EPV Performance University 2021

- z/OS Performance Analysis
- WLM Update
- Db2 Performance Analysis

15th June 2021

WLM basics - Part 1

Agenda

- Introduction
- Service Definition options
- Service Definition coefficients
- Service Class periods
- Service Class importance

Introduction

Introduction

- One of the z/OS major strengths is the ability to run multiple workloads at the same time
- The system function that makes this possible is dynamic workload management, which is implemented in the Workload Manager component (WLM) of the z/OS operating system
- WLM was initially a kind of intelligent interface between humans (users) and the z/OS operative system built on top of SRM
- WLM/SRM (WLM) rapidly evolved providing many new capabilities in addition (but correlated) to workload management

Introduction

- WLM functions:
  - Workload management
  - Workload balancing (not discussed in this course; no direct info in EPV)
  - 4HRA and soft capping
  - IRD support
  - HiperDispatch support
  - zIIP SMT support
  - Mobile workload support
  - Container pricing

Introduction

- Through the WLM definitions a kind of contract between the user and z/OS is established
- This contract uses understandable concepts and goals
- Basically the user:
  - classifies the work running on the z/OS operating system in distinct classes by assigning them meaningful names
  - defines for each class its importance for the business
  - defines a goal for each class expressing the expectation of how the work should perform
- WLM uses these definitions to manage the work across all the systems of a sysplex environment

Introduction

- Incoming work is assigned to WLM service and report classes by using user defined classification rules
- WLM manages the following unit of work:
  - Address Spaces
  - Enclaves
  - CICS and IMS transactions

Introduction

- Multiple sub classes called periods can be defined
- WLM managed resources are assigned to service classes, based on their importance and “happiness” in order to let them meet their goals
- A service class period is “happy” when it is achieving the goal
- Report classes are used for reporting purposes only

Introduction

- WLM managed resources:
  - CP priority
  - MPL, Working set
  - I/O priority
  - PAV aliases
  - Initiators
  - Server address spaces
  - DB2 Buffer Pool (autosize=yes)
  - CPU weight and Channels together with IRD
  - CPUs and zIIPs together with HiperDIspatch
  - zIIP threads with SMT

Introduction

- A WLM service definition needs to be defined and installed in the WLM couple dataset
- It has to contain all the information needed for workload management
- It also includes at least one base service policy
- There is one service definition for the entire sysplex

Introduction

- A service definition consists of:
  - one or more service policies, to eventually override part of the service definition
  - service classes
  - workloads, which aggregate a set of service classes for reporting purposes
  - report classes
  - classification rules
  - resource groups
  - tenant resource groups and tenant report classes
  - application environments, which are groups of application functions that execute in server address spaces and can be requested by a client; WLM manages the work according to the defined goal, and automatically starts and stops server address spaces as needed
  - scheduling environments, which are lists of resource names along with their required states; if a z/OS image satisfies all of the requirements in a scheduling environment, then units of work associated with that scheduling environment can be assigned to that z/OS image

Introduction

- In a service policy you can override the following settings of the service definition:
  - a goal for a service class period
  - number and duration of periods
  - a service class to resource group assignment
  - other properties of the service class such as CPU protection
  - resource group attributes
  - tenant resource group attributes

Introduction

- SMF records including the most relevant WLM related measurements:
  - SMF 72 subtype 3; for service and report class activity, resource usage, using and delay states, resource group capping, mobile work, container pricing, etc;
  - SMF 70 subtype 1; for 4HRA, HiperDispatch activity, SMT activity, container pricing
  - All details about SRM/WLM activity in 14 subtypes of SMF 99

Introduction

- Most of the information you normally need are in: *“MVS Planning: Workload Management”*
- If you want to become a WLM “guru”, this is for you: *“z/OS MVS Programming: Workload Management Services”*

Service Definition options

Service Definition options

```
  Coefficients/Options  Notes  Options  Help
------------------------------------------------------------------------------
                Service Coefficient/Service Definition Options

Enter or change the Service Coefficients:

CPU  . . . . . . . . . . . . . _______  (0.1-99.9)
IOC  . . . . . . . . . . . . . _______  (0.0-99.9)
MSO  . . . . . . . . . . . . . _______  (0.0000-99.9999)
SRB  . . . . . . . . . . . . . _______  (0.0-99.9)

Enter or change the service definition options:

I/O priority management  . . . . . . . . NO   (Yes or No)
Enable I/O priority groups . . . . . . . NO   (Yes or No)
Dynamic alias management  . . . . . . . NO   (Yes or No)
Deactivate Discretionary Goal Management NO   (Yes or No)
```

Service Definition options

I/O Priority Management

- Most customers run with I/O Priority Management enabled (IBM advice)
- It means that I/O priorities should be managed separately from dispatching priorities, according to the goals of the work providing performance benefits (I/O queuing reduction) especially if you don't have HyperPAV and SuperPAV
- It also means that I/O using and delay samples are included in velocity calculation; **more details in next sessions**
- When set to yes, it enables I/O Priority Groups (very few customers use it)
- It also to be set to yes if Dynamic Alias Management is enabled (not useful if using HyperPAV and SuperPAV)

Service Definition options

Is I/O Priority Management really needed?

- I/O requests associated with the system-provided service classes SYSTEM, SYSSTC, or SYSSTC1 - SYSSTC5 are not managed by the I/O priority manager
- It may happen that some service classes have unnaturally high velocities due to I/O Using
- In these cases, the work may suffer significant CPU delay before WLM decides to help the service class
- If you use modern disk subsystems, HyperPAV or SuperPAV you may consider the possibility to disable I/O Priority Management

Service Definition options

I/O Priority Groups:

- Most customers run with I/O Priority Management disabled
- It can be used to protect work which is extremely I/O-sensitive
- When you assign a service class to I/O priority group HIGH, you ensure that work managed by this service class always has a higher I/O priority than work managed by service classes assigned to I/O priority group NORMAL which is the default for service classes
- You also need to specify yes for I/O Priority Management to enable it

Service Definition options

Dynamic Alias Management

- Specifying YES will cause workload management to dynamically reassign parallel access volume aliases to help work meet its goals and to minimize IOS queueing
- Only useful when Dynamic PAV is used
- Not useful with HyperPAV or SuperPAV
- If I/O priority management is set to NO, WLM will make alias moves that minimize overall IOS queueing, but these moves will not take service class goals into consideration

Service Definition options

Deactivate Discretionary Goal Management

- Most customers run with Deactivate Discretionary Goal Management disabled (my advice ☺)
- Certain types of work, when overachieving their goals, can potentially have their general purpose processor resources “capped” in order to give discretionary work a better chance to run
- If you specify yes, you deactivate this kind of resource donation and workload management can not cap processor resources in order to help discretionary work
- **More details in next sessions**

Service Definition coefficients

Service Definition coefficients

- One of the basic functions of WLM/SRM is to monitor the rate at which a workload is receiving service
- This service is measured in service units
- Four service unit types are used:
  - CPU
  - SRB
  - I/O
  - Storage

Service Definition coefficients

- Service units (1/2):
  - CPU service units; TCB and preemptible SRB execution time, multiplied by an SRM constant; it includes the time used by the address space while executing in cross-memory mode; this time is not counted for the address space that is the target of the cross-memory reference
  - SRB service units; non preemptible SRB execution time for both local and global SRBs, multiplied by an SRM constant; it includes the time used by the address space while executing in cross-memory mode; this time is not counted for the address space that is the target of the cross-memory reference

Service Definition coefficients

- Service units (2/2):
  - I/O service units; measurement of data set I/O activity and JES spool reads and writes for all data sets associated with the address space; SRM calculates I/O service using I/O block (EXCP) counts; when an address space executes in cross-memory mode, the EXCP counts or the device connect time will be included in the I/O service units; these I/O service units are not counted for the address space that is the target of the cross-memory reference
  - Storage service units; calculated as (central page frames) x (CPU service units) x 1/50; the central storage page frames used by an address space while referencing the private virtual storage of another address space in cross-memory mode are not included in the calculation; these frames are counted for the address space whose virtual storage is being referenced

Service Definition coefficients

- Independent enclaves:
  - CPU service units are charged to the enclave
  - SRB service units are always 0 for the enclave
  - I/O service units are charged to the server address spaces (where the enclave runs) not to the enclave
  - Storage service units are always 0 for the enclave

Service Definition coefficients

- Service Definition Coefficients (CPU, SRB, IOC and MSO) are used to weight the different service units types and calculate the total service unit consumptions of a UOW

Weighted service units =

*(CPU * CPU service units) + (SRB * SRB service units) +*

*(IOC * I/O service units) + (MSO * storage service units)*

- Weighted service units will be used for service class period switching using the ‘duration’; **more details in next slides**

Service Definition coefficients

```
  Coefficients/Options  Notes  Options  Help
------------------------------------------------------------------------------
                Service Coefficient/Service Definition Options

Enter or change the Service Coefficients:

CPU  . . . . . . . . . . . . . _______  (0.1-99.9)
IOC  . . . . . . . . . . . . . _______  (0.0-99.9)
MSO  . . . . . . . . . . . . . _______  (0.0000-99.9999)
SRB  . . . . . . . . . . . . . _______  (0.0-99.9)

Enter or change the service definition options:

I/O priority management  . . . . . . . . NO   (Yes or No)
Enable I/O priority groups . . . . . . . NO   (Yes or No)
Dynamic alias management  . . . . . . . NO   (Yes or No)
Deactivate Discretionary Goal Management NO   (Yes or No)
```

Service Definition coefficients

- CPU and SRB coefficients should be equal; default is 10 for both of them
- Setting them to 1 is strongly adviced
- It will eliminate the distinction (and confusion) between weighted and not weighted service units
- These settings make WLM duration more understandable and performance analysis easier
- CPU and SRB are also used to weight zAAP and zIIP SU

Service Definition coefficients

- IOC optimal value depends on how strong is the I/O infrastructure; the stronger the lower IOC; default value is 5
- A value of 0,5 is adviced by IBM in the WLM manual
- Many customers use lower values such as 0,1

Service Definition coefficients

- Can you set IOC = 0 ?
- Limitations of I/O service units
  - Trivial transactions, such as most of the TSO transactions, are doing few or zero I/Os, so including or not I/O service units makes no difference
  - DDF service class periods (and other service classes based on independent enclaves) don’t accumulate I/O service units
- Workloads doing a lot of I/Os, such as batch jobs, give up the CPU regularly so their CPU impact to other workloads, because of a duration not considering I/O service units should be limited

Service Definition coefficients

- MSO optimal value is 0; default value is 0
- Higher values risk to make the CPU, SRB and I/O service units irrelevant

Exercise 1

Let's suppose an address space used 1 CPU second (about 100.000 CPU service units with modern machines), had a 10.000 frame footprint and MSO is set to 0,1. Calculate the number of storage service units.

Storage service units = (10.000) * (100.000) * 1/50 * 0,1 = 2.000.000

Service Definition coefficients

- Setting CPU and SRB coefficients to 1 and excluding I/O and Storage service units would make it easier to relate period durations to CPU consumptions
- IBM will probably eliminate the Service Definition Coefficient in the near future
- WLM will use only CPU and SRB internally fixed to 1

Service Class periods

Service Class periods

- A service class is a named group of work within a workload with similar performance characteristics:
  - Business importance to the installation
  - Performance goals
  - Resource requirements
- Some work may have variable resource needs so you can define multiple periods for a service class
- By using periods you can define different goals for work depending on the amount of used resources

Service Class periods

- Typically, periods are used to give shorter transactions more aggressive goals and to give longer running work, of the same type, less aggressive goals
- If you have multiple periods, each period except the last has a duration
- Duration is the amount of weighted SU (including zIIP) that a UOW (unit of work) may consume before it is switched to the goals of the next period

Service Class periods

- Example of a service class for batch jobs scheduled by TWS
- Two periods with the same goal
- Duration of first period is 200.000 weighted SU

```
 Service-Class  Xref  Notes  Options  Help
------------------------------------------------------------------------------
                          Modify a Service Class              Row 1 to 3 of 3
Command ===> _____________________________________________________________

Service Class Name . . . . . : BJOBTWS
Description  . . . . . . . . . Jobs scheduled by TWS
Workload Name  . . . . . . . . BJOB     (name or ?)
Base Resource Group  . . . . . ________  (name or ?)
Cpu Critical . . . . . . . . . NO       (YES or NO)
I/O Priority Group . . . . . . NORMAL   (NORMAL or HIGH)
Honor Priority . . . . . . . . DEFAULT  (DEFAULT or NO)

Specify BASE GOAL information. Action Codes: I=Insert new period,
E=Edit period, D=Delete period.

          -- Period --  ------------------- Goal -------------------
Action    #  Duration    Imp.  Description
  ___     _  ________    _     ____________________________________
  ___     1  200000      4     Execution velocity of 30
  ___     2  _______     5     Execution velocity of 30
******************************* Bottom of data ********************************
```

Service Class periods

- Exercise 2

Service Coefficients: CPU=1, SRB=1, MSO=0, IOC=0,5

SU rate: 40.000 – Period 1 duration = 200.000

Which of the following sentences are correct?

  - A job may read or write 400.000 blocks before moving to period 2
  - A job may use up to 5 CPU or zIIP seconds before moving to period 2
  - A job may use 2 CPU or zIIP seconds and read or write 200.000 blocks without moving to period 2

Service Class periods

- Exercise 2
  - A job may read or write 400.000 blocks before moving to period 2
  - **not correct because it's impossible not using any CPU when performing I/Os**
  - A job may use up to 5 CPU or zIIP seconds before moving to period 2
  - **theoretically correct but it's unlikely a job (unless in loop) will use CPU or zIIP only**
  - A job may use 2 CPU or zIIP seconds and read or write 200.000 blocks without moving to period 2
  - **correct**

Service Class periods

- What WLM manages are service class periods
- You may define up to 100 service classes
- A maximum of 8 periods can be defined for each service class
- Normally it doesn’t make sense having more than 3 periods
- Service classes for CICS and IMS transactions can’t be multi period

Service Class periods

- Every period has its own importance and goal

```
 Service-Class  Xref  Notes  Options  Help
------------------------------------------------------------------------------
                          Modify a Service Class              Row 1 to 3 of 3
Command ===> _____________________________________________________________

Service Class Name . . . . . : BJOBTWS
Description  . . . . . . . . . Jobs scheduled by TWS
Workload Name  . . . . . . . . BJOB      (name or ?)
Base Resource Group  . . . . . ________  (name or ?)
Cpu Critical . . . . . . . . . NO        (YES or NO)
I/O Priority Group . . . . . . NORMAL    (NORMAL or HIGH)
Honor Priority . . . . . . . . DEFAULT   (DEFAULT or NO)

Specify BASE GOAL information. Action Codes: I=Insert new period,
E=Edit period, D=Delete period.

        -- Period --  -------------------- Goal ------------------
Action  #  Duration    Imp.  Description
  __    _  _________   _    ____________________________________________
  __    1  200000      4    Execution velocity of 30
  __    2  _________   5    Execution velocity of 30
******************************* Bottom of data ********************************
```

Service Class importance

Service Class importance

- Importance reflects how important it is for your business to achieve the service class period goal
- Importance indicates the order in which work should receive resources when not achieving its goal
- As shown in the previous slide importance applies to service class periods and you can change importance from period to period
- Importance can be set by the user in five levels: 1 to 5; 1 being the highest importance

Service Class importance

- EPV conventionally shows 7 importance levels
  - 1-5 User
  - 0, 6 System
- 0 is for system critical Address Spaces (running in the system provided SYSTEM and SYSSTC service classes)
- 6 is for least important workloads:
  - with no assigned goal; they run in SYSOTHER service class with a Discretionary goal (if not in STC)
  - with an explicit Discretionary goal; least important workloads

Service Class importance

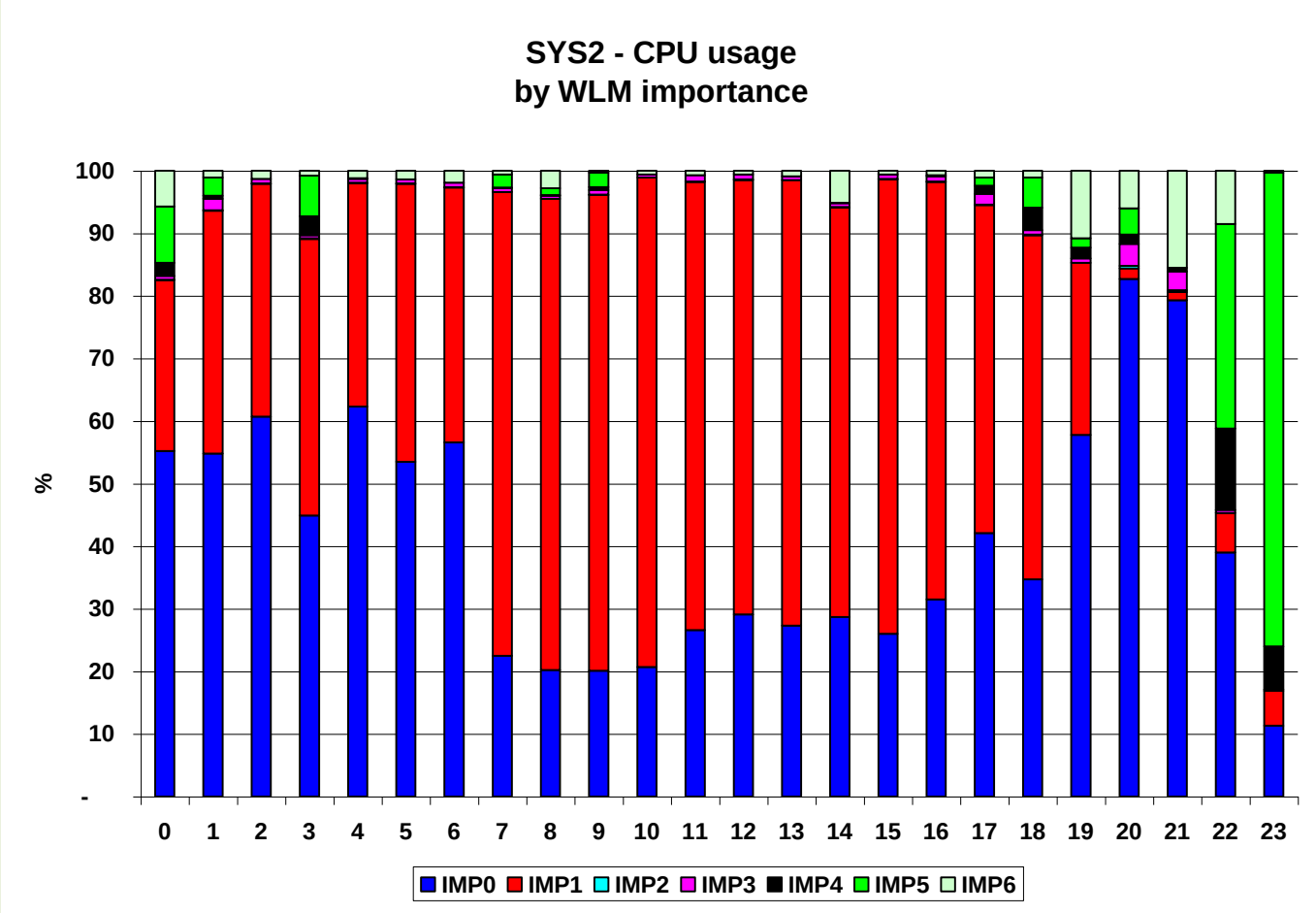
- Generally speaking, components providing services should be at higher importance than components using their services
- To set the right importance value to your service class periods you should ask yourself this question:
  - How much is important to the business that the goal will be achieved ?
- If you can't answer, you could try this other one:
  - If you had enough resources in the system for only one service class period to reach the goal, which service class period would you choose to serve ?

Service Class importance

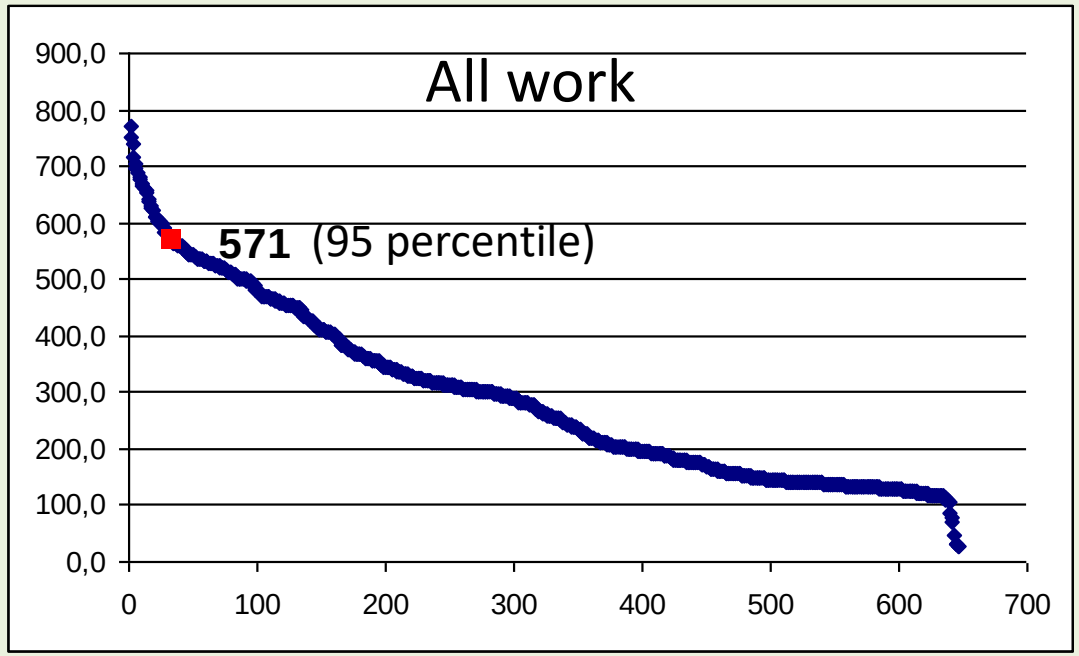
- You should try to use all the available user importance levels in order to differentiate work and make WLM job possible
- Remember than when in contention, workloads will be helped in importance order
- If everything is at the same importance the result is that nothing is important
- Plotting CPU usage by importance can help you understand if your settings are reasonable

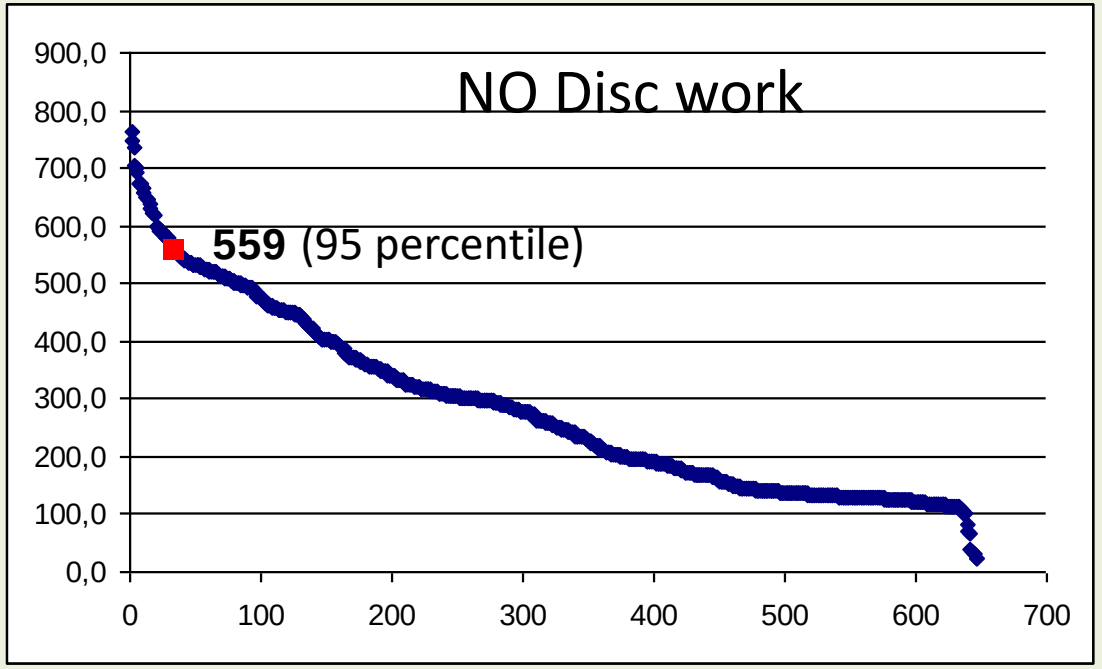
Service Class importance

- Sum of percentages is 100%
- Workload is using only importance 0 and 1
- Almost no work in discretionary
- WLM has a hard life with this workload



Service Class importance

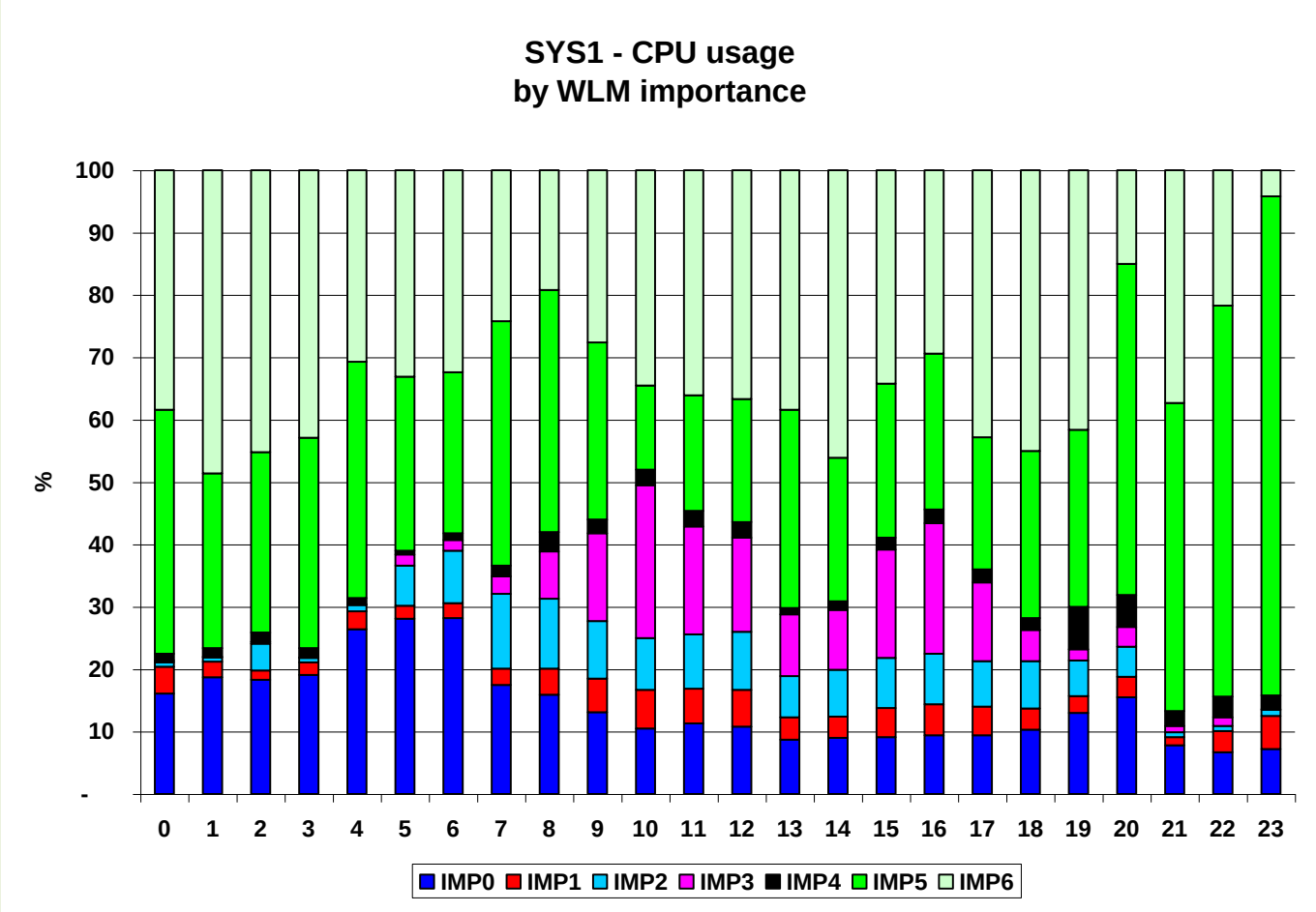




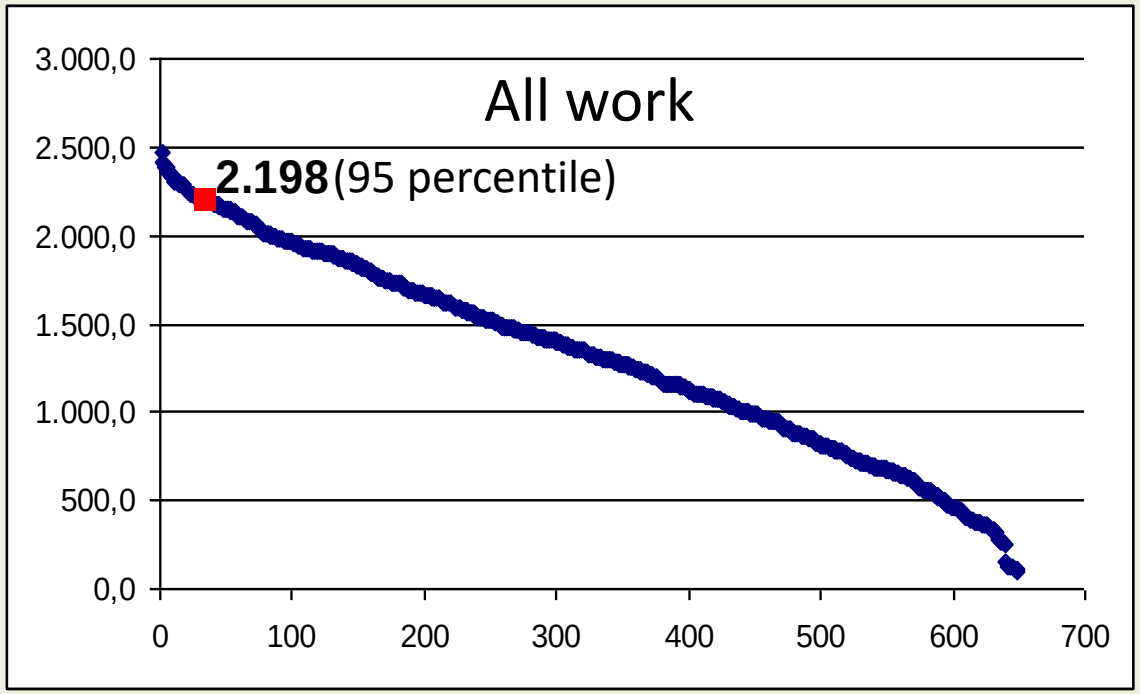
- MIPS used by hour in peak time in the last 60 days
- No space to delay an upgrade

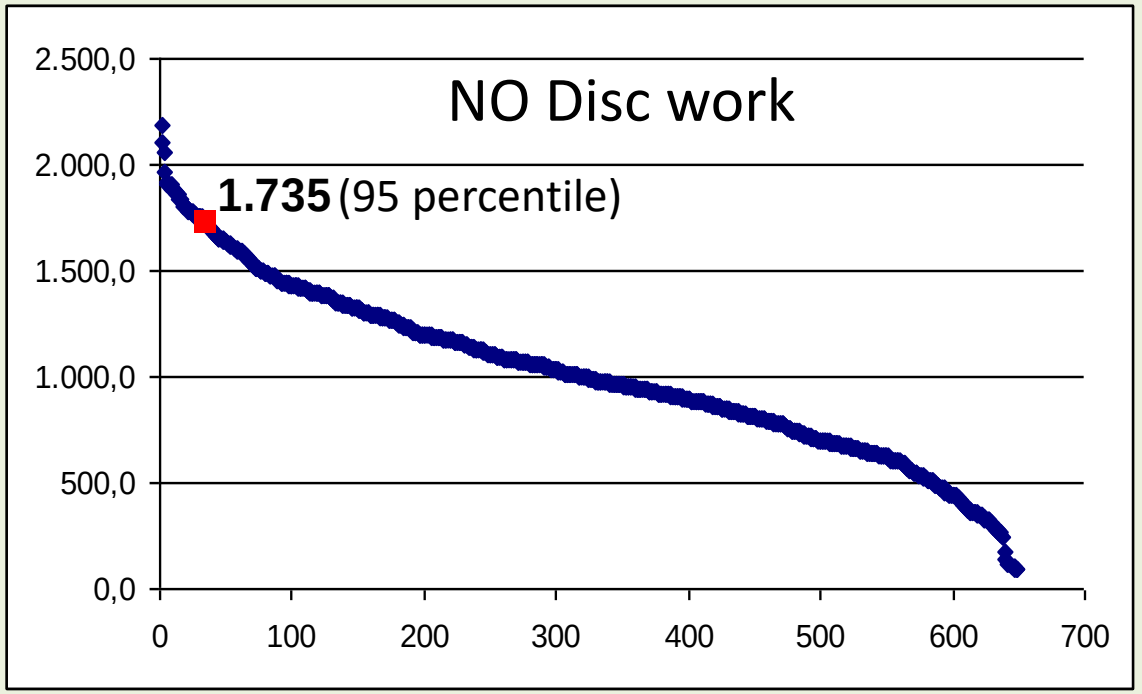
Service Class importance

- Sum of percentages is 100%
- Workload is distributed across all importance levels
- A good amount of work is running in discretionary
- Easier for WLM to manage this workload



Service Class importance





- MIPS used by hour in peak time in the last 60 days
- Opportunity to delay an upgrade

Service Class importance

IMPORTANCE/WORKLOAD	SYSTEM STC1	SYSTEM STC2	DB	NET/MQ	Transactions	Batch	Started Tasks
SYSTEM	High priority system task						
SYSSTC		Privileged system task	Db2 IRLM	TCP/IP	TSOHIGH		Monitors
1			Db2 DBM1 Db2 MSTR IMS DB	MQ MSTR MQ CHIN Session Manager TN3270	CICS TOR CICS CTG IMS CTRL IMS CONNECT		
2			Db2 DIST	Web server LDAP	CICS AOR IMS MPP TSO 1st period		
3					DDF 1st period CICS trx IMS trx TSO 2nd period	Batch High	STC High
4					TSO 3rd Period DDF 2nd Period	Batch Medium 1st Period	STC Medium
5						Batch Medium 2nd Period Batch Low 1st Period	
DISCRETIONARY						Batch Low 2nd Period	STC Low

Adapted from ‘WLM Policy Review at Gothaer Systems‘, Hans Gerd Schneider, EPV User Group 2017

Questions?