Rule WLM255: Service class was in Active state but server was denied access to CPU

Finding: CPExpert has determined that the transaction service class that missed its performance goal was in Active state, but the server service class was denied access to a CPU.

Impact: This finding has a HIGH IMPACT on performance of your computer system.

Logic flow: The following rule causes this rule to be invoked:
Rule WLM120: Significant transaction time was in Active state

Discussion: When CPExpert produces Rule WLM104 or Rule WLM105 to indicate that a subsystem service class did not achieve its performance goal, the logic of these rules tries to identify the cause of the delay. The cause of the delay initially is analyzed from the "served" service class view. Rule WLM120(series) and Rule WLM130(series) describe the results from this analysis.

After analyzing the subsystem transaction delays, CPExpert identifies the service classes that serve the transactions. The subsystem transactions typically are CICS transactions, and the servers are the CICS regions. Alternatively, the transactions could be IMS transactions and the servers could be the IMS control regions or transaction processing regions.

Address spaces executing in the system can be in a variety of states from the perspective of the Workload Manager: using the CPU, delayed for an identifiable reason, or delayed for some unknown reason.

The System Resources Manager (SRM) periodically samples the state of each address space in each service class. These samples are accumulated into variables that are recorded by RMF in the "Service Class Period Data Section" of SMF Type 72 (Subtype 3) records[1].

CPExpert produces Rule WLM120 when a significant cause of delay to a subsystem transaction was that the transaction was in Active state. The Active state indicates that a task was executing on behalf of the transaction, from the perspective of CICS or IMS.

[1]Please see Section 4 for a discussion of these states and the sampling process.

CPExpert analyzes the CPU requirements of the server service class to determine if the server was denied access to a CPU.

As the System Resources Manager takes its samples of the state of address spaces, it examines whether a TCB or SRB associated with the address space is waiting for dispatching to a CPU, or whether a TCB is waiting for a local lock.

If an address space is waiting for dispatching, it is being denied access to a CPU because processors are active with higher priority address spaces or with address spaces at the same dispatching priority as the address space waiting for dispatching. Samples reflecting the time address spaces are denied access to a CPU are recorded by RMF in the SMF Type 72 delay samples, as CPU Delay (R723CCDE).

For a service class consisting of CICS or IMS regions and managed to transaction goals, the CICS or IMS regions compete with each other at the same dispatching priority **if they are providing service to the same set of transaction service classes**. This is because the Workload Manager will create a dynamic internal service class ($SRMSnnn) consisting of all address spaces that provide service to the transactions service classes. The Workload Manager will manage these address spaces collectively from a CPU dispatching view (that is, all address spaces will have the same CPU dispatching priority)[2].

CPExpert computes the percent of CPU Delay for the server service class, as a function of the response time of the subsystem transaction service class missing its performance goal. CPExpert produces Rule WLM255 if the percent of CPU Delay for the subsystem transaction service class is greater than the significance value specified in the **WLMSIG** guidance variable in USOURCE(WLMGUIDE).

With Rule WLM255, CPExpert provides the total CPU service units consumed by the service class in the RMF measurement interval, the percent of active time when transactions in the service class were denied access to a CPU, and the average multiprogramming level of the server service class. The average multiprogramming level is provided so you can assess whether address spaces might have competed with each other.

The following example illustrates the output from Rule WLM255:

[2]The address spaces will be managed **individually** from a processor storage access policy view.

```
RULE WLM255:  SERVICE CLASS WAS ACTIVE BUT SERVER WAS DENIED CPU

   During the above measurement intervals, the CICUSRTX Service Class was
   in the READY STATE during a significant portion of its response time.
   However, at least one address space in the CICSRGN server was denied
   access to a CPU for a significant percent of this time.  During the
   following RMF measurement intervals, CICSRGN had a TCB or SRB waiting
   to be dispatched, or a TCB was waiting for a local lock.  The below
   information shows the CPU used by CICSRGN during the measurement
   interval, and the "PERCENT DENIED CPU" value represents the percent
   of CICSRGN's ACTIVE time when at least one address space was waiting
   for access to a CPU.  CPExpert will produce a report at the end of this
   analysis that shows the CPU time used by all service class periods.

                                                                  AVERAGE
                                  CPU  USED      PCT SERVER     SERVER
   MEASUREMENT INTERVAL           BY SERVER      DENIED CPU       MPL
   13:07-13:12,21JUN1994           0:02:00          60.4          4.0
   13:17-13:22,21JUN1994           0:02:00          57.7          4.0
```

Suggestion; Please refer to the suggestions in Rule WLM250 for a discussion of the alternatives that can be implemented to improve access to a CPU.

Please note that, at present, CPExpert cannot determine whether specific address spaces were denied access to a CPU. The address spaces, *while acting as servers*, can be managed separately (depending on the WLM's topology assessment).

For example, If a number of address spaces (CICS regions) are handled by the same service class (CICSRGN, for example), and if the regions in the CICSRGN service class process different transaction service classes with different goals and importance, the WLM can group various regions into dynamic internal service classes. The regions in these dynamic internal service classes will be managed collectively from a CPU dispatching view.

There can be more than one dynamic internal service class that consist of different CICS regions. All (regions in any particular dynamic internal service class will have the same CPU dispatching priority. However, if there are multiple CICS regions assigned to the CICSRGN service class and if there are multiple dynamic internal service classes, the regions in different internal service classes can have different CPU dispatching priority. This could mean that different regions (in the same CICSRGN service class) could be denied or not denied access to a CPU.

However, the service class period data in SMF Type 72 is an accumulation of all samples taken of address spaces in the CICSRGN service class. There is no way to determine which address spaces were provided or denied access to a CPU, since the data is an accumulation of samples.

Consequently, the information provided by Rule WLM255 must be viewed with caution if a service class (such as CICSRGN) consists of multiple address spaces and if these address spaces provide service to different transaction service classes. The finding means that some address spaces in the server service class were denied access to CPU, but the "denied CPU" information is "averaged" over all address spaces handled by the server service class.

These comments do not apply, of course, if the server service class (CICSRGN) is managed to the goals of the region rather than to the goals of the transactions.

