
Rule WLM107: Response time distribution for service class with percentile response performance goal

Finding: This rule provides information about the distribution of response times during those intervals when the identified service class missed its performance goal.

Impact: This finding has NO IMPACT on performance of your computer system. The finding is provided to allow you to assess the overall performance of service classes having a percentile response time performance goal.

Logic flow: The following rule causes this rule to be invoked:
Rule WLM102: Service Class did not achieve average response goal

Discussion: For service classes with response goals, RMF includes in SMF Type 72 records a count of transactions that completed in varying percentages of the response goal. These transaction counts are recorded by RMF as the "Response Time Distribution Count Table" contained in SMF Type 72(Subtype 3) records. Section 4 describes the percentages recorded by RMF.

When CPExpert produces Rule WLM102, CPExpert automatically produces Rule WLM107 to provide a summary distribution of the response information. The purpose of Rule WLM107 is to allow you to assess whether the response is meaningful, or whether there are some transactions that skew the finding.

The following example illustrates the output from Rule WLM107.

```

RULE WLM107:  RESPONSE TIME DISTRIBUTION FOR SERVICE CLASS

Service Class TSOUSERS (Period 1) did not achieve its response goal
during the measurement intervals shown below. The response goal was
80.00 percent of the transactions completing within 0.500 seconds.
The below information shows the distribution of response times:

                                --PERCENT COMPLETIONS RELATIVE TO GOAL--
                                50-   90-  100-  110-  200-
                                TOTAL <50%  90%  100%  110%  200%  400% >400%
MEASUREMENT INTERVAL  TRANS GOAL GOAL GOAL GOAL GOAL GOAL GOAL
10:45-11:00,07DEC1994   63  52.4  4.8  1.6  0.0  0.0  12.7  28.6
11:15-11:29,07DEC1994   32  40.6  31.3  3.1  3.1  0.0  6.3  15.6
11:29-11:30,07DEC1994    1   0.0  0.0  0.0  0.0  0.0  0.0  100.0
11:45-12:00,07DEC1994   14  64.3  0.0  7.1  0.0  7.1  14.3  7.1

```

Suggestion: If you find that some transactions skewed the findings, you may wish to consider the following alternatives:

- In the example shown above, there seemed to be a bimodal distribution of response: many transactions experienced a response time of less than 50% of the goal while many transactions experienced a response time of greater than 200% of the goal.
- The bimodal distribution may indicate that the service class contains transactions with dissimilar characteristics. In this case, perhaps you can use Workload Categorization to place the transactions into a different service class if you can identify the transactions.

You may wish to specify a different importance and different performance goal for this new service class. Other findings by CPEXpert may bolster this conclusion if (for example) CPEXpert notes that the service class required a significant amount of CPU per average transaction (see Rule WLM200 for a discussion of this situation).

- The bimodal distribution may indicate that there are system problems that cause the poor response of some transactions in the service class.

Other service classes may interfere with the service class missing its performance goal. This situation would typically be identified by a subsequent finding by CPEXpert that address spaces in the service class was "denied CPU" by other address spaces¹ (see Rule WLM255 for a discussion of this situation).

Alternatively, CPEXpert might identify DASD-related problems that cause elongated DASD I/O times for the transactions experiencing excessively long response times.

- C CPEXpert might identify DASD disconnect (DISC) time as a likely cause of delay (see Rule WLM355 for a discussion of this situation). DASD disconnect time normally² is caused by missed DASD reads (that is, the required records were not in the controller's cache and had to be fetched from the device).

¹Address spaces in the service class could be "denied CPU" by address spaces in other service classes, by system functions, or by address spaces in the service class itself competing with each other.

²With legacy DASD configurations (e.g., IBM-3380 devices attached to IBM-3390 controllers), DASD disconnect time is primarily composed of seek time or missed RPS reconnect delay. Seek time or missed RPS reconnect delays often are caused by I/O activity by other address spaces referencing the I/O subsystem.

-
- C CPEXpert might identify DASD pending (PEND) time as a likely cause of delay (see Rule WLM356 for a discussion of this situation).
 - C CPEXpert might identify DASD connect (CONN) time as a likely cause of delay (see Rule WLM357 for a discussion of this situation).
 - C CPEXpert might identify DASD I/O queuing in the MVS I/O Supervisor (IOSQ) time as a likely cause of delay (see Rule WLM358 for a discussion of this situation).
- You can simply ignore the findings that CPEXpert made associated with this service class for the interval. You may decide that the poor transaction response is an anomaly and not take any further action.