Planning for any System z processor, especially image upgrades, consideration must be given to handling the impacts of PR/SM where the following factors should be considered:
- The number of partitions
- The CP usage; either general purpose, traditional CPs or the use of IFL/ICF CPs and z/OS specialty processors.
- The partition mode, either dedicated or shared
- The ratio of logical CPs to physical CPs defined
- The number of logical CPs defined to the partitions
- The type of system control program (z/OS, z/VM, Linux, etc.), and the workload characteristics in each partition.
- The effect of the partitions shared weight and any impact of capping on the partition.

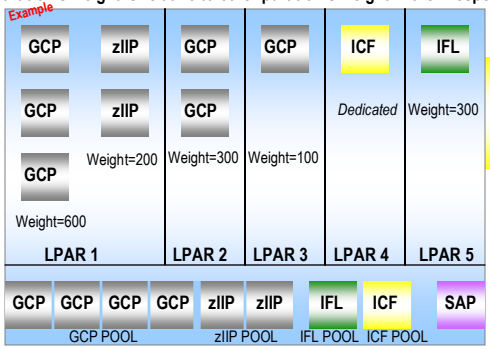
Important Concepts
* LPAR weights become important only when the processor is extremely busy (approaching 100%) or are capped
* There are two dispatchers involved in making resource allocations:
  - PR/SM
  - Operating System

Every partition must have a relative weight, the relative weight is also known as the initial processing weight **(hard cap).** This relative weight is used to determine which LPAR should get resources. However, the relative weight is only considered when there is LPAR contention.

Partitioning Control Basics
* Number of partitions, their relative weights, and CP mode (dedicated or shared)
* Number of logical CPs defined to the partitions
* Horizontal or Vertical CP Management (Hiperdispatch) – *See below*
* Capping Controls
  – Initial Capping (Hard Caps)
  – Defined Capacity (Soft Capping)
  – Group Capacity Controls
* Ratio of logical CPs to physical CPs
* CP usage; either general purpose, or specialty CP (IFL / ICF / zAAP / zIIP) CPs
* Type of system control program (z/OS, z/VM, Linux, etc.)

Soft-Capping is a feature which allows controlling the WLC bill by setting a limit for each LPAR of a MAXIMUM for the average consumption of the MSU. **Group capacity** allows soft capping across multiple LPARs defined in the same LPAR group.

Partition's weight is relative to other partition's weight in their respective pools

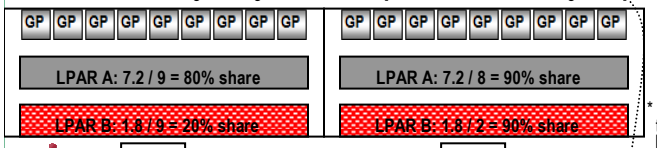


Typically in most cases PR/SM will manage the processing weights to within 1% of the LPAR's physical per CP share.

System z Virtualization
* 1 to 60 LPARs per CEC
* Number of CPs is CEC dependent
  – 1 to 64 for the 2097-E64 (z10)
  – 1 to 80 for the 2817-M80 (z196)
  – 1 to 101 for the 2827-HA1 (zEC12)
* Number of logical CPs is operating system dependent
* Operating System doesn't know it is not running on the hardware
  – More integration is happening over time i.e. HiperDispatch (see right side).
* Dispatching can be done event driven (typical) or time sliced
* Dispatch interval is based on a heuristic method which depends upon the logical to physical ratio.

Relative weight is *unit-less*, that is, they have no unit of measure where the term 'relative' means that the weights are compared against each other.

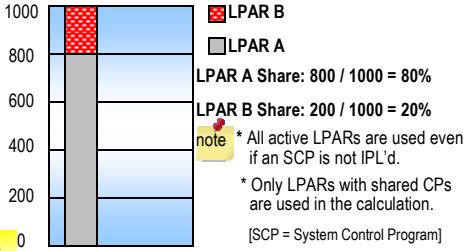
Determine Per CP Share - Horizontal CP Management
* PR/SM guarantees an amount of CPU service to a partition based on weights
* PR/SM distributes a partition's share evenly across the logical processors
* Additional logicals are required to receive extra service which is left by other partitions.
  - Any extra service is also distributed evenly across the logicals
* The OS must run on all logicals to gather all its share [z/OS Alternate Wait Management]



note Biggest per CP share possible is "best" when processor is busy

Calculate LPAR Share

$$\text{SHARE} = \frac{\text{LPAR Weight}}{\text{Sum of Weight}}$$



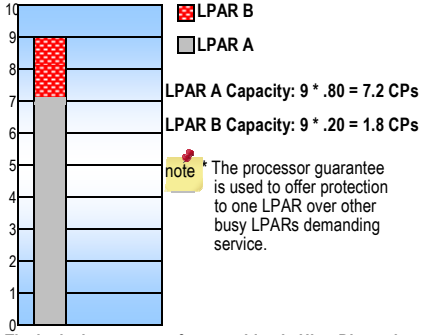
LPAR A Share: 800 / 1000 = 80%

LPAR B Share: 200 / 1000 = 20%

note * All active LPARs are used even if an SCP is not IPL'd.
* Only LPARs with shared CPs are used in the calculation.

[SCP = System Control Program]

Processor guarantee = # of General Purpose Physical (GCP) * LPAR Share



LPAR A Capacity: 9 * .80 = 7.2 CPs

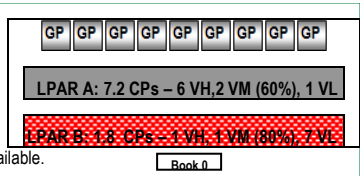
LPAR B Capacity: 9 * .20 = 1.8 CPs

note * The processor guarantee is used to offer protection to one LPAR over other busy LPARs demanding service.

The logical processors for a partition in HiperDispatch mode fall into one of the following categories:
* Some of the logical processors for a partition may receive a 100% processor share, meaning this logical processor receives an LPAR target for 100% share of a physical processor. This is viewed as having a high processor share. Typically, if a partition is large enough, most of the logical partition's share will be allocated among logical processors with a 100% share. PR/SM LPAR establishes a strong affinity between the logical processor and a physical processor, and these processors provide optimal efficiencies in HiperDispatch mode.
* Other logical processors may have a medium amount of physical processor share. The logical processors would have a processor share greater than 0% and up to 100%. These medium logical processors have the remainder of the partition's shares after the allocation of the logical processors with the high share. LPAR reserves at least a 50% physical processor share for the medium processor assignments, assuming the logical partition is entitled to at least that amount of service.
* Some logical processors may have a low amount, or 0%, of physical processor share. These "discretionary" logical processors are not needed to allow the partition to consume the physical processor resource associated with its weight. These logical processors may be parked. In a parked state, discretionary processors do not dispatch work; they are in a long term wait state. These logical processors are parked when they are not needed to handle the partition's workload (not enough load) or are not useful because physical capacity does not exist for PR/SM to dispatch (no time available from other logical partitions).

[If Alternate Weight Management is activated, SRM and LPAR cooperate to reduce low utilization effects and overhead].

Determine Per CP Share - Vertical CP Management
* Logical processors are classified as *vertical high*, *medium* or *low*
* PR/SM quasi-dedicates *vertical high* logicals to physical processors
  - The remainder of the share is distributed to the *vertical medium* processors
* *Vertical low* processors are only given service when other partitions do not use their entire share and there is demand in the partition
* *Vertical low* processors are parked by z/OS when no extra service is available.



System z Partitioning Controls
* *Example* CEC 2827-704: Access to resources is relative to other partitions.

Pool	LPAR Name	Weight	Logicals Defined	Logicals by Weight	Logicals to Physical Ratio	HiperDispatch=YE S
GCP	LPAR1	600	3	2.4		1 VH, 2 VM
GCP	LPAR2	300	2	1.2		2 VM, 60% share
GCP	LPAR3	100	1	0.4		1 VM, 40% share
		1000	6		1.5 : 1	
zIIP	LPAR1	200	2	2		2 VH
		200	2		1:1	
IFL	LPAR5	300	1	1		
		300	1		1:1	
ICF	LPAR4	DED	1	1		
			1	1	1:1	

Logical Processor Utilization
* Measurement which states the busy of the logical CPs
  - Independent measure of capacity
  - Can run out of logical CP capacity before the processor is 100% busy
  - More logical CPs than weight means the utilization is artificially low

Physical Processor Utilization
* Differs from effective time when the number of logical CPs defined to the partition does not match the number of GCPs
* It is this metric which is used in Capacity Planning exercises

Installations have the right to enforce the relative weight, also known as **HARDCAP** or PR/SM CAP. This means never allowing the LPAR to use more that its share of resources as specified in the relative weight. Ex, if you choose to enforce an LPAR at 20% (i.e. 2827-705), than the LPAR could never use more than 20% of the overall system capacity even if the other LPARs are not using their weight.

HiperDispatch and Shared Logical Partitions
* With the introduction of the z990, the zSeries LPAR hypervisor became aware of machine CP topology and began optimizing the allocation of logical partition CP resources to the physical resources accordingly.
  - This support was enhanced in the subsequent family of processors to dynamically re-optimize logical partition resource assignments as configuration changes were made.
  - This support optimized physical CP selection for dispatching logical partition CPs in order to minimize disruption to the various levels of internal caches on the machine.
* zEnterprise, IBM provides a higher level of synergy between the LPAR hypervisor and Linux or z/OS software for managing logical CP resource allocations, which is called *HiperDispatch*.
* The HiperDispatch mode aligns work to a smaller subset of processors to maximize the benefits of the processor cache structures, and thereby, reduce the amount of CPU time required to execute work. Access to processors has changed with this mode, and as a result, prioritization of workloads via WLM policy definitions becomes more important.
* Without HiperDispatch, for all levels of z/OS, a TCB or SRB may be dispatched on any logical processor of the type required (standard, zAAP or zIIP).
  - A unit of work starts on one logical processor and subsequently may be dispatched on any other logical processor.
  - The logical processors for one LPAR image will receive an equal share for equal access to the physical processors under PR/SM LPAR control.
    > Example, if the weight of a logical partition with four logical processors results in a share of two physical processors, or 200%, the LPAR hypervisor will manage each of the four logical processors with a 50% share of a physical processor. All logical processors will be used if there is work available, and they typically have similar processing utilizations.
* With HiperDispatch mode, work can be managed across fewer logical processors.

note A concept of maintaining a working set of processors required to handle the workload is available..

* In the previous example of a logical partition with a 200% processor share and four logical processors, two logical processors are sufficient to obtain the two physical processors worth of capacity specified by the weight; the other two logical processors allow the partition to access capacity available from other partitions with insufficient workload to consume their share. z/OS limits the number of active logical processors to the number needed based on partition weight settings, workload demand and available capacity.
  - z/OS also takes into account the processor topology when dispatching work, and it works with enhanced PR/SM microcode to build a strong affinity between logical processors and physical processors in the processor configuration.