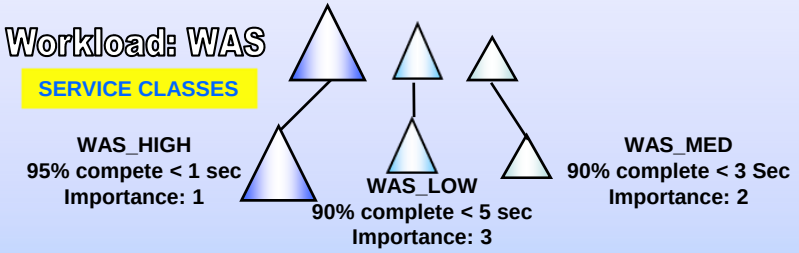
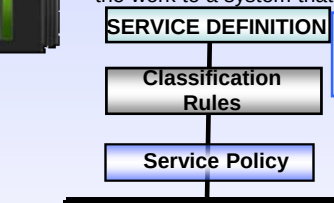
CheatSheet #31 zTidBits z/OS' WLM 101

- The basis of **z/OS' Workload Manager (WLM)** is to make a contract between the installation (user) and the operating system.
  - The installation classifies the work running on the z/OS operating system in distinct service classes and defines goals for them that express the expectation of how the work should perform.
  - WLM uses these goal definitions to manage the work across all systems of a sysplex environment.
  - WLM provides a solution for managing workload distribution, workload balancing, and distributing resources to competing workloads
- With WLM, you define performance goals and assign a business importance to each goal.
- You define the goal for work in business terms and the system decides how much resource such as CPU and storage should be given to that work to meet its goal.
- The WLM uses a **Service Policy**: a named set of overrides to the performance goals and processing boundaries.
  - A policy (rule based) applies to all of the work running across a sysplex, not just an LPAR
  - As processing requirements change at different times, service levels requires change at different times as well.
  - There can be several policies, but there is ONLY *one* policy active at a time.
- Within a workload, you group work with similar performance characteristics into what is called a SERVICE CLASS (SC) such as:
  - **Performance Goals** You assign a goal to each service class period such as a response time goal and a level of *importance* which is how important this work is relative to other work running to achieve its response time goal.
    There are 3 types of goals:
    1. Response time goals indicate how quickly you want your work to be processed. This will be appropriate for a CICS Transactions or web transactions entered from the terminal as a client request. Since response time goals are not appropriate for all kinds of work, such as long running batch jobs or started tasks (STC), the next category is…
    2. Velocity goals which define how fast work should run when it is ready without being delayed for processor, storage, I/O, or queue delay. Most subsystem's address spaces are defined with this goal, CICS, IMS, DB2, WAS , MQ, etc.
    3. Discretionary goals are for low priority work which you do not have any particular performance goal (a.k.a. SYSOTHER). It is under the discretion of z/OS when resources are not being used by any other work will then this goal be provided resources to run.
       NOTE: Generally, all work should be classified and never fall under this category.
  - **Resource Requirements** You can group work into a SC based on resource requirements. As such if you have a group of batch work that can consume vast amounts of resources, and you want to limit it, you can define a service class and assign it to a resource group (RG) with a maximum amount of capacity. If the work exceeds that capacity, WLM slows the execution rate. Also, if a certain group of work needs a minimum amount of processor capacity, you can set up a service class and assign it to a resource group.
    In addition, because some work may have variable resource requirements, you can define multiple *periods* for a SC. Periods are another way of defining different goals for work depending on the amount of resources that work consumes.
    NOTE: Typically, periods are used to give shorter transactions more aggressive goals and to give longer running work of the same type less aggressive goals. If you have multiple periods, each period except the last has a *duration*. A duration is the amount of resources (in service units), that the work consumes before switching the next goal of the next period (P1->P2->P3).
  - **Business Importance** When there is not sufficient capacity for all work in the system to meet its goals business importance is used to determine which work should give up resources and which work should receive more resources. You assign an importance to a service class period which indicates how important it is (relative to other work) that the goal be met.
    NOTE: Importance plays a role '*only*' when a SC period is not meeting its goal.
    There are 5 levels of importance: Lowest(5), Low, Medium, High and Highest (1).

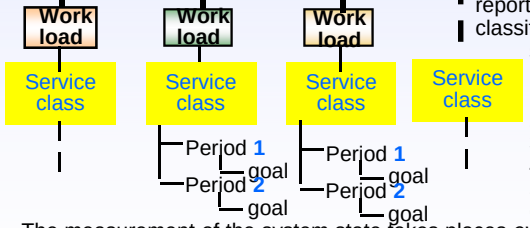


- Scheduling Environment(s) (SE) helps ensure that units of work are sent to the system that has the appropriate resources to handle the workloads.
  - A SE is a list of resource names along with their required status representing physical entities such as a data base or peripheral device or even a period of time (2nd shift or weekend).
  - JES checks the scheduling environment associated with each arriving job and then assigns the work to a system that matches the scheduling environment resources.



You define a minimum of one policy in a service definition. A service definition contains workloads, service classes and classification rules.

- Arriving work is associated with a service class , therefore a goal and potentially a Resource Group (RG). You can get feedback from the Resource Measurement Facility (RMF) on how well the system is doing achieving those goals.



RMF provides sysplex wide workload management reporting displaying any delays. Optionally, classification rules can assign incoming work to a specific report class. Data available for report classes include:

1. The number of transactions completed
2. Average response time
3. Resource usage data
4. State samples
5. Response time distribution buckets
6. Workload delays

- The measurement of the system state takes places *every 250 milliseconds*. At these points, WLM captures the *delay* and *using* states of all dispatchable units in the system and resources that manages. WLM basically manages CPU, DASD I/O devices, the system storage, server address spaces, and to some extent, resource serialization such as GRS enqueues. The states for these resources can be quite varied. While for CPU, there is currently one using and one delay state, it is possible to observe multiple cases why work is delayed because of storage, for example, paging delays, shortages of the page data sets, and much more.
- *Every 10 seconds,* WLM summarizes system state information and combines it with measured data such as CPU service or transaction end times. At this point, WLM *learns the behavior* of the work and the system resources potentially are adjusted. It keeps a certain amount of historical data in storage and maintains graphs learning the relationship between dependent functions in a system.
- The achieved response times and achieved execution velocities are compared with the goal definitions. The result is called the **performance index (PI),** which is the basis for all resource decisions in the system. There are local PIs per LPAR and a Sysplex wide PI.
- The movement of resource adjustment in each service class period is based on **donors/receivers**.
- In order to meet the goals and requirements of the workload requests, the policy adjustment function determines:
  - Which service class periods require more resources to attain their goals.
    > WLM selects one receiver from these candidates. The receiver is the most important work in terms of service class period that needs help meeting its goal.
  - Which resource bottlenecks (delays) are affecting the receiver. A service class period that needs help must already have detected delays. A service class period might be missing its goal, but if the delays are for resources that the WLM does not control, such as locks or enqueue then the service class cannot be helped, and is not considered a receiver.
  - Which donors can donate resource to the receiver or can exchange service level priorities.

In the policy adjustment loop, WLM checks:

- If PI>1, then the service class period is eligible for help and is a candidate to be a RECEIVER.
- If PI<1, then the service class period is a potential DONOR.

Note: If PI = 1 the Service Class period is meeting it's objective.

- **Enclaves** are transactions that can span multiple dispatch able Units (TCBs/SRBs) in one or more address spaces and in the case of multisystem enclaves, one or more address spaces on multiple systems in a sysplex.

Note: A response time objective can be placed on an enclave