WLM and SMF 99.1 – System Measurements Deeper Dive

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

Creators 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 or http://www.pivotor.com.

Copyright Notice:

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

- **WLM and SMF 99.1 – System Measurements Deeper Dive**

- The SMF 99.1 record has been around since the first availability of WLM. It contains system level data, the trace of SRM actions, data about resource groups, and a whole host of other measurements. The SMF 99.1 records system levels details used by the WLM policy and resource adjustment algorithms.

- During this webinar, Peter Enrico will walk through the measurements of the SMF 99.1 record and show how these measurements can be used. During this webinar, the attendees will also gain insights into the WLM policy adjustment algorithms and how they work.

EPS: We do z/OS performance…

- **Pivotor** – z/OS performance reporting and analysis software and services
  - Not just SMF reporting, but analysis-based reporting based on expertise
  - www.pivotor.com

- **Education and instruction**
  - We teach our z/OS performance workshops all over the world
  - Want a workshop in your area? Just contact me.

- **z/OS Performance War Rooms**
  - Intense, concentrated, and highly productive on-site performance group discussions, analysis and education
  - Amazing feedback from dozens of past clients

- **MSU Reduction Exercises**
  - The goal is to reduce the MSU consumption of your applications and environment

- **Information**
  - We present around the world and participate in online forums
  - https://www.pivotor.com/content.html
    https://www.pivotor.com/webinar.html

z/OS Performance workshops available

During these workshops you will be analyzing your own data!

- WLM Performance and Re-evaluating Goals
  - February 19-23, 2024

- Parallel Sysplex and z/OS Performance Tuning
  - August 20-21, 2024

- Essential z/OS Performance Tuning
  - October 7-11, 2024

- Also… please make sure you are signed up for our free monthly z/OS educational webinars! (email contact@epstrategies.com)

Like what you see?

- Free z/OS Performance Educational webinars!
  - The titles for our Summer / Fall 2024 webinars are as follows:
    - ✓ *What a z/OS Guy Learned About AWS in 10 Years*
    - ✓ *Advantages of Multiple Period Service Classes*
    - ✓ *Understanding z/OS Connect Measurements*
    - *WLM and SMF 99.1 – System Measurements Deeper Dive*
    - *WLM and SMF 99.2 – Service Class Period Measurements Deeper Dive*
    - *Optimizing Performance at the Speed of Light: Why I/O Avoidance is Even More Important Today*
    - *Understanding MVS Busy % versus LPAR Busy % versus Physical Busy %*
    - *Rethinking IBM Software Cost Management Under Tailored Fit Pricing*
    - *Understanding Page Faults and Their Influence on Uncaptured Time*
    - *Response Time Goals: Average or Percentiles?*
    - *Understanding and Using Enclave*

- If you want a free cursory 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

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!

History of the SMF 99 Record

- When WLM was first being developed, lots of trace records were embedded in the code so the WLM algorithm developers could understand and debug WLM actions
  - Before WLM was released, all these trace record were turned off

- When WLM was shipped in 1994, it was realized that that there would be a need for these trace records to assist in understanding and debug of customer situations
  - So, at the last minute, via an APAR, the SMF 99 records were coded and introduced

  - Originally only consisted of SMF 99 subtypes 1 through 6
  - Later releases of MVS introduced additional subtypes

Overview of SMF 99 Subtypes

- Subtype 1
  - System level measurement data used for decision input
  - Trace of WLM actions
  - Written every 10 seconds (i.e. policy adjustment interval)
- Subtype 2
  - Service class period measurement data used for decision input
  - Written every 10 seconds (i.e. policy adjustment interval)
- Subtype 3
  - Service class period plot data
  - Written every 10 seconds (i.e. policy adjustment interval)
- Subtype 4
  - Service class device cluster information
  - Written every 10 seconds (i.e. policy adjustment interval)
- Subtype 5
  - Data about monitored address spaces
  - Written every 10 seconds (i.e. policy adjustment interval)

Overview of SMF 99 Subtypes cont…

- Subtype 6
  - Service class period settings and measurements
  - Written every 10 seconds (i.e. policy adjustment interval)
- Subtype 7
  - Enterprise Storage Server ® (ESS) with Parallel Access Volumes (PAVs)
  - Written every 30 seconds (i.e. 3 policy adjustment intervals)
- Subtype 8
  - Information about LPAR CPU management
  - Written every 10 seconds (i.e. policy adjustment interval)
- Subtype 9
  - Information about dynamic channel path management
  - Written every 10 seconds (i.e. policy adjustment interval)
- Subtype 10
  - Information about dynamic processor speed changes
  - Written when speed changes

Overview of SMF 99 Subtypes cont…

- Subtype 11
  - Information about Group Capacity Limits
  - Written every 5 minutes
- Subtype 12
  - HiperDispatch interval data
  - Written every 2 seconds (i.e. policy adjustment interval)
- Subtype 13
  - HiperDispatch IBM internal use only (so undocumented)
  - *And very voluminous!*
- Subtype 14
  - HiperDispatch topology data
  - Written every 5 minutes

SMF 99 Recommendations

- Consider regularly collecting the following SMF 99 subtypes
  - Subtype 6 - Service class period settings and measurements
  - Subtype 11 - Information about Group Capacity Limits
  - Subtype 12 - HiperDispatch interval data
  - Subtype 14 - HiperDispatch topology data

  - Collectively these records typically produce about 40MiB/system/day
  - They contain the most interesting and useful data of the 99s

- Records to collect for problem periods of time, or when doing a study to better understand WLM decision making
  - Subtype 1 - System level measurement and trace data used for decisions
  - Subtype 2 - Service class period measurement data used for decision input
  - Subtype 3 - Service class period plot data
  - Subtype 5 - Data about monitored address spaces

  - Then call Peter Enrico and / or Scott Chapman to process with Pivotor

SMF 99.1 Record

Note: This presentation is meant to give a taste of the contents of the SMF 99.1 record. In no way is it a complete lesson, and in no way are all fields and value understood

Key Lesson about the SMF 99.1 Records

- The original intention of the SMF 99 records was to help IBM debug WLM algorithm decisions
  - These records are cut every 10 seconds
- While practical use has been found by regularly processing an analyzing certain subtypes:
  - Such as the SMF 99.6, SMF 99.12, SMF 99.14
- The other SMF 99 subtypes, such as SMF 99.1 and SMF 99.2:
  - Are cryptic, not well documented
  - Contain values that are only meaningful to WLM algorithms
  - Are not very useful for our day-to-day use by general performance analyst
- At EPS, we use the SMF 99.1 and SMF 99.2 by searching for unusual patterns during periods of time customers call us asking to analyze a situation

Overview of SMF 99, subtype 1 record

- Subtype 1
  - Contains system level data, the trace of SRM actions, and data about resource groups. The SRM actions are recorded in trace codes.

- Triplets include:
  - System state data section
  - Trace table entries section
  - System paging plot section
  - Priority table entry section
  - Resource group entry section
  - Generic resource group entry section
  - Software licensing information section
  - Software licensing table information

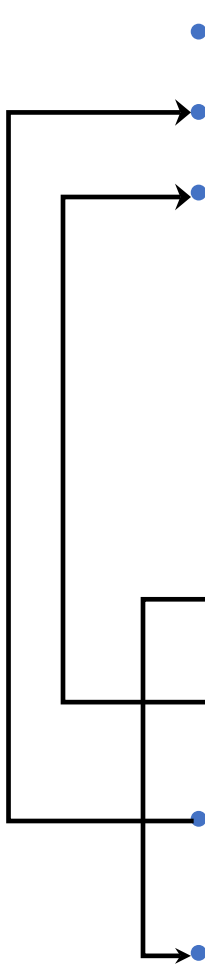
SMF 99.1– PA and RA Decisions

- There are two primary phases of WLM algorithms

- Policy Adjustment (PA)
  - Done approximately every 10 seconds (AKA 'PA interval')
  - Objectives include:
    - Summarize state of system and resources
    - Help work meet goals by setting resource controls
    - Housekeep resource controls that may be out of date

- Resource Adjustment (RA)
  - Done approximately every 2 seconds (AKA 'RA interval')
  - Objectives include:
    - improve efficiency of system resources
    - avoided if at the expense of goals

WLM Policy Adjustment – 'The Loop'



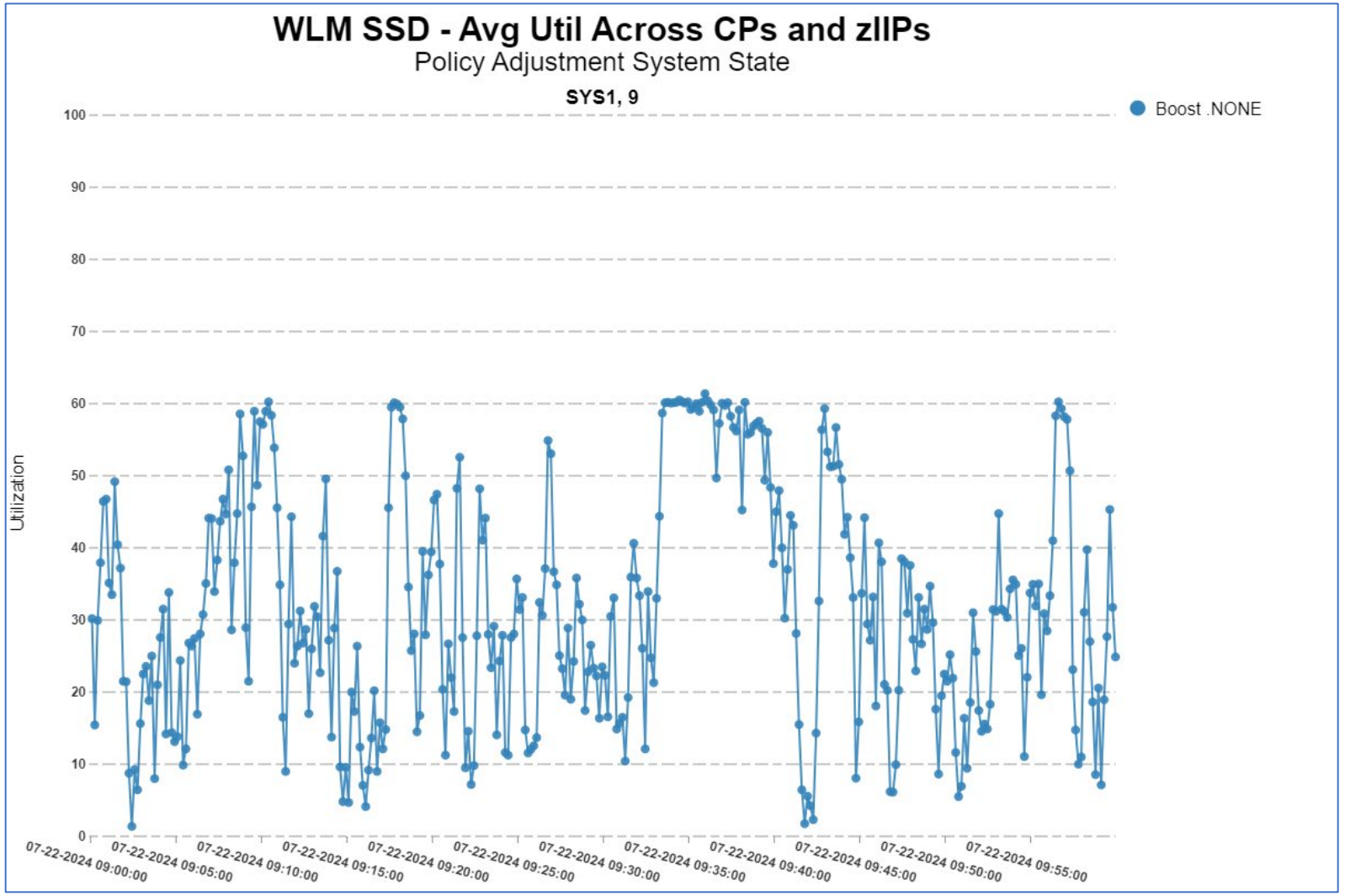
- Summarize data for state of the system and workloads
- Select a receiver period (highest importance missing goal the most)
- Find the receiver's largest bottleneck
  - Determine fix for receiver's bottleneck
    - Determine if needed resources can be gotten from unused resources
    - Find donor(s) of resource that receiver needs
    - Assess effect of reallocating resources from donor(s) to receivers
    - If allocation has both net and receiver value
      Then commit change
      Else don't make change
  - If reallocation was done
    then jump to Exit and allow change to be absorbed
  - If reallocation was not done
    then try to fix receiver's next largest bottleneck
- If cannot help receiver
  then look for next receiver (highest importance missing goal the most)
- Exit
  - Housekeep current set of controls

SMF 99.1 System State Data

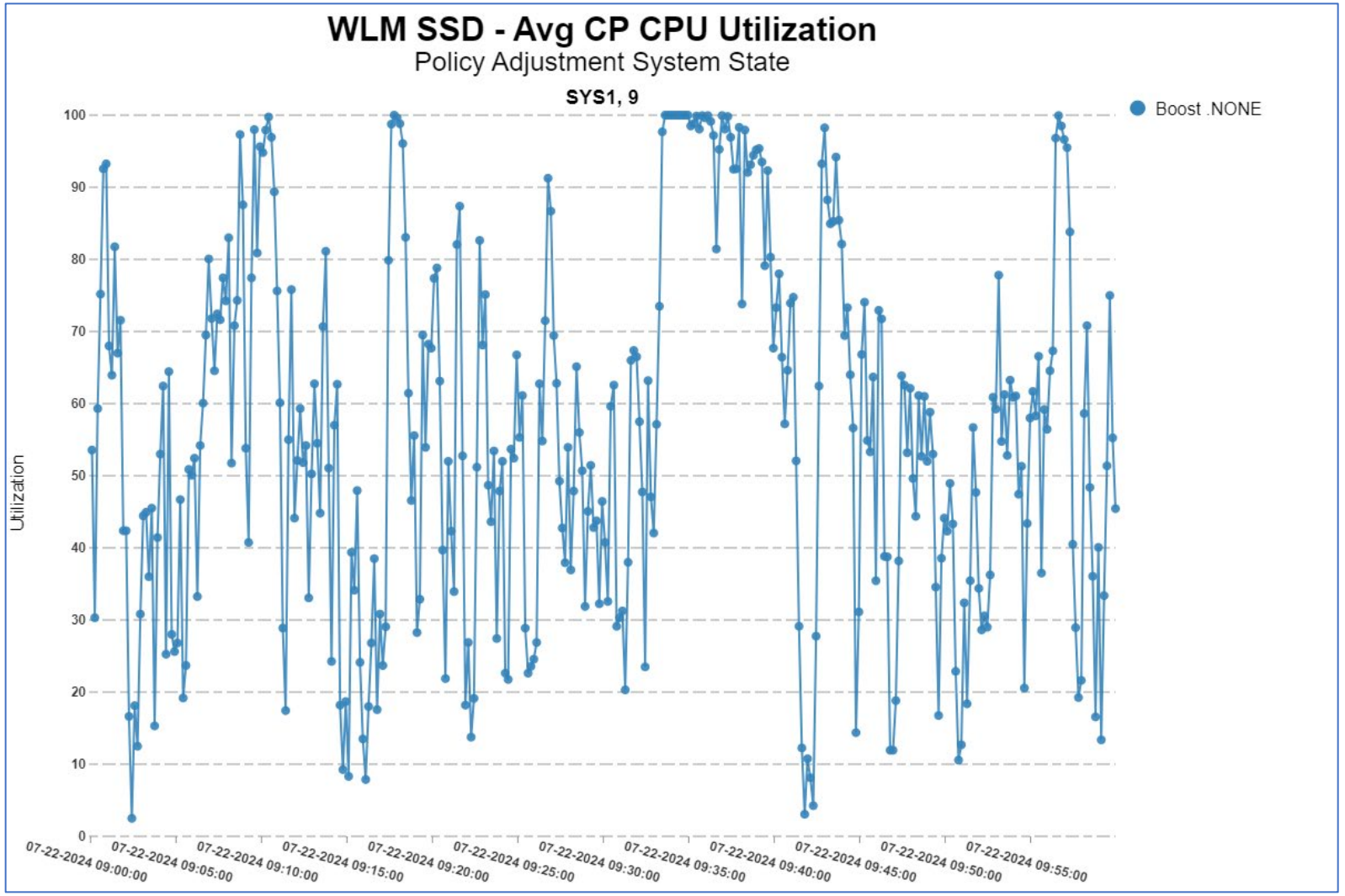
- System state data contains the Systemwide WLM algorithms need to make decisions
  - Think of this as system level monitor data to help WLM understand what is going on at the system level
  - The SMF 99.2 contains data at the period level

- Contains data such as:
  - CPU Data for CPs and zIIPs
    - Processor utilization
    - CPs and zIIPs online
    - Free LPAR capacities always available based on weights
    - Capacity based on weights and online processors
    - Number of times CPUs hit 100% utilization
    - Arrays of values based on importance levels (to see how much each importance level is consuming)
    - Boost information
  - Storage Data
    - Types of paging and paging costs
    - Common storage targets
    - Frames owned by logical swapped address spaces
  - Lots of counters
    - Number of internal and external service class periods
    - Various velocity values
    - Housekeeping clocks

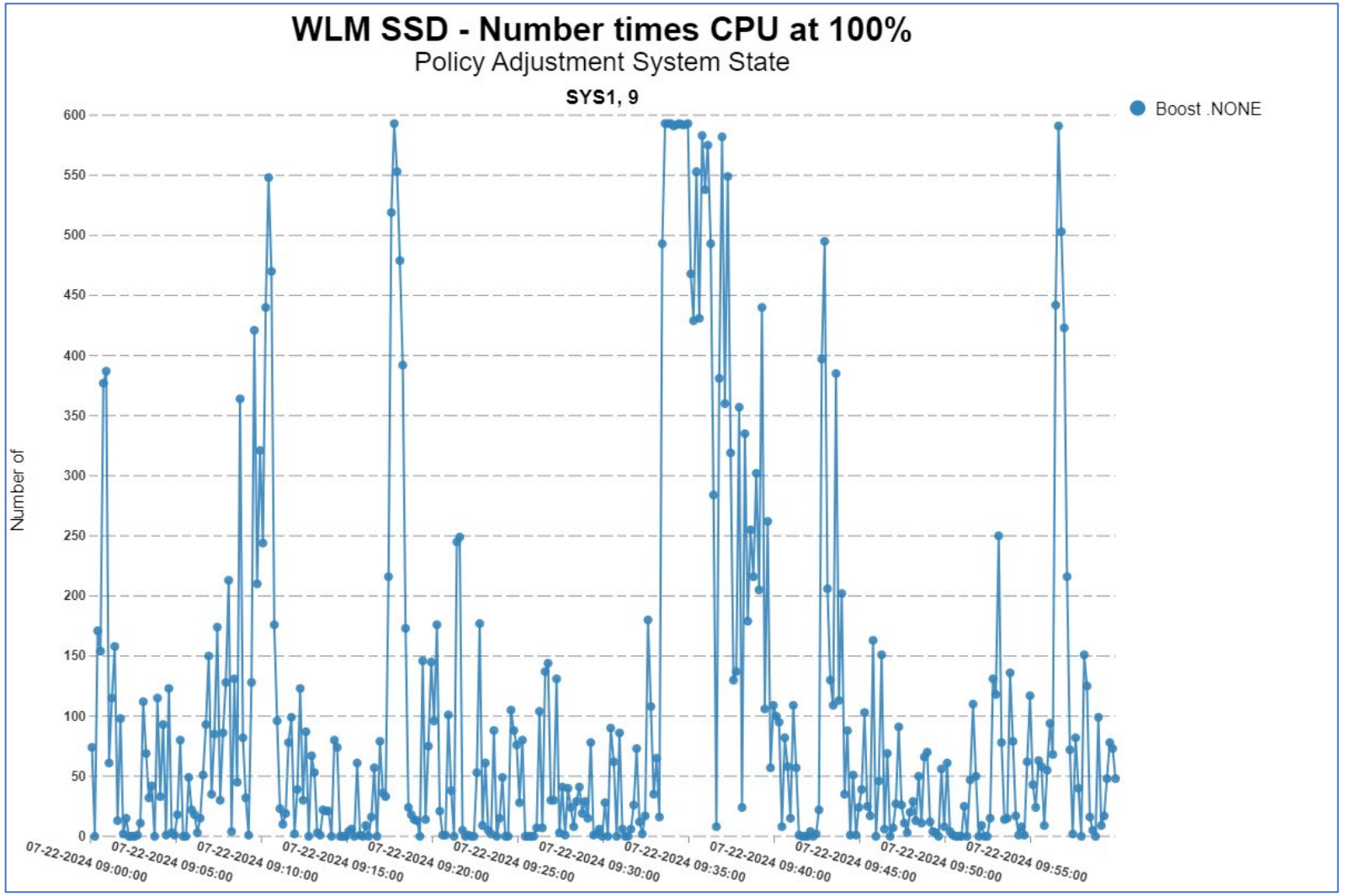
Example: SMF 99.1 System State Data



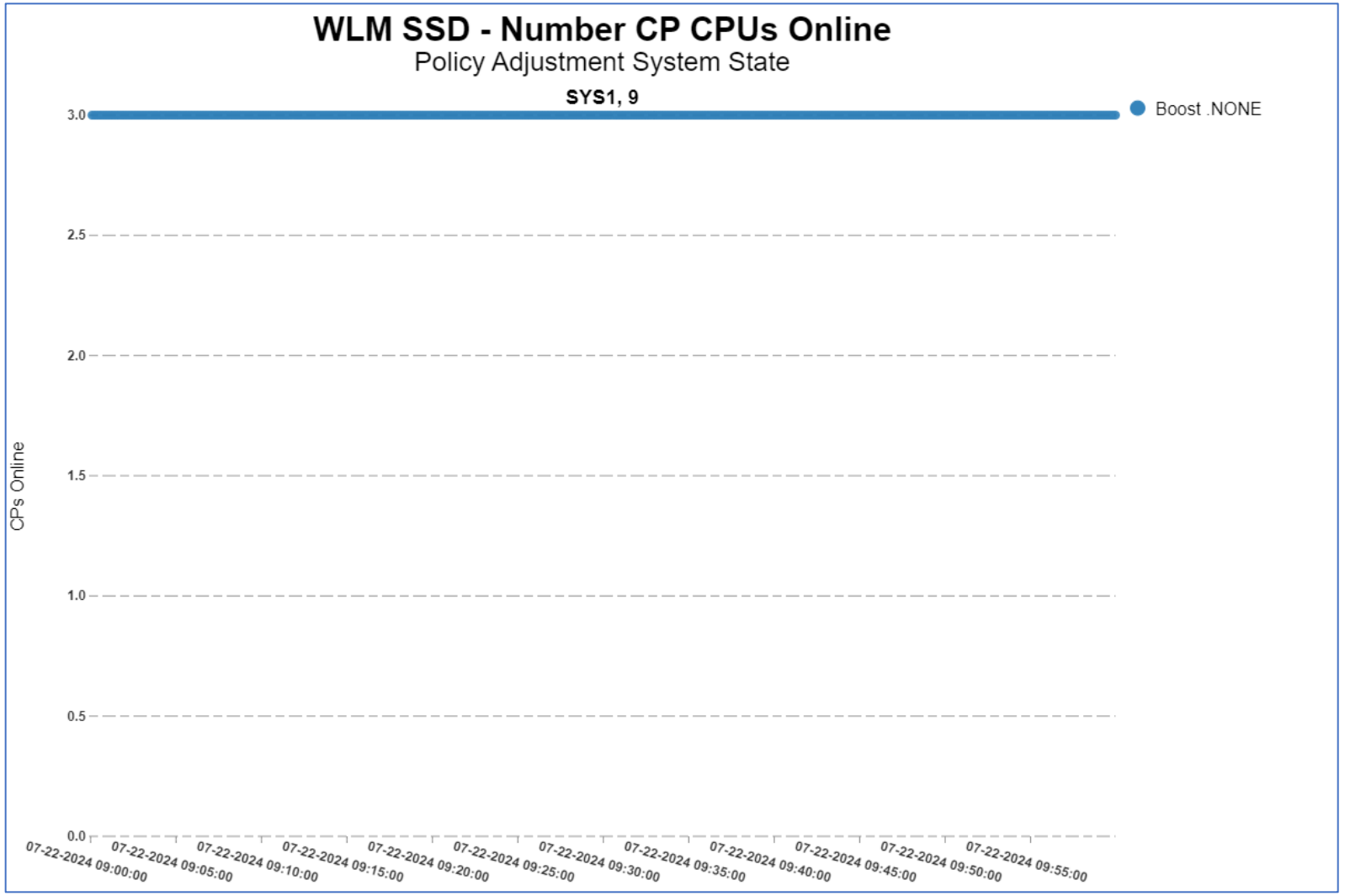
Example: SMF 99.1 System State Data



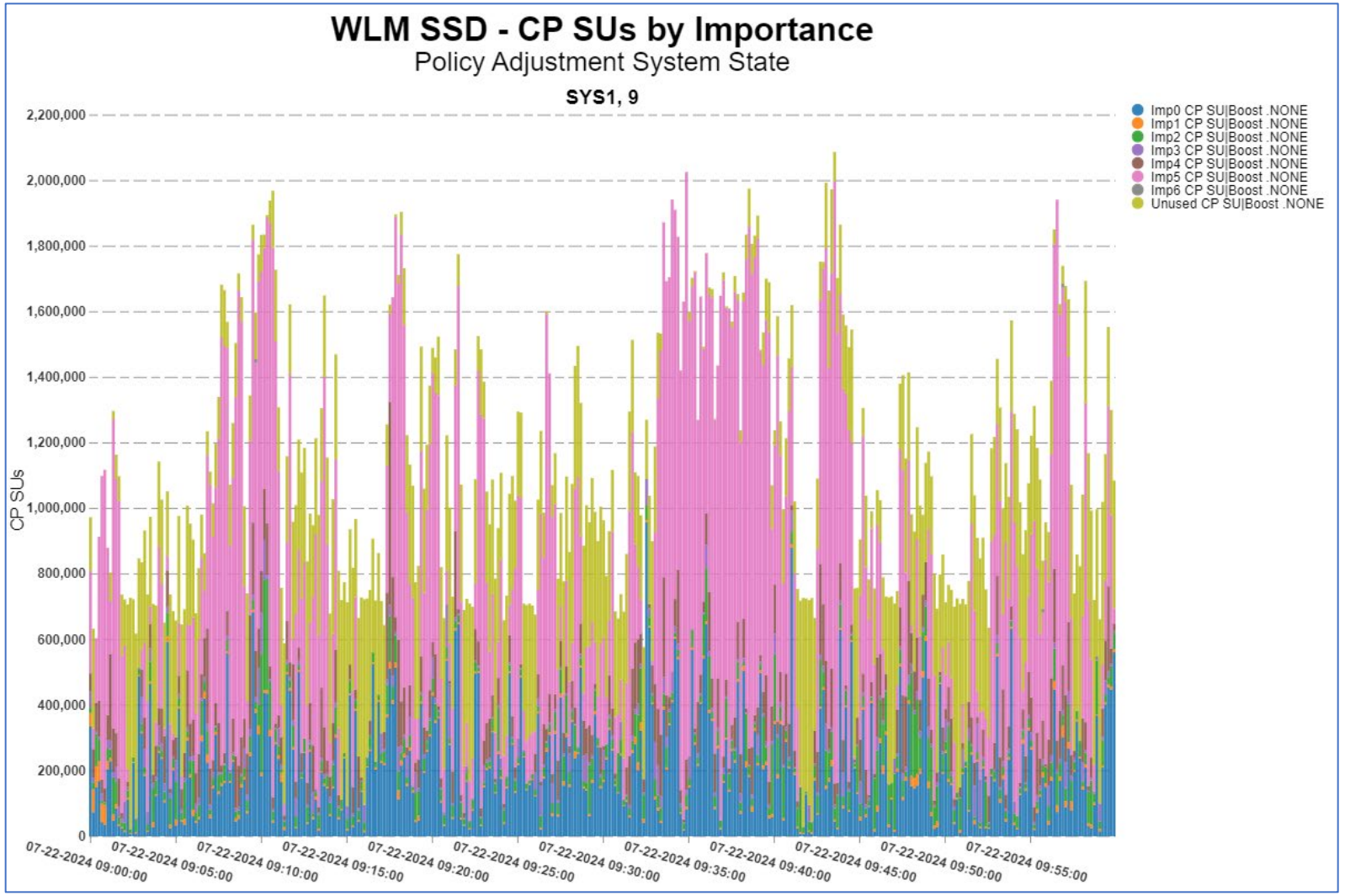
Example: SMF 99.1 System State Data



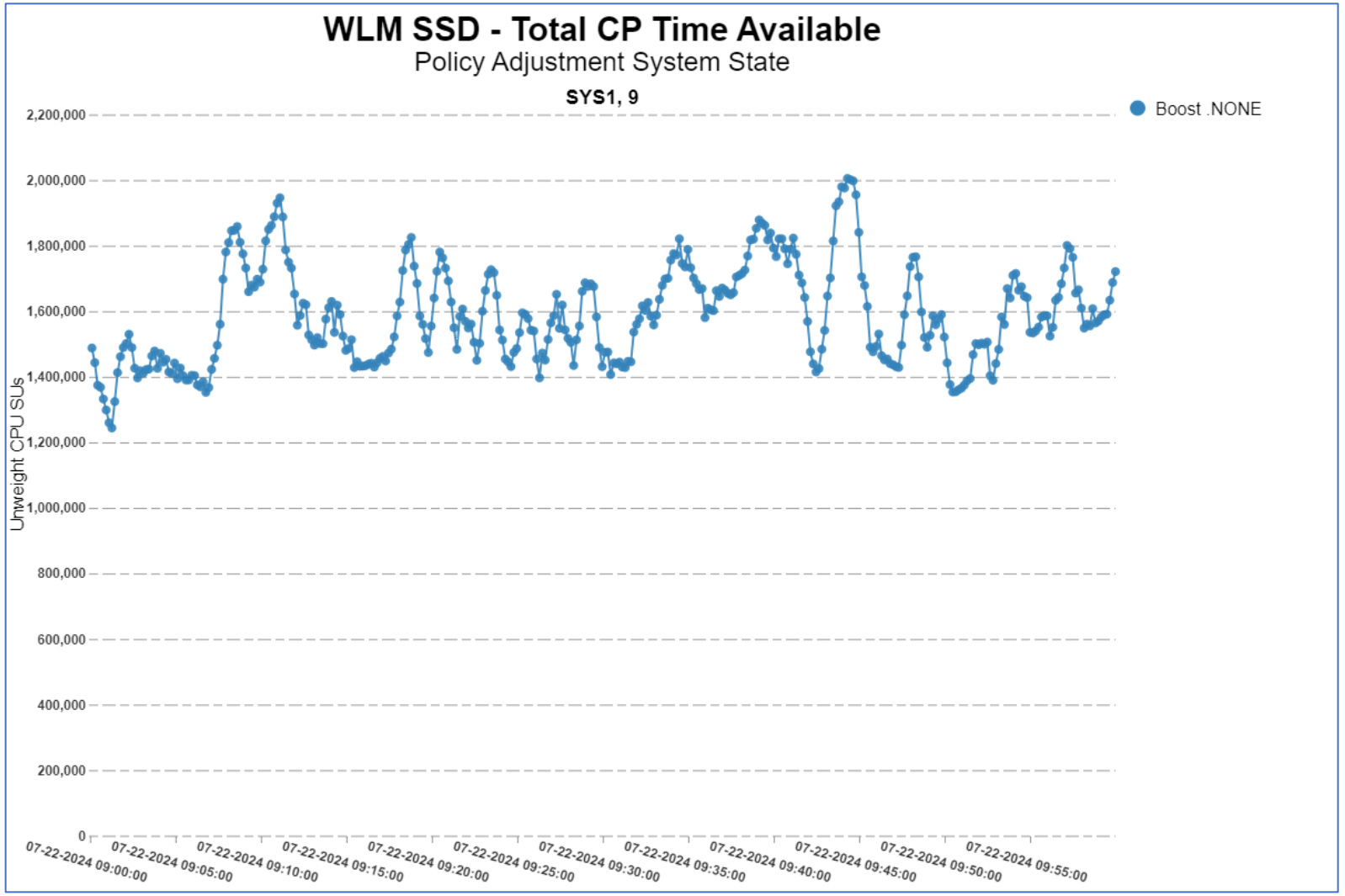
Example: SMF 99.1 System State Data



Example: SMF 99.1 System State Data



Example: SMF 99.1 System State Data



SMF 99.1 Trace Table Entry Section

- Trace Table Entry Section contains entries for each action

- Contains data such as:
  - Policy Adjustment Interval ID – Each 10 second period has its own PA interval number
  - Resource Adjustment interval ID – Each 2 second period has its own RA interval number
  - Trace Code – indicates the action or consideration
  - Address Space name – if an individual address space is considered for an action
  - Performance Indexes
    - Local PI – PI based on data just from the system being analyzed
    - Sysplex PI – PI based on data from all systems in WLMplex
  - Service Class Name
  - Service Class Period
  - Resource Group Name – If an RG is affected by action

- There are other values, but these are the key values

Receivers and Donors

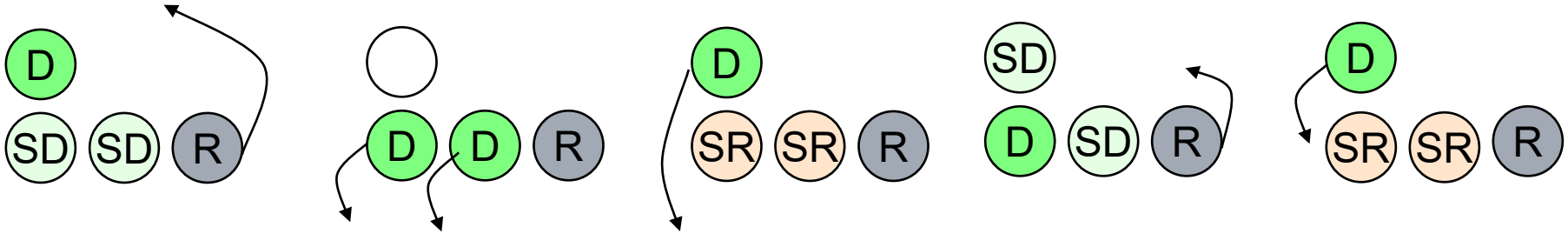
- Receiver
  - Service class period to potentially 'receive' resources
  - WLM will help only one receiver during each policy adjustment interval

    - Goal Receiver - Period with goal that needs help
    - Resource Receiver - Period to give the resources to in order to help the goal receiver
    - Secondary Receiver - Period helped indirectly due to an action to help the goal receiver

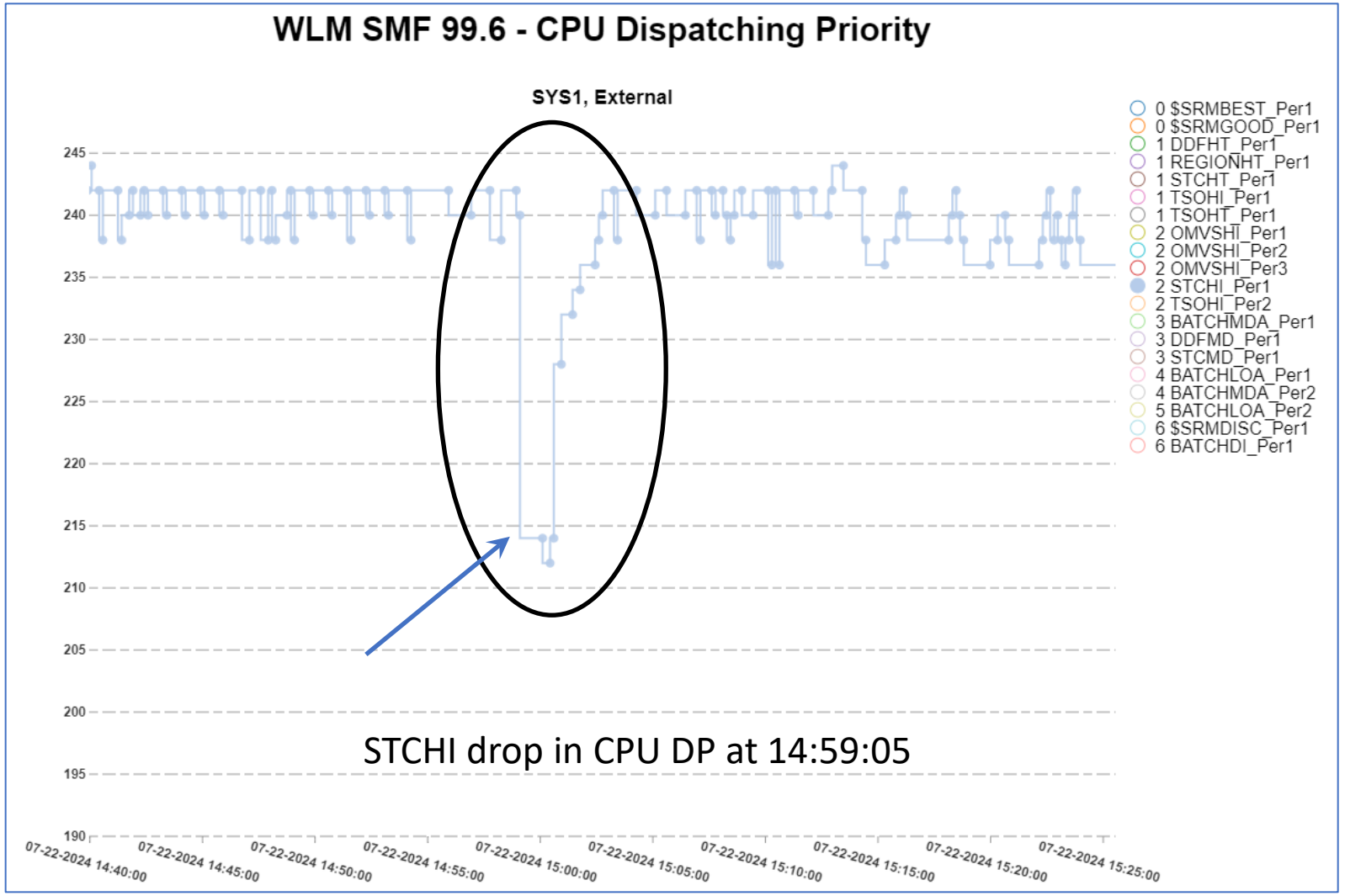
- Donor
  - Service class period to potentially 'donate' resources to help receiver
  - WLM may take from multiple donors during each policy adjustment interval
    - Goal Donor - Period whose goals may be impacted by resource donation
    - Resource Donor - Period to donate resources
    - Secondary Donor - Period that donates indirectly when receiver is helped

Example of WLM Decisions – CPU DP

- Dispatching priority adjustments
  - Objective: Increase Receiver's CPU using, or decrease Receiver's CPU delay
  - Interesting concepts:
    - Wait-to-Using ratio - ratio of CPU delay samples to CPU using samples (change in ratio used to determine change in CPU delay)
    - Maximum demand
      - Theoretical maximum percentage of total processor time a period can consume if it had no CPU delay
    - Achievable maximum demand
      - Percentage of total processor time a service period is projected to consume, taking into account demand of all higher work
  - Some possible actions



SMF 99.6 CPU Dispatching Priority – Every 10 Seconds



SMF 99.6 CPU Dispatching Priority – Every 10 Seconds

smftime	SMF99_Trace_Event_Number	SMF99_TPID	SMF99_TRID	SMF99_TCOD	SMF99_TCNM	SMF99_TPER	SMF99_TLPI	SMF99_TSPI	SMF99_TJOB	SMF99_description
14:59:05	18	180	68	9330		0	0	0		Logical partition is to be capped with a non-zero pricing management adjustment weight in order to enforce a soft cap.
14:59:05	19	180	68	9332		0	0	0		Logical partition was capped as part of enforcing a soft cap.
14:59:05	20	180	68	4747	$SRMS039	1	1.1	0.5	CICSREM1	The protective processor target cannot be increamented the new target would be below threshold for storage critical address spaces.
14:59:05	21	180	68	4747	$SRMS03B	1	5	5	CICSONT4	The protective processor target cannot be increamented the new target would be below threshold for storage critical address spaces.
14:59:05	22	180	68	4747	$SRMS03B	1	5	5	CICSTST1	The protective processor target cannot be increamented the new target would be below threshold for storage critical address spaces.
14:59:05	23	180	68	3620		0	0	0		Balancing of initiators; no action.
14:59:05	24	180	68	997		0	0	0		When Tuning alias adjustment was entered.
14:59:05	25	180	68	240	REGIONHT	1	50	50		Policy adjustment; goal receiver candidate selected.
14:59:05	26	180	68	975	DDFHT	1	3.5	3.5		Policy adjustment; select donor failed selecting period as the donor because period is small consumer.
14:59:05	27	180	68	850	REGIONHT	1	50	50		Policy adjustment; no processor action because no donor selected.
14:59:05	28	180	68	9348	REGIONHT	1	50	50		Lpar weight management; was skipped because one or several required conditions were not met.
14:59:05	29	180	68	280	$SRMS03A	1	50	50		Policy adjustment; resource receiver candidate selected.
14:59:05	30	180	68	9280	$SRMS03A	1	50	50		Policy adjustment; no qmpl action because no requests queued.
14:59:05	31	180	68	270	STCHT	1	3.33	3.33		Policy adjustment; receiver candidate selected.
14:59:05	32	180	68	975	DDFHT	1	3.5	3.5		Policy adjustment; select donor failed selecting period as the donor because period is small consumer.
14:59:05	33	180	68	850	STCHT	1	3.33	3.33		Policy adjustment; no processor action because no donor selected.
14:59:05	34	180	68	9348	STCHT	1	3.33	3.33		Lpar weight management; was skipped because one or several required conditions were not met.
14:59:05	35	180	68	270	OMVSHI	3	5	5		Policy adjustment; receiver candidate selected.
14:59:05	36	180	68	975	DDFHT	1	3.5	3.5		Policy adjustment; select donor failed selecting period as the donor because period is small consumer.
14:59:05	37	180	68	975	STCHT	1	3.33	3.33		Policy adjustment; select donor failed selecting period as the donor because period is small consumer.
14:59:05	38	180	68	308	STCHI	1	15	15		Policy adjustment; donor period.
14:59:05	39	180	68	880	STCHI	1	15	15		Policy adjustment; processor resource donor candidate selected.
14:59:05	42	180	68	620	OMVSHI	3	5	5		Policy adjustment; assess moving primary processor receiver up to occupied priority.
14:59:05	43	180	68	528	STCHT	1	3.33	3.33		Policy adjustment; CPU critical; blocker has been moved to a new priority.
14:59:05	44	180	68	640	$SRMS037	1	3.5	3.5		Policy adjustment; assess moving secondary processor receiver up.
14:59:05	45	180	68	680	$SRMS037	1	3.5	3.5		Trace table action code description not found
14:59:05	46	180	68	651	DDFHT	1	3.5	3.5		Policy adjustment; move up small processor consumer to next priority.
14:59:05	49	180	68	613	OMVSHI	3	5	5		Policy adjustment; processor move up; rejected for no net value; receiver trace.
14:59:05	50	180	68	590	REGIONHT	1	50	50		Policy adjustment; processor move up; rejected for no net value; donor trace.
14:59:05	53	180	68	530	STCHI	1	15	15		Policy adjustment; assess moving primary processor donor down to occupied priority.
14:59:05	54	180	68	530	STCHI	1	15	15		Policy adjustment; assess moving primary processor donor down to occupied priority.
14:59:05	55	180	68	530	STCHI	1	15	15		Policy adjustment; assess moving primary processor donor down to occupied priority.
14:59:05	58	180	68	950	OMVSHI	3	2.5	2.5		Policy adjustment; unchanged receiver.
14:59:05	59	180	68	660	STCHI	1	15	15		Policy adjustment; decrease priority for donor.
14:59:05	60	180	68	660	STCHI	1	15	15		Policy adjustment; decrease priority for donor.
14:59:05	61	180	68	660	STCHI	1	15	15		Policy adjustment; decrease priority for donor.
14:59:05	64	180	68	970	$SRMS039	1	1.1	0.5		Policy adjustment; unchanged secondary receiver.
14:59:05	65	180	68	936	ONLINEMD	1	1.1	0.5		Policy adjustment; served goal receiver selected.

SMF 99.6 CPU Dispatching Priority – Every 10 Seconds

- First set of codes is for resource adjustment

smftime	Event_Numbe	SMF99_TPID	SMF99_TRID	SMF99_TCOD	SMF99_TCNM	SMF99_TPER	SMF99_TLPI	SMF99_TSPI	SMF99_TJOB	SMF99_description
14:59:05	18	180	68	9330		0	0	0		Logical partition is to be capped with a non-zero pricing management
14:59:05	19	180	68	9332		0	0	0		Logical partition was capped as part of enforcing a soft cap.
14:59:05	20	180	68	4747	$SRMS039	1	1.1	0.5	CICSREM1	The protective processor target cannot be increamented the new targ
14:59:05	21	180	68	4747	$SRMS03B	1	5	5	CICSONT4	The protective processor target cannot be increamented the new targ
14:59:05	22	180	68	4747	$SRMS03B	1	5	5	CICSTST1	The protective processor target cannot be increamented the new targ
14:59:05	23	180	68	3620		0	0	0		Balancing of initiators; no action.
14:59:05	24	180	68	997		0	0	0		When Tuning alias adjustment was entered.

SMF 99.6 CPU Dispatching Priority – Every 10 Seconds

- Then WLM starts the policy adjustment algorithms
  - It starts looking at periods to help
  - In this example it selects several receiver candidates but fails each one

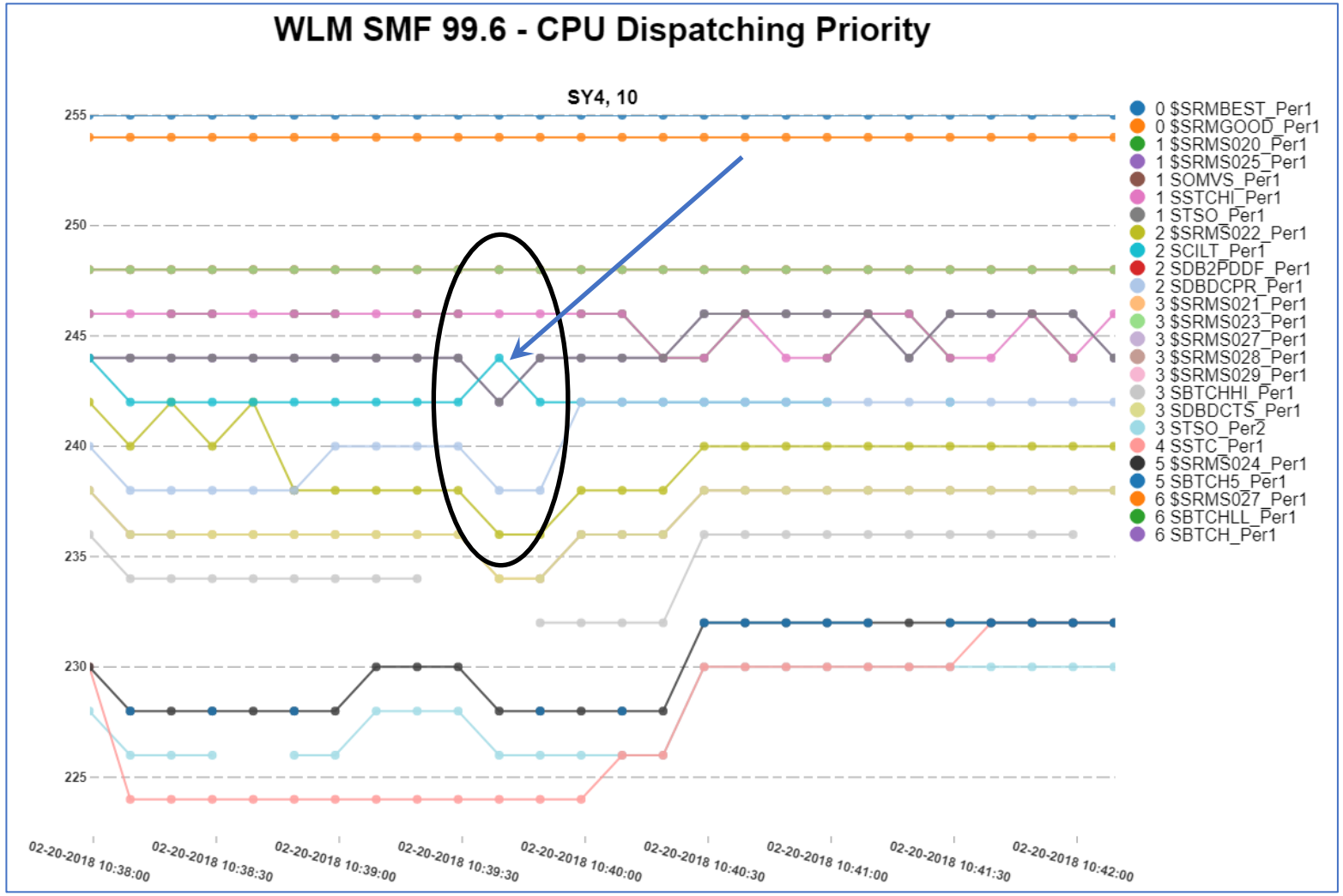
mftime	Event_Numbe	SMF99_TPID	SMF99_TRID	SMF99_TCOD	SMF99_TCNM	SMF99_TPER	SMF99_TLPI	SMF99_TSPI	SMF99_TJOB	SMF99_description
14:59:05	25	180	68	240	REGIONHT	1	50	50		Policy adjustment; goal receiver candidate selected.
14:59:05	26	180	68	975	DDFHT	1	3.5	3.5		Policy adjustment; select donor failed selecting period as the donor because period is sma
14:59:05	27	180	68	850	REGIONHT	1	50	50		Policy adjustment; no processor action because no donor selected.
14:59:05	28	180	68	9348	REGIONHT	1	50	50		Lpar weight management; was skipped because one or several required conditions were n
14:59:05	29	180	68	280	$SRMS03A	1	50	50		Policy adjustment; resource receiver candidate selected.
14:59:05	30	180	68	9280	$SRMS03A	1	50	50		Policy adjustment; no qmpl action because no requests queued.
14:59:05	31	180	68	270	STCHT	1	3.33	3.33		Policy adjustment; receiver candidate selected.
14:59:05	32	180	68	975	DDFHT	1	3.5	3.5		Policy adjustment; select donor failed selecting period as the donor because period is sma
14:59:05	33	180	68	850	STCHT	1	3.33	3.33		Policy adjustment; no processor action because no donor selected.
14:59:05	34	180	68	9348	STCHT	1	3.33	3.33		Lpar weight management; was skipped because one or several required conditions were n

SMF 99.6 CPU Dispatching Priority – Every 10 Seconds

- It then selects our period of interest and analyzes it
  - Here we see it actually takes an action by moving the CPU DP
  - Move STCHI from CPU DP 240 to CPU DP 214

smfdate	smftime	Event_Numbe	SMF99_TPID	SMF99_TRID	SMF99_TCOD	SMF99_TCNM	SMF99_TPER	SMF99_TLPI	SMF99_TSPI	SMF99_TJOB	SMF99_description
7/22/2024	14:59:05	35	180	68	270	OMVSHI	3	5	5		Policy adjustment; receiver candidate selected.
7/22/2024	14:59:05	36	180	68	975	DDFHT	1	3.5	3.5		Policy adjustment; select donor failed selecting period as the donor because period is small consumer.
7/22/2024	14:59:05	37	180	68	975	STCHT	1	3.33	3.33		Policy adjustment; select donor failed selecting period as the donor because period is small consumer.
7/22/2024	14:59:05	38	180	68	308	STCHI	1	15	15		Policy adjustment; donor period.
7/22/2024	14:59:05	39	180	68	880	STCHI	1	15	15		Policy adjustment; processor resource donor candidate selected.
7/22/2024	14:59:05	42	180	68	620	OMVSHI	3	5	5		Policy adjustment; assess moving primary processor receiver up to occupied priority.
7/22/2024	14:59:05	43	180	68	528	STCHT	1	3.33	3.33		Policy adjustment; CPU critical; blocker has been moved to a new priority.
7/22/2024	14:59:05	44	180	68	640	$SRMS037	1	3.5	3.5		Policy adjustment; assess moving secondary processor receiver up.
7/22/2024	14:59:05	45	180	68	680	$SRMS037	1	3.5	3.5		Trace table action code description not found
7/22/2024	14:59:05	46	180	68	651	DDFHT	1	3.5	3.5		Policy adjustment; move up small processor consumer to next priority.
7/22/2024	14:59:05	49	180	68	613	OMVSHI	3	5	5		Policy adjustment; processor move up; rejected for no net value; receiver trace.
7/22/2024	14:59:05	50	180	68	590	REGIONHT	1	50	50		Policy adjustment; processor move up; rejected for no net value; donor trace.
7/22/2024	14:59:05	53	180	68	530	STCHI	1	15	15		Policy adjustment; assess moving primary processor donor down to occupied priority.
7/22/2024	14:59:05	54	180	68	530	STCHI	1	15	15		Policy adjustment; assess moving primary processor donor down to occupied priority.
7/22/2024	14:59:05	55	180	68	530	STCHI	1	15	15		Policy adjustment; assess moving primary processor donor down to occupied priority.
7/22/2024	14:59:05	58	180	68	950	OMVSHI	3	2.5	2.5		Policy adjustment; unchanged receiver.
7/22/2024	14:59:05	59	180	68	660	STCHI	1	15	15		Policy adjustment; decrease priority for donor.
7/22/2024	14:59:05	60	180	68	660	STCHI	1	15	15		Policy adjustment; decrease priority for donor.
7/22/2024	14:59:05	61	180	68	660	STCHI	1	15	15		Policy adjustment; decrease priority for donor.
7/22/2024	14:59:05	64	180	68	970	$SRMS039	1	1.1	0.5		Policy adjustment; unchanged secondary receiver.
7/22/2024	14:59:05	65	180	68	936	ONLINEMD	1	1.1	0.5		Policy adjustment; served goal receiver selected.

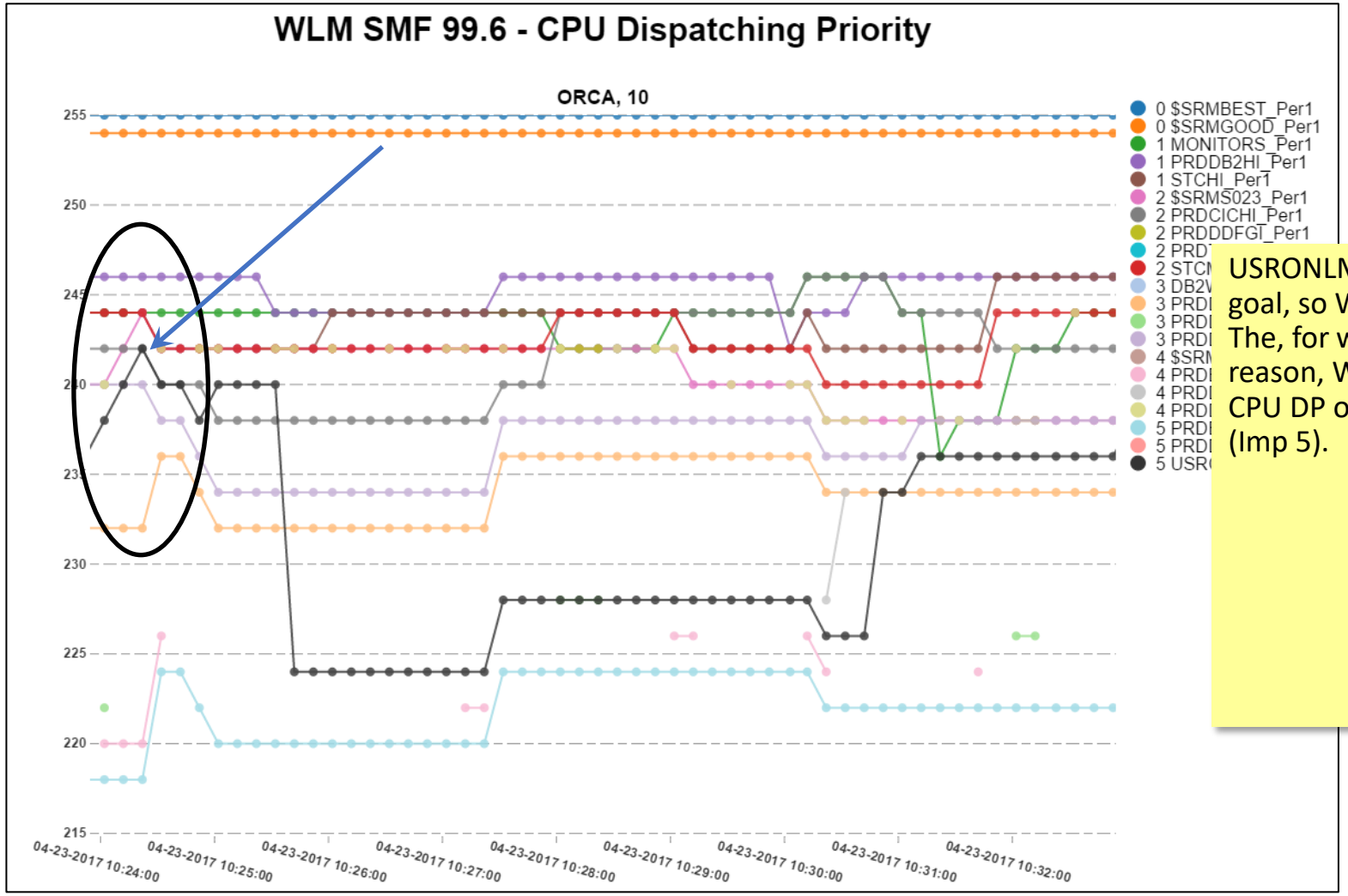
SMF 99.6 CPU Dispatching Priority – Every 10 Seconds



SMF 99.1
- Example of WLM Actions Trace

system	Time	PA Int	RA Int	Code	Code	Explain	Local P	Syspl	Perio	Service Class
SY4	10:39:39	142	173	270	PA_REC_CAND	Policy adjustment, receiver candidate selected.	50	50	1	SCILT
SY4	10:39:39	142	173	308	PA_DONOR_PERIOD	Policy adjustment, donor period.	0.64	0.94	1	STSO
SY4	10:39:39	142	173	308	PA_DONOR_PERIOD	Policy adjustment, donor period.	0.78	1.09	1	SSTCHI
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	0.01	0.14	1	$SRMS024
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	0.01	0.14	1	SBTCH5
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	0.5	5.5	1	SOMVS
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	0.54	0.9	2	STSO
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	0.64	0.94	1	$SRMS025
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	0.64	0.94	1	STSO
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	0.74	1.09	1	$SRMS021
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	0.74	1.09	1	SDBDCTS
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	0.76	0.97	1	SBTCHHI
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	1.1	0.8	1	$SRMS022
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	50	50	1	SCILT
SY4	10:39:39	142	173	525	HSK_UNBUNCH_PRTY	Housekeeping, unbunch priorities.	60	0.85	1	SDBDCPR
SY4	10:39:39	142	173	530	PA_PMDO_DON	Policy adjustment, assess moving primary processor donor down to occupied priority.	0.64	0.94	1	STSO
SY4	10:39:39	142	173	530	PA_PMDO_DON	Policy adjustment, assess moving primary processor donor down to occupied priority.	0.78	1.09	1	SSTCHI
SY4	10:39:39	142	173	531	PA_PCC_DON_VIOLTN	Policy adjustment, moving the donor to the receivers priority violates CPU critical rules.	0.78	1.09	1	SSTCHI
SY4	10:39:39	142	173	532	PA_PCC_BLKR_IS_DON	Policy adjustment, cannot move the blocker up because it is the donor.	0.78	1.09	1	SSTCHI
SY4	10:39:39	142	173	580	PA_PMD_SEC_DON	Policy adjustment, assess moving secondary processor donor down.	0.64	0.94	1	$SRMS025
SY4	10:39:39	142	173	580	PA_PMD_SEC_DON	Policy adjustment, assess moving secondary processor donor down.	0.78	1.09	1	$SRMS020
SY4	10:39:39	142	173	620	PA_PMUO_REC	Policy adjustment, assess moving primary processor receiver up to occupied priority.	50	50	1	SCILT
SY4	10:39:39	142	173	635	PA_PMUUB_REC	Policy adjustment, assess moving primary processor receiver up to unoccupied priority between donor and receiver's current priorities.	50	50	1	SCILT
SY4	10:39:39	142	173	750	PA_PRO_INCP_REC	Policy adjustment, increase priority for receiver.	50	50	1	SCILT

SMF 99.6 CPU Dispatching Priority – Every 10 Seconds



USRONLME misses goal, so WLM helps it. The, for what ever reason, WLM raises CPU DP of USRONLME (Imp 5).

SMF 99.1
- Example of WLM Actions Trace

SMFDateTime	PA Inteval	RA Interval	Trace Code	Code	Job	Local PI	Sysplex PI	Service Class	Period
4/23/17 10:24:03 AM	175	124	270	PA_REC_CAND		131	131	USRONLME	1
4/23/17 10:24:03 AM	175	124	975	PA_SDO_DONFAIL_SPC		110	110	PRDDDFGI	1
4/23/17 10:24:03 AM	175	124	975	PA_SDO_DONFAIL_SPC		70	70	PRDDDFOM	1
4/23/17 10:24:03 AM	175	124	975	PA_SDO_DONFAIL_SPC		27	27	STCME	1
4/23/17 10:24:03 AM	175	124	975	PA_SDO_DONFAIL_SPC		110	110	PRDDDFPD	1
4/23/17 10:24:03 AM	175	124	308	PA_DONOR_PERIOD		40	40	STCHI	1
4/23/17 10:24:03 AM	175	124	880	PA_PRO_RDON_CAND		40	40	STCHI	1
4/23/17 10:24:03 AM	175	124	620	PA_PMUO_REC		131	131	USRONLME	1
4/23/17 10:24:03 AM	175	124	620	PA_PMUO_REC		131	131	USRONLME	1
4/23/17 10:24:03 AM	175	124	620	PA_PMUO_REC		131	131	USRONLME	1
4/23/17 10:24:03 AM	175	124	651	PA_PMU_SPC_NXT_DP		110	110	PRDDDFPD	1
4/23/17 10:24:03 AM	175	124	940	PA_PRO_UNC_DON		40	40	STCHI	1
4/23/17 10:24:03 AM	175	124	940	PA_PRO_UNC_DON		40	40	STCHI	1
4/23/17 10:24:03 AM	175	124	940	PA_PRO_UNC_DON		40	40	STCHI	1
4/23/17 10:24:03 AM	175	124	740	PA_PRO_INCP_DON		110	110	PRDDDFPD	1
4/23/17 10:24:03 AM	175	124	740	PA_PRO_INCP_DON		110	110	PRDDDFPD	1
4/23/17 10:24:03 AM	175	124	740	PA_PRO_INCP_DON		110	110	PRDDDFPD	1
4/23/17 10:24:03 AM	175	124	780	PA_PRO_INCP_SC		110	110	PRDDDFPD	1
4/23/17 10:24:03 AM	175	124	780	PA_PRO_INCP_SC		110	110	PRDDDFPD	1
4/23/17 10:24:03 AM	175	124	780	PA_PRO_INCP_SC		110	110	PRDDDFPD	1
4/23/17 10:24:03 AM	175	124	750	PA_PRO_INCP_REC		113	113	USRONLME	1
4/23/17 10:24:03 AM	175	124	750	PA_PRO_INCP_REC		113	113	USRONLME	1
4/23/17 10:24:03 AM	175	124	750	PA_PRO_INCP_REC		113	113	USRONLME	1

SMF 99.1 Priority Table Entry

- Trace Table Entry Section contains entries for each occupied CPU DP
  - Value are before the PA action (i.e. so input into the decision)

- For each occupied CPU DP, contains data such as:
  - New unbunched priority
  - Initial maximum demand percentage
  - Projected maximum demand percentage
  - CPU and CPU delay samples
    - Samples at priority in the last 10 seconds
    - and Samples based on sampling
  - Wait to using ratio
  - Actual measured processor used at the priority in CPU SUs
  - Projected processor time to be used at the priority in CPU SUs
  - Achievable cumulative maximum demand percentage
  - Initial and projected MTTW in CPU SUs

Example of fields in the Priority Table Entries

- SMF99_PTPRTY : Dispatching priority, after policy adjustment.
- SMF99_PTNP : New dispatching priority, after un-bunching.
- SMF99_PTIMDP : Initial maximum percentage of processor demanded at priority, initial value before any priority moves.
- SMF99_PTPMDP : Projected maximum percentage of processor demanded at priority.
- SMF99_PTCPUU : CPU using samples at priority found in the last 10 seconds.
- SMF99_PTCPUD : CPU delay samples at priority found in the last 10 seconds.
- SMF99_PTW2UR : Wait-to-using ratio at priority scaled by 16.
- SMF99_PTAPU : Actual measured processor used at priority in unweighted CPU service units per second.
- SMF99_PTPPU : Projected processor time to be used at priority in unweighted CPU service units per second.
- SMF99_PTACMD : Achievable cumulative maximum demand percentage scaled by 10.
- SMF99_PTIMAXD : Initial cumulative maximum demand percentage scaled by 10.
- SMF99_PTWMAXD : Projected cumulative maximum demand percentage scaled by 10.
- SMF99_PTIAMTW : Initial average mean time to wait in unweighted CPU service units per second scaled by 1000.
- SMF99_PTWAMTW : Projected average mean time to wait in unweighted CPU service units per second scaled by 1000.
- SMF99_PTSCPUU : Sample based CPU using samples at priority.
- SMF99_PTSCPUD : Sample based CPU delay samples at priority.

SMF 99.1 Priority Entry Table

- Earlier we gave an example of STCHI moving from CPU DP 240 to CPU DP 214
  - This is the CPU priority table before the action and after the action

SMF99_PTPRTY	SMF99_PTIMDP	SMF99_PTPMDP	SMF99_PTCPUU	SMF99_PTCPUD	SMF99_PTW2UR	SMF99_PTAPU	SMF99_PTPPU	SMF99_PTACMD	SMF99_PTIMAXD	SMF99_PTWMAXD	SMF99_PTIAMTW	SMF99_PTWAMTW	SMF99_PTSCPUU	SMF99_PTSCPUD
192	59	59	0	37	232	6069	6069	829	882	882	4400	4400	6	37
214	384	384	13	50	232	359602	359602	772	823	823	14900	14900	12	50
216	169	169	7	51	232	183346	183346	439	439	439	4400	4400	18	51
230	2	2	0	0	232	0	0	270	270	270	7497	7497	0	0
234	1	1	0	0	232	82	82	268	268	268	10000	10000	0	0
236	6	6	2	29	232	19256	19256	267	267	267	7100	7100	2	18
238	3	3	8	37	144	39330	39330	261	261	261	14900	14900	9	37
240	35	100	3	27	144	70624	72056	258	258	258	3100	3000	6	27
242	93	28	0	7	0	1796	364	158	223	158	800	2800	0	7
244	6	6	0	3	0	12313	12313	130	130	130	5800	5800	1	3
246	6	6	0	15	0	775	775	124	124	124	3500	3500	3	15
254	94	94	20	36	0	191797	191797	118	118	118	14900	14900	47	36
255	24	24	3	3	0	49465	49465	24	24	24	2300	2300	20	3
14:59:16														
192	74	74	4	37	160	15234	15234	841	894	894	4200	4200	7	37
214	363	363	15	15	160	316185	316185	769	820	820	14900	14900	20	15
216	164	64	1	7	160	147972	1635	457	457	457	4900	200	7	7
217	0	163	0	0	160	0	147970	393	293	393	7497	4900	0	0
230	2	2	0	0	160	0	2406	230	293	230	7497	2400	0	0
234	1	1	0	0	160	82	82	228	291	228	10000	10000	0	0
236	6	6	0	16	160	17677	17677	227	290	227	3100	3100	4	16
238	4	4	4	9	160	46236	46236	221	284	221	14900	14900	5	9
240	98	35	1	10	160	71822	70189	217	280	217	5300	5500	1	10
242	31	31	0	1	0	360	360	182	182	182	2500	2500	0	1
244	10	10	0	4	0	11976	11976	151	151	151	3600	3600	0	4
246	6	6	0	4	0	895	895	141	141	141	5400	5400	0	4
254	112	112	23	13	0	236038	236038	135	135	135	14900	14900	34	13
255	23	23	1	2	0	45129	45129	23	23	23	2600	2600	8	2

Key Lesson about the SMF 99.1 Records

- The original intention of the SMF 99 records was to help IBM debug WLM algorithm decisions
  - These records are cut every 10 seconds
- While practical use has been found by regularly processing an analyzing certain subtypes:
  - Such as the SMF 99.6, SMF 99.12, SMF 99.14
- The other SMF 99 subtypes, such as SMF 99.1 and SMF 99.2:
  - Are cryptic, not well documented
  - Contain values that are only meaningful to WLM algorithms
  - Are not very useful for our day-to-day use by general performance analyst
- At EPS, we use the SMF 99.1 and SMF 99.2 by searching for unusual patterns during periods of time customers call us asking to analyze a situation

Questions?