z/OS Workload Manager: Intelligent Resource Director (IRD)

G. Tom Russell
IBM Canada Ltd.
Tom_Russell@ca.ibm.com
CMG Canada
April 24, 2002
Toronto, Ontario

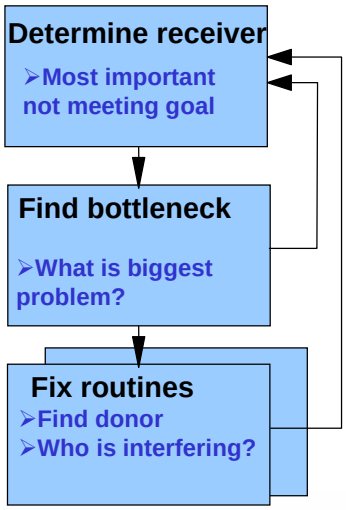
Agenda

- Brief Introduction to MVS Workload Management (WLM) Externals
  - Priority I/O queueing
  - LPAR CPU management
  - Dynamic channel management (DCM)
- Summary

WLM Vocabulary

- Service Definition
- Policy
- Classification Rules
- Service Class (SC)
  - Period (SCP)
  - Goal

WLM Policy Adjustment



Intelligent Resource Director

- Extension of WLM management of Goals
  - Manage hardware resources of Logical Partitions
    - Number of CPUs
    - Weight
    - Number of Channel Paths
    - I/O Priority

IRD – CSS I/O Priority Queueing

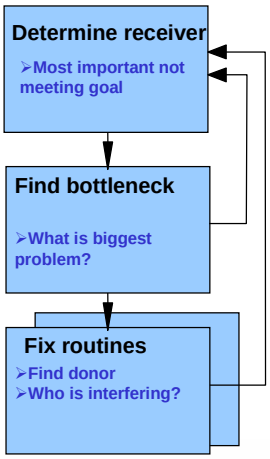
- Already have priority I/O queueing on UCB and inside ESS
- CSS IOPQ now in channel subsystem
- Works with DCM to ensure 'right' user gets additional capacity when paths are added to LCU
- All device types

WLM role in I/O priority queueing

- WLM assigns UCB and CU I/O priority so that:
  - System-related SCP is assigned highest priority
  - An SCP missing its goal because of I/O delay gets helped
  - An SCP competing for the same devices with a more important, and not happy (PI > 1), SCP is the donor
- WLM assigns CSS I/O priorities so that:
  - System-related SCP is assigned highest priority
  - High Importance SCP missing goals has next highest
  - SCPs meeting goals are managed so that light I/O SCPs have a higher I/O priority than heavy I/O users
  - Discretionary work has the lowest priority
- I/O weight = ratio of Connect Time to Elapsed Time

I/O Priority Management

- Managing I/O Delay
  - Donor/Receiver Logic
- Device Clustering Algorithm
  - Usage helps determine Donor
  - New I/O priority broadcast to other systems
- CSS priority based on I/O weight



Assigning UCB and CU I/O priorities

Priority	Type of work
FF	SYSTEM service class
FE	SYSSTC service class
F9-FD	Managed by WLM Policy Adjustment
F8	Discretionary

Assigning CSS I/O priorities

Priority	Type of work
FF	**System work**
FE	**Importance 1 and 2 missing goals**
FD	**Importance 3 and 4 missing goals**
F9-FC	**Meeting goals - adjusted by ratio of connect time to elapsed time**
F8	**Discretionary**

Managing I/O priorities

- The following conditions must be met to implement WLM CSS I/O Priority Management:
  - IBM 2064 or 2066 running z/OS in z/Architecture mode
  - Basic or LPAR mode
  - z/OS running in WLM Goal mode
  - I/O priority management set to YES in WLM policy
  - CSS I/O priority management enabled in CPC Reset Profile
  - Valid range of CSS I/O priorities specified in Image Profile

IRD – LPAR CPU Management

- Two parts to WLM LPAR CPU Management:
  - WLM LPAR Weight Management
    - Automatically change the weight of a Logical Partition
    - Based on analysis of the current workloads by WLM
  - WLM Vary CPU Management
    - Automatically vary a logical CP online or offline in an LP
    - Based on analysis and requests from WLM
- Software managing hardware resources:
  - Software - WLM Goal mode
  - Hardware - Shared CPs and Logical Partition weights
  - Parallel Sysplex - Used to share WLM information between the systems

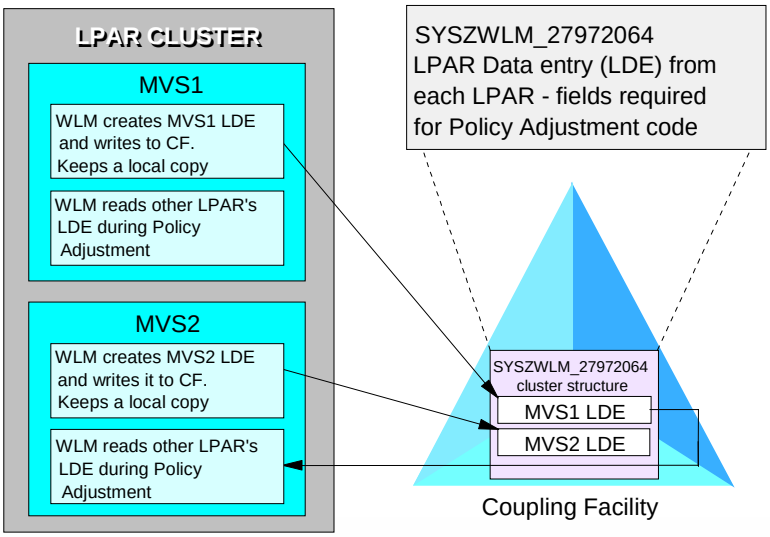
Prerequisites - CPU Management

- WLM CPU management only applies to:
  - Z/OS in 64-bit mode
  - 2064 or 2066 in LPAR mode
  - Shared CPs
  - Not capped using traditional LPAR capping
  - Goal mode
  - Access to a Coupling Facility

Coupling Facility prerequisites

- Requires a CF running CFLEVEL=9
- Each LPAR cluster requires its own LPAR cluster structure:
  - The structure name is in the format SYSZWLM_ssssttt, where ssssttt = last four digits of serial number and model number
    - serial=5142797, model=2064, ssssttt=27972064
  - The structure is a cache structu
  - WLM automatically connects when structure is made available
    - Does not disconnect until the system is shut down
- A sysplex may contain a number of LPAR clusters and therefore a number of LPAR cluster structures

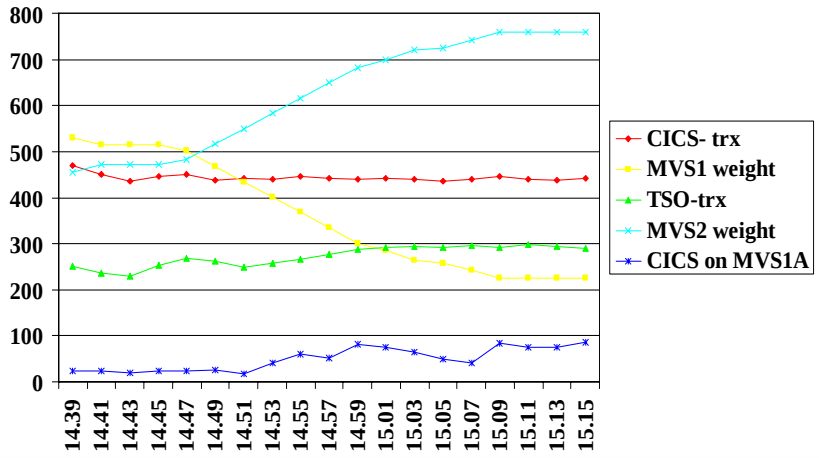
Use of CF Structures



IRD - WLM Weight Management

- Standard Receiver/Donor processing
  - Most important work not meeting goal because of CPU delay
    - Dprty change does not fix
    - Donor of weight in other LP in same cluster
  - Sum of Weights remains constant
  - No more than one adjustment per minute

IRD – WLM Weight Management



IRD – Vary CPU Management

- Ensures the number of logical CPs is appropriate to deliver the weight
  - Ensures fastest UP speed available
  - Minimizes LPAR overhead
  - Not directly driven by Policy Adjustment
- Can be turned off by specifying VARYCPU=NO in IEAOPTxx
- Starts with “Initial” CPs in image profile
- Recommend set Initial + Reserved CPs to maximum possible for CEC

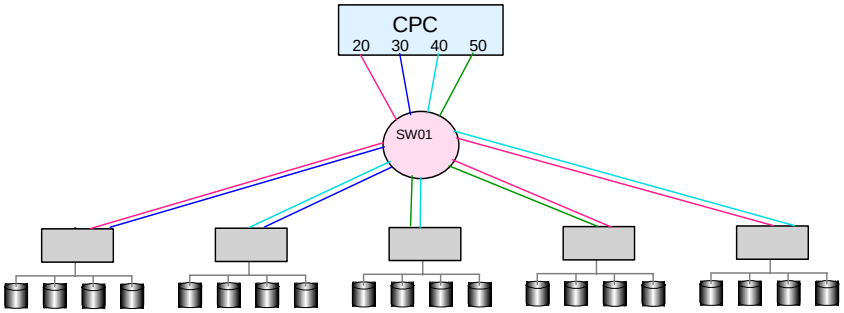
Dynamic Channel Path Management

- DCM
  - Does not move channels between LPARs
  - Does add or remove channels from the configuration for a DCM-capable control unit - similar to a systems programmer modifying HCD and doing dynamic activates - it is not just varying paths on and offline
  - Must be z/OS on a z900. Goal mode is optional, and CPC can be in basic or LPAR mode. Control Unit must be attached via ESCON Director
  - Works with ESS and RVA, and some OEM DASD - must be ESCON or FICON-Bridge attached

Goal mode with DCM

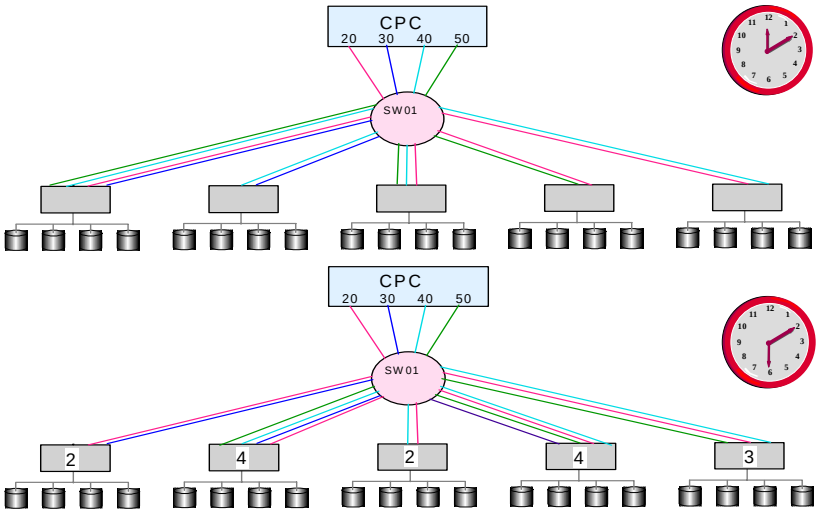
- Donor/Receiver Logic
- I/O pend time is bottleneck
  - Select LCU having largest impact and project new I/O velocity
  - If increased I/O velocity provides benefit to SCP, set new I/O velocity target for all selected LCUs
- IOS actually makes change in config to achieve new I/O velocity
- If not goal mode, I/O velocities are balanced
- I/O velocity = Connect Time / (Connect Time + Pend Time – (CU busy + Dev Busy))

Configuration without DCM



IBM @server. For the next generation of e-business.

Configuration with DCM



IBM @server. For the next generation of e-business.

DCM recommendations

- Use DCM to cater for the difference between average and peak workloads
- Configuration must be symmetric - all Logical Partitions in the LPAR Cluster must have access to the same devices through the same number of shared channels
- Must have more than one managed LCU behind an ESCON director for DCM to be effective
- Spread managed channels across processor cages and SAPs, and across multiple ESCON directors so DCM can select path that will provide the best availability

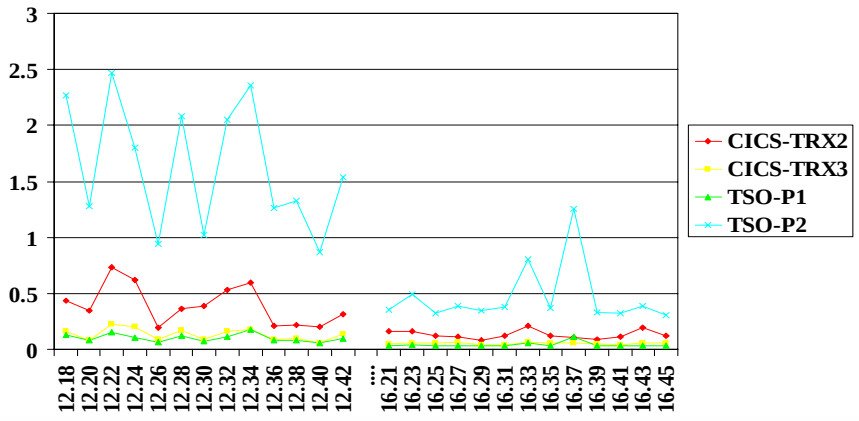
DCM restrictions

- CU must be ESS or RVA or supported OEM
- ESCON only
- CU must be attached to managed channels via a switch
- Managed channels cannot be shared outside the LPAR cluster
- All ’affected’ LCUs must also support DCM
- Cannot mix FC and FCV/ESCON on a managed LCU

Managed Channels

- Control Unit Ports must be defined for each Escon director
- Define set of channels as managed Channels
- Define managed ports on control units
- If managed channels and managed control units are available at IPL, system will come up with DCM active if this is the first system in the LPAR cluster. If not the first, it comes in the same mode as the existing systems.

Dynamic Channel Management



IRD - Documentation

- z/OS MVS Planning: Workload Management, SA22-7602
- Redbook z/OS Intelligent Resource Director, SG24-5952
- WLM Web site:
  - http://www.ibm.com/servers/eserver/zseries/zos/wlm/documents/ird/ird.html
- IRD Bucket - 2064DEVICE subset IRD

Summary

- IRD operates the same way as other WLM functions
  - Donor/Receiver logic for policy adjustment
    - I/O priority queueing for UCB and CU
    - LPAR weight management
    - Dynamic Channel management
  - Resource adjustment for performance optimization
    - I/O priority queueing for CSS for work meeting goals
    - LPAR vary CPU processing

Thank you – Any Questions?

Tom_Russell@ca.ibm.com

IBM @server. For the next generation of e-business.