

Exploring z/OS WLM CPU Measurements:

SUs vs. CPU Seconds vs. APPL% vs. Workload%

Scott Chapman

Enterprise Performance Strategies, Inc.

Scott.Chapman@EPStrategies.com



Contact, Copyright, and Trademarks



Questions?

Send email to performance.questions@EPStrategies.com, or visit our website at <https://www.epstrategies.com> or <http://www.pivotor.com>.

Copyright Notice:

© Enterprise Performance Strategies, Inc. All rights reserved. No part of this material may be reproduced, distributed, stored in a retrieval system, transmitted, displayed, published or broadcast in any form or by any means, electronic, mechanical, photocopy, recording, or otherwise, without the prior written permission of Enterprise Performance Strategies. To obtain written permission please contact Enterprise Performance Strategies, Inc. Contact information can be obtained by visiting <http://www.epstrategies.com>.

Trademarks:

Enterprise Performance Strategies, Inc. presentation materials contain trademarks and registered trademarks of several companies.

The following are trademarks of Enterprise Performance Strategies, Inc.: **Health Check[®], Reductions[®], Pivotor[®]**

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries: IBM[®], z/OS[®], zSeries[®], WebSphere[®], CICS[®], DB2[®], S390[®], WebSphere Application Server[®], and many others.

Other trademarks and registered trademarks may exist in this presentation

Abstract



As always, z/OS processor measurements are critical during any z/OS performance analysis and tuning exercise. During this webinar, **Scott Chapman** will explore the WLM workload processor measurements in the SMF 72 record. Specifically, we will look at some of the key processor measurements available in the SMF 72, subtype 3 record and then explore calculating and using processor measurements at the service unit level, at the CPU seconds level, at the application percentage (APPL%) level, and finally at the workload utilization level.

Agenda



- What CPU measurements are in the SMF 72.3 record?
- What is being measured?
- How do you equate all the various representations of CPU consumption?

EPS: We do z/OS performance...



- Pivotor - Reporting and analysis software and services
 - Not just reporting, but analysis-based reporting based on our expertise
- Education and instruction
 - We have taught our z/OS performance workshops all over the world
- Consulting
 - Performance war rooms: concentrated, highly productive group discussions and analysis
- Information
 - We present around the world and participate in online forums

z/OS Performance workshops available



During these workshops you will be analyzing your own data!

- Essential z/OS Performance Tuning
 - October 3-7, 2022
- WLM Performance and Re-evaluating Goals
 - September 12-16, 2022
- Parallel Sysplex and z/OS Performance Tuning
 - August 8-12, 2022
- Also... please make sure you are signed up for our free monthly z/OS educational webinars! (email contact@epstrategies.com)

Like what you see?



- The z/OS Performance Graphs you see here come from Pivotor™
- If you don't see them in your performance reporting tool, or you just want a free cursory performance review of your environment, let us know!
 - We're always happy to process a day's worth of data and show you the results
 - See also: <http://pivotor.com/cursoryReview.html>
- We also have a **free** Pivotor offering available as well
 - 1 System, SMF 70-72 only, 7 Day retention
 - That still encompasses over 100 reports!

All Charts (132 reports, 258 charts)

All charts in this reportset.

Charts Warranting Investigation Due to Exception Counts (2 reports, 6 charts, [more details](#))

Charts containing more than the threshold number of exceptions

All Charts with Exceptions (2 reports, 8 charts, [more details](#))

Charts containing any number of exceptions

Evaluating WLM Velocity Goals (4 reports, 35 charts, [more details](#))

This playlist walks through several reports that will be useful in while conducting a WLM velocity goal an.



SMF 72.3

SMF 72.3: Workload Activity



- Contains interval measurements by:

- Service Class Period
- Report Class Period
- Resource Group (not discussed here)

- To try to confuse us, contains CPU measurements three ways:

- Service Units

R723CCPU	8	floating	Total TCB service units. This value includes SUs on general purpose CPs and normalized SUs on zIIPs and zAAPs.
----------	---	----------	--

- CPU Time

R723TPDP	8	floating	CPU time consumed while dispatching priority of work with low importance was temporarily raised to help blocked workloads (in 1024 microsecond units).
----------	---	----------	--

- CPU Samples

R723CCUS	4	binary	CPU using samples. These are included in R723CTOU.
----------	---	--------	--

These are measures of CPU consumption

This is not a measure of CPU consumption!
(But may be indicative of such.)

CPU terminology can be confusing



MIPS Percent Utilization Appl %

MSUs Percent Busy CPU Using

SUs Workload % MVS Busy

CPU Seconds

Most of the raw SMF measurements are going to be CPU seconds or SUs

MT1ET

Less confusing: beer!



- Let's say beer flows out of this tap at the rate of 1 ounce every 1.875 seconds
- So it dispenses at the speed of 0.53 oz/sec
- It has capacity for dispensing 1920 oz/hour
- If last hour we served up 20 beers how much of the tap's capacity did we use?
 - How big were the beers?
 - Pints = 16 oz = $20 * 16 / 1920 = 16.667\%$
 - Liters = 33.8 oz = $20 * 33.8 / 1920 = 35.2\%$
- We could also rate the tap as 120 pints/hour
- How long did it take to dispense 20 pints?
 - $20 * 16 * 1.875 = 600$ seconds

Capacity vs. Consumption



- Time is beer (1.875 oz per second)
- Beer is time (30 seconds to pour a pint)
- The tap has a speed (1.875 oz/second) which equates to a capacity to do work (120 PPH – Pints Per Hour)
- If we know how many pints were poured over the course of some interval we can calculate the utilization of the tap for that interval
 - From there, we can calculate how much of the capacity was used
 - I poured 20 liters in an hour, how many PPH of my tap did I use?
 - $\text{Liters} = 33.8 \text{ oz} = 20 * 33.8 / 1920 = 35.2\%$ of the tap's capacity
 - $35.2\% * 120 \text{ PPH} = \text{A capacity of } 42.24 \text{ PPH}$ was used
 - Note that you always need an interval! “How many PPH is 4 liters?” doesn't make sense without an interval to compare against.

Beer and CPUs



- Like our CPUs (ignoring SMT) our tap can only pour one beer at a time
- But if we're properly pouring Guinness we might multi-thread!
 - I.E. while we're letting one partial pint settle, we could start on the next
 - Just like we'll let the CPU work on another program when one program starts waiting for an I/O
- Adding more taps may not increase capacity linearly if we don't add bartenders too (and they'll start tripping over each other)
 - Similar contention happens as you add more CPUs

Coming back to CPUs...



- For the beer tap, the work is dispensing beer
 - We can measure the work done in ounces, pints, liters, or even time that the tap was dispensing beer
- For CPUs, the work is fetching, transforming, storing data
 - At the lowest level, the work is measured in time that the CPU spent working
- As pints can be converted to time, so CPU time can be converted to other measurements
 - We just need the conversion factors

CPU Time



- CPU time = total time that a CPU has spent performing work for task
 - Time that a workload is dispatched to a CPU
- z/Architecture provides CPU timers with nominal resolution of 1 μ s
 - Yes, that's one millionth of a second, although the times aren't usually externalized to that precision
- Instruction EXTRACT CPU TIME (ECTG) can be used by problem-state programs to determine the amount of CPU time consumed by the current task
- When a CPU is interrupted to process something else, the CPU timer is readjusted once the interrupted task is dispatched again
- CPU timers are not dependent on the time-of-day clock because (e.g.) the time of day clock may be steered to remain in sync with a time source
 - One of the reasons why use of system time for performance analysis is problematic

IBM Conversion Factors



- E.G. a z16 3931-604 is said to deliver 70,484 SUs/second and has a rated capacity of 675 MSUs or 5511 PCI (IBM’s “MIPS” by another name)
- Note that machine with 4 CPUs is not 4x the capacity of machine with 1 CPU

Processor	STIDP Type	STSI Model	#CP	SU/SEC	SRMsec/RealSec
3931-601	3931	601	1	76555.0239	1780.3494
3931-602	3931	602	2	73394.4954	1780.3494
3931-603	3931	603	3	71748.8789	1780.3494
3931-604	3931	604	4	70484.5815	1780.3494

<https://www-40.ibm.com/servers/resourcelink/lib03060.nsf/pages/srmindex#ibmz16>

Processor	#CP	PCI**	MSU***	Low*	Average*	High*
3931-601	1	1496	185	2.68	2.67	2.56
3931-602	2	2873	355	5.22	5.13	4.81
3931-603	3	4211	519	7.71	7.52	6.97
3931-604	4	5511	675	10.15	9.85	9.03

<https://www-40.ibm.com/servers/resourcelink/lib03060.nsf/pages/lsprITRzOSv2r4#ibmz16>

Service Units



- Service Units intended to translate resource utilization to a value that can be used across different machine generations and configurations
 - WLM policy uses SUs to measure resource consumption
- Actual recorded values based on service coefficients defined in the WLM service definition (except when they're "unweighted")
 - Recommended values today (enforced in z/OS 2.5):
 - CPU: **1**, SRB: **1**, I/O: **0**, MSO: **0**
 - CPU Service units = CPU time * SRM Constant * CPU Coefficient
 - SRM Constant recorded in SMF data, available in control blocks, & online here: <https://www-304.ibm.com/servers/resourcelink/lib03060.nsf/pages/srmindex>
 - SRM constant is based on number of GCPs online to the LPAR and the machine model
 - So on a 604, an LPAR with 3 online GCPs gets the SRM constant for a 603
 - Even if one of those 3 is parked

(Ignoring details about types of CPU time)

SUs <-> CPU Time



- Just do the math
 - SUs/second and CPU time expressed in seconds
- If our LPAR has 3 CPUs, and the SU/sec is recorded as 71748.8789
- R723CCPU = 165,400 for SCP GUINNESS_1, then the first period of GUINNESS_1 consumed $165400 / 71748.8789 = 2.305$ CPU seconds
- R723TPDP = 5120 CPU microseconds for SAMADAMS_1 then it used $71748.8789 * .005120 = 367$ service units while promoted to help blocked workloads

R723CCPU	8	floating	Total TCB service units. This value includes SUs on general purpose CPs and normalized SUs on zIIPs and zAAPs.
R723TPDP	8	floating	CPU time consumed while dispatching priority of work with low importance was temporarily raised to help blocked workloads (in 1024 microsecond units).

MIPS (PCI) and MSUs



- We need to know the interval
 - If BLLXPDTN_3 used 360 CPU seconds in a 900 second interval on our z16-604:
 - $4 \text{ CPUs} * 900 = 3600 \text{ CPU seconds available}$
 - $360 / 3600 = 0.1 = 10\%$ of available capacity (aka 10% Physical Machine Utilization)
 - $675 \text{ MSUs} * 0.1 = 67.5 \text{ MSUs}$
 - $5511 \text{ PCI} * 0.1 = 551.1 \text{ PCI}$
- } Over the 15 minute interval
- Note we could also say that BLLXPDTN_3 used $360 / 900 = 40\%$ of one CPU
 - Sometimes expressed as 40 APPL% (percent of one CPU)
 - Sometimes expressed as 0.4 CPUs

When are MSUs not MSUs?



- When you're measuring for Tailored Fit Pricing!
 - Well, they are, but...
- MSUs as I've used on the previous charts are based on the idea that MSUs and MIPS are a capacity rating for the machine
 - So just as we said we used 10% of the capacity of the machine, we used 10% of the MIPS or MSUs and so do the easy math
- But for TFP, you are effectively billed in MSUs which are a measure of your CPU consumption, not a percentage of your capacity over the interval
 - I.E. it is equivalent to saying how many CPU seconds have you consumed, not what percent of your capacity in this interval was consumed
- MSUs stands for Millions of Service Units *per hour*
 - Even though they no longer are that specific calculation
- It would be clearer if IBM said they charge you in "MSU-hours"
 - Presumably they felt that was redundant (surely they weren't trying to be confusing!)

CPU Time to MSUs for TFP



- Basically... normalize to an hour
 - Do everything based on hour interval
 - Scale the MSUs by the percent the interval is of an hour
- Looking at 4 interval of BLLXPDTN_3:

Interval	CPU Secs	MSUs	MSU-hrs
10:14	360	67.5	16.875
10:29	120	22.5	5.625
10:44	0	0	0
10:59	560	105	26.25
Hour 10	1040	48.75	12.1875

Logical vs. Physical



- All z/OS CPU time is physical
 - One actual physical CPU was busy for some time doing some work
 - The fact that the physical CPU is represented by a logical CPU in z/OS is immaterial
 - Now two logical CPUs can be using the same physical CPU at the same instant
- If you have 4 physical CPUs on a machine the total available CPU time per hour is 4 hours
 - If you add up all the CPU time across all LPARs, it will be ≤ 4 hours
- But... we *could* calculate a logical LPAR utilization level as well
 - Today, in most environments, this is of somewhat limited use
 - Basically, it lets you see how close you're getting to the limit imposed by the number of logical CPs the LPAR has assigned to it

Workload Percent



- CPU time / (logical CPs * interval length)
 - This work was what percent of the LPAR's potential capacity based on the number of logical processors?
- This and MVS Busy could be a useful of your LPAR reaching the limit of its logical CPs
 - This used to be a bigger concern before HiperDispatch when it was more important to closely control the number of logical processors
 - But today having 1-2 more logical CPs than the LPAR needs (i.e. have 1 or 2 low pool processors) is fine and even a good practice
 - And such LPARs may find it impossible to use 100% of their logical CPs' capacity
- Workload Percent is dependent on the number of CPs online so can't always be compared between LPARs even running on the same machine
 - 200 CPU seconds, 900s interval with 2 logical CPs = $200 / (2 * 900) = 11.1\%$
 - 200 CPU seconds, 900s interval with 3 logical CPs = $200 / (3 * 900) = 7.4\%$



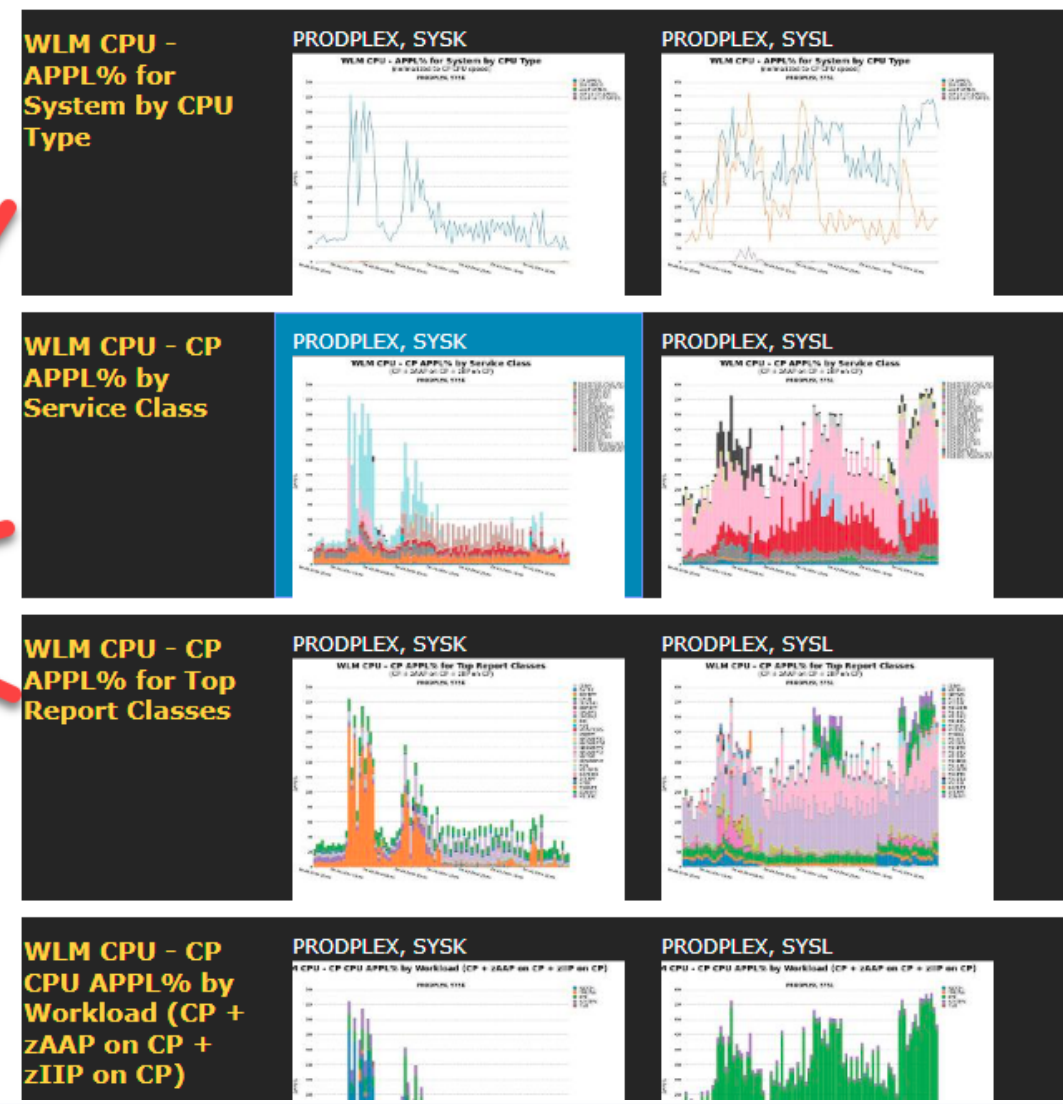
Pivotor reports



Search Reports Titles Tags Values All
Playlist 0/630 View Day 20220620 unzoom Hold No series Alter Chart Report Help
Chart Table Image CSV Show All series

- z/OS Health Check
 - Processor Analysis
 - MSU Usage Analysis
 - Storage / Memory Analysis
 - SMF 30 Address Space Analysis
 - WLM Analysis
 - WLM Policy Analysis
 - WLM PI Analysis
 - WLM Importance Analysis
 - WLM Velocity Goal Analysis
 - WLM Response Time Goal Analysis
 - Discretionary Goal Analysis
 - WLM Multi-Period Analysis
 - WLM Enclave Analysis
 - WLM CPU APPL Analysis
 - WLM CPU - Captured APPL% CP CPU
 - General Purpose Engines**
 - WLM CPU - APPL% for System by CPU Type
 - WLM CPU - CP APPL% by Service Class
 - WLM CPU - CP APPL% for Top Report Classes
 - WLM CPU - CP CPU APPL% by Workload (CP + zAAP on CP + zIIP on CP)
 - Specialty Engines
 - WLM CPU APPL Trx Analysis
 - WLM Storage Analysis
 - WLM I/O Analysis
 - WLM Contention Analysis
 - Resource Group Analysis
 - WLM Throughput Analysis
 - WLM Samples Detailed Analysis
 - Virtual Storage Activity Analysis
 - Environment Summary

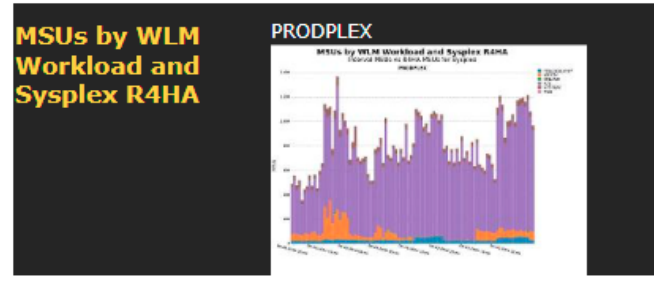
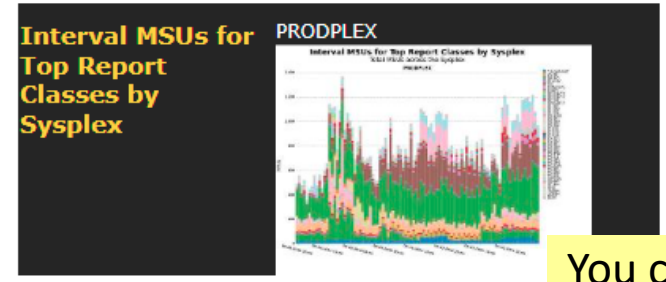
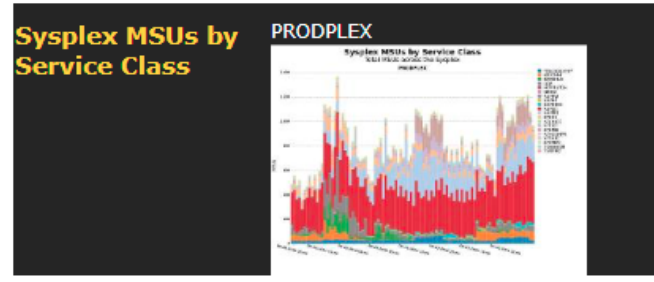
General Purpose Engines



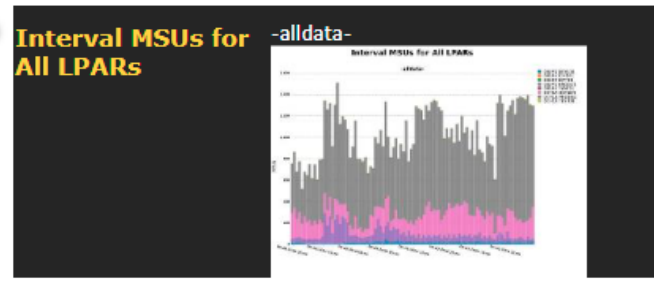
This is where I generally go to look at CPU consumption by service class, report class, or workload.

- z/OS Health Check
 - Processor Analysis
 - MSU Usage Analysis
 - LPAR - R4HA, Actual MSUs, Image Capacity
 - Long-term (R4HA) Average MSUs for All LPARs
 - CEC Group Cap R4HA Average MSUs for LPARs
 - MSU Usage by CEC
 - MSU Usage by Sysplex
 - Sysplex MSUs by Service Class
 - Interval MSUs for Top Report Classes by Sysplex
 - MSUs by WLM Workload and Sysplex R4HA
 - MSU Usage by LPAR
 - Interval MSUs for All LPARs
 - CEC Capping Actually Limited Percentage
 - MSUs by Service Class and LPAR R4HA
 - Interval MSUs for Top Report Classes by LPAR
 - MSUs by WLM Workload and LPAR R4HA
 - Storage / Memory Analysis
 - SMF 30 Address Space Analysis
 - WLM Analysis
 - Virtual Storage Activity Analysis
 - Environment Summary

MSU Usage by Sysplex



MSU Usage by LPAR



You can also see how many MSUs the service and report classes are using each interval here.

Note these are capacity based, not MSU-hours for TFP purposes. I.E. in each 15 minute interval if you add up all your SCs you could get to (potentially) 100% of your capacity.

Search Reports Titles Tags Values All

Playlist 0/630 View Day 20220620 unzoom Hold No series Alter Chart

Home Search Exact Chart Table Image CSV Show All series Report Help

- z/OS Health Check
 - Processor Analysis
 - Machine Level Analysis
 - LPAR Level Analysis
 - Processor Cache Counters (SMF 113)
 - SMT Analysis
 - HiperDispatch Analysis
 - WLM CPU Workload% Analysis
 - General Purpose Engines**
 - WLM CPU - CP CPU Workload% by Service C
 - WLM CPU - CP CPU Workload% for Top Repo
 - Specialty Engines
 - Crossover Analysis
 - MSU Usage Analysis
 - Storage / Memory Analysis
 - SMF 30 Address Space Analysis
 - WLM Analysis
 - WLM Policy Analysis
 - WLM PI Analysis
 - WLM Importance Analysis

General Purpose Engines

WLM CPU - CP CPU Workload% by Service Class

PRODPLEX, SYSK

WLM CPU - CP CPU Workload% by Service Class

PRODPLEX, SYSL

WLM CPU - CP CPU Workload% for Top Report Classes

PRODPLEX, SYSK

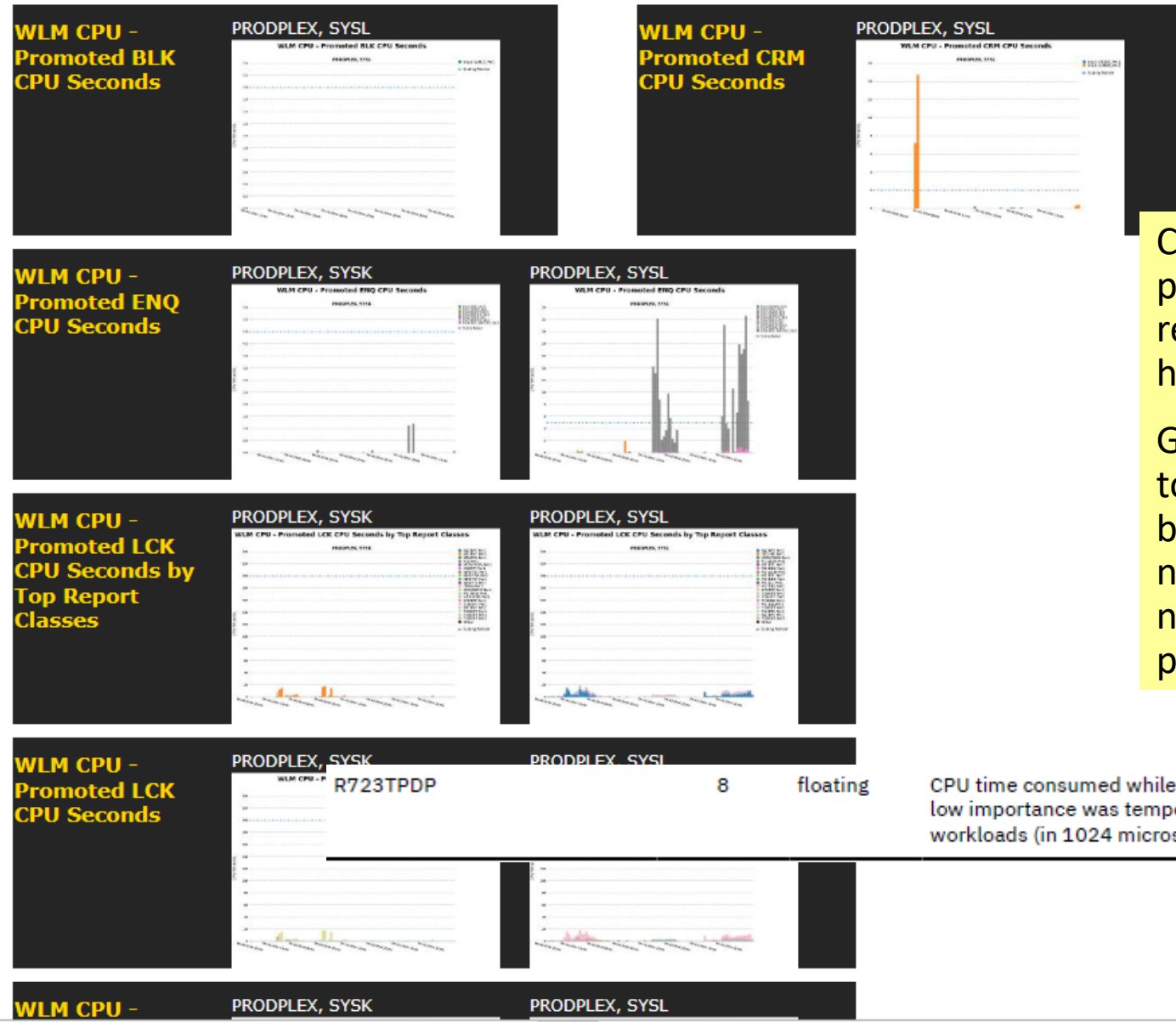
WLM CPU - CP CPU Workload% for Top Report Classes

PRODPLEX, SYSL

If you really want the logical workload percent value it's here.

- z/OS Health Check
 - Processor Analysis
 - MSU Usage Analysis
 - Storage / Memory Analysis
 - SMF 30 Address Space Analysis
 - WLM Analysis
 - WLM Policy Analysis
 - WLM PI Analysis
 - WLM Importance Analysis
 - WLM Velocity Goal Analysis
 - WLM Response Time Goal Analysis
 - Discretionary Goal Analysis
 - WLM Multi-Period Analysis
 - WLM Enclave Analysis
 - WLM CPU APPL Analysis
 - WLM CPU APPL Trx Analysis
 - WLM Storage Analysis
 - WLM I/O Analysis
 - WLM Contention Analysis
 - WLM CPU - Promoted BLK CPU Seconds
 - WLM CPU - Promoted CRM CPU Seconds
 - WLM CPU - Promoted ENQ CPU Seconds
 - WLM CPU - Promoted LCK CPU Seconds
 - WLM CPU - Promoted LCK CPU Seconds
 - WLM CPU - Promoted SUP CPU Seconds
 - Resource Group Analysis
 - WLM Throughput Analysis
 - WLM Samples Detailed Analysis
 - Virtual Storage Activity Analysis
 - Environment Summary

WLM Contention Analysis



CPU Time while promoted for various reasons can be found here.

Generally expect these to be smallish numbers, but somewhat larger numbers are not necessarily a significant problem.

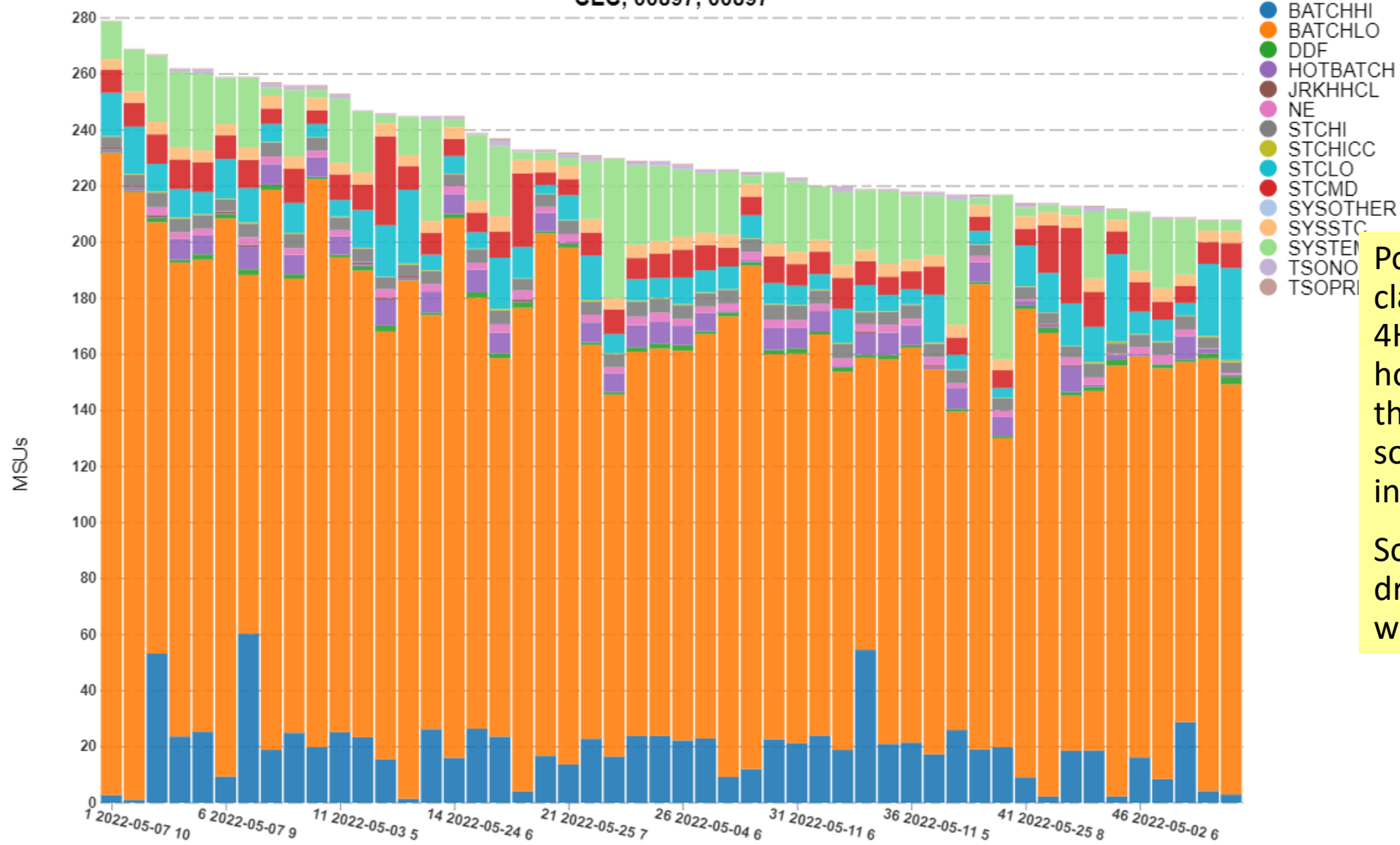
CPU time consumed while dispatching priority of work with low importance was temporarily raised to help blocked workloads (in 1024 microsecond units).



Top Service Classes

Estimated Billable R4HA MSUs

CEC, 00897, 00897



Portion each service class contributed to the 4HRA MSUs for the peak hours over the course of the month can sometimes be quite interesting.

So if "BATCHLO" is driving our peaks, could we cap and save money?

Summary



- Beer is time, time is beer
- CPU time is MSUs is MIPS, MIPS is MSUs is CPU time
- Converting between the different ways of expressing CPU time is generally quite useful
- But Scott prefers to avoid expressions that are based on logical CPs

Bonus: what about those CPU samples??

R723CCUS	4	binary	CPU using samples. These are included in R723CTOU.
----------	---	--------	--

This is not a
measure of CPU
consumption!
(But may be
indicative of such.)

CPU Samples are not CPU Time

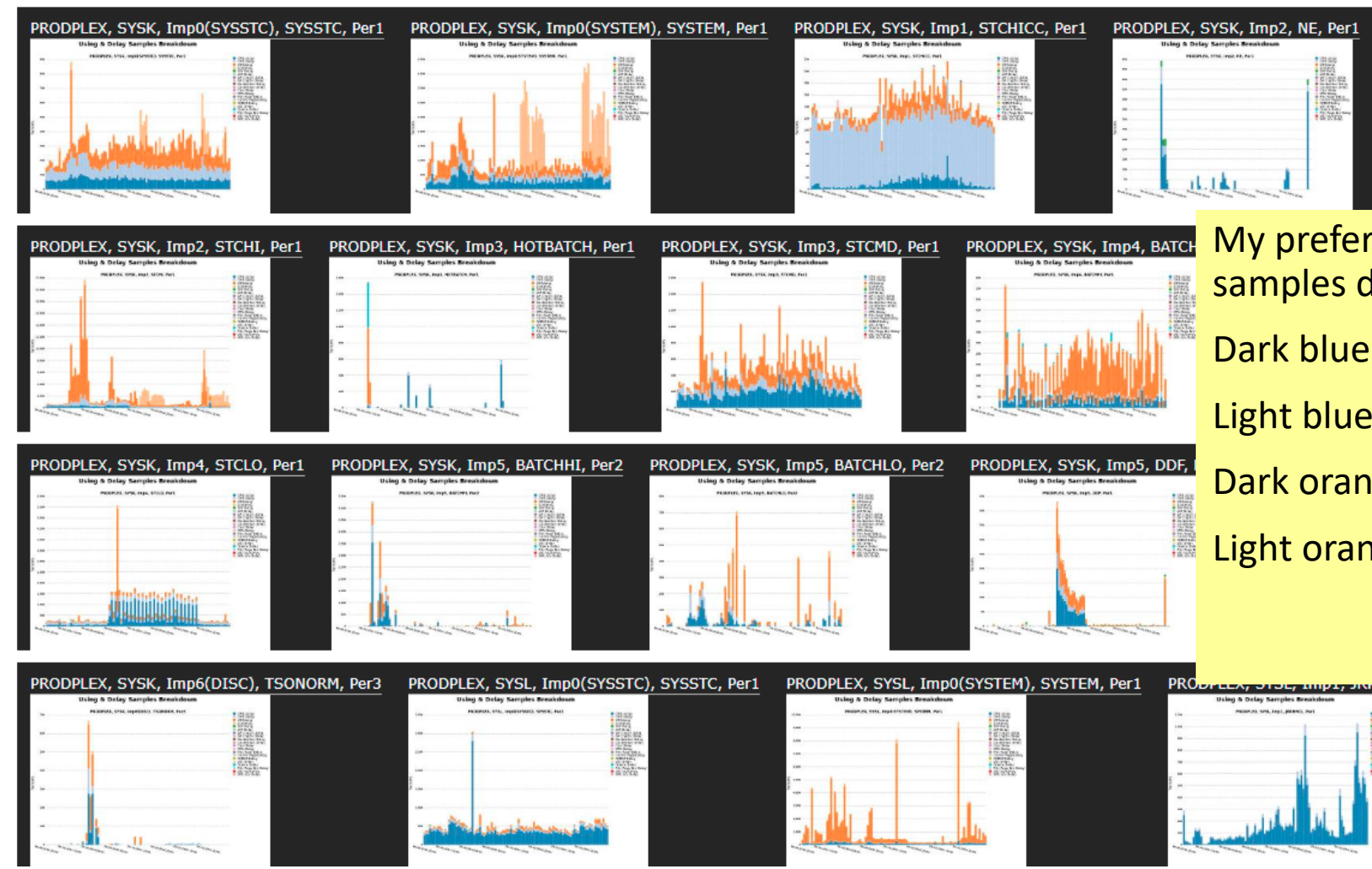


- The CPU using and delay samples come from SRM sampling the state of the system every 250ms
 - Samples are adjusted for the fact that one CP is sampling the system!
- Higher CPU using samples are probably indicative of more CPU consumption, but you don't know the magnitude
- If we check the beer tap every 5 minutes and note whether the tap is dispensing beer or not, can we determine how many pints were poured over the hour?
 - Of course not, although an hour that we note the tap was in use in 6 of our checks probably means that more beer was dispensed than when there were only 2.
 - But what if those 6 were all 4 oz pours for flights and the 2 samples in the other hour were for liters?
- Nonetheless, the samples have value, especially relative to the other sample states
 - Just can't use them for understanding specifically how busy the CPUs are



- z/OS Health Check
 - Processor Analysis
 - MSU Usage Analysis
 - Storage / Memory Analysis
 - SMF 30 Address Space Analysis
 - WLM Analysis
 - WLM Policy Analysis
 - WLM PI Analysis
 - WLM Importance Analysis
 - WLM Velocity Goal Analysis
 - WLM Response Time Goal Analysis
 - Discretionary Goal Analysis
 - WLM Multi-Period Analysis
 - WLM Enclave Analysis
 - WLM CPU APPL Analysis
 - WLM CPU APPL Trx Analysis
 - WLM Storage Analysis
 - WLM I/O Analysis
 - WLM Contention Analysis
 - Resource Group Analysis
 - WLM Throughput Analysis
 - WLM Samples Detailed Analysis
 - WLM Using Samples by Period
 - WLM Delay Samples by Period
 - Using & Delay Samples Breakdown
 - PRODPLEX, SYSK, Imp0(SYSSTC), SYSSTC, Per1
 - PRODPLEX, SYSK, Imp0(SYSTEM), SYSTEM, Per1
 - PRODPLEX, SYSK, Imp1, STCHICC, Per1
 - PRODPLEX, SYSK, Imp2, NE, Per1
 - PRODPLEX, SYSK, Imp2, STCHI, Per1
 - PRODPLEX, SYSK, Imp3, HOTBATCH, Per1
 - PRODPLEX, SYSK, Imp3, STCMD, Per1
 - PRODPLEX, SYSK, Imp4, BATCHHI, Per1
 - PRODPLEX, SYSK, Imp4, BATCHLO, Per1
 - PRODPLEX, SYSK, Imp4, STCLO, Per1
 - PRODPLEX, SYSK, Imp5, BATCHHI, Per2
 - PRODPLEX, SYSK, Imp5, BATCHLO, Per2
 - PRODPLEX, SYSK, Imp5, DDF, Per2
 - PRODPLEX, SYSK, Imp6(DISC), TSONORM, Per3
 - PRODPLEX, SYSL, Imp0(SYSSTC), SYSSTC, Per1
 - PRODPLEX, SYSL, Imp0(SYSTEM), SYSTEM, Per1
 - PRODPLEX, SYSL, Imp2, STCHICC, Per1

Using & Delay Samples Breakdown



My preferred view of the samples data. Note:

- Dark blue: CPU using
- Light blue: CPU delay
- Dark orange: IO using
- Light orange: IO delay



Thanks!

Questions?

Email: scott.chapman@epstrategies.com