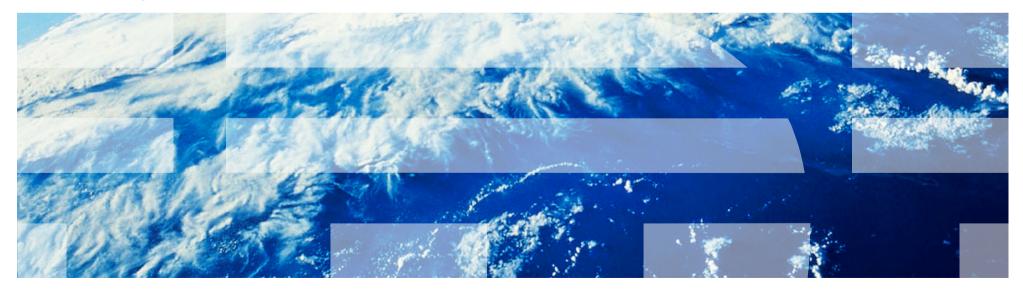


Configuring LPARs for Performance

Session:13101



Kathy Walsh
IBM Corporation





Agenda

Overview of Terms and Partitioning Controls

- Per CP Share
- Short CPs

Managing Capacity

- Intelligent Resource Director
- Initial Capping
- Soft Capping
- Group Capacity



Important Terms to Understand

- LPAR weight and per CP share
- Effective Dispatch Time
- Partition Dispatch Time
- Short CPs

Important Concepts to Understand

- LPAR weights become important only when the processor is very busy or capped
- There are two dispatchers involved in making resource allocations
 - PR/SM
 - Operating System



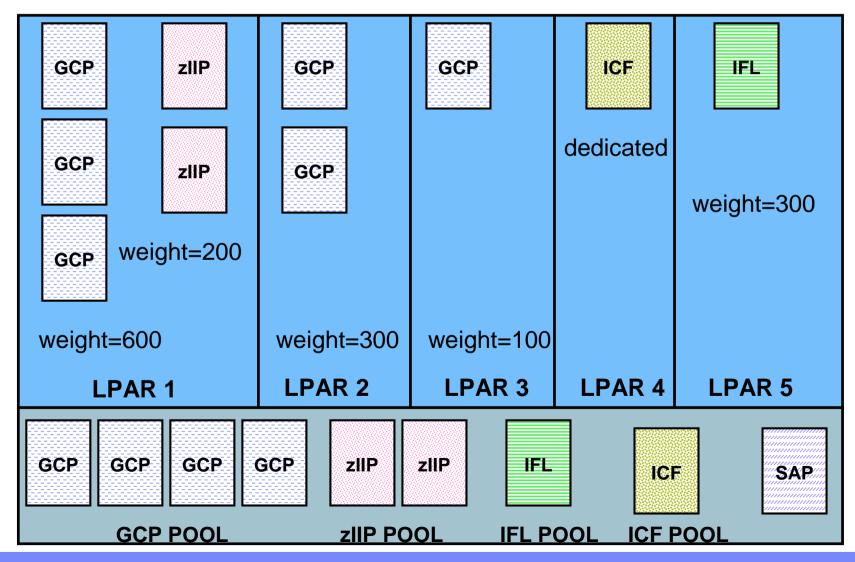
Partitioning Controls

- 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)
- 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.)



Partitioning Controls

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



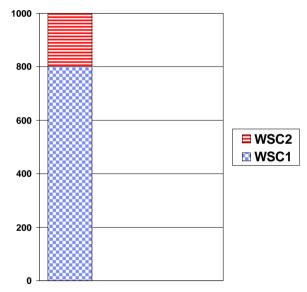


System z Virtualization

- 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
- 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



Calculate LPAR Share



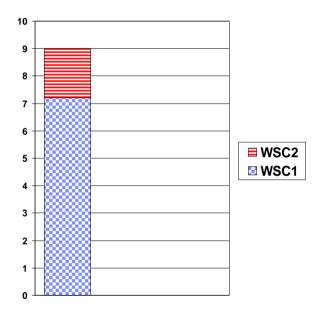
WSC1 Share: 800 / 1000 = 80%

WSC2 Share: 200 / 1000 = 20%



- All active LPARs are used even if an SCP is not IPL'ed
- Only LPARs with shared CPs are used in the calculation

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



WSC1 Capacity: 9 * .80 = 7.2 CPs

WSC2 Capacity: 9 * .20 = 1.8 CPs

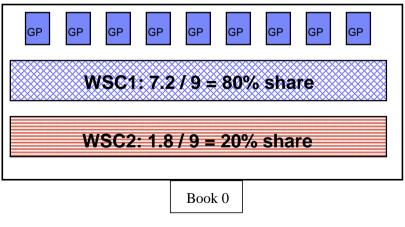


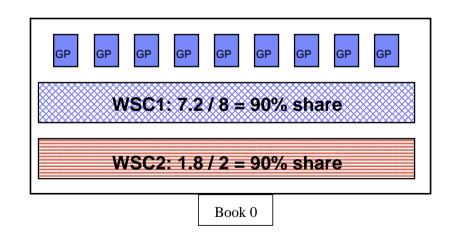
 The processor guarantee is used to offer protection to one LPAR over other busy LPARs demaning service



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 <u>evenly across</u> the logical processors
- Additional logicals are required to receive extra service which is left by other partitions. The extra service is also distributed <u>evenly across</u> the logicals
- The OS must run on all logicals to gather all its share [z/OS Alternate Wait Management]





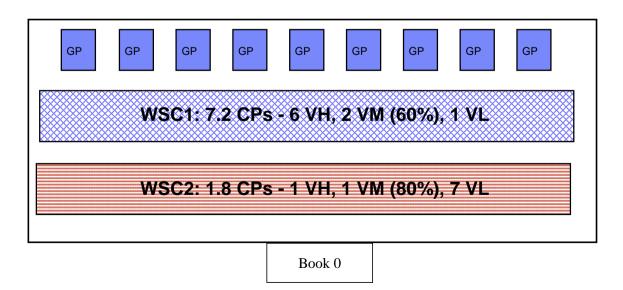


Biggest Per CP Share possible is best when processor is busy



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

 Access to resources is relative to other partitions on the CEC 2827-704

| Pool | LPAR Name | Weight | Logicals Defined | Logicals by Weight | Logical to Physical Ratio | HD=YES |
|------|--------------|--------|---------------------|--------------------------|---------------------------------|-----------------|
| 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

| | - PAR | TITION | I DATA | | | | L | OGICAL | PARTITION PROC | ESSOR DATA | AVERAGE | PROCESSO | R UTILIZATI | ON PERCENTA | AGES |
|----------|------------|--------|--------|-----|------|------|------|--------|----------------|--------------|-------------|----------|-------------|-------------|-------|
| | | | MS | U | -CAF | PING | PROC | ESSOR- | DISPATCH | TIME DATA | LOGICAL PRO | CESSORS | PHYSIC | AL PROCESSO | ORS |
| NAME | S | WGT | DEF | ACT | DEF | WLM% | NUM | TYPE | EFFECTIVE | TOTAL | EFFECTIVE | TOTAL | LPAR MGMT | EFFECTIVE | TOTAL |
| WSC1 | A | 370 | 0 | 700 | NO | 0.0 | 15.0 | CP | 01.45.57.466 | 01.46.19.021 | 47.09 | 47.25 | 0.10 | 28.26 | 28.35 |
| WSC2 | A | 315 | 0 | 288 | NO | 0.0 | 15.0 | CP | 00.43.23.443 | 00.43.46.035 | 19.28 | 19.45 | 0.10 | 11.57 | 11.67 |
| WSC3 | А | 315 | 0 | 178 | NO | 0.0 | 15.0 | CP | 00.26.39.732 | 00.27.00.535 | 11.85 | 12.00 | 0.09 | 7.11 | 7.20 |
| WSC4 | А | 25 | 45 | 4 | NO | 0.0 | 2.0 | CP | 00.00.32.779 | 00.00.34.362 | 1.82 | 1.91 | 0.01 | 0.15 | 0.15 |
| PHYSICAL | L * | | | | | | | | | 00.01.05.674 | | | 0.29 | | 0.29 |
| | | | | | | | | | | | | | | | |
| TOTAL | | | | | | | | | 02.56.33.422 | 02.58.45.630 | | | 0.59 | 47.08 | 47.67 |



RMF Partition Report

| MVS PARTITION NAME | WSC1 | NUMBER OF PHYSICAL PROCESSORS | 31 |
|---------------------------------|---------|-------------------------------|----|
| IMAGE CAPACITY | 2469 | CP | 25 |
| NUMBER OF CONFIGURED PARTITIONS | 17 | IFL | 1 |
| WAIT COMPLETION | NO | ICF | 2 |
| DISPATCH INTERVAL | DYNAMIC | IIP | 3 |

| | PART | TITION | DATA | | | | LO | OGICAL |
|-----------|------|--------|------|-----|------|------|-------|--------|
| | | | MS | SU | -CAP | PING | PROCI | ESSOR- |
| NAME | S | WGT | DEF | ACT | DEF | WLM% | NUM | TYPE |
| WSC1 | А | 370 | 0 | 700 | NO | 0.0 | 15.0 | CP |
| WSC2 | А | 315 | 0 | 288 | NO | 0.0 | 15.0 | CP |
| WSC3 | А | 315 | 0 | 178 | NO | 0.0 | 15.0 | CP |
| WSC4 | А | 25 | 45 | 4 | NO | 0.0 | 2.0 | CP |
| *PHYSICAL | * | | | | | | | |
| | | | | | | | | |
| TOTAL | | | | | | | | |
| | | | | | | | | |
| CF01 | A | DED | | | | | 2 | ICF |
| *PHYSICAL | * | | | | | | | |
| | | | | | | | | |
| TOTAL | | | | | | | | |
| | | | | | | | | |
| WSC1 | А | 10 | | | | | 3 | IIP |
| WSC2 | Α | 10 | | | | | 3 | IIP |
| WSC3 | Α | 10 | | | | | 3 | IIP |
| WSC4 | Α | 10 | | | | | 1 | IIP |
| *PHYSICAL | * | | | | | | | |

Processor Running Time

- Default is limited to a range of 12.5-25 ms
- Dynamically calculated

25 ms * (Number of Physical Shared CPs) Total # of Logical CPs for all LPARs

- Vertical Highs get run time of 100 ms
- Recalculated when LPARs are stopped or started or CPs are Configured on/off
- When a logical CP does not go into a wait state during its run time, it loses the physical CP when it reaches the end of its run time



Managing Capacity on System z and z/OS

- Intelligent Resource Director
- PR/SM Initial Capping Hard Capping
- Defined Capacity Soft Capping
- Group Capacity
- Other Methods of Changing Capacity
 - WLM Resource Groups
 - Discretionary Goal Management
 - Config CPU Command
 - Customer Initiated Power Save Mode
 - OOCoD



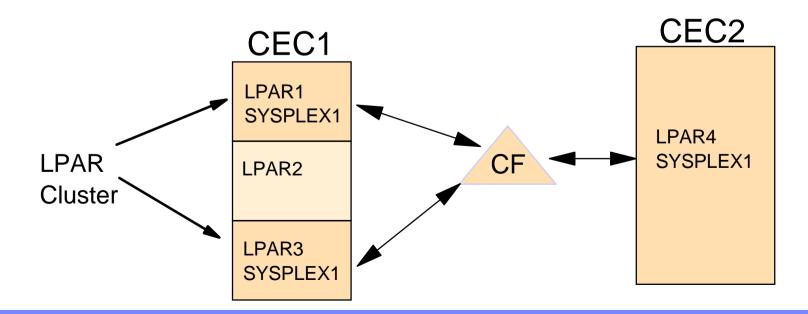
Intelligent Resource Director

WLM Function which:

- Manages LPAR Weights
- Varies logical CPs On and Off Disabled and replaced with Hiperdispatch=YES
- Manages CHPIDs
- Manages I/O Priorities

Scope is an LPAR Cluster

All MVS images on the same physical processor, in the same sysplex





IRD Management

WLM manages physical CPU resource across z/OS images within an LPAR cluster based on service class goals

- LPAR Weight Management
 - Dynamic changes to the LPAR weights
 - Sum of LPAR weights can be redistributed within the cluster
 - Partition(s) outside of the cluster are not affected
 - Moves CP resource to the partition which requires it
 - Reduces human intervention

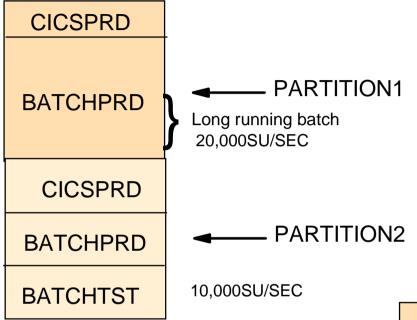
LPAR Vary CP Management

- Dynamic management of online CPs to each partition in the cluster
- Optimizes the number of CPs for the partition's current weight
- Prevent 'short' engines
- Maximizes the effectiveness of the MVS dispatcher
- Has an IEAOPTxx option (VARYCPUMIN) to set minimum number of CPs regardless of LPAR's weight
- Reduces human intervention
- Replaced by Hiperdispatch=yes

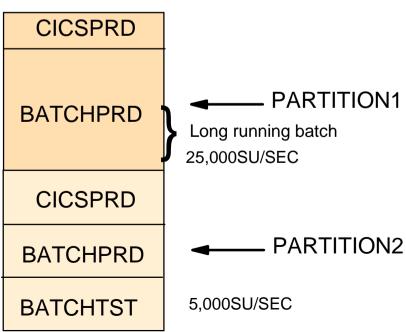


Benefit of LPAR IRD Management

This



Becomes





LPAR Weight Management

- Enabled using PR/SM server processor definition panels
 - WLM managed
 - Initial processing weight
 - Minimum processing weight
 - Maximum processing weight
- Weights should be 2 or 3 digit values to provide optimal results
 - Weight is increased by 5% of average LPAR weight value

| LPAR | Weight | | |
|----------|--------|--|--|
| Receiver | 400 | | |
| Donor | 600 | | |

Weight Donation

| LPAR | Weight |
|----------|--------|
| Receiver | 425 |
| Donor | 575 |



LPAR Weight Mgmt Requirements

- LPARs have shared CPs
- LPARs do not use Initial capping (Hard Caps)
- WLM LPAR CF structure is defined and connected
- LPAR CPU management is enabled on LPAR panels or Hiperdispatch=YES
- System must be running z/OS
- Processor must be a zSeries or System z processor



IRD Performance Management

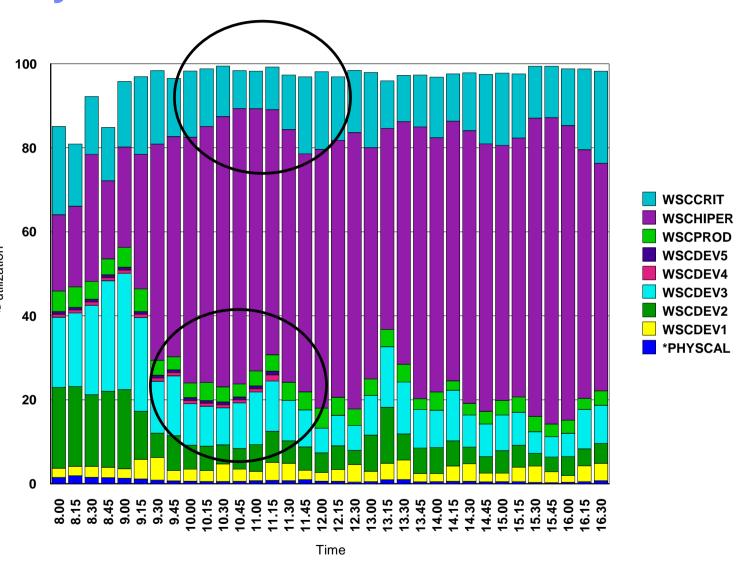
- Need to have a multi system perspective when looking at overall throughput in an LPAR Cluster
 - WLM Policy and Goal Attainement
- Need to examine CEC demand within and outside the cluster
 - Whitespace

19



Overall CEC Busy

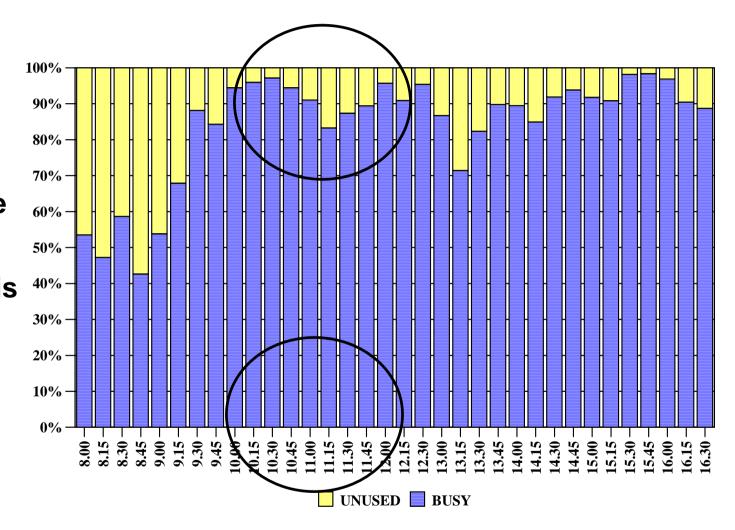
- IRD is active
- WSCCRIT,
 WSCHIPER and
 WSCPROD are in a LPAR Cluster called WSCPLEX
- When WSCDEV5 are and WSCDEV5 are there the LPAR Cluster gets 82% of the CEC, when they are stopped the LPAR Cluster gets 89%
- CEC is very busy





WSCPLEX Cluster View of Capacity

- This chart represents the fair share of the CEC that the WSCPLEX cluster should have access to
- WSCPLEX Cluster is not using all of its capacity so is donating white space to the other LPARs





WSCCRIT Performance Degradation

| HOUR | SCLASS | IMP | СР | PINDX |
|------|----------|-----|------|-------|
| 10 | SYSTEM | 0 | 10.5 | 1 |
| 10 | SYSSTC | 0 | 17.9 | 1 |
| 10 | TSOL1 | 1 | 1.5 | 0.5 |
| 10 | TRANONE | 1 | 2.6 | 0.1 |
| 10 | DMGMT | 1 | 4 | 1.4 |
| 10 | SERVERS | 1 | 23.3 | 25.5 |
| 10 | CICSL2 | 1 | 0.1 | 3.4 |
| 10 | STCHI | 1 | 1.2 | 2 |
| 10 | TSOHI | 1 | 1.3 | 0.6 |
| 10 | TRANTWO | 1 | 8.9 | 0.1 |
| 10 | TRNMULT | 1 | 6.5 | 0.3 |
| 10 | TRNMULT | 2 | 28.6 | 1.4 |
| 10 | STC2 | 2 | 46.7 | 1.1 |
| 10 | TRANFIVE | 2 | 3.5 | 3.4 |
| 10 | TSOHI | 2 | 0.2 | 1.9 |
| 10 | TSOI1 | 2 | 1.4 | 13.3 |
| 10 | TRANONE | 2 | 12.5 | 3.6 |
| 10 | HOTPROD | 2 | 0 | 0.1 |
| 10 | DBASE | 2 | 0 | 0.1 |
| 10 | BATCHL1 | 3 | 52.2 | 2 |
| 10 | TSOL1 | 3 | 1.6 | 1.3 |
| 10 | DBASE | 3 | 0 | 1.2 |
| 10 | BATCHL2 | 5 | 2.7 | 3.7 |

| HOUR | SCLASS | IMP | CP | PINDX |
|-------|---------|-----|------|--------|
| 11.15 | SYSTEM | 0 | 14 | 1 |
| 11.15 | SYSSTC | 0 | 21.7 | 1 |
| 11.15 | TSOL1 | 1 | 1.4 | 1.3 |
| 11.15 | TSO1 | 1 | 1 | 0.3 |
| 11.15 | TRANONE | 1 | 5.1 | 1.1 |
| 11.15 | SERVERS | 1 | 17.5 | 2.9 |
| 11.15 | TRANTWO | 1 | 1.1 | 3.4 |
| 11.15 | CICSL2 | 1 | 0 | 72.1 |
| 11.15 | DMGMT | 1 | 3 | 3.9 |
| 11.15 | STCHI1 | 1 | 3.7 | 5.3 |
| 11.15 | TRNMULT | 1 | 9.8 | 1.7 |
| 11.15 | STC2 | 2 | 25.2 | 12.7 |
| 11.15 | TRNMULT | 2 | 24.7 | 107.2 |
| 11.15 | TRANTWO | 2 | 2.7 | 236.5 |
| 11.15 | TSOL1 | 2 | 1 | 4.2 |
| 11.15 | TSOHI | 2 | 0.8 | 30.9 |
| 11.15 | TRANONE | 2 | 9.7 | 198.9 |
| 11.15 | DBASE | 2 | 0 | 14.5 |
| 11.15 | BATCHL1 | 3 | 0 | 5138.8 |
| 11.15 | DBASE | 3 | 0 | 160.3 |



WSCPLEX LPAR Share Within the Cluster

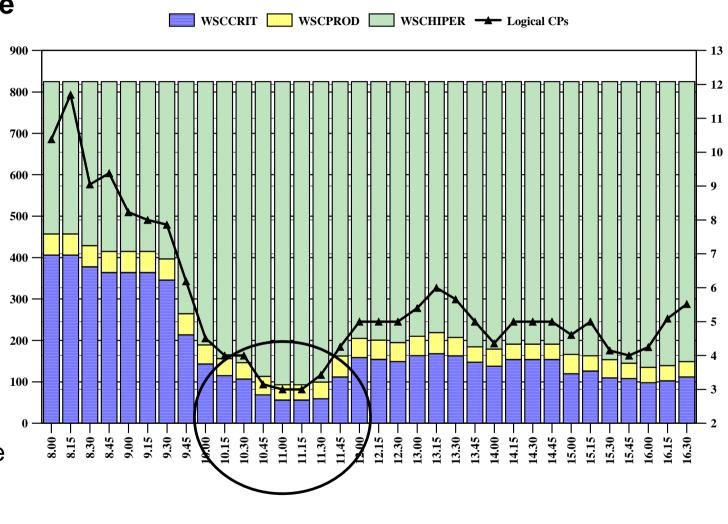
 This chart shows the change in LPAR weight over time

WSCCRIT is losing weights and logical CPs

Two issues to examine:

Why did we donate white space?

WSCCRIT suffers
 performance
 problems but did the
 benefit to
 WSCHIPER
 outweigh the costs?





WSCHIPER Gets Additional Weight

LPAR Weight

| | 10.00 | 11.15 |
|----------|-------|-------|
| WSCHIPER | 636 | 732 |
| WSCCRIT | 143 | 56 |

LPAR Busy

| | 10.00 | 11.15 |
|----------|-------|-------|
| WSCHIPER | 58.5 | 62.5 |
| WSCCRIT | 15.7 | 8.93 |

- WSCHIPER gets more weight but doesn't do more work
- High PIs makes IRD hesitant to move weight back
- High CEC Busy means no additional logicals can be added to WSCCRIT
- Low number of logical CPs means WSCCRIT can't schedule the work and hence the whitespace is donated

| HOUR | SCLASS | IMP | СР | PINDX |
|-------|---------|-----|-------|-------|
| 10 | SYSSTC | 0 | 34.1 | 1 |
| 10 | SYSTEM | 0 | 20.2 | 1 |
| 10 | CICSL1 | 1 | 307.9 | 1 |
| 10 | CICSL2 | 1 | 182.4 | 1.1 |
| 10 | CICSL3 | 1 | 81.6 | 1.2 |
| 10 | SERVERS | 1 | 59.6 | 1.3 |
| 10 | STCHI | 1 | 12.7 | 1.4 |
| 10 | OMVS | 2 | 0.1 | 0.4 |
| 10 | STC2 | 2 | 33.9 | 1.3 |
| 10 | BATCHL1 | 3 | 135.2 | 1.4 |
| 10 | STCLO | 3 | 1.3 | 2.4 |
| 10 | TSOL1 | 3 | 0.2 | 0 |
| 10 | BATCHL2 | 5 | 5 | 2.2 |
| 11.15 | SYSSTC | 0 | 35.9 | 1 |
| 11.15 | SYSTEM | 0 | 31.3 | 1 |
| 11.15 | CICSL1 | 1 | 315.8 | 1 |
| 11.15 | CICSL2 | 1 | 193.7 | 1 |
| 11.15 | CICSL3 | 1 | 78.2 | 1.1 |
| 11.15 | SERVERS | 1 | 53.4 | 1.3 |
| 11.15 | STCHI | 1 | 20.7 | 1.2 |
| 11.15 | OMVS | 2 | 0.8 | 0.3 |
| 11.15 | STC2 | 2 | 5 | 1.1 |
| 11.15 | BATCHL1 | 3 | 118.3 | 1.5 |
| 11.15 | STCLO | 3 | 1.4 | 1.5 |
| 11.15 | TSOL1 | 3 | 0.1 | 0 |
| 11.15 | BATCHL2 | 5 | 9.4 | 1.2 |



What are the Tuning Options

- If Hiperdispatch = No then use VaryCPUMIN to keep sufficient logicals available
- Update the WLM policy so the goals are more reasonable
- Provide protection with IRD Minimum values



Providing Protection in an IRD Environment

- Decisions across LPARs are based on WLM Policy
 - Ensure WLM definitions are well defined and accurate
 - Review Performance data at the LPAR Cluster level
- Protection comes from the use of MIN weights
 - Special protection for LPARs with high percentage of work which can be donated

Initial Weight = 6 CPs

| IMP | CPs |
|-----|-----|
| 0 | |
| 1 | |
| 2 | |
| 3 | |
| 4 | 1.3 |
| 5 | 0.8 |
| 6 | 0.4 |

Initial Weight = 3 CPs

Min Weight = 0.5 CPs

| IMP | CPs | | | | | |
|---------|-----|--|--|--|--|--|
| 0 | 0.4 | | | | | |
| 4 | 0.5 | | | | | |
| 5 | 1.3 | | | | | |
| 6 | 0.8 | | | | | |
| 2.6 CPs | | | | | | |

Initial Weight = 10 CPs

| IMP | CPs |
|-----|-----|
| 0 | |
| 1 | |
| 2 | |
| 3 | |
| 4 | 2.3 |
| 5 | 1.2 |
| 6 | 0.9 |

2.5 CPs _____

Min Weight = 3.5 CPs

4.5 CPs

Min Weight = 5.5 CPs



PR/SM Initial Capping – Hard Cap

- HiperDispatch=No
 - The LPAR's relative weight per CP is the share for each logical CP and the goal of the LPAR dispatcher is to give each logical CP its share of the total relative weight
 - Capping is done on a logical CP basis
- Hiperdispatch=YES
 - Vertical High's will be capped at 100% of the logical
 - Vertical Mediums and Vertical Lows will share the allowed weight on a per CP basis



PR/SM – Weight Enforcement

Weight Enforcement Depends Upon LPAR definitions

- LPAR with Initial Capping
 - Enforces processing weights to within 3.6% of the LPAR's physical per CP share for logical CPs entitled to <u>1/10 or more</u> of one physical CP
- LPAR is Uncapped
 - Enforces the processing weights to within 3.6% of the LPAR's physical per CP share for logical CPs entitled to 1/2 or more of one physical CP
- LPAR Logical CP fails enforcement levels
 - Enforce the processing weights to within 3.6% of the total capacity of the shared physical CP resources
- Typically in most cases PR/SM will manage the processing weights to within 1% of the LPAR's physical per CP share



PR/SM Initial Capping – Weight Allocations

- An LPAR's Hard Capped Capacity is relative to the other LPARs
 - If an LPAR is started or stopped on a CEC with a hard cap a weight change must be done concurrently or the capped LPAR's allowed capacity will change
 - With Hiperdispatch you need to deactivate the LPAR so the VHs are reallocated correctly otherwise VLs will be used
 - WSC2 needs to go from 4 VH, 2 VM to 12 VH, 1 VM

2817-718

| Name | Status | Weight | Capped | Weight in CPs | Status | Weight | Capped | Weight in CP |
|------|--------|--------|--------|---------------|--------|--------|--------|--------------|
| WSC1 | A | 400 | NO | 7.2 | D | | | |
| WSC2 | Α | 300 | NO | 5.4 | Α | 300 | NO | 9 |
| WSC3 | Α | 300 | YES | 5.4 | A | 300 | YES | 9 |



Defined Capacity – Soft Capping

Specified via LPAR definitions

- Provides sub-CEC pricing by allowing definition of LPAR capacity in MSUs
 - Allows a defined capacity smaller than the total capacity of the LPAR
 - Provides 1 MSU of granularity
- Only way to get a soft cap
- Initial Capping (PR/SM Hard Cap) and Defined Capacity cannot be defined for the same partition
 - Rule applies to GCPs only, Specialty CPs can be hard capped while the GCPs are soft capped
- LPAR must be using Shared CPs (Dedicated CPs are not allowed)

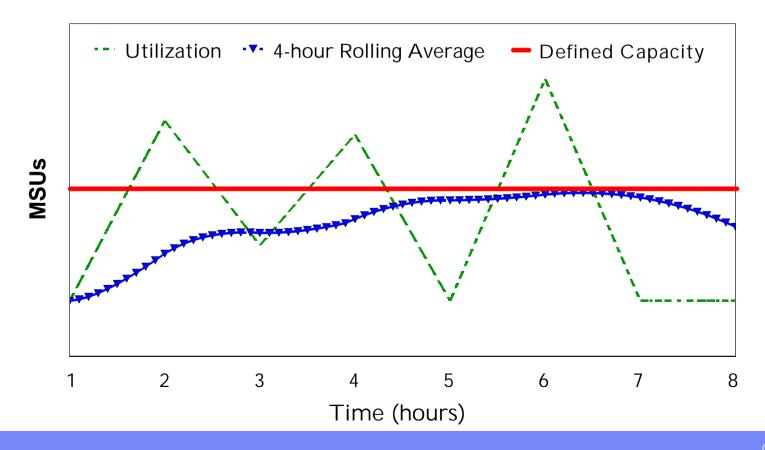
All sub-capacity software products in an LPAR have the same capacity

- LPAR partition capacity based pricing not product usage based pricing
- Regardless of actual CPU seconds used by the product



Rolling 4 hr Average & Defined Capacity

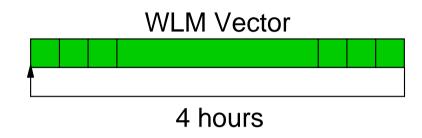
- Rolling 4-hour average tracked by Workload Manager
 - Rolling 4-hour average is not permitted to exceed defined capacity
 - May exceed during early capping intervals
 - If 4-hour average exceeds defined capacity, LPAR gets soft capped



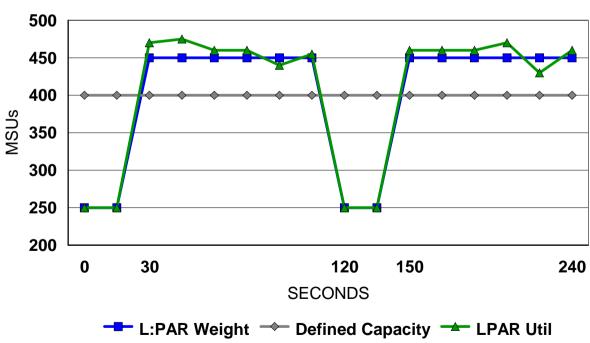


Managing to the Rolling 4hr Average

- LPAR Effective dispatch time for partition averaged in 5 min intervals
 - 48 entry vector
 - Every 10 seconds WLM issues Diagnose command to hardware to get effective time
 - Vector wraps after 4 hours
- Calculate a capping pattern
 - Control ratio of capped versus non-capped periods to keep partition usage at defined capacity
 - Capping state should change no more than once per minute
 - Limit partition to it's weight



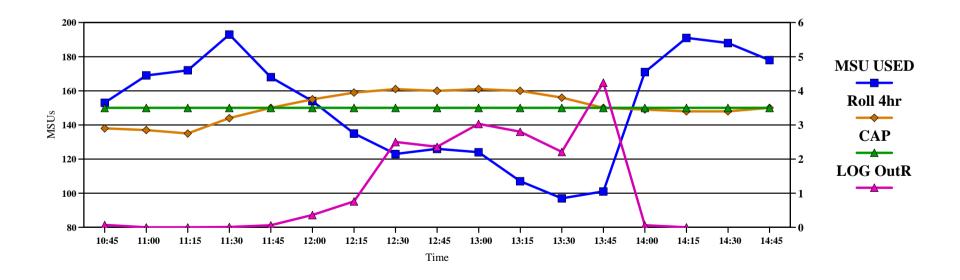






Managing to the Rolling 4hr Average

- When softcapped the LPAR is allowed to continually use the amount of capacity defined
- Work is not stopped to "make up" for time period when rolling 4hr average exceeds the defined capacity





LPAR Group Capacity Basics

Manage CPU for a group of z/OS LPARs on a single CEC

- Limit is set to total usage by all LPARs in group
 - Level of granularity is 1 MSU
 - Members which don't want their share will donate to other members
- Independent of sysplex scope and IRD LPAR cluster
- Works with defined capacity limits on an LPAR
 - Target share will not exceed defined capacity
- Works with IRD
- Can have more than one group on a CEC but an LPAR may only be a member of one group
- LPARs must share engines and specify WAIT COMPLETION = NO

Capacity groups are defined on the HMC Change LPAR Group Controls panels

- Specify group name, limit in MSUs, and LPARs in the group
- Members can be added or removed dynamically



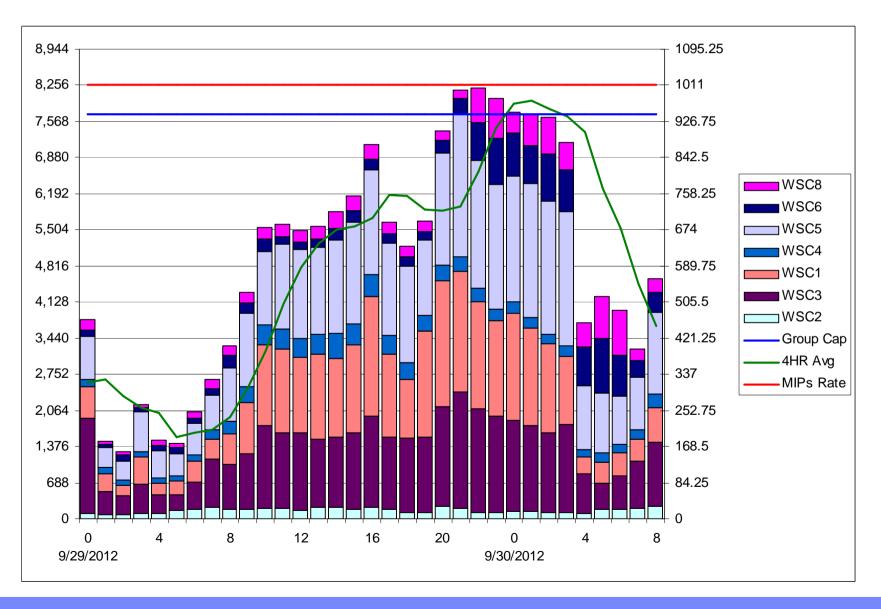
LPAR Group Capacity Basics

- Uses WLM rolling 4 hr avg in MSUs to manage the Group Capacity limit
 - Cap enforced by PR/SM if group rolling 4 hr avg exceeds limit
 - Each member is aware of other members' usage and determines its share based on its weight as a percentage of total weight for all members in group
 - NOTE: When using IRD the weights can change and therefore the target MSU value can change
 - The defined capacity limit, if also specified, is never exceeded
- Until members "learn" about the group and build a history, the cap is not enforced
 - May take up to 4 hours (48 measurements at 5 minute intervals are maintained for rolling 4 hour average) for capping to start
 - Similar to the bonus period with defined capacity
 - When new member joins the group, it has to build up its history and during this time the group usage may exceed the capacity limit
 - Capping is removed when the group rolling 4 hour average drops below group limit
- Example shows how many MSUs each LPAR would get if they all wanted their share. Target MSUs based on a group limit of 200. Total group weight is 500.

| LPAR | WEIGHT | SYSPLEX | CAPACITY GROUP | TARGET MSU |
|-------|--------|---------|----------------|------------|
| LPAR1 | 150 | PLEX1 | GROUPA | 60 |
| LPAR2 | 300 | PLEX2 | GROUPA | 120 |
| LPAR3 | 500 | PLEX1 | n/a | n/a |
| LPAR4 | 50 | PLEX1 | GROUPA | 20 |



Example of Group Capacity





RMF Group Capacity Reporting

- CAPPING WLM% (percentage of time WLM capped the partition) is insufficient when the partition is member of a capacity group:
 - WLM% only tells to what extent a partition is <u>subject to capping</u> but not whether the partition was actually capped
 - WLM% is more a matter of <u>how</u> WLM caps the partition instead of <u>how much</u> it is being capped
- CAPPING ACT% displays the percentage of time where the partition was actually capped
 - Users of Capacity Groups can determine the available (unused) capacity for their group and whether the partition was actually capped:

| NUMBER OF PHYSICAL PROCESSORS | | | | | 6 | | | GROUP | NAME | ATS | | |
|-------------------------------|----------|-----------|--------|-----|-----|-----|---------------|---------|-------|----------|---------|--|
| CP | | | | | | 5 | | | LIMIT | | 141 | |
| IIP | | | | | | 1 | 1 AVAILABLE 1 | | | | 1 | |
| GROUP-0 | CAPACITY | PARTITION | SYSTEM | MS | SU | WGT | | CAPPING | G | - ENTITL | EMENT - | |
| NAME | LIMIT | | | DEF | ACT | | DEF | WLM% | ACT% | MINIMUM | MAXIMUM | |
| ATS | 141 | WSC1 | WSC1 | 0 | 0 | 25 | NO | 0.0 | 0.0 | 7 | 141 | |
| | | WSC2 | WSC2 | 0 | 85 | 380 | NO | 87.5 | 13.1 | 119 | 141 | |
| | | WSC3 | WSC3 | 0 | 24 | 25 | NO | 0.0 | 0.0 | 7 | 141 | |
| | | WSC4 | WSC4 | 0 | 2 | 20 | NO | 0.0 | 0.0 | 6 | 20 | |
| | | | | | | | | | | | | |
| | | | TOTAL | | 111 | 450 | | | | | | |



Intersection of IRD and Group Capacity

- OA29314 DOC IRD and Group Capacity
- WLM only manages partitions in a Group Capacity which meet the following conditions:
 - Partition must not be defined with dedicated processors
 - Partition must run with Shared processors and Wait Complete=No must be set
 - Operating System must be z/OS 1.8 and above
 - z/OS cannot be running as a z/VM Guest
 - PR/SM Hard Capping is not allowed
- Any LPAR not meeting the conditions is removed from the Group and the remaining LPARs are managed to the Group Limit
- Group Capacity will function with IRD weight management as long as the partitions in the Group are not subject to capping
 - No Weight moves will take place as long as the Group is being capped



Enhanced SMF Recording

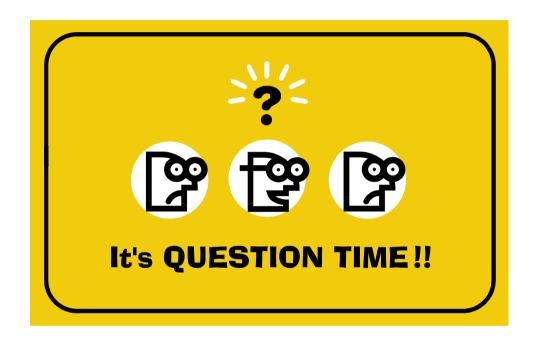
- It is recommended to turn on recording of SMF 99 subtype 11 when you start to exploit group capping
 - The collected data is small and only written every 5 minutes
 - Size is about 1300 bytes fixed + 240 bytes per LPAR on a CEC
 - Approximately 3k for a CEC with 8 partitions
 - The data is crucial for all analysis done by IBM therefore recommend the data be collected unconditionally



Summary

- LPAR controls are important in controlling capacity available to workloads
- IRD weight management is still valuable if you have the right environment
 - Measure and manage at the LPAR Cluster level
- Capping controls are inter-related and can be used to control overall CEC capacity
 - Be aware of the impacts on performance









Trademarks

The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*BladeCenter®, DB2®, e business(logo)®, DataPower®, ESCON, eServer, FICON, IBM®, IBM (logo)®, MVS, OS/390®, POWER6®, POWER6+, POWER7®, Power Architecture®, PowerVM®, S/390®, Sysplex Timer®, System p®, System p5, System x8, System z9, System z10®, Tivoli®, WebSphere®, X-Architecture®, zEnterprise®, z9®, z10, z/Architecture®, z/OS®, z/VM®, z/VSE®, zSeries®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are registered trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

All other products may be trademarks of registered trademarks of their respective companies

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.



Notice Regarding Specialty Engines (e.g., zIIPs, zAAPs and IFLs):

Any information contained in this document regarding Specialty Engines ("SEs") and SE eligible workloads provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs). IBM authorizes customers to use IBM SEs only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at:

www.ibm.com/systems/support/machine_warranties/machine_code/aut.html ("AUT").

No other workload processing is authorized for execution on an SE.

IBM offers SEs at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.