Introduction to z/OS Performance Measurement and Tuning Tips

Instructor: Peter Enrico

Email: Peter.Enrico@EPStrategies.com

Enterprise Performance Strategies, Inc.
3457-53rd Avenue North, #145
Bradenton, FL 34210
http://www.epstrategies.com
http://www.pivotor.com

Voice: 813-435-2297
Mobile: 941-685-6789

z/OS Performance
Education, Software, and
Managed Service Providers

Creator of Pivotor®

Contact, Copyright, and Trademark Notices

Questions?

Send email to Peter at Peter.Enrico@EPStrategies.com, or visit our website at http://www.epstrategies.com.

Copyright Notice:

Trademarks:

Abstract and Offer

- **Abstract**
  - For a zNextGen person, z/OS performance measurement and tuning can be a big and intimidating area to explore. In the z/OS environment there are so many measurements available, and there are so many areas to be tuned. Where should one start to become quickly productive?

    During this presentation, Peter Enrico will discuss a variety of z/OS performance measurements and performance tuning recommendations to get any zNextGen person quickly productive and started on their journey to z/OS performance optimization.

Current 2011 Class Schedule

- WLM Performance and Re-evaluating of Goals
  - Instructor: Peter Enrico
  - September 12 – 16, 2011 Baltimore, Maryland, USA
- Parallel Sysplex and z/OS Performance Tuning
  - Instructor: Peter Enrico
  - September 19 – 23, 2011 Dallas, Texas, USA
- z/OS Capacity Planning and Performance Analysis
  - Instructor: Ray Wicks
  - August 15 – 17, 2011 Columbus, Ohio, USA

Presentation Overview

- Quick Tip – Learn to Write!
- Quick Tip – Learn A Top Down Approach to Performance Analysis
- Quick Tip – Understand Your Workloads
- Quick Tip – Evaluate LPAR Weight Enforcement
- Quick Tip – Verify WLM Address Space Classification
- Quick Tip – Understand Coupling Facility Sync CPU Spin Seconds
- Quick Tip – Understand Which DASD I/O Logical Volumes to Concentrate On

Quick Tip – Learn to Write!

Documentation and communication are key to a performance assignment!

Write reports for your management and your team to let them know what you are doing.

Documentation and Communication

- One of the biggest mistakes made by those responsible for performance is failure to properly communicate
- Performance personnel need to…
  - let others know what performance work is being done
  - let others know about performance findings, and to put those findings in context
  - let customers and management know about savings and cost
  - let everyone know the valuable job they are doing
- Document
  - The problem or area of concern
  - A proposed solution
  - An estimate in savings
    - Try to put this value in real monetary terms
- Let others know you accomplishments!
  - Let others know the work, accomplishments, and value being delivered by the company's investment in performance personnel

Quick Tip – Learn A Top Down Approach to Performance Analysis

Learn a logical approach to attack any performance problem:

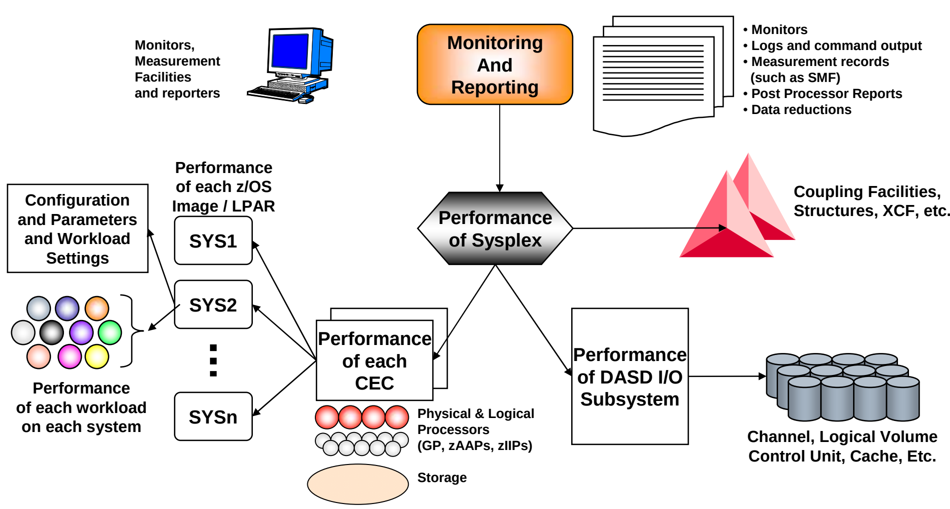
Long term analysis
Capacity planning
Real time performance management
Etc.

Performance Analysis Fundamentals

- Each should be understood in the context of computer performance
  - Understanding workloads and their requirements
  - Understanding managed resources
  - Understanding and managing performance objectives
  - Knowledge and usage of performance controls
  - Ability to measure, monitor, and report
  - Common methodologies and techniques
    - For Capacity Planning
    - For Performance Management
    - For Performance Analysis
    - For Performance Tuning
  - Documentation and communication

The z Performance Environment

The following is a simplified view of the z/OS performance environment



Cookbook Approach to Performing a System Performance Analysis

- High level steps for revisiting your WLM setup and service definition
  - Step 1: Inventory Your Managed Resources
  - Step 2: Inventory System Workloads
  - Step 3: Understand Current WLM Definitions and System Parameters
  - Step 4: Learn How to Interpret Measurements
  - Step 5: Analyze the basic system resources
    - Processor, Storage, DASD I/O
  - Step 6: Analyze your WLM Service Definition and goals
  - Step 7: Analyze Sysplex communication and resources
    - XCF, Coupling Facility and Coupling Facility Structures
  - Step 8: Analyze your Subsystem work managers
    - CICS, IMS, DB2, WebSphere, MQ, etc, etc, etc…
  - Step 9: Analyze your applications
  - Step 10: Write a report

Quick Tip – Understand Your Workloads

Spend a day or two (or three (or four)) really learning about your workloads and their performance requirements, objectives, and expectations.

Who are your customers?

Workload Investigation Questions

- On z/OS, when investigating workload performance, at a very high level the following broad questions need to be asked:
  - What is the workload
  - What type of address space is used by this workload?
  - What is the definition of the workload's transaction?
  - Why does this workload exist?
  - Who / what make up this workload?
  - When does the workload run?
  - Where does the workload run
  - How does the workload run?
  - What are the performance challenges of the workload?

Key Workload Subsystem Types

- On z/OS there are many different types of workloads that run
  - Some workloads are common to most installations
  - Other workloads are specific to certain industries
  - Other workloads are very customer specific
- But in general, when investigating workloads, the following the primary groupings to be investigated
  - System and system support workloads
  - Interactive workloads
  - Batch workloads
  - Database management workloads
  - Legacy online transaction processing (OLTP) workloads
  - e-business transaction processing workloads
  - Distributed request workloads
- Each of these workloads have a number of different variations
  - Most interact some how with the others

Example: System and System Support Workloads

- Who / What make up this workload?
  - Required and usual address spaces that provide some operating system related function
  - Tools, facilities, and vendor products for system support purposes
  - Unix System Services Daemons providing operating system functions
  - Monitors – performance or other
- Why is performance a concern for this workload?
  - Much of the overall system performance and the performance of most other workloads directly depend on sufficient performance of these system workloads
- Performance challenges
  - Lots of miscellaneous address spaces that need investigation
  - When some of these address spaces need system resources, they need the resource right away

Example: System and System Support Workloads

- Examples of system and system support workloads on z/OS include:
  - Required high priority address spaces
    - MASTER, ALLOCAS, ANTMAIN, BPXOINIT, CATALOG, CONSOLE, DUMPSRV, GRS, IOSAS, IXGLOGR, OMVS, PCAUTH, RASP, SMSPDSE, TRACE, XCFAS, WLM, etc.
  - Privileged address spaces for operating system functions
    - ANTASxxx, APPC, ASCH, JES2, RACF, ZFS, LLA, SMS, VLF, RRS, ENF, etc.
  - UNIX System Services daemons doing operating system work
    - INETD, TN3270, SYSLOGD, etc.
  - Monitors
    - Performance monitors by BMC, CA, IBM, and many other vendors
    - Other non-performance monitors by many vendors
  - Other
    - Many different z/OS products used for system support purposes

Example: Batch Workloads

- Who / What make up this workload?
  - Work requests that perform some function or data processing request.
  - These requests are run in a 'background' mode
    - Means that units of work generally have no human interaction
    - Work is 'scheduled' and run asynchronously
- Why is performance a concern for this workload?
  - There are many different types of batch workload and each has it own unique performance considerations and requirements
- Performance challenges
  - In most installations batch is one of the largest, if not the largest workloads
    - Typically consumes great amounts of system resources
  - Completion of many batch workloads is required within very strict windows of time

Example: Batch Workloads

- Examples of batch workloads on z/OS include:
  - Normal Production Jobs submitted by a Job Scheduler
  - Critical Path Jobs submitted through a Job Scheduler
  - Ad-hoc Jobs (possibly submitted by a Job Scheduler)
  - Development Jobs
  - Normal System Support Jobs
  - High-Priority System Support Jobs
  - Logs, Archival, Backup, and D/R Jobs (possibly submitted by a Job Scheduler or an appropriate Subsystem)
  - Quick Utility Jobs
  - Emergency of Hot Jobs

Example: Batch Workloads

- Jobs requiring setup include
  - Batch jobs that require input from media that is not normally online to a system
  - Batch jobs that require offline media to generate output have setup requirements that involve operator intervention
- May Include:
  - Manual tape mounts (reels or cartridges)
  - ATL/Silo Mounts- normally very quick unless media is not in the ATL
  - Virtual Tape Servers (VTS)- normally very quick unless data is not staged in the disc Cache
  - Direct SYSOUT Writers (not used much)- not normally quick if forms or print trains need to be inserted, or device is particularly slow

Quick Tip – Evaluate LPAR Weight Enforcement

Understand when weight enforcement is happening.

Investigation: If any LPAR is regular getting more than its weight because another LPAR is regularly demanding less than its weight

Tuning: If the LPAR regularly running below its weight suddenly demands its weight (CPU) then it will hurt the LPAR that is regularly enjoying more than his weight since that LPAR will have its weight enforced.

Maybe shift weight so assigned weights are based on regular running state.

Quick Tuning Tip – PR/SM Weight Enforcement

- What you need to get started:
  - Basic understanding of PR/SM, the LPAR configuration, LPAR weights, etc
  - Basic understanding PR/SM weight enforcement
  - Understanding of available SMF 70 PR/SM and z/OS CPU measurements
    - Several reporting packages
    - Or RMF (or CMF) Partition Data Report and RMF (or CMF) CPU Activity Report
  - A calculator or program to measure
    - Guaranteed LPAR share
    - Percentage of weight consumed
    - Delta between LPAR Utilization % and MVS Utilization %
  - Report to your manager
    - Pattern of weight enforcement

Example of RMF Partition Data Report

```
                                        P A R T I T I O N   D A T A   R E P O R T
                                                                                                                     PAGE    2
          z/OS V1R9                SYSTEM ID SYP1          DATE 08/05/2009           INTERVAL 15.00.006
                                   RPT VERSION V1R9 RMF    TIME 02.29.00             CYCLE 1.000 SECONDS

MVS PARTITION NAME                 LPARSYP1      NUMBER OF PHYSICAL PROCESSORS          9                GROUP NAME       N/A
IMAGE CAPACITY                        422                  CP                           6                LIMIT            N/A
NUMBER OF CONFIGURED PARTITIONS         5                  AAP                          0
WAIT COMPLETION                        NO                  IFL                          0
DISPATCH INTERVAL                  DYNAMIC                 ICF                          0
                                                           IIP                          3

--------- PARTITION DATA -----------------  -- LOGICAL PARTITION PROCESSOR DATA --   -- AVERAGE PROCESSOR UTILIZATION PERCENTAGES --

                  ----MSU----  -CAPPING--  PROCESSOR-  ----DISPATCH TIME DATA----   LOGICAL PROCESSORS  --- PHYSICAL PROCESSORS ---
NAME        S  WGT  DEF     ACT  DEF    WLM%  NUM   TYPE   EFFECTIVE        TOTAL        EFFECTIVE    TOTAL   LPAR MGMT  EFFECTIVE  TOTAL

LPARSP1     A   54    0     260  NO      0.0    5   CP    00.55.27.732  00.55.30.468      73.95     74.01      0.05       61.62  61.67
LPARSD1     A   12    0      35  YES     0.0    1   CP    00.07.26.536  00.07.28.786      49.61     49.86      0.04        8.27   8.31
LPARSM1     A    3    0       0  YES     0.0    1   CP    00.00.00.000  00.00.00.000       0.00      0.00      0.00        0.00   0.00
LPARSM2     A    3    0       0  YES     0.0    1   CP    00.00.00.000  00.00.00.000       0.00      0.00      0.00        0.00   0.00
LPARSP4     A   28    0     122  YES     0.0    2   CP    00.26.01.838  00.26.02.148      86.77     86.79      0.01       28.92  28.93
*PHYSICAL*                                                              00.00.08.517                          0.16               0.16
                                                          ------------  ------------                        ------      ------ ------
  TOTAL                                                   01.28.56.106  01.29.09.922                          0.26       98.82  99.07


LPARSP1     A   54                                  1   IIP   00.04.47.344  00.04.47.989      31.93     32.00      0.02       10.64  10.67
LPARSD1     A   12                                  1   IIP   00.00.01.969  00.00.02.039       0.22      0.23      0.00        0.07   0.08
LPARSM2     A    3                                  1   IIP   00.00.00.000  00.00.00.000       0.00      0.00      0.00        0.00   0.00
LPARSP4     A   28                                  1   IIP   00.00.06.040  00.00.06.123       0.67      0.68      0.00        0.22   0.23
*PHYSICAL*                                                              00.00.02.329                          0.09               0.09
                                                          ------------  ------------                        ------      ------ ------
  TOTAL                                                   00.04.55.355  00.04.58.482                          0.12       10.94  11.05
```

LPAR Configuration, LPAR Weight% and Guaranteed Share CPs

Partition Configuration and Setup - CP Processors

Partition Number	Processor Type	LPAR Name	Partition Mode	Num Logical Processors	Partition Weight	% Weight	Guaranteed CPs of Capacity	Capping	Def MSU Capacity
1	CP	LPARSP1	SHR	5	54	54.0%	3.24		0
2	CP	LPARSD1	SHR	1	12	12.0%	0.72	Y	0
3	CP	LPARSM1	SHR	1	3	3.0%	0.18	Y	0
4	CP	LPARSM2	SHR	1	3	3.0%	0.18	Y	0
5	CP	LPARSP4	SHR	2	28	28.0%	1.68	Y	0
				10	100	100%	6.00		

Partition Configuration and Setup - zIIP Processors

Partition Number	Processor Type	LPAR Name	Partition Mode	Num Logical Processors	Partition Weight	Partition Weight	Guaranteed CPs of Capacity	Capping
1	zIIP	LPARSP1	SHR	1	54	55.7%	0.557	
2	zIIP	LPARSD1	SHR	1	12	12.4%	0.124	Y
4	zIIP	LPARSM2	SHR	1	3	3.1%	0.031	Y
5	zIIP	LPARSP4	SHR	1	28	28.9%	0.289	Y
				4.00	97	100%	1.00	

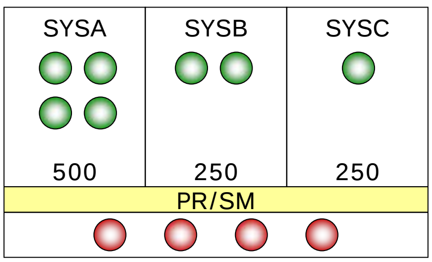
Percentage of LPAR weight consumed

- Value is not reported on any post processor report and few monitors
  - Need to manually calculate it
- Percentage of weight consumed helps us understand:
  - If a partition regularly consumes more or less than its guaranteed share
    - If > 100% then LPAR consumed more than its guaranteed share
      - Example: 150% means LPAR consumed 1.5 times its guaranteed share
    - If < 100% then LPAR had less demand for CPU than guaranteed share
      - Example: 50% means LPAR only consumed 50% of its weight
  - Insights into if weights are being enforced
    - Which LPARs are competing for CPU
- Reminder, when processor demand of LPARs is greater than physical CPU capacity
  - Weights are enforced

What was percentage of LPAR weight consumed?

- Scenario 2:
  - Assume SYSA had 100% demand for its 4 logical capacity
  - SYSB, and SYSC each had 0% demand for their logical capacity

  - SYSA guaranteed 50% of 4 physicals but used 100% of 4 physicals
    - So consumed 200% of its weight
  - SYSB guaranteed 25% of 4 physicals but used 0% of 4 physicals
    - So consumed 0% of its weight




```
Assume the following utilizations:
           Physical | Logical | %Weight
• SYSA   : 100.0%   |  100.0% |  200%
• SYSB   :   0.0%   |    0.0% |    0%
• SYSC   :   0.0%   |    0.0% |    0%
           ======   |         |
• Total  : 100.0%   |         |

Reminder:
25% of 4 CPUs = 100% of 1 CPU
```


Calculating percentage of weight consumed

- Calculating the percentage of weight consumed

- % LPAR Weight Consumed =

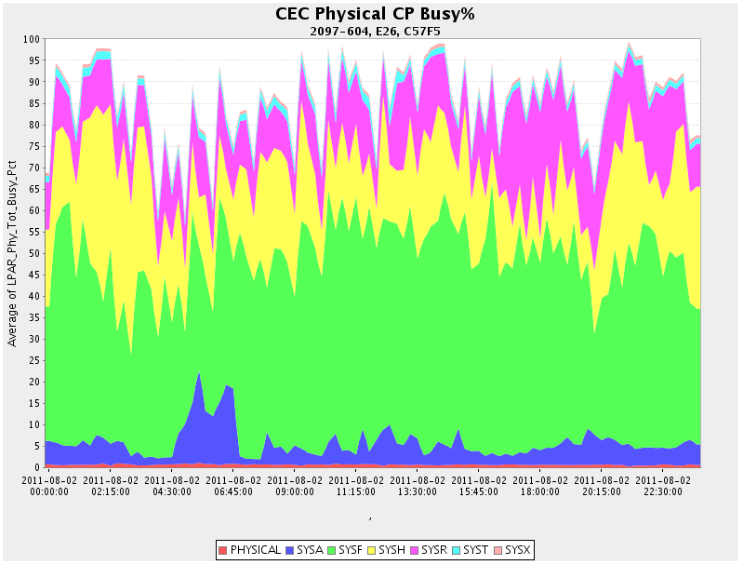
  ((Total Physical CPU Busy %) / LPAR Guaranteed Share)) * 100

- Where:
  - Total Physical CPU Busy %
    - From RMF Partition Data report

  - LPAR Guaranteed Share
    = ((LPAR's Weight)/(Sum all LPAR's Weights))*100

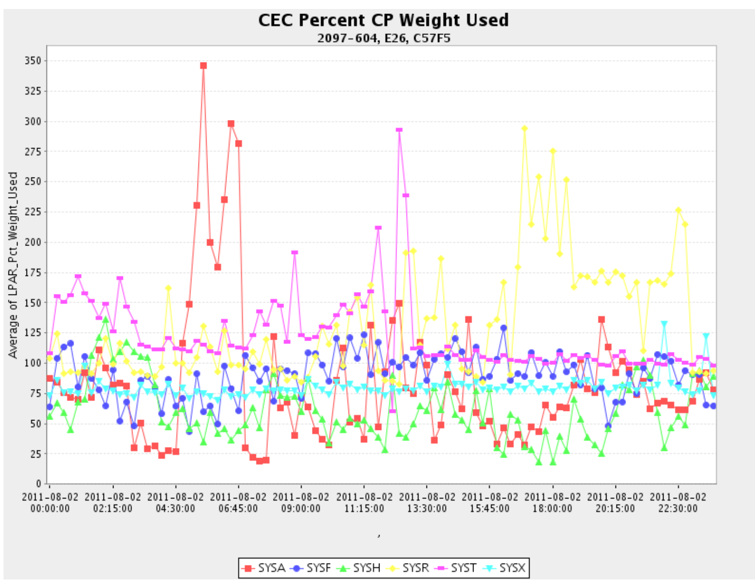
- Example: LPARSP1:
  - If LPARSP1 is 61.67% Total Physical Processor Busy
  - And is guaranteed share is 54%
  - Then %Weight Consumed = ((61.67%)/54%)*100 = 114.2%

LPAR Utilization for CEC by LPAR Over Time



Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 27

Percentage Weight Enforcement



Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 28

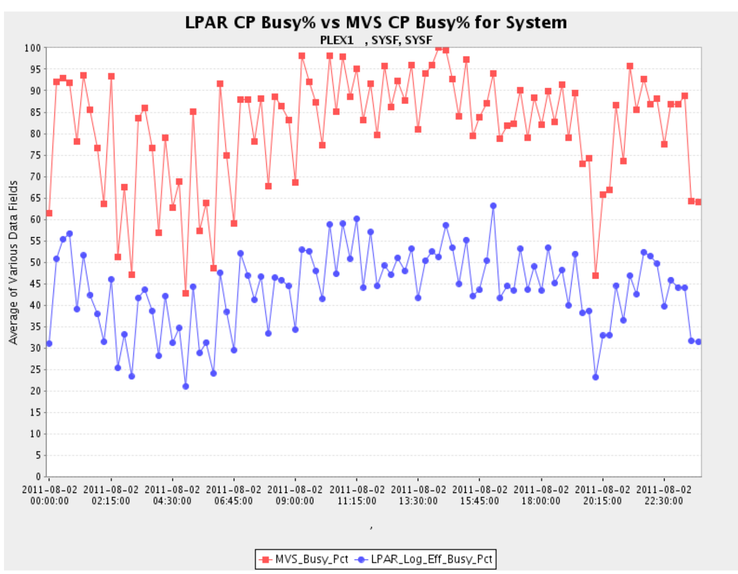
RMF CPU Activity Report - Page 1

- Provides information on active processors
  - Configuration information
    - Processor Model, number of CPUs, CPU serial number
  - Resource usage and load calculations
    - Processor online time percentage
    - CPU Busy % - LPAR view and MVS view
    - I/O Interrupt activity - interrupt rate, interrupts handled by TPI

```
                                                  C P U   A C T I V I T Y

            z/OS V1R9                SYSTEM ID P01             DATE 08/05/2009             INTERVAL 15.00.
                                     RPT VERSION V1R9 RMF       TIME 02.29.00              CYCLE 1.000 SEC
 CPU  2094    MODEL  706    H/W MODEL  S18   SEQUENCE CODE 00000000000D63FC    HIPERDISPATCH=N/A

---CPU---    ---------------- TIME % ----------------      LOG PROC      --I/O INTERRUPTS--
 NUM  TYPE    ONLINE    LPAR BUSY     MVS BUSY   PARKED     SHARE %        RATE     % VIA TPI
  0    CP     100.00    74.03         100.0      ------      64.8          7.01     98.99
  1    CP     100.00    74.03         100.0      ------      64.8          6.80     99.00
  2    CP     100.00    74.02         100.0      ------      64.8          6.49     99.02
  3    CP     100.00    74.02         100.0      ------      64.8          6.69     99.04
  4    CP     100.00    73.95         100.0      ------      64.8          8402     23.06
 TOTAL/AVERAGE          74.01         100.0                 324.0          8429     23.30

  C    IIP    100.00    32.00         31.96      ------     100.0
 TOTAL/AVERAGE          32.00         31.96                 100.0
```

Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 29

LPAR Busy % vs MVS Busy %



Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 30

Quick Tip – Verify WLM Address Space Classification

Understand your WLM service definition and settings.

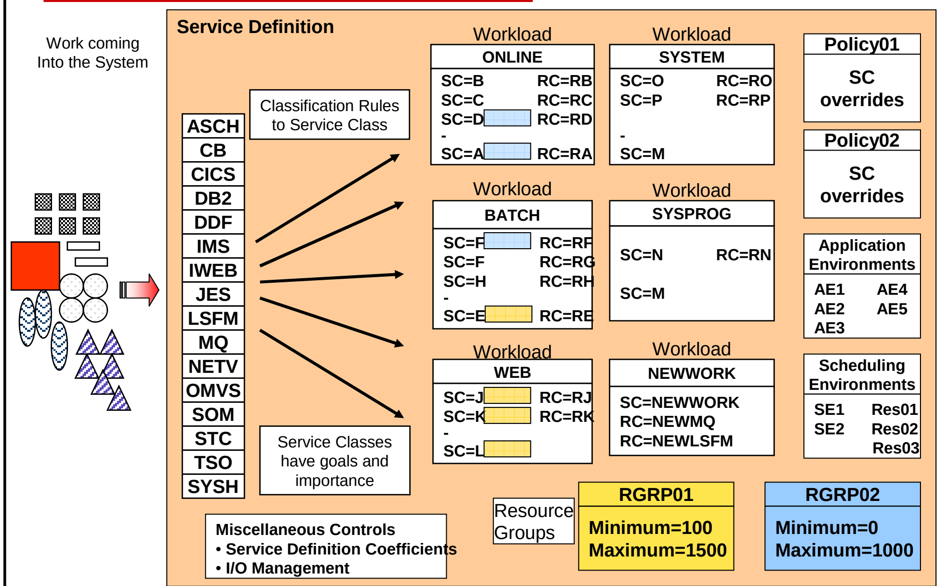
Investigation: Understand where your work is being classified.

Tuning: Verify where things are being classified
- You will learn lots about your workloads
- You will learn lots about your WLM configuration

Quick Tuning Tip – Verify WLM Classification

- What you need to get started:
  - Basic understanding of Workload Manager, service classes, importance level, goals, etc.
  - Basic understanding of WLM classification rules
  - Understand your WLM service definition
    - Visit my website (www.epstrategies.com) to convert your WLM service definition to HTML format
  - Understanding of available SMF 30 Address Space SMF records
    - You will need a SMF data reductions program
  - SMF 30 records to verify the classification of address spaces
    - Note: Different exercise to verify classification of enclaves
  - Make sure work is being classified as expected
    - If nothing else, you will learn lots about your workloads and WLM management of your workloads

Overview of WLM Service Definition



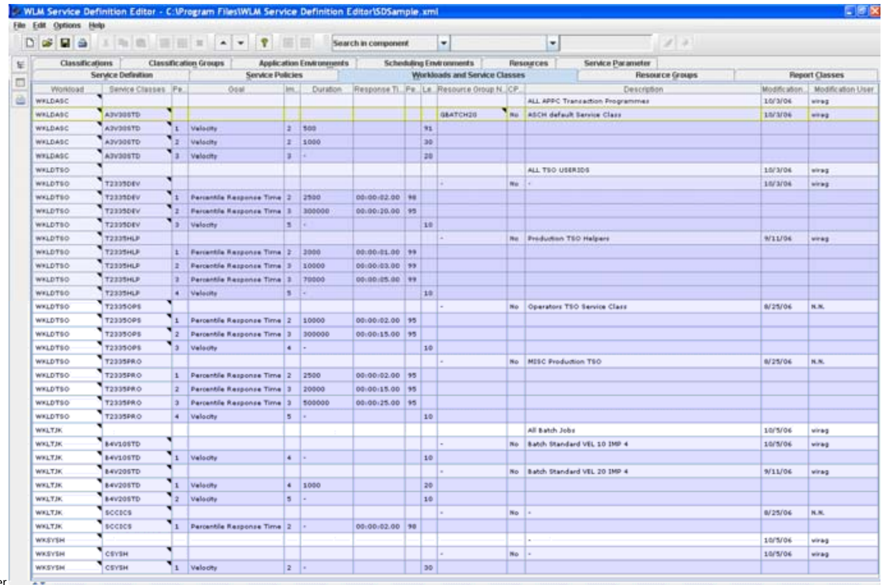
Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 33

Understand Current WLM Definition

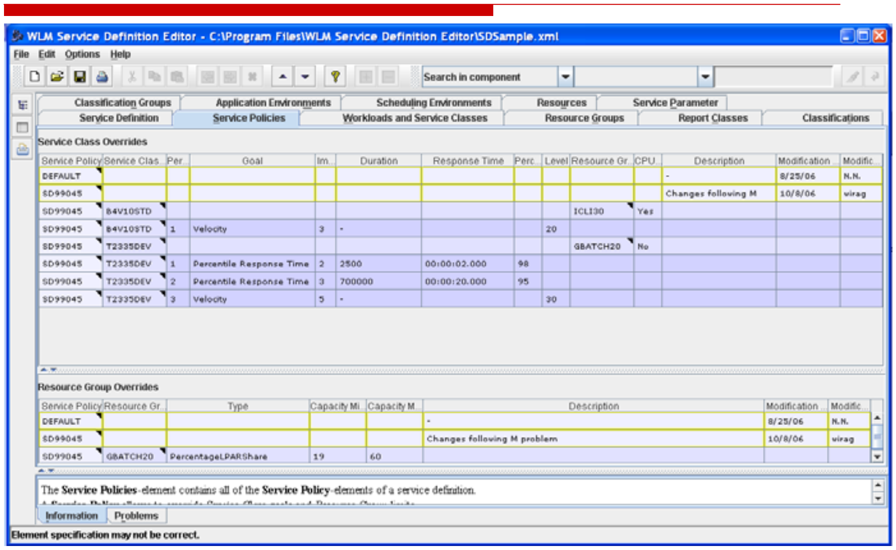
- Understand your current WLM service definition
  - It is your starting point for reevaluation of goals
- Items to consider:
  - How many and what service classes and service class periods defined
  - What are the goals and importance levels of the work of the periods
    - How do all these relate to each other?
  - Are report classes being used and if so, how many report classes are there?
  - Are there application environments or scheduling environment defined?
  - What are the classification rules?
    - Do all subsystems have classification rules
  - Are resource groups being used?
    - For what purpose
  - Are critical controls (such as CPU or storage being used)
  - Etc..

Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 34

WLM Policy Editor



WLM Policy Editor



Easier Way: Convert Your Service Definition to HTML

- When cleaning up your service definition it may help to have it in an easy to read and analyze format
- Convert your WLM service definition to HTML for readability and reference

1. Using the WLM ISPF application:
   - 'File' option at top of screen
   - Sub option 'Print as GML'
   - Creates a flat file of service definition with GML formatting tags

HTML

2. Download GML version of WLM Service Definition to workstation as text file

2. Go to www.epstrategies.com and select WLM Tool button to convert

5. Follow instructions - select file and fill in email address

6. Presto! HTML file will be emailed to you within minutes

WLM Service Definition in HTML Format

Workload Manager Service Definition

PLEX50 Silver Scripted: 2003-03-24.

Table of Contents

Use SMF 30 Records to Verify Classification

- SMF 30 job / address space identification information

Name	Description
SMF30JBN	Job or session name.
SMF30PGM	Program name (taken from PGM= parameter on EXEC card).
SMF30STM	Step name (taken from name on EXEC card).
SMF30UIF	User-defined identification field
SMF30JNM	JES job identifier.
SMF30STN	Step number (first step = 1, etc.).
SMF30CLS	Job class (blank for TSO/E session or started tasks)
SMF30SSN	Substep number. This field is set to zero for non-z/OS UNIX System Services steps. When the z/OS UNIX System Services exec function is requested, a new substep is begun and this value is incremented.
SMF30EXN	Program name. For a z/OS UNIX program, this contains the UNIX program that was run or the 8 character name of an MVS program that was run.
SMF30ASI	Address Space identifier

SMF 30 Measurements to Correlate to SMF 72.3

- Can also use SMF 30 WLM information to correlate measurements to the SMF 72.3 records

Name	Description
SMF30TRS	Number of system resources manager (SRM) transactions.
SMF30WLM	Workload name.
SMF30SCN	Service class name.
SMF30GRN	Resource group name.
SMF30RCN	Report class name.
SMF30ETC	Independent enclave transaction count.

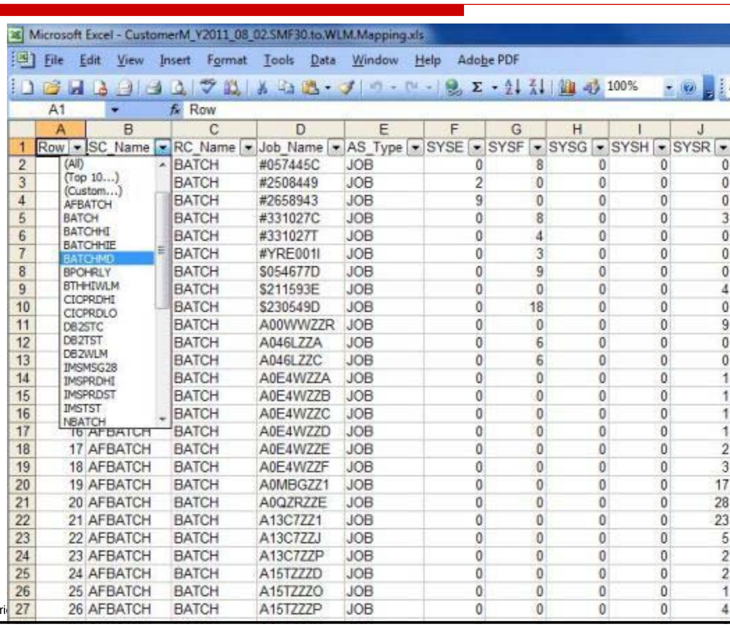
- Use Service Class name and Report Class name to correlate measurements to the SMF 72.3 records

Great Exercise (for Any Workload)

- Using the SMF 30 measurements, it is easy to create a cross reference spreadsheet to gain insights into the following:
  - Mapping of which address spaces were active on which system and how many intervals of time
    - For every address space SMF 30.2 or 30.3 record, create a CSV file that contains Address Space Name, Service Class name, Report Class Name, and system where address space ran.
  - Mapping of what programs ran
    - z/OS programs
    - Could expand to include Unix System Services programs
    - Etc.
  - Mapping of address spaces to WLM Service Classes and Report Classes
    - Helps with review and verification of WLM classification rules
- See following slides for some examples.

Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 41

Create a SMF30.2 to WLM Mapping



	A	B	C	D	E	F	G	H	I	J
1	Row	SC_Name	RC_Name	Job_Name	AS_Type	SYSE	SYSF	SYSG	SYSH	SYSR
2			BATCH	#057445C	JOB	0	8	0	0	0
3			BATCH	#2508449	JOB	2	0	0	0	0
4			BATCH	#2658943	JOB	9	0	0	0	0
5			BATCH	#331027C	JOB	0	8	0	0	3
6			BATCH	#331027T	JOB	0	4	0	0	0
7			BATCH	#YRE001I	JOB	0	3	0	0	0
8			BATCH	$054677D	JOB	0	9	0	0	0
9			BATCH	$211593E	JOB	0	0	0	0	4
10			BATCH	$230549D	JOB	0	18	0	0	0
11			BATCH	A00WWZZR	JOB	0	0	0	0	9
12			BATCH	A046LZZA	JOB	0	6	0	0	0
13			BATCH	A046LZZC	JOB	0	6	0	0	0
14			BATCH	A0E4WZZA	JOB	0	0	0	0	1
15			BATCH	A0E4WZZB	JOB	0	0	0	0	1
16			BATCH	A0E4WZZC	JOB	0	0	0	0	1
17	16	AFBATCH	BATCH	A0E4WZZD	JOB	0	0	0	0	1
18	17	AFBATCH	BATCH	A0E4WZZE	JOB	0	0	0	0	2
19	18	AFBATCH	BATCH	A0E4WZZF	JOB	0	0	0	0	3
20	19	AFBATCH	BATCH	A0MBGZZ1	JOB	0	0	0	0	17
21	20	AFBATCH	BATCH	A0QZRZZE	JOB	0	0	0	0	28
22	21	AFBATCH	BATCH	A13C7ZZ1	JOB	0	0	0	0	23
23	22	AFBATCH	BATCH	A13C7ZZJ	JOB	0	0	0	0	5
24	23	AFBATCH	BATCH	A13C7ZZP	JOB	0	0	0	0	2
25	24	AFBATCH	BATCH	A15TZZZD	JOB	0	0	0	0	2
26	25	AFBATCH	BATCH	A15TZZZO	JOB	0	0	0	0	1
27	26	AFBATCH	BATCH	A15TZZZP	JOB	0	0	0	0	4

Peter Enri[...] Performance Tuning Tips - 42

Example:
Investigate Address Spaces in PS_BATHI

Microsoft Excel - CustomerM_Y2011_08_02.SMF30.to.WLM.Mapping.xls

	Row	SC_Name	RC_Name	Job_Name	AS_Type	SYSE	SYSF	SYSG	SYSH	SYSR
3917	3916	PS_BATHI	RPSFIPRD	M8441054	STC	206	350	0	0	0
3918	3917	PS_BATHI	RPSFIPRD	M8441058	STC	128	150	0	0	0
3919	3918	PS_BATHI	RPSFIPRD	M8441099	STC	0	0	0	10	0
3920	3919	PS_BATHI	RPSFIPRD	M8441211	STC	0	0	13	13	0
3921	3920	PS_BATHI	RPSFIPRD	M8441213	STC	0	0	22	0	0
3922	3921	PS_BATHI	RPSFIPRD	M8441232	STC	0	0	53	0	0
3923	3922	PS_BATHI	RPSFIPRD	M8441233	STC	0	0	0	11	0
3924	3923	PS_BATHI	RPSFIPRD	M8441314	STC	0	0	0	40	0
3925	3924	PS_BATHI	RPSFIPRD	M8441332	STC	0	0	0	10	0
3926	3925	PS_BATHI	RPSFIPRD	M8441373	STC	0	0	0	10	0
3927	3926	PS_BATHI	RPSFIPRD	M84413FA	STC	0	0	20	10	0
3928	3927	PS_BATHI	RPSFIPRD	M84413FB	STC	0	0	0	40	0
3929	3928	PS_BATHI	RPSFIPRD	M84413FC	STC	0	0	10	20	0
3930	3929	PS_BATHI	RPSFIPRD	M84413FD	STC	0	0	30	10	0
3931	3930	PS_BATHI	RPSFIPRD	M84413FE	STC	0	0	0	30	0
3932	3931	PS_BATHI	RPSFIPRD	M84413FF	STC	0	0	20	20	0
3933	3932	PS_BATHI	RPSFIPRD	M84413YS	STC	0	0	0	20	0
3934	3933	PS_BATHI	RPSFIPRD	M8441501	STC	0	0	13	0	0
3935	3934	PS_BATHI	RPSFIPRD	M8441502	STC	0	0	13	0	0
3936	3935	PS_BATHI	RPSFIPRD	M8441504	STC	0	0	0	26	0
3937	3936	PS_BATHI	RPSFIPRD	M8441511	STC	0	0	13	0	0
3938	3937	PS_BATHI	RPSFIPRD	M8441513	STC	0	0	52	0	0
3939	3938	PS_BATHI	RPSFIPRD	M8441518	STC	0	0	13	13	0
3940	3939	PS_BATHI	RPSFIPRD	M8441540	STC	0	0	13	0	0
3941	3940	PS_BATHI	RPSFIPRD	M8444200	STC	0	0	30	0	0
4464										
4465										
4466										

Sheet1 / Sheet2 / Sheet3

Filter Mode

Example:
Investigate Address Spaces in STCDEF

Microsoft Excel - CustomerM_Y2011_08_02.SMF30.to.WLM.Mapping.xls

	Row	SC_Name	RC_Name	Job_Name	AS_Type	SYSE	SYSF	SYSG	SYSH	SYSR
3942	3941	STCDEF	RSTCDFLT	AD10DSPC	STC	23	0	0	0	0
3943	3942	STCDEF	RSTCDFLT	ANTAS000	SYS	23	24	23	24	23
3944	3943	STCDEF	RSTCDFLT	ASTX	STC	0	24	0	0	0
3945	3944	STCDEF	RSTCDFLT	AXR	SYS	23	24	23	24	23
3946	3945	STCDEF	RSTCDFLT	AXR03	STC	0	24	0	0	0
3947	3946	STCDEF	RSTCDFLT	AXR04	STC	24	24	23	0	24
3948	3947	STCDEF	RSTCDFLT	CCITCPGW	STC	23	24	0	0	0
3949	3948	STCDEF	RSTCDFLT	DFS	STC	0	0	0	0	23
3950	3949	STCDEF	RSTCDFLT	DFSKERN	STC	0	0	0	0	23
3951	3950	STCDEF	RSTCDFLT	EOSARC	STC	8	0	0	0	0
3952	3951	STCDEF	RSTCDFLT	FDRUPS	STC	23	0	0	0	0
3953	3952	STCDEF	RSTCDFLT	FICPGBSP	STC	0	0	3	0	0
3954	3953	STCDEF	RSTCDFLT	FICPGBST	STC	0	0	3	0	0
3955	3954	STCDEF	RSTCDFLT	FIDVGBSP	STC	0	0	3	0	0
3956	3955	STCDEF	RSTCDFLT	FIDVGBST	STC	0	0	3	0	0
3957	3956	STCDEF	RSTCDFLT	FIDVGRSP	STC	0	0	3	0	0
3958	3957	STCDEF	RSTCDFLT	FIDVGRST	STC	0	0	3	0	0
3959	3958	STCDEF	RSTCDFLT	FIQAGBSP	STC	0	0	3	0	0
3960	3959	STCDEF	RSTCDFLT	FIQAGBST	STC	0	0	3	0	0
3961	3960	STCDEF	RSTCDFLT	FISYGRSP	STC	0	0	3	0	0
3962	3961	STCDEF	RSTCDFLT	FISYGRST	STC	0	0	3	0	0
3963	3962	STCDEF	RSTCDFLT	GSS	STC	24	24	23	24	23
3964	3963	STCDEF	RSTCDFLT	HRDVUESP	STC	0	0	3	0	0
3965	3964	STCDEF	RSTCDFLT	HRDVUEST	STC	0	0	3	0	0
3966	3965	STCDEF	RSTCDFLT	HRQAUESP	STC	0	0	3	0	0
3967	3966	STCDEF	RSTCDFLT	HRQAUEST	STC	0	0	3	0	0
3968	3967	STCDEF	RSTCDFLT	HZSPROC	STC	23	24	23	24	23
3969	3968	STCDEF	RSTCDFLT	IMSRDR	STC	0	1	0	0	0

Sheet1 / Sheet2 / Sheet3

36 of 4462 records found

Example:
Investigate Address Space Names Ending in MSTR

Microsoft Excel - CustomerM_Y2011_08_02.SMF30.to.WLM.Mapping.xls

1	Row	SC_Name	RC_Name	Job_Name	AS_Type	SYSE	SYSF	SYSG	SYSH	SYSR
1968	1967	DB2STC	RDB2RGNS	D2DAMSTR	STC	0	0	0	0	23
1971	1970	DB2STC	RDB2RGNS	D2DBMSTR	STC	0	0	24	0	0
1974	1973	DB2STC	RDB2RGNS	D2DCMSTR	STC	0	0	0	0	23
1977	1976	DB2STC	RDB2RGNS	D2DDMSTR	STC	0	0	24	0	0
1980	1979	DB2STC	RDB2RGNS	D2PAMSTR	STC	0	24	0	0	0
1983	1982	DB2STC	RDB2RGNS	D2PBMSTR	STC	23	0	0	0	0
1986	1985	DB2STC	RDB2RGNS	D2PCMSTR	STC	0	24	0	0	0
1989	1988	DB2STC	RDB2RGNS	D2PDMSTR	STC	23	0	0	0	0
1992	1991	DB2STC	RDB2RGNS	D2PPMSTR	STC	0	0	0	24	0
1995	1994	DB2STC	RDB2RGNS	D2PSMSTR	STC	0	0	23	0	0
3982	3981	STCHI	UQP	UQPAMSTR	STC	0	24	0	0	0
3986	3985	STCHI	UQP	UQPBMSTR	STC	23	0	0	0	0
3990	3989	STCHI	UQP	UQPCMSTR	STC	0	0	23	0	0
3994	3993	STCHI	UQT	UQTEMSTR	STC	0	0	0	0	24
3996	3995	STCHI	UQT	UQTFMSTR	STC	0	0	0	0	23

15 of 4462 records found

Quick Tip – Understand Coupling Facility Sync CPU Spin Seconds

What is the cost to z/OS for doing Coupling Facility synchronous requests?

Investigation: Determine how many CPU seconds were spent 'spinning' for the processing of coupling facility synchronous requests.

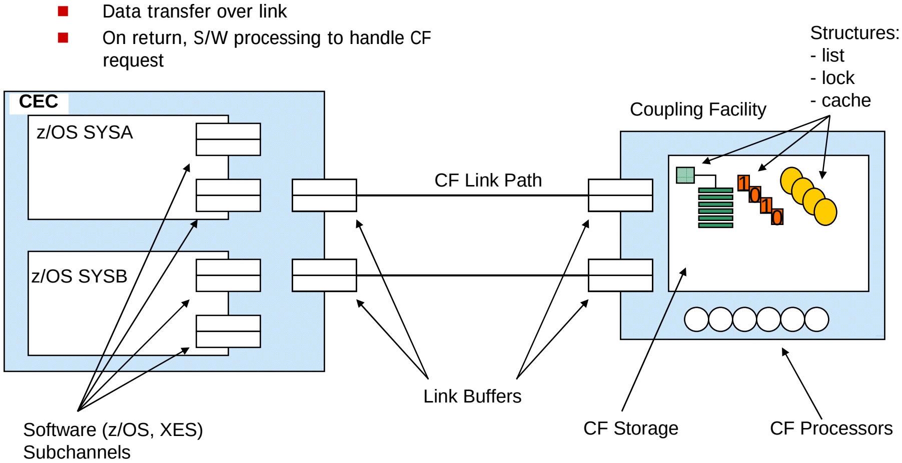
Tuning: Ensure within guidelines; determine most expensive structure if there is a problem.

Quick Tuning Tip – CF Host Effect Spin Seconds

- What you need to get started:
  - Basic understanding of the coupling environment, structures, and exploiters
  - Basic understanding of coupling facility synchronous request processing and its effects on processor consumption
  - Understanding of available SMF 74.4 Coupling Facility measurements
    - Several reporting packages
    - Or RMF (or CMF) Coupling facility report
  - A calculator or program to measure the CPU cost to z/OS for executing Coupling Facility synchronous requests
  - Report to your manager the cost due to Coupling Facility Synchronous spin CPU seconds
    - By measurement interval for the system
    - By structure

Performance Analyst View of CF Resource

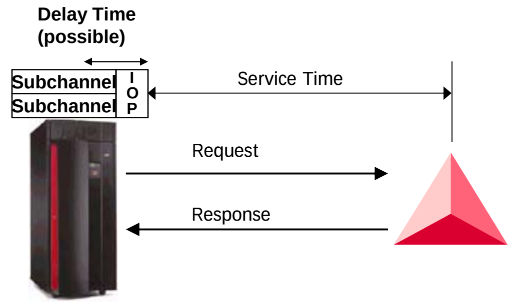
- z/OS Processing
  - S/W processing to make CF request
  - Request a sub-channel
  - Request a path
  - Data transfer over link
  - On return, S/W processing to handle CF request
- Coupling Facility Processing
  - Link time (i.e. time on path)
  - CF busy processing request



CF Synchronous Request Processing

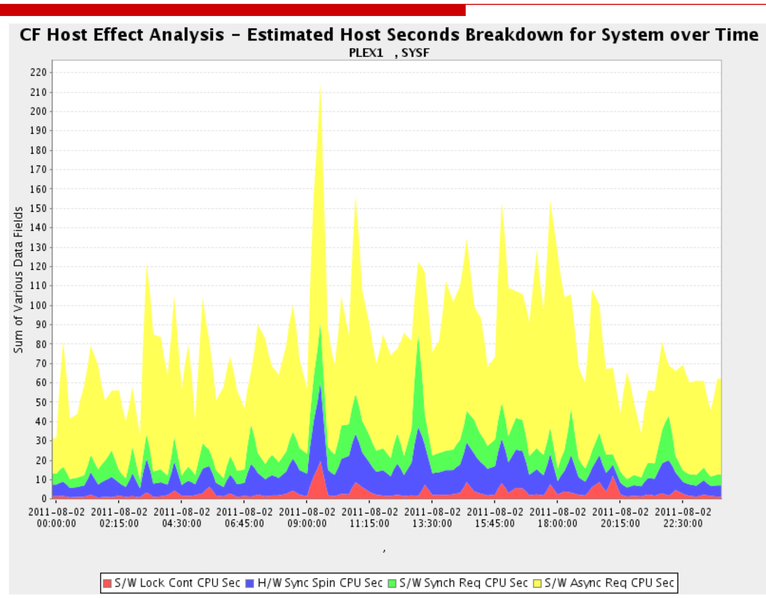
- Requesting processor spins waiting for CF request to complete
- Two types of sync requests
  - Those that must continuously run as synchronous
    - Lock requests - XES spins
  - Those that start out as sync
    - But converted to async if doing so helps performance
    - Sync cache/list requests - XES changes to async

Impacts response time of sender, and performance of sending system



Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 49

Host Effect Seconds Over Time



Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 50

CF Synchronous Request Processing

- Synchronous requests cause the requesting CPU to spin waiting for a response from the CF. Thus, performance is heavily dependant on a number of factors:
  - Speed of requesting CPU
    - Larger processor will 'waits faster' for a response
  - Subchannel busy conditions
  - Path busy conditions
  - Time it takes to transmit data to the CF
    - CF link performance
    - Speed of data over link
    - Distance - Geographically dispersed parallel Sysplex?
  - Speed of CF processor
    - Shared LPAR?
    - Dedicated CF?

CF Subchannel Activity

- CPU seconds consumed due to sync immediate Spin
  - Sync Immediate requests cause processor issuing the request to 'spin'
  - How many CPU seconds did the sending LPAR spend spinning?
    - Logical processor unavailable to other work running in the same LPAR
    - Physical processor that logical processor is dispatched to is unavailable to other LPARs
  - It is helpful to understand capacity consumed to these spinning conditions

$$\text{CPU Seconds Spinning} = \frac{(\#\text{REQ Sync})*(\text{Sync Service Time})}{1{,}000{,}000}$$

```
          # REQ                                 ----------- REQUESTS -----------   ------------------ DELAYED REQUESTS ----
SYSTEM    TOTAL    -- CF LINKS --   PTH                  #   -SERVICE TIME(MIC)-               #     % OF  ------ AVG TIME(M
NAME      AVG/SEC TYPE   GEN  USE   BUSY                 REQ      AVG    STD_DEV               REQ    REQ   /DEL    STD_DEV

SYSD       29203K CBP      2    2      0   SYNC        9047K   14.6        10.9   LIST/CACHE 9714   0.0  574.3       5937
           16224  SUBCH   14   14          ASYNC      20201K   47.4       215.1   LOCK        587   0.0    9.9       19.0
                                           CHANGED      6321  INCLUDED IN ASYNC   TOTAL        10K  0.0
                                           UNSUCC          0     0.0         0.0
```

CF Subchannel Activity

- % Effective Utilization due to sync immediate Spin
  - Sync Immediate requests cause processor issuing the request to 'spin'
  - What % CPU utilization did the sending system spend spinning?
    - When a sync immediate request encounters a busy condition it results in a 'spin' condition
    - It is helpful to understand the percent of an engine (processor busy) that was devoted to these spinning conditions

$$\%\ \text{Proc Spent Spinning} = \frac{(\#\ \text{REQ Sync}) * (\text{Sync Service Time})}{\text{Interval Seconds} * \#\ \text{LP} * 1\text{Million}} * 100$$

SYSTEM NAME	# REQ TOTAL AVG/SEC	-- CF LINKS -- TYPE	GEN	USE	PTH BUSY	----------- REQUESTS -----------	# REQ	-SERVICE TIME(MIC)- AVG	STD_DEV	------------------ DELAYED REQUESTS ----	# REQ	% OF REQ	------ AVG TIME(M /DEL	STD_DEV
SYSD	29203K	CBP	2	2	0	SYNC	9047K	14.6	10.9	LIST/CACHE	9714	0.0	574.3	5937
	16224	SUBCH	14	14		ASYNC	20201K	47.4	215.1	LOCK	587	0.0	9.9	19.0
						CHANGED	6321	INCLUDED	IN ASYNC	TOTAL	10K	0.0		
						UNSUCC	0	0.0	0.0					

Same exercise could be done on a structure by structure basis

COUPLING FACILITY STRUCTURE ACTIVITY

STRUCTURE NAME = DSNDB3G_LOCK1 TYPE = LOCK STATUS = ACTIVE

SYSTEM NAME	# REQ TOTAL AVG/SEC	-------------- REQUESTS	# REQ	% OF ALL	-SERV TIME(MIC)- AVG	STD_DEV	REASON	------------- DELAYED REQUESTS # REQ	% OF REQ	---- AVG TIME(MIC) /DEL	STD_DEV	----- /ALL	EXTERNAL REQUEST CONTENTIONS	
SYSA	1946K	SYNC	1946K	6.7	15.6	8.4	NO SCH	0	0.0	0.0	0.0	0.0	REQ TOTAL	2226K
	1081	ASYNC	0	0.0	0.0	0.0	PR WT	0	0.0	0.0	0.0	0.0	REQ DEFERRED	67K
		CHNGD	0	0.0	INCLUDED	IN ASYNC	PR CMP	0	0.0	0.0	0.0	0.0	-CONT	67K
													-FALSE CONT	16K
SYSB	3471K	SYNC	3471K	11.9	12.8	7.5	NO SCH	38	0.0	11.4	4.5	0.0	REQ TOTAL	3617K
	1928	ASYNC	0	0.0	0.0	0.0	PR WT	0	0.0	0.0	0.0	0.0	REQ DEFERRED	77K
		CHNGD	0	0.0	INCLUDED	IN ASYNC	PR CMP	0	0.0	0.0	0.0	0.0	-CONT	77K
													-FALSE CONT	14K
SYSC	9728K	SYNC	9725K	33.5	11.9	7.4	NO SCH	10	0.0	9.2	4.6	0.0	REQ TOTAL	9643K
	5404	ASYNC	2999	0.0	75.4	89.9	PR WT	0	0.0	0.0	0.0	0.0	REQ DEFERRED	80K
		CHNGD	0	0.0	INCLUDED	IN ASYNC	PR CMP	0	0.0	0.0	0.0	0.0	-CONT	80K
													-FALSE CONT	18K
SYSD	4975K	SYNC	4975K	17.1	12.5	7.5	NO SCH	583	0.0	9.8	19.1	0.0	REQ TOTAL	5161K
	2764	ASYNC	0	0.0	0.0	0.0	PR WT	0	0.0	0.0	0.0	0.0	REQ DEFERRED	58K
		CHNGD	0	0.0	INCLUDED	IN ASYNC	PR CMP	0	0.0	0.0	0.0	0.0	-CONT	58K
													-FALSE CONT	13K
SYSE	8928K	SYNC	8927K	30.7	13.8	7.4	NO SCH	154	0.0	8.7	4.0	0.0	REQ TOTAL	7229K
	4960	ASYNC	1498	0.0	93.7	82.4	PR WT	0	0.0	0.0	0.0	0.0	REQ DEFERRED	94K
		CHNGD	0	0.0	INCLUDED	IN ASYNC	PR CMP	0	0.0	0.0	0.0	0.0	-CONT	94K
													-FALSE CONT	20K
TOTAL	29048K	SYNC	29M	100	12.9	7.6	NO SCH	785	0.0	9.7	16.6	0.0	REQ TOTAL	28M
	16138	ASYNC	4497	0.0	81.5	87.9	PR WT	0	0.0	0.0	0.0	0.0	REQ DEFERRED	376K
		CHNGD	0	0.0			PR CMP	0	0.0	0.0	0.0	0.0	-CONT	375K
													-FALSE CONT	81K

Quick Tip – Understand Which DASD I/O Logical Volumes to Concentrate On

What is the cost to z/OS for doing Coupling Facility synchronous requests?

Investigation: Determine how many CPU seconds were spent 'spinning' for the processing of coupling facility synchronous requests.

Tuning: Ensure within guidelines; determine most expensive structure if there is a problem.

Quick Tuning Tip – Calculate I/O Intensity

- What you need to get started:
  - Basic understanding of DASD I/O subsystem environment
  - Basic understanding of primary I/O response time component measurements
    - Reported on a logical volume basis
    - WLM Service Class and Report Class basis
    - Etc.
  - Understanding of I/O Intensity and Queue Intensity
    - Several reporting packages
  - A calculator or program to calculate I/O intensity values
    - Note: Must be calculated on a DASDplex basis
    - In other words, never look at I/O performance a system at a time. Any logical volume must be examined at the DASDplex level (by merging all measurements)
  - Concentrate on logical volumes with the highest I/O queuing intensity

Results of a Typical DASD I/O Analysis

- It is always critical to do a I/O subsystem analysis
  - Logical volumes impacting the workloads the most
  - Evaluation of I/O technologies
  - I/O balancing
- Typical SMF measurements analyzed for an I/O analysis include:
  - SMF 74.1 – Device Activity measurements
  - SMF 74.5 – Cache Control Unit measurements
- These give you a DASD I/O logical volume analysis from the logical volume point-of-view
- A detailed analysis provides exceptional benefit

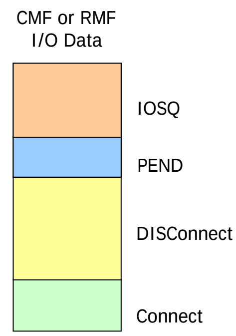
Ways of Looking at I/O Performance from an I/O Subsystem Point-of-View

- DASDplex level for all data
- Logical Volume level
  - Problem logical volumes (severe and warnings)
  - Response Time and Queue Time component analysis, I/O Intensity and Queuing Intensity analysis
  - Caching statistics
- Logical Control Unit (LCU) Level
  - Problem LCUs (severe and warnings)
  - Response Time and Queue Time component analysis, I/O Intensity and Queuing Intensity analysis
  - Caching Statistics
- Control Unit (CU)
  - Problem CUs (severe and warnings)
  - Response Time and Queue Time component analysis, I/O Intensity and Queuing Intensity analysis
  - Caching Statistics
- Other

Overview of Components of I/O Response Time

- DASD I/O requests and their response times
  - SSCH Rate - Total number of start sub-channel instructions
    Can think of this as the number of DASD I/O requests
  - RESP - Average DASD I/O response time (ms)

- I/O Response Time =
  IOSQ time
  + Pend time
  + Disconnect time
  + Connect time



Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 59

DASD I/O Response Time Measurements

- Measures shown are the I/O responses times broken into their four major components
  - IOSQ - IOSQ time is a delay time accumulated while the I/O is still in MVS and is waiting for a UCB to allow the I/O against the device
  - PEND - PEND is a delay time that usually occurs on devices with allegiance to multiple systems, and I/O are delayed waiting for the other system to complete its I/O
  - DISC - Disconnect time occurs when an I/O is being setup to complete
  - CONN - Connect times occurs as the data is being written out to the device. High disconnect times occur for a variety of reasons.

DASD I/O response time = IOSQ + PEND + DISC + CONN

Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 60

Intensity

- When it comes to I/O many analysts make the mistake on concentrating on the following:
  - Poor I/O response times
  - Logical volumes with the most activity
- It is better to look at the calculated *Intensity* values
  - A great measurement to gain insight into what component / workload to concentrate your analysis on
- Intensity = (Rate) * (Average Response Time)
  - Example: I/O Intensity = (I/O Rate) * (CONN + DISC + PEND + IOSQ)
  - Example: Queuing Intensity = (I/O Rate) * (DISC + PEND + IOSQ)

Calculating I/O Intensity and Queue Intensity

- Evaluating I/O performance based mainly on I/O response times is insufficient
  - Must take I/O rate into consideration
  - Since frequency of I/O influences the impact of the I/O response times
- intensity is a useful measure of the impact I/O may be having on the system

I/O Intensity = (I/O Rate) * (Average I/O Response Time)
Queue Intensity = (I/O Rate) * (Average I/O Response Time – CONN Time)

- Intensities can be calculated for
  - Logical Volumes (on DASDplex basis)
  - Control Units
  - LCUs
  - Service Class Period, Service Class, Report Class, Workload, WLM Service Policy (or system)
  - Etc.
- Regularly monitor

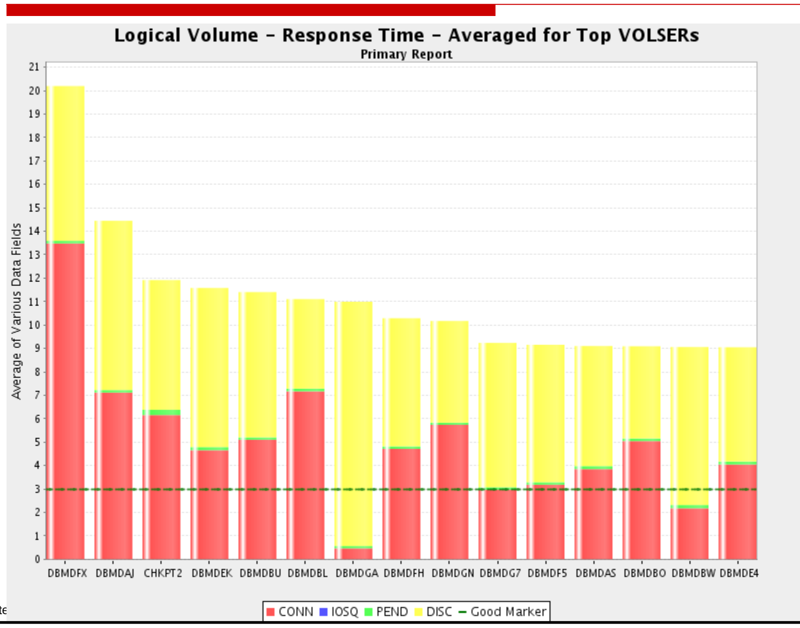
Why calculate intensity values?

- Common I/O guideline: DASD I/O response times should be 3 milliseconds or below (on average)

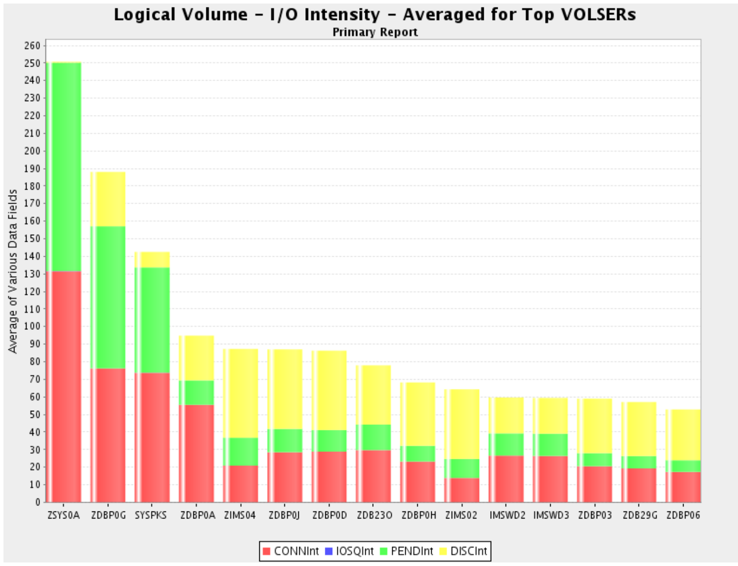
- Example 1:
  - Logical volume TSO001 has a 18 ms I/O response time with 10 I/Os per second
    - CONN = 12, DISC = 5, PEND = 2, and IOSQ = 1
    - I/O Intensity = (18 ms) * (10 I/O per second) = 180
    - I/O Queue Intensity = (6 ms) * (10 I/O per second) = 60

- Example 2:
  - Logical volume DBA022 has a 4 ms I/O response time with 1000 I/Os per second
    - CONN = 1, DISC = 2.5, PEND = 0.5, and IOSQ = 0
    - I/O Intensity = (4 ms) * (1000 I/O per second) = 4000
    - I/O Queue Intensity = (3 ms) * (1000 I/O per second) = 3000

Example: Note Top LVs based on I/O Response Time

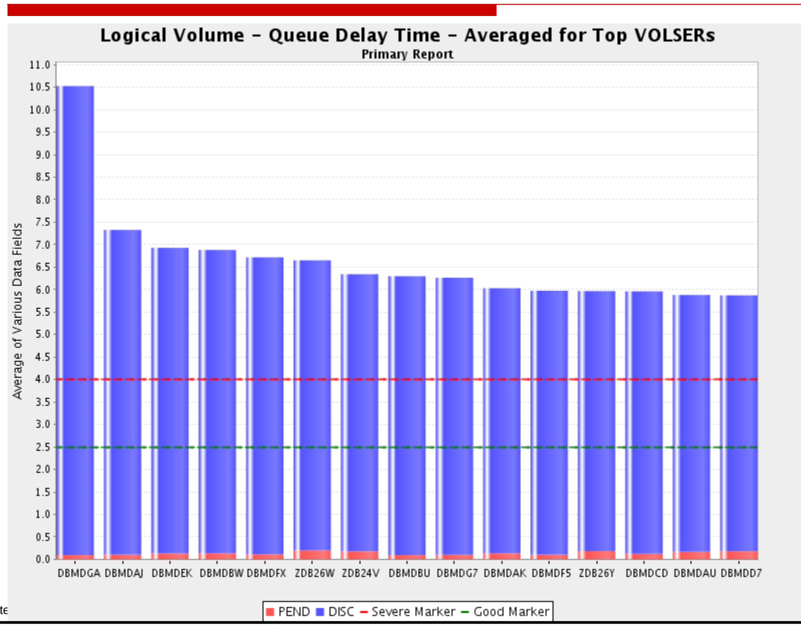


Example: Note Top LVs based on I/O Intensity



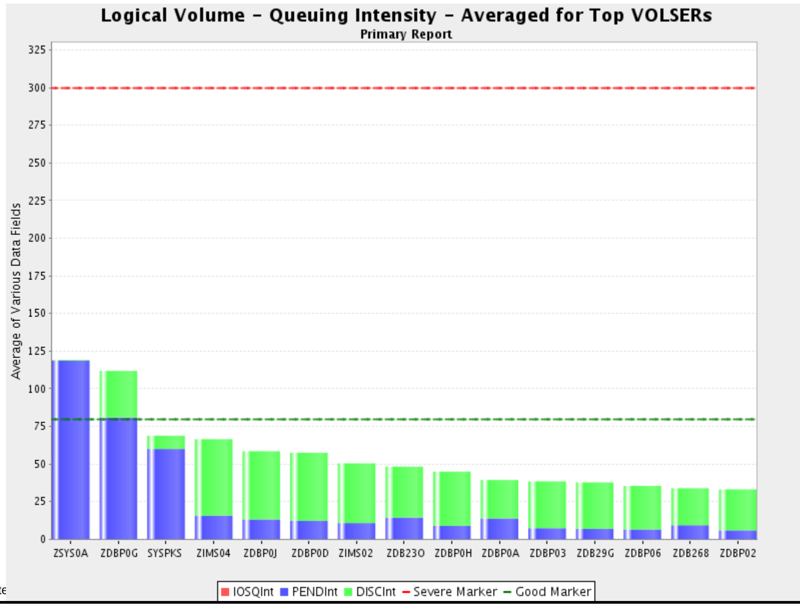
Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 65

Example: Note Top LVs based on I/O Queue Time



Pete ance Tuning Tips - 66

Example: Note Top LVs based on I/O Queue Intensity



Pete ... ance Tuning Tips - 67

Presentation Overview

- Quick Tip – Learn to Write!
- Quick Tip – Learn A Top Down Approach to Performance Analysis
- Quick Tip – Understand Your Workloads
- Quick Tip – Evaluate LPAR Weight Enforcement
- Quick Tip – Verify WLM Address Space Classification
- Quick Tip – Understand Coupling Facility Sync CPU Spin Seconds
- Quick Tip – Understand Which DASD I/O Logical Volumes to Concentrate On

Current 2011 Class Schedule

- WLM Performance and Re-evaluating of Goals
  - Instructor: Peter Enrico
  - September 12 – 16, 2011 Baltimore, Maryland, USA
- Parallel Sysplex and z/OS Performance Tuning
  - Instructor: Peter Enrico
  - September 19 – 23, 2011 Dallas, Texas, USA
- z/OS Capacity Planning and Performance Analysis
  - Instructor: Ray Wicks
  - August 15 – 17, 2011 Columbus, Ohio, USA

Peter Enrico : www.epstrategies.com © Enterprise Performance Strategies, Inc. zNextGen Performance Tuning Tips - 69