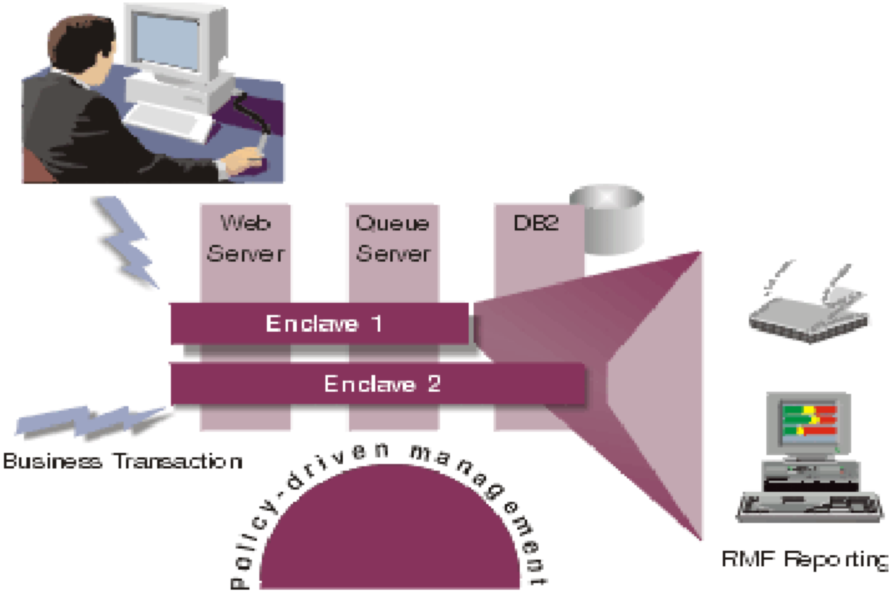
A more precise method of transaction management became necessary to handle the complex internal flows of some of the new OS/390 transaction managers. Examples of transaction managers that use enclaves are:

- DB2 V4 Distributed Data Facility (DDF)

- DB2 V5 Stored Procedures
- IBM HTTP Server for OS/390
- MQSeries Workflow

The individual enclaves active in an address space are assigned their own unique dispatching and I/O priorities, determined by the goals the customer has assigned to them. Thus more important enclave transactions executing in an address space will *not* miss their goals because of less important transactions running in the same address space.

Moreover, less important transactions will not get a "free ride" when they happen to be executing in the same address space as more important work. When important work is protected in this way, *it is possible to consolidate diverse types of work onto a single system without the critical work suffering*.



Managing and monitoring enclave transactions

If your OS/390 system is running one of the work managers listed above, then you can take advantage of enclaves. But how does WLM manage the enclaves? The answer is: You assign unique performance goals to enclave transactions, based on their attributes (such as user ID, transaction name, URL, and subsystem instance).

These attributes are communicated to WLM when an enclave transaction arrives in the system, and they provide great flexibility in assigning goals. By filtering on one or more of these attributes in the WLM policy, you can differentiate transactions critical to the business from those that are less urgent and can be deferred.

This differentiation is accomplished by mapping attributes to a WLM *service class*. Each service class has one or more periods, and each period has a goal associated with it. The goal reflects the business value of a transaction.

Multiple periods in a service class allow changing the goal over time (to favor shorter transactions over long-running transactions, for example). The RMF Monitor I Workload Activity report and the Monitor III Sysplex Summary report give you feedback on how well each service class period performs, compared to its goal.

If you need detailed reports of activity and resource consumption for subsets of transactions, you can define report classes in the WLM policy. In this way, you can have WLM collect granular data for multiple report classes while still managing the work with a single set of goals. The RMF Monitor I Workload Activity report includes resource data for both service classes and report classes.

Just as with address space-oriented transactions such as batch or TSO, enclave transactions consume processor and I/O resources. Since there are no limits to the complexity of an enclave transaction, it is mandatory to be able to monitor enclaves and their processor consumption online in order to support problem diagnosis and workload characterization. This becomes even more important as the number of subsystems exploiting enclaves grows and as use of these applications increases.

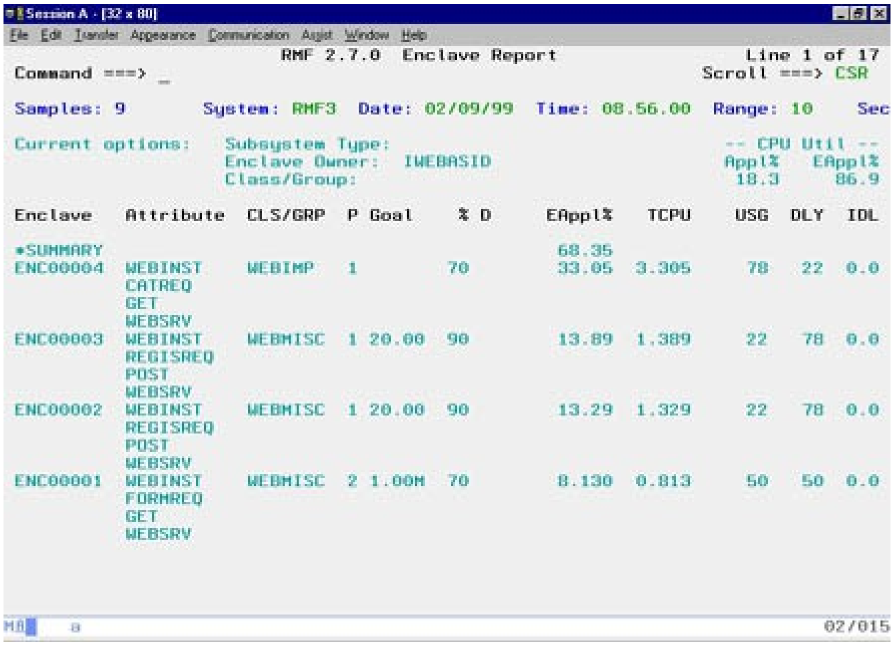
Most of today's performance monitors, however, can handle only address space transactions. *One of their major deficiencies is that they do not report on the resource consumption of individual enclaves.*

RMF Monitor III Enclave report

With the new RMF Monitor III Enclave report, you are able to monitor a single enclave transaction during its lifetime. The report shows you the CPU time consumed by an enclave during the sampled range and the total CPU consumption since the enclave was created. This allows you to identify business transactions that are high CPU consumers. It also enables you to distinguish transactions currently consuming significant CPU (possibly looping) from those whose lifetime CPU usage is high (candidates for performance analysis). Resource usage and delay percentages show you the progress of a single enclave and indicate possible resource constraints.

The report can be tailored to list only those enclaves which belong to the same subsystem, have the same business goals, or are owned by the same address space. Transaction attributes can be selected to further limit the transactions being analyzed. This allows you to drill down to a single user ID or a specific database request. In the following figure, for example, only those enclaves owned by address space *IWEBASID* are shown.

In addition to the new Enclave report, captured processor time reported by RMF Monitor III is more complete, since the RMF Processor Delay report now includes processor service consumed by enclaves. It is aggregated to and reported with the address spaces that created the enclaves.


Session A - [32 x 80]
File Edit Transfer Appearance Communication Assist Window Help
MA a
02/015

```
                              RMF 2.7.0  Enclave Report                  Line 1 of 17
Command ===> _                                                         Scroll ===> CSR

Samples: 9        System: RMF3  Date: 02/09/99  Time: 08.56.00  Range: 10     Sec

Current options:    Subsystem Type:                                  -- CPU Util --
                    Enclave Owner:   IWEBASID                        Appl%   EAppl%
                    Class/Group:                                      18.3     86.9

Enclave   Attribute  CLS/GRP  P Goal     % D      EAppl%   TCPU     USG  DLY  IDL

*SUMMARY                                          68.35
ENC00004  WEBINST    WEBIMP   1          70       33.05  3.305      78   22  0.0
          CATREQ
          GET
          WEBSRV
ENC00003  WEBINST    WEBMISC  1 20.00   90        13.89  1.389      22   78  0.0
          REGISREQ
          POST
          WEBSRV
ENC00002  WEBINST    WEBMISC  1 20.00   90        13.29  1.329      22   78  0.0
          REGISREQ
          POST
          WEBSRV
ENC00001  WEBINST    WEBMISC  2 1.00M   70        8.130  0.813      50   50  0.0
          FORMREQ
          GET
          WEBSRV
```

Summary

Enclaves are a flexible mechanism enabling you to manage business transactions. Several products on OS/390 already use enclaves, and you can expect more to use them in the future. The more precise transaction management possible with enclaves allows diverse workloads to be safely consolidated onto a single OS/390 system or sysplex. This allows you to better deploy and monitor your business applications and meet the objectives defined in your service level agreement.

Additional information

For additional information, visit the WLM home page.

You also can refer to these IBM publications:

- *MVS Planning: Workload Management*, GC28-1761
- *MVS Programming: Workload Management Services*, GC28-1773
- *RMF Report Analysis*, SC28-1950
- *DB2 for OS/390 Version 5 Administration Guide*, SC26-8957 (for classifying Distributed Data Facility transactions)