Rule WLM892: LPAR could not use the defined logical zIIP processors

Finding: HiperDispatch was specified for one or more LPARs within a CPC, and this LPAR used at least one high polarity zIIP processor. However, an LPAR that was not operating in HiperDispatch Management Mode could not use all of its defined logical zIIP processors because of the high polarity zIIP processor.

Impact: This finding has a MEDIUM IMPACT, HIGH IMPACT, or VERY HIGH IMPACT on the performance of the LPAR that could not use all of its defined logical zIIP processors.

Logic flow: This is a basic finding based on analysis of the LPARs that were defined to Processor Resource/System Manager (PR/SM). The finding applies only to LPARs that were not running in HiperDispatch Management Mode.

Discussion: If LPARs are assigned shared processors, there is a mechanism that is used to indicate what percent of physical processor capacity in the shared pool should be assigned to each specific LPAR. In the LPAR definition process, each LPAR is assigned a *weight* which is simply a numerical value, ranging from 1 to 999.

The weights for all LPARs that are activated[1] are summed, and used as the divisor in a calculation of the percent of processor capacity that any specific LPAR is to be guaranteed. The following algorithm illustrates the general concept:

Before HiperDispatch was introduced with z10, each *logical processor* assigned to an LPAR was allocated an even share of CPC capacity. The share of CPC capacity to which the LPAR was entitled was simply divided by the number of logical processors assigned to the LPAR, and each logical processor was allocated the resulting percent of a physical processor.

This “even share” concept nominally distributed the work in the LPAR across the number of physical processors corresponding to the number of logical processors assigned to the LPAR.

Beginning with z990, PR/SM attempted to assign a logical processor to the same physical processor last assigned to the logical processor, and attempted to assign the logical processor to a physical processor that was

[1] LPARs can be deactivated or can be activated, based on operator command. The “share” calculation depends on the LPARs that are activated.

in the same book as the previous assignment. However, PR/SM would not delay assignment of a logical processor to a physical processor if the desired assignments could not be made (that is, the previous physical processor was already assigned to a different logical processor or there were no otherwise idle physical processors in the same book as the previous assignment). In this case, there would be hardware degradation caused by L1 and L2 cache misses. Considering the high speed of current processors, this degradation could be a significant percent of the potential hardware speed of the processors.

HiperDispatch was introduced with z10, HiperDispatch was designed (1) to minimize the degradation caused by L1 and L2 cache misses, (2) minimize the hardware conflicts inherent in a multiprocessor environment, and (3) to maximize the amount of CPU processing power associated with any single logical processor. To achieve these design objectives, HiperDispatch implemented features within z/OS and within PR/SM, and implemented an exchange of information between PR/SM and z/OS. With the exchange of information, z/OS would be aware of the topology between logical processors and physical processors.

With HiperDispatch, z/OS dynamically manages the weight given to *each logical processor* assigned to an LPAR. This weight is communicated to PR/SM, with the result that PR/SM can treat each logical processor as having a potentially unique share of a physical processor. This dynamic adjusting of the weights is done instead of PR/SM having a fixed weight associated with each logical processor and each logical processor having an equal share of CPC capacity,

This information is exchanged for central logical processors, for zAAP logical processors, and for zIIP logical processors. If an LPAR is operating under HiperDispatch Management Mode, zAAP and zIIP logical processors are automatically managed by HiperDispatch. It is not possible to operate only central logical processors in HiperDispatch if zAAP or zIIP logical processors are assigned to the LPAR.

To accomplish this, z/OS Dispatcher manages work in multiple *affinity dispatch queues*. Work units in an affinity dispatch queue are assigned an *affinity identifier*, and work units with the same affinity identifier normally will be assigned to the same affinity dispatch queue. z/OS monitors workflow performance, and can switch work units among affinity dispatch queues if necessary to achieve better performance.

One or more logical processors are associated with an affinity dispatch queue. z/OS assigns a PR/SM weight to each logical processor assigned to an affinity dispatch queue (and implicitly, the weight is associated with the work in the associated affinity dispatch queue). The weight that z/OS assigns to each logical processor falls into one of three categories:

- A weight that causes PR/SM to assign 100% of an equivalent physical processor to the logical processor. These logical processors are called **high polarity** logical processors. There can be multiple affinity dispatch queues with high polarity logical processors.

  z/OS can create an affinity dispatch queue that has one or more *high* polarity processors only if:

  - The LPAR has a share of CPC capacity that results in at least 1.5 equivalent physical processors.

  - Two or more logical processors were assigned to the LPAR.

  - The workload requires 1.5 or more equivalent physical processors (that is, the CPU processing requirements of the workload result in using the capacity of at least 1.5 physical processors).

  An affinity dispatch queue that has one or more *high* polarity processors cannot exist unless these conditions exist.

- A weight that causes PR/SM to assign more than 0% but less than 100% of an equivalent physical processor to a logical processor. These logical processors are called **medium polarity** logical processors.

  z/OS will always maintain the medium affinity dispatch queue with at least 0.5 equivalent physical processor share, but less than1.5 equivalent physical processor share (a high polarity logical processor would be created if the workload demand in the medium polarity affinity dispatch queue was at least 1.5 equivalent physical processors).

  The capacity share in the medium polarity affinity dispatch queue essentially is the processor share that remains after any high polarity logical processors have been assigned their 100% share, with the restriction that the medium polarity affinity dispatch queue must have at least 0.5 equivalent physical processor share.

- A weight that causes PR/SM to assign 0% of an equivalent physical processor to the logical processor. These logical processors are called **low polarity** processor. Low polarity logical processors are not associated with an affinity dispatch queue.

Without proper planning, running HiperDispatch Management Mode in one LPAR can have serious performance implications for LPARs that are not running in HiperDispatch Management Mode. This is because HiperDispatch can create high polarity logical processors (with 100% share of an equivalent physical processor) and PR/SM will essentially dedicate a physical processor to be used by the logical processor (this is the pseudo-

dedicated processor described earlier). This assignment of a pseudo-dedicated physical processor to a logical processor can have serious implications for any LPAR not running in HiperDispatch Management Mode, particularly with small systems.

For example, consider a z10 Model 703 with two LPARs, each LPAR having three logical zIIP processors assigned. Suppose that HiperDispatch is turned on for LPAR1, and that LPAR1 has one high polarity zIIP logical processor, one medium polarity zIIP logical processor, and one parked zIIP logical processor. The high polarity processor is pseudo-dedicated. That leaves two physical zIIP processors which can be used by LPAR2.

LPAR2 is not participating in HiperDispatch Management Mode, so its share would be split equally among three logical zIIP processors. However, these three logical zIIP processors would no longer have access to three physical zIIP processors since one of the three physical zIIP processors would be dedicated to the high polarity zIIP processor on LPAR1!

LPAR2 would have access to only two zIIP physical processors (assuming that the workload on LPAR1 is sufficient to drive the high polarity zIIP processor to 100% for LPAR1). The three logical zIIP processors assigned to LPAR2 would be dispatched to two physical zIIP processors in a circular fashion, but no more than two logical zIIP processors could be active at any time. LPAR2 could effectively see its available zAP capacity reduced by 1/3. Any work running on LPAR2 could experience significant delays because z/OS still dispatch work to the three logical zIIP processors, but each logical zIIP processor would queue waiting its turn for a physical zIIP processor.

If APAR OA24074 has been applied, SMF Record Type 70 (CPU Activity) records contain information describing whether HiperDispatch is supported for the LPAR, whether HiperDispatch is active for the LPAR, and whether HiperDispatch status of the LPAR changed during the recording interval.

Additionally, SMF Record Type 70 (PR/SM Logical Processor Data Section) contains the *polarization weight* assigned to each logical processor by z/OS when the LPAR is operating in HiperDispatch Management Mode. These SMF records also contain information describing the PR/SM weight that was assigned to zIIP processors assigned to the LPAR. For LPARs not operating in HiperDispatch Management Mode, this weight will be the actual weight under which the LPAR is operating

CPExpert produces Rule WLM892 for an LPAR if:

- The LPAR was not running in HiperDispatch Management Mode

- At least one LPAR was running HiperDispatch Management Mode and established one or more high polarity zIIP logical processors.

- The number of physical zIIP processors in the CPC, minus the number of high polarity logical zIIP processors, was less than the number of logical zIIP processors assigned to the LPAR not running in HiperDispatch Management Mode.

Suggestion: If Rule WLM892 is produced, CPExpert suggests that you consider the following actions:

- You should review the work executing on the LPAR identified by Rule WLM892 to assess the importance of this work. If the LPAR is a test LPAR, it is possible that Rule WLM892 is not as important as it might be if the LPAR is a production LPAR.

- You should consider reducing the number of logical zIIP processors assigned to the LPAR identified by Rule WLM892. This reduction in the number of logical zIIP processors should reduce the number to where the LPAR can actually use the assigned logical zIIP processors. The number of logical zIIP processors should be no more than the total number of physical zIIP processors purchased with the z10, minus the total number of high polarity zIIP processors established by LPARs running in HiperDispatch Management Mode.

- If the above alternative is not feasible, you should consider reducing the weight of the LPAR running in HiperDispatch Management Mode so that the weight results in a smaller share of the total zIIP processor capacity. If a smaller share were available, HiperDispatch Management Mode could not establish as many high polarity zIIP processors for this LPAR.

- If the above alternatives are not feasible, you should consider not running any LPAR in HiperDispatch Management Mode, particularly for a small z10 CPC.

- If you decide to take no action based on this finding, you should "turn off" Rule WLM892. Section 2 of this document describes how to "turn off" rules. In general, **you should not turn of Rule WLM892**, since this finding could report a VERY HIGH performance impact.

Reference: *System z9 109 PR/SM Planning Guide* (SB10-7041-01)
Chapter 3. Determining the Characteristics of Logical Partitions

System z10 PR/SM Planning Guide (SB10-7153-00)
Chapter 3. Determining the Characteristics of Logical Partitions

zEnterprise System (z196) PR/SM Planning Guide (SB10-7155)
Chapter 3. Determining the Characteristics of Logical Partitions

zEnterprise System (EC12) PR/SM Planning Guide (SB10-7156)
Chapter 3. Determining the Characteristics of Logical Partitions

z/OS Intelligent Resource Director (SG24-5952)
Chapter 3. How WLM LPAR CPU Management works (Section 3.5: LPAR Weights)