z/OS CPENABLE Recommendations

New Option for CPENABLE

With the introduction of z/OS 3.1 there is a new option for CPENABLE. The new option is CPENABLE=SYSTEM. The setting of SYSTEM ensures the low and high thresholds are set to the IBM recommended value for each processor family in the table below. The use of the SYSTEM option removes the need to find and update the parameter with a hardware migration.

IBM Recommendations for CPENABLE

Set **CPENABLE=SYSTEM** for all z/OS LPARs running z/OS V3.1.

Recommended CPENABLE Settings for z/OS V2R5 and Earlier

Processor Family	LPAR Dedicated	LPAR Shared
IBM z14 through IBM z17	**5,15**	**5,15**
IBM z13 and z13s	10,30	10,30

The enablement of z/OS logical processors as being marked as enabled for I/O interrupts has changed for the IBM Z Systems z14 and later processors. This change affects how logical processors in a wait state handle I/O interrupts. Previously all logical processors in a no-work wait were enabled for I/O interrupts. The enablement support changed to rely solely upon WLM to discover and set the appropriate number of logical processors to enable for interrupts. Logical processors in a no work wait are not enabled for I/O interrupts. In certain situations, the number of logical processors enabled for I/O interrupts has proven to be too low with these enhancements. Because of these changes in z/OS, LPARs running on z14 and later processors an updated recommendation is being made to best maintain I/O responsiveness for LPARs on z14 and later processors.

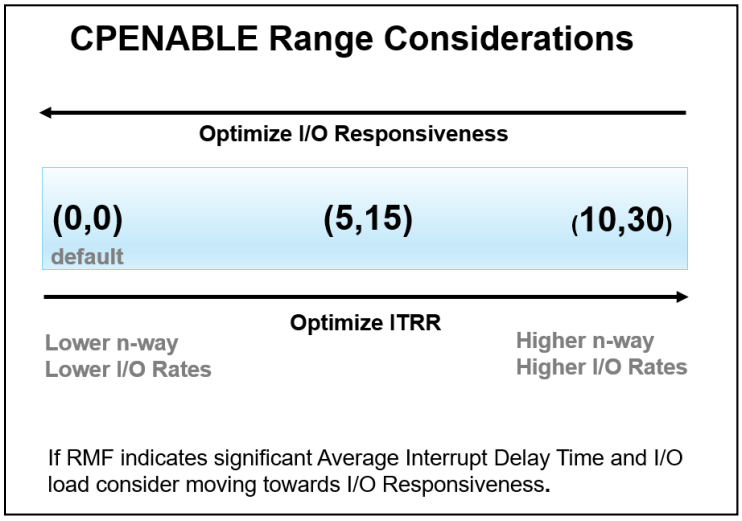
What is CPENABLE?

CPENABLE is a parameter in the IEAOPTxx parmlib which allows z/OS to dynamically manage the number of logical processors which field I/O interrupts. The original benefit of CPENABLE was to provide increased efficiency in a Basic mode environment by limiting the code used to process I/O interrupts to fewer CPs. By limiting the I/O enablement to fewer CPs the I/O interrupt code didn't have to be in every CP's hardware cache. This freed up cache space for the engine and helped to improve the amount of hardware cache useable by other non I/O functions and could then drive efficiency benefits.

The z/OS system monitors the percentage of I/O interrupts are handled via Test Pending Interrupts (TPI). This measured value is compared to the values set in the CPENABLE parameter. When the measured value is higher than the range specified, the system will enable an additional CP for I/O interrupts. When the measured value is lower than the specified range the system will disable one of the processors already handling I/O interrupts.

Measurements on current models of processors show improved ITRR benefits of 0% - 5% when running with CPENABLE settings other than the default of (0,0). Environments with high n-way and/or high I/O rates showed greater improvement, whereas lower n-way and lower I/O rates showed no difference.

The following shows the CPENABLE range considerations.



History of CPENABLE Settings

As stated above CPENABLE was first introduced to help provide efficiencies

when the number of physical CPs was small, and the hardware cache size was very small. With the introduction of PR/SM and the ability of LPARs to share processors the CPENABLE recommendations needed to change.

When an LPAR was using shared CPs, z/OS would enable logical CPs for I/O but had no control over which physical CP would be dispatched to the logical processor. Thus, the CPENABLE efficiency algorithm had limited impact because there was no method to control the physical CP dispatching and reap the benefits of assigning the I/O processing to fewer CP caches. Therefore, CPENABLE was originally set to (0,0) for a shared LPAR environment. It remains the default to this day.

But since the introduction of PR/SM there have been significant changes in PR/SM capabilities, in the number of LPARs, and the number of logical processors in an LPAR. There have also been z/OS enhancements like WLM and HiperDispatch which have changed the processing of workloads.

The reason for the previous recommendation to set (10,30) for shared LPAR environments was made when it was discovered the I/O processing in modern Z processors provided better throughput when there were fewer CPs in the path due to serialization issues. The use of CPENABLE(10,30) provided this benefit. Other changes in this area, like HiperDispatch, have also limited I/O enablement to vertical high and vertical medium logical processors again reducing the impact of PR/SM dispatching

Moving forward in time what has been determined is the efficiency costs of bringing a waiting logical processor out of a no-work wait simply to handle an I/O interrupt has increased. Pressures came from the overall number of physical CPs and the high level of logical processor virtualization seen in a z/OS environment.

Having fewer logical processors handling I/O interrupts while dynamically reacting to changes in system loads has always been an efficiency item provided by CPENABLE. In order to continue to provide this efficiency changes were made in this area with the introduction of the z14. In this case, changes were made to limit the impact of waiting vertical high and vertical medium logical processors from being dispatched by PR/SM to only handle an I/O. There are efficiency costs in having to dispatch a physical engine to a logical processor to field an I/O interrupt and then have the logical processor return to a wait.

The changes introduced with the z14 have identified cases where insufficient

logical processors are enabled for interrupts causing an impact to workloads. There have been reports of the only logical processor(s) enabled for I/O interrupts all running disabled or doing other CPU intensive operations. In such a case no I/O interrupts could be fielded by the LPAR, and I/O would time out waiting in the channel. It has been identified more aggressive I/O enablement is needed in this environment. The recommended change to (5,15) will provide for more aggressive I/O enablement ensuring greater I/O responsiveness for z/OS workloads. The support will still be heuristic with WLM adjusting the values based on TPI load factors.

In certain cases, client systems have already determined the need for greater I/O enablement than the previously recommended value of (10,30). This recommendation is not intended to disrupt these changes. Client LPARs running with a more aggressive CPENABLE setting than (5,15) should continue with those settings.

Evaluating this Change

A new Average Interrupt Delay Time measurement was introduced on the IBM Z Enterprise EC12 (zEC12) and later servers. In z/OS RMF is designed to report on interrupt delay time to help determine whether I/O processing delays are occurring. This measurement is designed to measure the time between when primary status is presented to the channel subsystem and when the operating system clears the primary status to begin processing the interrupt.

Increases in this time measures the impact of LPAR virtualization such as an LPARs with a low weight running on a busy server. It also will measure the impact of insufficient logical CPs enabled for I/O interrupts.

RMF writes this information to fields in SMF type 74 subtype 1 and SMF 79 subtype 9 records and will display the averaged interrupt delay times in the Postprocessor Device Activity report. The higher numbers in this field (see highlighted example below) indicate the CPENABLE setting should be reviewed to determine if enabling more CPs for interrupt processing is warranted. An RMF CPU Activity report can be used to review the number of logical processors enabled for I/O Interrupts.

One warning about the Average Interrupt Delay Time. This field can report high values if the I/O rate is very low. Before acting on this field ensure the I/O activity level is greater than 2-3 per second.

STORAGE GROUP	DEV NUM	DEVICE TYPE	NUMBER OF CYL	VOLUME SERIAL	PAV	LCU	DEVICE ACTIVITY RATE	AVG RESP TIME	AVG IOSQ TIME	AVG CMR DLY	AVG DB DLY	AVG INT DLY	AVG PEND TIME	AVG DISC TIME	AVG CONN TIME
XTEST	02208	33903	3339	TRXSX9	1	0032	0.001	.384	.000	.128	.000	.123	.256	.000	.128
XTEST	02209	33903	3339	TRXSXA	1	0032	0.001	.256	.000	.000	.000	.135	.256	.000	.000
	0220A	33909	10017	TRXT01	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000
	0220B	33909	10017	TRXT02	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000
	0220C	33909	10017	TRXT03	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000
	0220D	33909	10017	TRXT04	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000
	0220E	33909	10017	TRXT05	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000
	0220F	33909	32760	TRXT06	1	0032	0.000	.000	.000	.000	.000	.000	.000	.000	.000

Summary

A new CPENABLE recommendation is being made for all systems running z/OS V3.1. The new option of setting **CPENABLE=SYSTEM** eases system migrations and management by following the IBM recommended values automatically.

Any systems running an earlier version of z/OS should follow the recommendations in the table on the first page of this document. The changes for z14 processors allow for more aggressive enablement of I/O processors in a z/OS environment. With the z14 server introduction z/OS no longer enables logical processors in a wait state for I/O. This change was done for efficiency reasons, but it has identified cases where insufficient logical processors were enabled for interrupts impacting workloads. The change to a more aggressive CPENABLE setting will better support most z/OS workloads.

Special Notice:

The information contained in this document has not been submitted to any formal IBM test and is distributed on an “as is” basis without any warranty either expressed or implied. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer’s ability to evaluate and integrate them into the customer’s operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Performance data contained in this document was determined in a controlled environment; therefore, the results which may be obtained in other operating environments may vary significantly. No commitment as to your ability to obtain comparable results in any way is intended or made by this release of information.