

z/OS 3.1 IBM Education Assistant

Solution Name: Implicit long-term CPU protection

Solution Element(s): MVS Workload Manager (WLM)

July 2023



Agenda

- Trademarks
- Objectives
- Overview
- Usage & Invocation
- Interactions & Dependencies
- Upgrade & Coexistence Considerations
- Installation & Configuration
- Summary
- Appendix

Trademarks

- See url <http://www.ibm.com/legal/copytrade.shtml> for a list of trademarks.
- Additional Trademarks:
 - None.

Objectives

- When you assign long-term CPU protection to critical work, you ensure that less important work will generally have a lower dispatch priority.
- Without long-term CPU protection it is possible that work in service classes of lower importance gets a higher dispatch priority than work of higher importance. This is often described as dispatch priority inversion. It may happen, for example, when work of higher importance has exceeded its goal definition while work at lower importance has not met its goal.
- Dispatch priority inversion may have a high potential of causing CPU constraints for your business-critical work.
- You can prevent this by specifying the CPU-critical option in the WLM policy for your CPU-sensitive work.
- With IBM z15 and z16 System Recovery Boost, while the boost is in effect, CPU protection is already assigned to critical work. That is, any single-period service class of importance 1 or 2 is implicitly set to CPU critical. When the boost period ends, the CPU-critical option is reset to the state as defined in the currently active WLM policy.
- Starting with z/OS 3.1, this approach is supplemented. CPU protection is also assigned to critical work when no boost is in effect. That is, any work of importance 1 is implicitly set to CPU critical.
- And furthermore, the first period of any service class is implicitly set to CPU critical. No matter whether this is a single-period or multi-period service class and no matter whether the system is boosted or not.

Objectives (Summary of long-term CPU protection)

Explicit CPU protection	Implicit CPU protection	Implicit CPU protection (boost in effect)
<p>As defined in the active WLM policy:</p> <pre>Service Class Name APPC9 (Required) Description Workload Name APPC (name or ?) Base Resource Group (name or ?) Cpu Critical YES (YES or NO) I/O Priority Group NORMAL (NORMAL or HIGH) Honor Priority DEFAULT (DEFAULT or NO)</pre> <p>Yes: CPU protection is assigned no matter of CCImp or CCImpBoost</p> <p>No: CPU protection is not assigned explicitly but implicitly as defined by CCImp and CCImpBoost</p> <p>Applies to single-period service classes of any importance no matter whether CCImp=0 or CCImpBoost=0 defined</p>	<p>Any work of importance 1 or as defined by OPT parameter CCImp</p> <p>CCImp=<u>1</u> 2 defines the level of importance</p> <p>CCImp=1 is the default</p> <p>CCImp=0 turns off implicit CPU critical</p> <p>Applies to the first period of any service class</p>	<p>Any work of importance 1 or 2 or as defined by OPT parameter CCImpBoost</p> <p>CCImpBoost=1 <u>2</u> defines the level of importance</p> <p>CCImpBoost=2 is the default</p> <p>CCImpBoost=0 turns off implicit CPU critical</p> <p>Applies to the first period of any service class</p>

Overview

- Who (Audience)
 - Senior z/OS system administrators which want automatically protect the business-critical work from CPU constraints
 - Early tenure z/OS system programmers which want a simplified WLM policy specification to avoid common pitfalls
- What (Solution)
 - CPU critical is automatically assigned to any first period of service classes with importance 1 (or 2 when system recovery boost is in effect)
 - The level of importance at which CPU Critical is automatically assigned can be controlled by a new OPT parameters
- Wow (Benefit / Value, Need Addressed)
 - Definition of the WLM service policy is simplified by automatically assigning CPU critical to any service class period running business critical (importance 1 or 2) work
 - This applies to single period service classes and to the first period of multi-period service classes

Usage & Invocation (OPT parameters)

- Two new OPT parameters are provided with z/OS 3.1. Their description is added to z/OS MVS Initialization and Tuning Reference.
- **CCImp=0|1|2** and **CCImpBoost=0|1|2**
 - Defines how CPU protection is implicitly assigned to work. CCImp applies when no boost is active, CCImpBoost applies during boost periods.
 - **0** = CPU protection is not assigned to any work unless CPU Critical has been explicitly assigned for the service class in the active WLM policy.
 - **1|2** = Defines the importance level up to which CPU protection is implicitly assigned to the work when CPU Critical has not been assigned for the service class in the active WLM policy.
 - **Default:** CCImp=1 and CCImpBoost=2
- **Notes:**
 - Implicit CPU protection is only assigned to single period service classes and the first period of multi period service classes.
 - To complete a CCImp change for an active system after the SET OPT command was issued, SRM issues a policy refresh for that active system. Workload Manager indicates the policy refresh with message IWM065I.
 - To complete a CCImpBoost change for an active system, SRM does not immediately issue a policy refresh but activates the change only when the next Recovery Process or Shutdown Boost starts or the VARY WLM command was used to activate a policy.
- **Example:**
 - CCImp=0 and CCImpBoost=1 suppresses implicit CPU protection in general but enables implicit CPU-protection during boost periods for importance 1 work.
 - CCImp=1 and CCImpBoost=2 assigns implicit CPU protection for any importance 1 work, and during a boost phase even for importance 2 work.

Usage & Invocation (WLM services)

- Following WLM/SRM reporting services are extended to provide information on CPU critical and the new OPT parameters:
 - **SYSEVENT REQASD/REQFASD** – Obtain address space related information
 - **IWMRQRY** – Collect AS delay information
 - **IWMRCOLL** – Collect workload activity data
 - **IWM4OPTQ** – Query IEAOPTxx parameters

Usage & Invocation (SYSEVENT REQASD/REQFASD)

- **SYSEVENT REQASD/REQFASD** – Obtain address space related information
 - The SYSEVENTs REQASD and REQFASD provide information about an address space’s workload activity.
 - The parameter list is described by the IRARASD mapping macro and is extended as follows (see in red):

OFFSET DECIMAL	OFFSET HEX	TYPE	LENGTH	NAME (DIM)	DESCRIPTION
=====	=====	=====	=====	=====	=====
...					
100	(64)	CHARACTER	12	RASDOU4	RASD output fields version 4
100	(64)	BIT(32)	4	RASDFLAGS1	Flags
100	(64)	BIT(8)	1	RASDFLG1	First flag byte
		1... ..		RasdCProtCpu	Service class assigned by classification or RESET SRVCLASS was designated CPU-critical in the active policy
...					
	 1...		RasdCpuProtected	CPU protection was assigned either to the address space (see RasdCProtCpu) or to transaction service classes being served by the space or implicitly (see RasdIProtCpu), and SRM is honoring the protection
102	(66)	BIT(8)	1	RASDFLG3	Third flag byte
		1... ..		RasdIProtCpu	CPU protection was implicitly assigned since since first period of service class and importance 1 (or 2 when boost is active). See also OPT parms CCImp and CCImpBoost for importance used.
...					
240	(F0)	CHARACTER	0	RASDEND	

Usage & Invocation (IWMRQRY service)

- **WLM service IWMRQRY** – Collect AS delay information
 - IWMRQRY is the interface reporting products should use to obtain address space related general execution delays. Enclave related information may optionally be requested.
 - The IWMWRQAA answer area mapping macro is extended as follows:

OFFSET DECIMAL	OFFSET HEX	TYPE	LENGTH	NAME (DIM)	DESCRIPTION
146	(92)	UNSIGNED	2	RQAERESOURCECONTENTIONUSING	Contention using samples. One sample is accumulated for each resource in use. Only resource users identified via IWMCNTN are reported.
148	(94)	BIT(16)	2	RQAEFLG3	Additional flags
	1... ..			RQAEINCLSPECIALTY	Same as RasdInclSpecialty
	.1..			RQAE TENANTRESOURCEGROUP	Same as RasdTenantResourceGroup
	..1.			RQAEHWCONTAINER	Same as RasdHwContainer
	...1			RQAEASBOOST	Same as RasdAsBoost
 1...			RQAEIProtCpu	Same as RasdIProtCpu

Usage & Invocation (IWMRCOLL service)

- **WLM service IWMRCOLL** – Collect workload activity data
 - With the IWMRCOLL macro, a performance monitor can get the following workload activity information.
 - The IWMWRCAA answer are mapping is extended as follows:

DECIMAL	HEX	TYPE	LENGTH	NAME (DIM)	DESCRIPTION
=====	=====	=====	=====	=====	=====
0	(0)	STRUCTURE	76	RCAEIHDR	RCAE - period header
0	(0)	UNSIGNED	1	RCAEPERI	Period number
...					
72	(48)	BIT(8)	1	RCAEIFLG	Period flags.
		1... ..		RCAECCIMP	1: Service class period implicitly designated CPU critical
		.111 1111		*	Reserved
73	(49)	CHARACTER	3	*	Reserved

Usage & Invocation (IWM4OPTQ service)

- **WLM service IWM4OPTQ** – Query IEAOPTxx parameters

- The IWM4OPTQ service queries the current IEAOPTxx settings in the system and returns a list of the IEAOPTxx parameters, including the actual value, unit, default value, and description.
- The list returned by the service is extended for the new OPT parameters CCImp and CCImpBoost.
- Thus, the RMF Monitor II OPT report displays the new parameters as follows:

RMF - OPT Settings				Line 1 of 39
Command ==>				Scroll ==> CSR
CPU= 3/ 0 UIC= 65K PR= 0				System= SYS1 Total
OPT: MG				Time: 09/01/22 04:56:06
-- Parameter --	-- Default --	-- Value --	Unit	Description
ABNORMALTERM	Yes	Yes	Y/N	Abnormal terminations in routing
ABSMSUCAPPING	No	No	Y/N	Absolute, permanent MSU capping
BLWLINTHD	20	20	sec	Time blocked work waits for help
BLWLTRPCT	5	5	0/00	CPU cap. to promote blocked work
CCCAWMT	3200	3200	usec	Alternate wait management time
CCCSIGUR	45	33	msec	Min. mean-time-to-wait threshold
CCIMP	1	1	#	Implicit CPU Crit importance
CCIMPBOOST	2	2	#	Implicit CPU Crit imp when boost
CNTCLIST	No	No	Y/N	Clist commands count individually
CPENABLE	0,0	SYSTEM	%	Threshold for TPI (low,high)
DVIO	Yes	Yes	Y/N	Directed VIO is active
ERV	500	500/CB	SU	Enqueue residency CPU Service/DP

Interactions & Dependencies

- Software Dependencies
 - None
- Hardware Dependencies
 - None
- Exploiters
 - z/OS Data Gatherer

Upgrade & Coexistence Considerations

- To exploit this solution, all systems in the Plex must be at the new z/OS level: No
- Migration considerations:
 - Upgrade action is included in the z/OS Upgrade Workflow for z/OS 3.1
 - *“Be aware that the WLM CPU Critical option is automatically assigned to importance 1 work”*
 - Migration guidance
 - Evaluate the current distribution of CPU at different importance levels especially those which are covered by CPU Critical to ensure that they have a consistent CPU demand
 - Implicit CPU critical for importance 1 work can impact the CPU distribution to lower importance work, so ensure that the goals are appropriate given their importance level
- If you are unsure about the impact of implicit long-term CPU protection, you may choose to disable the function by specifying CCImp=0 in the IEAOPTxx parmlib member

Installation & Configuration

- If you are fine with the defaults for CCImp and CCImpBoost, no further action is needed
- If you don't like implicit CPU critical for the first periods with importance 1 (and 2 during boost)
 - Update CCImp or CCImpBoost in your OPT parmlib member and activate it

Summary

- Definition of the WLM service policy is **simplified** by automatically assigning **CPU critical**
- Long-term CPU protection of your business-critical work is implicitly enabled
- Single period service classes and the first period of multi-period service classes are implicitly assigned CPU critical when their importance is 1
- With System Recovery Boost, while the boost is in effect, single period service classes and the first period of multi-period service classes are implicitly assigned CPU critical when their importance is 1 or 2
- These importance levels are configurable by 2 new OPT parameters (CCImp and CCImpBoost)

Appendix

- Publication references
 - SA23-1375 **z/OS MVS Programming: Authorized Assembler Services Reference SET-WTO**
 - SA23-1380 **z/OS MVS Initialization and Tuning Reference**
 - SC34-2662 **z/OS MVS Planning: Workload Management**
 - SC34-2663 **z/OS MVS Programming: Workload Management Services**