Rule WLM031: Dynamic Alias Management active but not I/O Priority

Finding: Dynamic alias management was active, but I/O priority management was not selected.

Impact: This finding should be viewed as LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of your computer system. The level of impact depend on the amount of DASD I/O activity, the importance of the work requiring DASD I/O operations, and whether wait on DASD I/O caused performance problems for important work.

Logic flow: This a basic finding. There are no predecessor rules.

Discussion: With OS/390 Release 3, IBM introduced the I/O Priority Management option for Goal Mode. With I/O Priority Management, the system manages I/O priorities in the sysplex based on service class goals. The Workload Manager (WLM) dynamically adjusts the I/O priority based on how well each service class is meeting its goals and whether non-paging DASD I/O can contribute to meeting the goal. I/O priority queuing is used to control non-paging DASD I/O requests that are queued because the device is busy.

The I/O Priority Management process adjusts the I/O priority of non-paging DASD I/O operations (and includes the I/O activity (using and delay) as a part of the execution velocity calculation as described later in this document). However, these WLM algorithms operate only at the system level, and had no direct effect on I/O priority at the subsystem (controller and device) level until the Parallel Access Volumes (PAV) concept was available with OS/390 Release 9 (or with APAR OW39854 installed on OS/390 Release 7 or 8).

The IBM Enterprise Storage Server (ESS) supports concurrent or parallel data transfer operations to or from the same volume from the same system or system image. One feature of the PAV design is that I/O queuing is normally moved from queuing in the I/O Supervisor (commonly viewed as IOSQ time), to queuing in the 2105 ESS DASD subsystem. The 2105 ESS DASD subsystem can uses the I/O priority assigned to the I/O operation to prioritize the I/O access to the device.

With the PAV concept, *multiple device addresses* of a single control-unit image are *associated* with a single device (or parallel access volume).

With PAV, a base address is assigned to a CKD[1] logical device, and *alias device addresses* can be configured for the CKD logical device. While the base address is the actual unit address of a given volume, there can by many alias addresses assigned to a base address, and any or all of those alias addresses can be reassigned to a different base address.

These logical devices are identified by device numbers and represented by unit control blocks (UCBs) in the system. When an I/O is requested to a parallel access volume, the request identifies the base UCB and the system uses the base UCB or one of its alias UCBs, depending on their availability, to initiate the request. The Hardware Configuration Definition (HCD) is used to define multiple device numbers for a control unit that provides parallel access volumes.

When PAV was introduced, IBM also introduced optional *dynamic alias management*. With dynamic alias management, the Workload Manager can automatically perform the alias address reassignments to help work meet its goals and to minimize IOS queuing.

- If dynamic alias management is not enabled, the number of aliases to be assigned to each base device is determined by the parallel access volume's configuration application ("ESS Specialist") at IPL time or at the time a base device is varied online.

- If dynamic alias management is enabled, the Workload Manager monitors the device performance and automatically reassigns alias addresses from one base to another to help work meet its goals and to minimize IOS queuing.

The efficiency of the Workload Manger in assigning aliases to base devices depends on whether I/O Priority Management is enabled.

- If I/O Priority Management is enabled, and if the Workload Manager determines that a workload is not meeting its goal due to IOSQ time, then the Workload Manager attempts to find alias devices that can be moved to help that workload achieve its goal. Even if all work is meeting its goals, the Workload Manager will attempt to move aliases to the busiest devices to minimize overall I/O queuing.

- If I/O Priority management is **not** enabled, the Workload Manager will not use information about whether service classes meet goals and whether I/O queuing causes service classes to miss goals. Consequently, the Workload Manager will make decisions about assignment of alias

[1] CKD (count key data) is a term for a logical device that specifies the format of the logical data units on the device. The logical data unit is a track that can contain one or more records, each consisting of a count field, an optional key field, and an optional data field.

devices based only on an attempt to minimize overall IOS queuing, and these alias movement decisions will not take service class goals into consideration.

IBM strongly recommends that I/O Priority Management be specified if Dynamic Alias Management is specified.

CPExpert examines variable R723MSCF (Service/Report class flags) in SMF Type 72 information to detect whether Dynamic Alias Management and I/O Priority Management have been selected. CPExpert produces Rule WLM031 when Dynamic Alias Management has been selected, but I/O Priority Management was not selected.

The following example illustrates the output from Rule WLM031:

```
RULE WLM031: DYNAMIC ALIAS MANAGEMENT BUT NOT I/O PRIORITY MANAGEMENT

   Service Policy WLMPOL01, effective 17AUG2000:15:35:4, specified Dynamic
   Alias Management, but did not specify I/O Priority Management.  IBM
   strongly recommends that I/O Priority Management be specified if
   Dynamic Alias Management is specified.  If Dynamic Alias Management is
   specified but I/O Priority Management is not specified, the Workload
   Manager will not consider goals while managing Parallel Access Volumes.
   This means that the Workload Manager will make alias moves to minimize
   overall IOS queuing, but the Workload Manager will not consider service
   class goals when deciding the alias moves.
```

Suggestion: CPExpert suggests that you consider the following alternatives:

- You should consider specifying I/O Priority Management so the Workload Manger will use information about goal achievement in its Dynamic Alias Management decisions.

  Additionally, I/O Priority Management is **required** with Channel Subsystem Priority Queuing available on z900 processors.

- There are, however, issues that might cause you to reconsider IBM's recommendation. These issues relate to how the Workload Manager creates and uses execution velocity.

With I/O Priority Management, samples of I/O activity[2] are used in the execution velocity formula. If I/O Priority Management is selected, non-paging DASD connect time and disconnect time are included in the "using" part of the execution velocity formula, and delays[3] waiting for non-paging DASD I/O operations are included in the "delay" part of the execution velocity formula.

There are two implications of the above algorithm:

- Execution velocity goals must be revised to include the performance index that would result once I/O activity is included in the execution velocity calculations. This revision requires some effort.

- The Workload Manager computes the execution velocity of a service class by applying the following algorithm:

$$\frac{using\ samples}{using\ samples + delay\ samples} * 100$$

where:

using samples include:

- The number of samples of work using the processor (CPU Using).
- The number of calculated samples of work using non-paging DASD I/O resources (DASD connect state or DASD disconnect state). I/O using samples are included only if the installation has elected to include WLM-managed I/O.

delay samples include:

- The number of samples of work delayed for the processor (Denied CPU Delay or CPU Capping delay).
- The number of samples of work delayed for processor storage. Delay for processor storage includes:
  - Paging delay
  - Swap-in delay
  - Swapped out for multi programming (MPL) reasons

[2] The "I/O samples" are not actually samples, but are numeric values computed by multiplying the amount of measured I/O activity by the WLM sample period of four samples per second. The WLM samples activity (such as CPU activity) and delays (such as CPU delay or processor delay) for the execution velocity formula. Since the DASD I/O activity and delay are measurements made at the controller level, it is not possible for the WLM to sample the activity or delay state of the device. However, Don Deese pointed out to IBM SRM/WLM developers that simply multiplying the DASD activity and delay measured time by the WLM sample rate would yield calculated sample counts that would be equivalent to actual sample counts. These calculated sample counts are used by I/O Priority Management and execution velocity calculations.

[3] Non-paging DASD I/O delays include IOS queue delays, SubChannel pending delays, and control unit queue delays. Note that DASD disconnect time is not included in the execution velocity delay calculations, but is including in the "using" component of the calculation!

- Server address space creation delay
- Initiation delays for batch jobs in WLM-managed job classes

The result from the algorithm is multiplied by 100, to yield an execution velocity ranging from 0 (when the address space did not use the CPU) to 100 (when the address space was not delayed for any reason controlled by the SRM).

Many analysts were dismayed that the DASD disconnect time is included in the "using" part of execution velocity. IBM's rationale for including disconnect in the "using" part is that the Workload Manager can take action to reduce delays. Since the WLM cannot take action to decrease disconnect, so there is little point in including disconnect in the "delay" part of the execution velocity calculations.

Unfortunately, the effect of including disconnect time in the execution velocity calculation meant that execution velocities were not a reliable indicator of performance. As the DASD disconnect time increases, the DASD response time would increase (which meant that DASD performance was getting worse). However, the execution velocity algorithm would calculate that performance was improving with higher disconnect time, since disconnect was included in the "using" part of execution velocity.

Even worse, high disconnect time and resulting large execution velocity would cause the Performance Index to be small. The low Performance Index would make it appear the service class is performing better than if disconnect time were not counted. This could allow a service class to become a donor of a I/O resource, when the devices it was using were performing poorly.

Additionally, since the DASD disconnect time often would be large relative to other potential causes of using or delay, the disconnect time would dominate the execution velocity algorithm.

IBM responded to this problem with APAR OW47667, in which disconnect time is no longer counted as productive I/O time. Also, Disconnect time is not counted as I/O delay, because there is nothing the Workload Manager can do to reduce disconnect time.

The APAR states "This change will affect achieved velocities for systems which have the I/O priority queuing option set to ON in the WLM policy. Achieved velocities will be the same or lower, depending on how much disconnect time the service class experiences. Customers should review velocity goals in their WLM policy and adjust downward if needed."

Consequently, CPExpert recommends that you install APAR OW47667 before turning on I/O Priority Management.

- You can "turn off" Rule WLM031 by specifying **%LET WLM031 = OFF.**; in USOURCE(WLMGUIDE) if decide not to implement I/O Priority Management.

- You can disable CPExpert's checking the service definition by modifying the CHKPLCY guidance variable in USOURCE(WLMGUIDE). If the CHKPLCY guidance variable is specified as **%LET CHCKPLCY=Y;**, CPExpert will not check the service definition for potential problems.

  Before you globally disable CPExpert's checking the service definition, you may wish to review other guidance variables. Many of the tests which CPExpert makes can be made inoperative by a guidance variable that applies to the specific test. The discussion of each finding describes the associated guidance variable.

Reference: MVS Planning: Workload Management

OS/390 (V2R9):	Chapter 11.3: Specifying Dynamic Alias Management
OS/390 (V2R10):	Chapter 11.3: Specifying Dynamic Alias Management
z/OS (V1R1):	Chapter 11.3: Specifying Dynamic Alias Management
z/OS (V1R2):	Chapter 11.3: Specifying Dynamic Alias Management
z/OS (V1R3):	Chapter 11.3: Specifying Dynamic Alias Management
z/OS (V1R4):	Chapter 11.3: Specifying Dynamic Alias Management

