# Workload Management Overview

WSC Performance Team IBM Washington Systems Center Advanced Technical Support



#### Disclaimer

The information in this document has not been submitted to any formal IBM test and is distributed on an "as is" basis without any warranty expressed or implied. Use of this information or the implementation of any of these techniques is a user responsibility and depends on the users ability to evaluate and integrate them into the users operational environment. While each item may have been reviewed for accuracy in a specific situation there is no guarantee the same or similar results may be achieved elsewhere. Users attempting to adapt these techniques do so at their own risk.

#### Trademarks

The following names are trademarks of the IBM corporation and may be used throughout this presentation:

DB2, MVS/ESA, MVS/XA, MVS/SP, ES/9000, ES/3090

ESA/370, ESA/390, PR/SM, IBM

Large Systems Performance Reference, LSPR

**RMF**, Resource Measurement Facility

#### Agenda

- Overview
- Enclave Support
- Application Environments
- Setting Goals
  - Servers
  - System Workloads
  - Batch and TSO
- WLM Dynamic Workload Management
  - WLM Adjustments
  - I/O Priority Management
  - Session Placement
  - Batch Management



## WLM Objectives

- Simplifies definition, control, and reporting of the performance requirements for MVS workloads
- Provides goal oriented, dynamic resource mgmt
- Goals are based on business importance
- WLM uses these goals to dynamically adjust access to processor and storage resources
- Co-ordinate MVS images in a sysplex to process work towards a single set of goals
- Share goals with other products
- Provide programming interfaces so all products can allow customers to set goals for their transactions

### **Performance Management**



#### **GOAL:** System Responsibilities



#### Advantages to Single Image Customers

- Simpler externals
- Work is managed and reported in business terms
- Resources are consistently applied to important work
- Dynamic, on-going tuning and workload characteristics
- Subsystem awareness
- Classification improvements
- Software capping
- Reporting improvements

#### Advantages to Multi-System Customers

- All of the single image advantages, plus:
- Sysplex management scope
- Sysplex scope software capping
- Workload balancing capabilities
  - CICS
  - DB2 DRDA
  - IMS
  - TSO
  - Web Server
  - Batch

#### Requirements

#### Performance Administration

- Definition of work, (goals), and business importance for all work
- Provide methods to allow others access to desired results
- Done via WLM Application
- Performance Monitoring
  - Capture of results for each workload
  - Where did the work spend it's time
  - Information on what got in the way of success
  - Provide methods to allow others access to acheived results
  - Done via WLM Sampling
- Performance Reporting
  - Single place where desired results and achieved results are reported
  - Done via RMF

#### WLM Fundamentals



#### WLM Constructs

- Service Definition = IPS /ICS
- Policy
- Service Class = Performance Group / Domain
  - Period switched
  - Expectation of arrival rate, resource use, and response time
- Report Class = Reporting Performance Group
- Resource Group new functionality



#### **Service Policies**



#### **Service Classes**

	NAME	GOAL TYPE	VALUE	IMPORTANCE
SHORT BATCH	"SHORTBAT"	VELOCITY	40	LOW (5)
APPL DEV TSO	"ADTSO" PERIODS	AVERAGE RESPONSE TIME	1.2 SECONDS	MEDIUM (3)
PAYROLL (CICS)	"PAYROLL"	PERCENTILE RESPONSE TIME	90% IN LESS THAN 1 SECOND	HIGH (1)
	"LONGBAT"	DISCRETIONARY	N/A	N/A

## **Defining Service Classes**

- Named collection of work within a workload
- WLM manages a service class period as a <u>single entity</u> when allocating resources
- A service class is associated with only 1 workload
- Can define up to 100 service classes
- Additional granularity is provided by report classes
  - Can define up to 999 report classes, with a maximum of one report class per work request or transaction

## **Defining Service Classes**

#### Keep it simple and small

- Sampling is used to allocate resources so WLM is more efficient when more work is present in any given service class. SRM can make better decisions more quickly and hence be more responsive to changing system conditions
- -Goal: 15-30 service class periods
- Greater detail above service class level use report classes
- Do not combine:
  - Distributed DB2 work with other service classes
  - OpenEdition/MVS work with other service classes
  - Transaction and server address spaces in the same service class
- SRM sampling data, plots and projections become distorted when diverse types of work are combined
- Avoid having work classified to SYSOTHER
- Do not try to force certain resource allocations to occur

## Service Class Types

- Percentile Resposne Time
- Average Resposne Time
- Velocity
- Discretionary
- WLM Defined Goals
  - -SYSTEM
  - -SYSSTC
  - SMALL Consumer

### **Measuring Goals**

- Introduction of the Performance Index (PI)
- Used to compare goals of different types
- A calculated value to determine how well work is meeting it's goal
- Performance Index shows:
  - PI = 1.0 meeting the goal
  - PI > 1.0 missing the goal
  - PI < 1.0 beating the goal

### Percentile Response Time

- Percentile of work to be completed in the specified amount of time
- Best goal type able to handle variability in transaction response times
- Boundaries are from 1 to 99
- Response times can range from 15 ms to 24 hours
- WLM will not delay work, or limit it, to acheive the response time goal when extra processing time exists

### Average Response Times

- Heavily influenced by the extremes of transaction distribution
- WLM will manage work based on worst behaving transactions
- When to use:
  - Current SLA's are defined in terms of averages
  - As a starting point when transaction information is unavailable
  - Once average response time goal is established, RMF will show on Workload Activity Report distribution of response times around the average

## Velocity Goal

- Measure of acceptable delays while work is capable of running
  - Delays measured for CPU and storage, optionally for I/O and JES Queue time
- Used for long running and non-interactive workloads
- High velocity can be interpreted as:
  - "When this work is ready to run, make sure it runs without delay"
- Low velocity can be interpreted as:
  - "When this work is ready to run, keep it plodding along to ensure it will eventually finish"
- Velocity is not equal to dispatch priority

## **Velocity Calculations**



OS/390 V2.R4 Velocity Definition (WLM Inits):

CPU Using

CPU Using + CPU Delay + Storage Delay + JES Q Delay

## Velocity Goals

- Velocity doesn't look at idle periods, so velocity for this type of work can be highly variable over short time intervals
  - High execution velocity goal for A/S with high idle time can make SRM work harder
  - Set workloads to velocities of 50-70 even if compat mode shows in the 90s
  - Calculated based on state samples, low # of samples can cause SRM to manage a problem which doesn't exist

Example: 3 State Samples cause velocity to swing from 40 to 80

CPU Using: 3 CPU Delay: 1 Stor Delay: 1 Velocity: 3 / (3 + 1 + 1)= 60 CPU Using: 4 CPU Delay: 0 Stor Delay: 1 Velocity: 4 / (4 + 0 + 1) = 80

CPU Using: 2 CPU Delay: 2 Stor Delay: 1 Velocity: 2 / (2 + 2 + 1) = 40

## **Discretionary Goals**

- Runs when all work with non-discretionary goals are absent
  - OS/390 R2V6 will introduce changes to discretionary workloads. Work with goals which is over-acheiving is capped, freed CPU now can be used by discretionary work.
    PI of capped work kept between .7 and .81
  - Only work with a R/T goal > 1 minute, or velocity <= 30</p>
- Equivalent to work in lowest mean time to wait (MTTW) group in the IPS whose MPL fluctuates based on available capacity
- Only WLM goal type managed by SRM with the MTTW algorithm
- Work with this goal is a candidate for individual storage control via working set management

### System Goals

- Used to handle some work by default
  - SYSTEM Dispatch Priority of 255
  - -SYSSTC Dispatch Priority of 254
  - -TSO RT goal of 80% between 0 and 2 seconds
  - -SYSOTHER WLM Discretionary Goal
- This is also the default service policy
- Workloads classified in this manner are excluded from storage isolation control
- SRM does not have to spend time looking at state samples for these workloads and comparing them to a goal
- Can be directly specified via classification rules

## Setting Goal Importance

- Set by range of Very Important (1) to Desirable (5)
- Relative value, not absolute value, sets importance
- Significance of meeting goal says nothing about how easy or difficult the goal is to achieve, e.g. a batch job with velocity of 5 but importance 1
- Used by WLM to:
  - Identify critical workloads
  - Make tradeoffs to protect critical workloads
  - React to changing capacity. Scarce resources will cause WLM to degrade equally work with goals of lower importance

#### Work Classification



## **Classification Rules**

- Performance is search order sensitive
- Place highest probability of hit first
- Work, like STCs and TSO, which is started infrequently the quantity of rules is not as critical as for CICS / IMS trans
- Make wise use of Default Service Class
- Groups are allowed, use them
- Each transaction will be classified when in GOAL mode. In Compat mode, CICS/IMS transactions are not classified by SRM
- Document classification rules with the service policy

### **Migration Aids**



#### Migration Reporting - Compat Mode

Subsystem Type . : CICS Fold qualifier names? Y (Y or N) Description . . . CICS Rules - Compat Mode Rules

	ion codes: A B			A-ALCEL	C=Copy	M=MOVe	I=Insert rule		
			B=Before	D=Delete row	R=Repeat	IS=Insert Sub-rul $\epsilon$			
				Qualifie	r		Class		
Actio	on		Type	Name	Start		Service	Report	
						DEFAULTS:	RCICS4		
	_	1	SI	CICS	PROD				
	_	2	TN	AB	C*		RCICS1		
	_	2	TN	EF	G*		RCICS1		
	_	2	TN	CE	M*		RCICS4		
	_	2	TN	IJ	K*		RCICS2		
	_	3	τ	Л	WALSH*		RCICS1		
	_	1	SI	CICS	TEST				
	_	2	TN	AB	C*		RCICS3		
	_	2	TN	EF	G*		RCICS3		
	_	2	TN	IJ	K*		RCICS3		
	_	2	$\mathbf{TN}$	CE	M*		RCICS4		

IEAICSxx

SUBSYS = CICS, SRVCLASS=RCICS1,RPGN=1001 SRVCLASS=RCICS2,RPGN=1002 SRVCLASS=RCICS3,RPGN=1003 SRVCLASS=RCICS4,RPGN=1004 Use RMF workload activity report and review report PGNs

Page 31

## Migration Reportng - Goal Mode

Subsyst	cem	Type	. :	CICS		Fold qu	alifier nam	es? Y (	Y or N)	
Descrip	pti	on .	••	CICS R	ules	- Report	Classes On	ly		
Action codes		des:	A=Af	ter C=Copy		M=Move	I=Inser	I=Insert rule		
			B=Be	efore	D=De	lete row	R=Repeat	IS=Inse	rt Sub-rule	
Qualifier							-	Class		
Action		Type		Name	1	Start		Service	Report	
							DEFAULTS	:	RCICS4	
	1	SI		CICS	PROD					
	2	$\mathbf{TN}$		AE	C*				RCICS1	
	2	TN		EF	'G*				RCICS1	
	2	TN		CE	М*				RCICS4	
	2	TN		IJ	К*				RCICS2	
	3	τ	JI		WALSH	*			RCICS1	
	1	SI		CICS	TEST					
	2	TN		AE	C*				RCICS3	
	2	TN		EF	'G*				RCICS3	
	2	TN		IJ	К*				RCICS3	
	2	TN		CE	:м*				RCICS4	

Use RMF workload activity and specify WLMGL(RCLASS)

### **Resource Groups**

- Tradeoff resource definition may prevent achievement of stated goals
- Defined by service units
- Work in service classes across sysplex are used when calculating service units consumed
- To manage IMS or CICS, the regions must be in the resource group not the transactions
- Maximum Resource Groups
  - Resource group maximums can override goals
  - Work in the resource group will be delayed based on goals and importance
  - Can apply to OLTP regions, e.g. test
- Minimum Resource Groups
  - Allows distribution of CPU resource to specific workloads

## Administrative Support



#### **Operator Controls**

#### V WLM, POLICY=name



F D/MML,MODE={GOAL|COMPAT} RESET jobname,{QUIESCE|RESUME|SRVCLASS=name}

#### New OS/390 Facilities


#### **Enclave SRBs**

Enclave - A territory surrounded by another country's territory

- New dispatching enhancement, representing a "business unit of work"
- Managed separately from the address space, they are individual SRM transactions
- Can include multiple TCBs / SRBs
- Can span multiple address spaces
- All enclaves are classified using an active WLM policy in both compat and goal mode
- Provides value in distributed DB2 environment
  - Customer control over individual requests
  - New logons not shut down by SRB activity
  - Individual requests don't interfere with other work

#### **Enclave Attributes**

- Collection of 0 or more TCB / SRBs treated as an aggregate
  - Management
  - Reporting
- SRBs are the pre-emptible type
- More TCB-like than SRB-like
  - Calculate service units at the CPU rate, not SRB rate
  - CPU time will be reported in new fields, and aggregated into CPU time of the ASID which created the enclave
- Control provide in both Goal and Compat modes
  - Goal Mode DDF will call WLM with a classify and receive a service class, no rules match then default service class will be discretionary
  - Compat Mode ICS will allow classification for SUBSYS=DDF, PGN=xx definition. If no rules match then if PGN specified, use that, else use PGN of server address space

#### **Enclave Work**

Products using enclaves include:

- DB2 Sysplex Query Parallelism (DB2 V5 or later)
- DDF Distributed DB2 Transactions (DB2 V4 or later)
- IWEB Domino Go Webserver OS/390
- -LSFM Lan Server for MVS
- -SOM SOMObjects
- In goal mode these transactions default to SYSOTHER, probably NOT where you want them. So classify the work.
- All three goal types are appropriate
- Multiple periods can be used

# **Application Environment**

- Named group of similar server programs which execute in a server address space
- Transactions are scheduled in the server address space as MVS enclave SRBs
- Available in WLM goal and compat modes
- WLM will
  - Classify transactions
  - Maintain transaction queues
  - Start/stop server address spaces (goal mode only). Manual or automated process in compat mode
- Sysplex or System Scope
  - Determined by Subsystem

# **Application Environment Example**



#### **Transaction Flow**

- Transaction reaches work manager address space
- Work manager creates an enclave and sends the request to WLM
- WLM uses classification rules to assign a service class. Which rules to modify is work manager dependent
- WLM enqueues transaction to an internal WLM application environment queue
- WLM starts / stops server address spaces as needed (goal mode only)
- Server address space selects transactions from the queue

#### **Application Environment Queues**



# Application Environment Enabled Subsystems

IBM Subsystems

SUBSYSTEM TYPE	SCOPE	
CB (Component Broker)	System	
DB2	System	
IWEB	System	
MQ	System	
SOM	Sysplex	

# **Application Environment Definition**

Application-Environment	Notes Options Help	
Create Command ===>	an Application Environment	
Application Environment . Description Subsystem Type Procedure Name Start Parameters	<pre>. AE1 . Application Environment #1 . DDF Required . PROCS1</pre>	Required

Limit on starting server address spaces for a subsystem instance:

- 1 1. No limit
  - 2. Single address space per system
  - 3. Single address space per sysplex



# How to Choose a Goal Type

- Response time goals should be used whenever possible
- Response time goals are not appropriate for all types of work
  - Frequency of completion
    - Requires 10 completions in 20 minutes for response time goals to be effective
    - Workloads which have few completions include started tasks, CICS pre V4.1, IMS pre 5.1, DB2 regions, and long running batch
  - Variability of queue time
- Ability to specify a goal does not mean resources are available to meet the goal

# Short vs Long Response Time Goals

- Work is managed differently if goal is considered short or long
- Short Goal less than or equal to 20 seconds
  - Assumes work won't be around long
  - No individual storage access controls or policies
  - Newly arrived transactions are controlled by period wide storage access policies
- Long Goal greater than 20 seconds
  - Assumes work will be around a long time
  - WLM looks at all address spaces in the period to review storage access

# Recommendations for Goal Setting

- What you will need:
  - Copy of IPS/ICS for each system
  - SLA
  - Inventory of all performance management / capacity planning reports
  - Inventory of all billing applications
  - Historical data

#### Revisit Service Definitions Coefficients

- Specified in WLM policy not IPS
- Review changing MSO from default of 3.0 to 0.1
- Reset CPU and SRB coefficients to 1.0 and IOC coefficient to 0.5
  - Provides the same ratios as previously
  - Definition of service units per second for WLM resource groups will be in the same terms
  - Impact billing programs which probably need to be changed anyway because of WLM
  - Must remember to change any multi-period durations to reflect these changes

### **Review Service Level Agreements**

- Assess validity of SLA, are they valid or something done 5 years ago. Is management measured against SLAs?
- WLM Goals don't include network time, though often SLAs do
- SLAs may define worst case though actuals are much better. If so use caution when setting goals, "You may get what you asked for, even if you never got it before"
- Keep SLA definition in sync with WLM goals



#### **TSO** Workloads



For each transaction:

P1	R/T	90% in 1 sec	IMP 2
P2	R/T	5 secs	IMP 3
P3	VEL	20	IMP 4

# **TSO Service Classes and Goals**

- Probably the easiest work to set
- Should have a response time goal
- May have a velocity goal
  - Less efficient
  - SRM will control swap protect time on an individual A/S basis rather than period wide
  - TSO transactions may incur working set management though the work may end quickly
- Discretionary goals should be used cautiously, except for a very complex last period
- HOTTSO service class for systems programming is still possible by creating service class with more aggressive goals

#### **Batch Jobs**



For each transaction:

P1	R/T	90% in 1 sec	IMP 3
P2	VEL	20	IMP 4

# Batch Service Classes and Goals

- Use response times, velocity, or discretionary goals
- Response Time Goals
  - Minimal JES Queue Time
  - Either average or percentile goal
- Velocity Goals
  - Appropriate for long running batch
  - Appropriate for CICS and IMS when started as batch job
- Discretionary Goals
  - Critical path production jobs should use velocity, unless CPU is plentiful
- Multiple periods are still possible
  - Later periods have less aggressive goals, lower importance, or both
- SWAP OUT service class not needed.
  - Use MODIFY WLM RESET, QUIESCE | RESUME

#### **CICS Servers and Trans**



### **IMS Server and Trans**



and JES Rules for the BMPs

# Server Address Spaces and Goals

- Assign a velocity goal
- Classify the regions to service classes different from the transactions they serve
- Even if WLM is managing transactions still need velocity goal for start-up / shut-down, idle periods, and restarts
- No classification rules for CICS / IMS work then transactions treated at servers goal, not SYSOTHER (discretionary)
- Create a report class for each region for reporting
- To cap CICS or IMS work, the regions must be in a resource group, not the transactions

# **OLTP** Transactions and Goals

- Only a response time goal can be assigned
- Must be a single period defined for the service class
- Define only a few service classes for CICS work, don't define by region.
- CICS dispatching doesn't use WLM goals so fine granularity is not critical. WLM controls only the address space
- Group transactions in service classes which have similar characteristics
  - Unlike types of work will impact response time data
  - May wish to put mission critical work in a separate service class

#### **Service Classes**



### Migration Considerations for OLTP Work

- Migration support to get information about the performance characteristics of OLTP work
- Can be done while in compat more
  - Define a couple data set, install and activate a WLM policy with classification rules
  - Define service classes with goals (ignored)
  - Use SRVCLASS, and RPGNs in ICS for appropriate subsystem
- RMF can report on number of transactions and avg response time by RPGN. No action taken on goals.
- Can also be done in goal mode. Classification rules specify only report classes, no service classes. OLTP will be viewed as having no rules, and ASIDs will run with velocity goal of the region.

## Started Tasks and System Goals

- Allocate STCs into a small number of service classes (1-3)
  - STCHIGH Velocity
  - STCMON Less Velocity
  - STCMISC Even Less Velocity
- May want to use transaction name groups
- Recommend don't classify System spaces
- Review requirements for use of report classes. TNG and report classes are mutually exclusive

### Unix Services Work

- Treat the OMVS Kernel and OMVS processes as all other started tasks
  - OMVS Kernel should be grouped with high priority tasks
  - OMVS daemons should be grouped with medium priority tasks
- Use multi-period service classes for OMVS transactions
  - Classify under the OMVS subsystem type
  - Many / most OMVS transactions have long response times but use little resources
    - First period should have a short response time goal
    - Next period should have a low velocity goal

# APPC / MVS Work

- Treat the ASCH address space as any other started task
- For APPC transactions, use single period velocity goals
  - Classify under the ASCH subsystem type
  - Most APPC transactions keep a permanent, or lengthy, network connection and repeatedly process individual conversations across the network
- Note: If you are sure all APPC transaction programs process a single network request and then complete, you can use multiple periods with response time goals

# **DDF Server and Trans**



## DDF Work

#### Transactions are subject to period switching

- All goal types are allowed
- THREADS=INACTIVE and RELEASE(COMMIT)
  - DDFcreates one enclave per active interval
  - Response times do not include think time
  - Response time goals and multiple periods can be used
- THREADS=ACTIVE or RELEASE(DEALLOCATE)
  - DDF creates one enclave for the life of the thread
  - Enclave response times include think times
  - Response time goals and multiple periods should not be used

# **Query Parallelism**

- Query Parallelism
  - I/O parallelism (DB2 V3)
  - CP parallelism (DB2 V4)
  - Sysplex query parallelism (Parallel Sysplex, DB2 V5.1, DB2 Data Sharing)
- Classification of originator doesn't change
  - Each local split query starts in the current period of the originator
- Classification of participants is under "DB2" rules
  - Classification attributes inherited from originator
  - If no rules supllied query runs in SYSOTHER
  - Each remote split query starts in 1st period
- Accounting is done on the system where the enclave runs

### **DB2 Stored Procedure Example**



#### **DB2 Stored Procedures**

- Introduced in DB2 V4.1
- Execute in the dsnxSPAS address space
- dsnxSPAS is a single address space implementation
  - Multiple TCBs
  - DSNTIPX=Max number of TCBs
- DB2 V5 and OS/390 V1.3
  - DB2 schedules enclave SRBs instead of using multiple TCBs
  - WLM provides support for Application Environments

#### Domino Go Webserver

- HTTP server for OS/390
- System level servers
- Standalone server
- Scalable server
- Multiple servers

#### Standalone Server

- Independent server address space
- Unique TCP/IP port
- WLM manages at the address space level
- Classify
  - SYSSTC
  - High Velocity, High Importance
## Scalable Server Subsystem

- Uses WLM application environment services
- Domino Go Webserver address spaces
  - Queue manager
  - Queue server(s)
- Request processing:
  - A match with the application environment directives causes use of WLM application environment services
  - Otherwise processed in queue manager address space

## Scalable Go Webserver



## **Goal Setting Summary**

- Re-evaluate goals over time
- Keep it simple
- Watch number of service classes
- Pay attention to enclave transactions (high growth)
- Keep work managed as address spaces (STC, batch, OMVS, etc.) in service classes separate from other work
- Use response time goals whenever possible
- Use resource groups only when there is a special need
- Avoid monitoring how WLM is setting dispatch priority
- Watch for work classified to SYSOTHER