z Processor Consumption Analysis, or What Is Consuming All The CPU?

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®

© Enterprise Performance Strategies, Inc.

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 and Reports Offer

- **Abstract**

  – The first step to any processor analysis is to understand your processor configuration and settings. The second step is to understand what workloads, address spaces, and transactions are consuming the fixed processor resource. It is only after understanding what and how the processor is being consumed can you conduct any sort of processor tuning or optimization exercise.

  – During this presentation Peter Enrico will show you how to conduct a processor resource consumption analysis. You will be provided with a top down approach to better understand processor measurements available to help you gain a drilldown insight into how the CPU resource is being consumed, and by what LPARs, Workloads, and transactions. Shown is what is known as a drill down approach for a processor performance analysis.

Performance Workshops Available

During these workshops you will be analyzing your own data!

- WLM Performance and Re-evaluating of Goals
  - Instructor: Peter Enrico and Scott Chapman
  - September 28 – October 2, 2015 – Columbus, Ohio, USA

- Parallel Sysplex and z/OS Performance Tuning
  (Web / Internet Based!)
  - Instructor: Peter Enrico and Scott Chapman
  - July 21 – 23, 2015
  - November 17 – 19, 2015

- Essential z/OS Performance Tuning Workshop
  - Instructors: Peter Enrico, Scott Chapman, Tom Beretvas
  - October 19 - 23, 2015 – Dallas, Texas, USA

- z/OS Capacity Planning and Performance Analysis
  - Instructor: Ray Wicks

EPS Sessions at Share

Peter Enrico

Day	Time	Location	Presentation
Wed	11:15	Asia 3	SMF 113 Processor Cache Counter Measurements – Overview, Update, and Usage
Wed	1:45	Asia 3	WLM – Effective Setup and Usage of WLM Report Classes
Thu	11:15	Asia 3	zProcessor Consumption Analysis (including z13), or What is Consuming All the CPU?

Scott Chapman

Day	Time	Location	Presentation
Tue	11:15	Asia 3	Memory Management in the TB Age
Tue	3:15	Southern Hemisphere 4	Lessons Learned from implementing an IDAA
Fri	11:15	Asia 3	WLM in One Page

Presentation Overview

- Many areas need to be examined when decomposing CPU consumption
  - This presentation just discussions some of the many areas

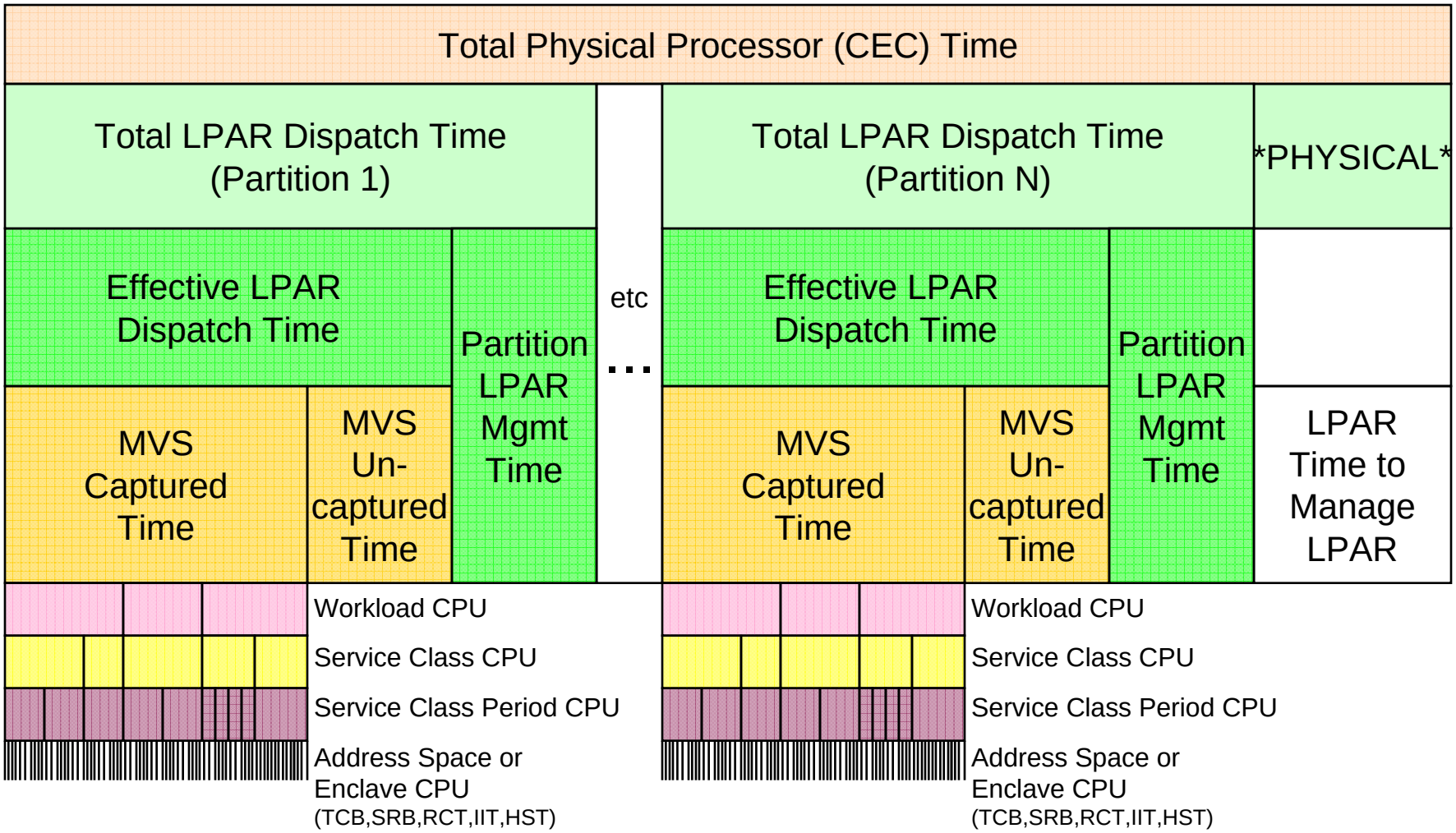
- Basic Processor Consumption Analysis
  - Decomposing CPU Consumption
    - By importance level
    - Displaced workloads
    - By Service Class and Report Class
  - Looking at CPU Dispatching Priorities
  - Looking at Latent Demand

CPU Measurement Reports Processing/Discussion Offer !!!

- Special Reports Offer!
  - See your Coupling Facility records in chart and table format
  - Please contact me, Peter Enrico for instructions for sending raw SMF data
    - Send an email to peter.enrico@epstrategies.com
  - Deliverable: Dozens of coupling facility based reports (charts and tables)
    - CPU - Machine Level Analysis
    - CPU - LPAR Level Analysis
    - CPU - HiperDispatch CPU Activity
    - CPU - SMF 113 Processor Counters
    - WLM - Workload Utilization Analysis
    - Coupling Facility Host Effect
    - And much more!
    - One-on-one phone call to explain your coupling facility measurements

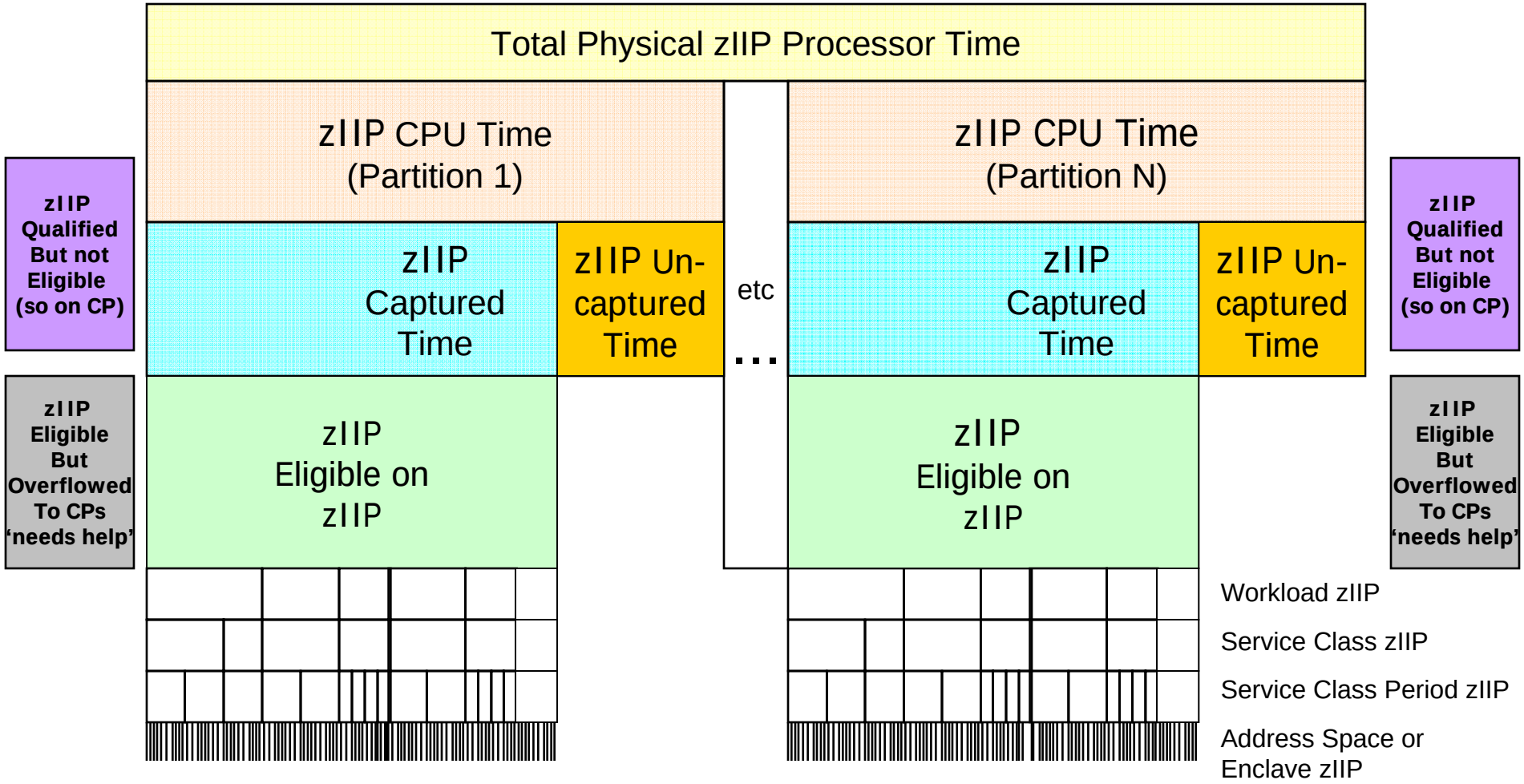
Breakdown of General Purpose Processor

- We always needed to understand the break down of CP CPU consumption



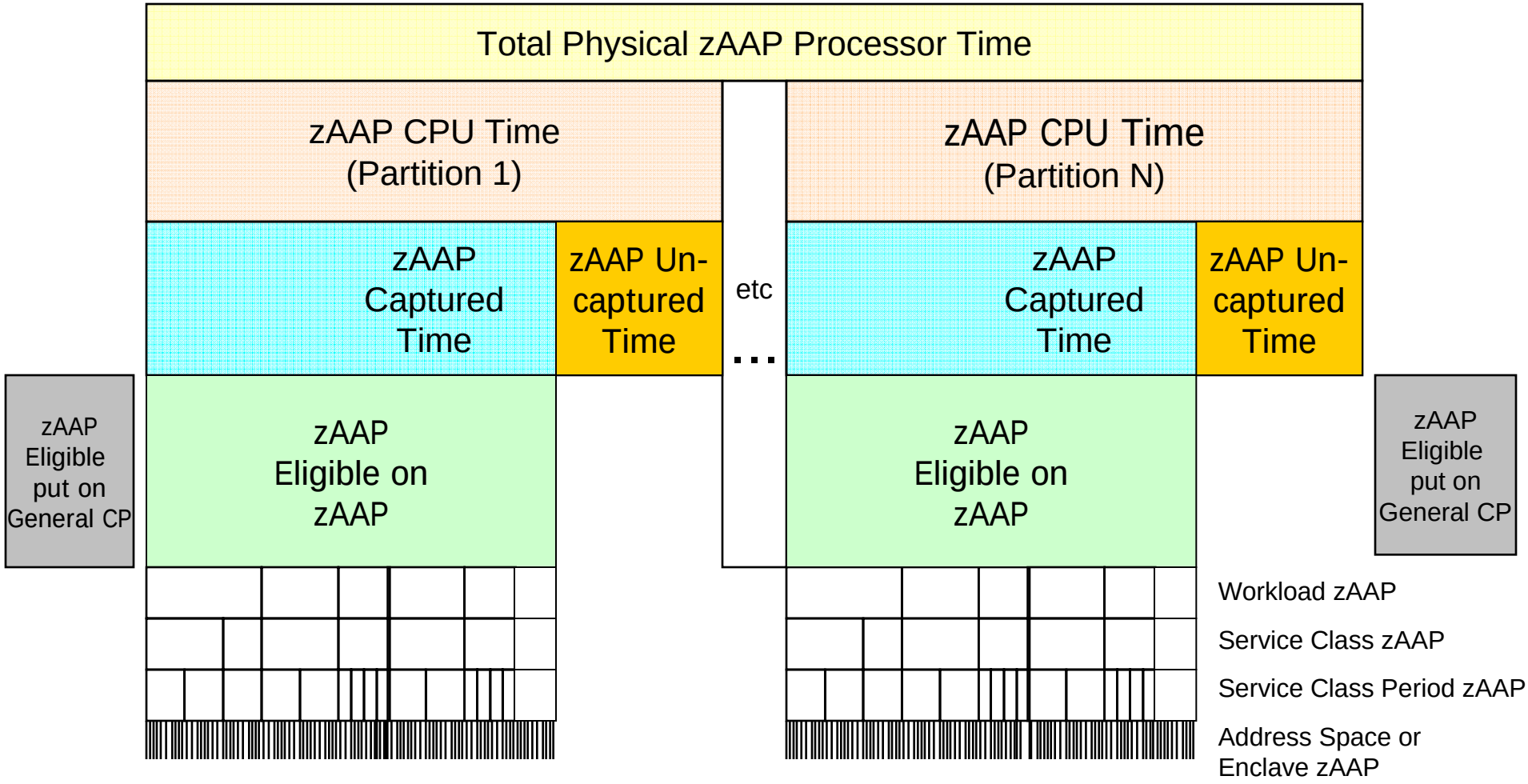
Breakdown of zIIP Engine Time

- We need to understand how PR/SM allocates the zIIP processor resource
  - In all measurements zIIPs



Breakdown of zAAP Engine Time

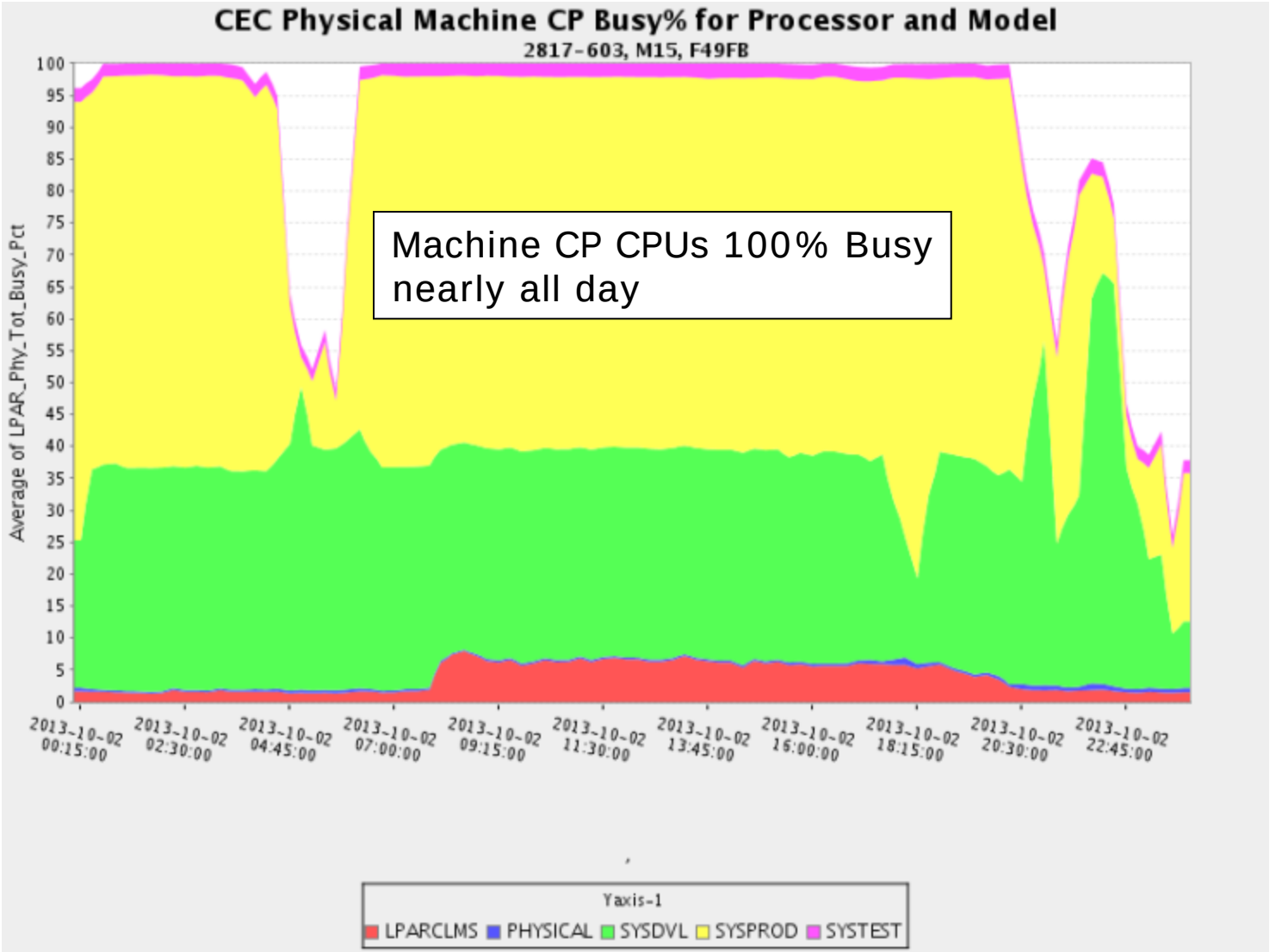
- We now need to understand where the zAAP CPU time is consumed



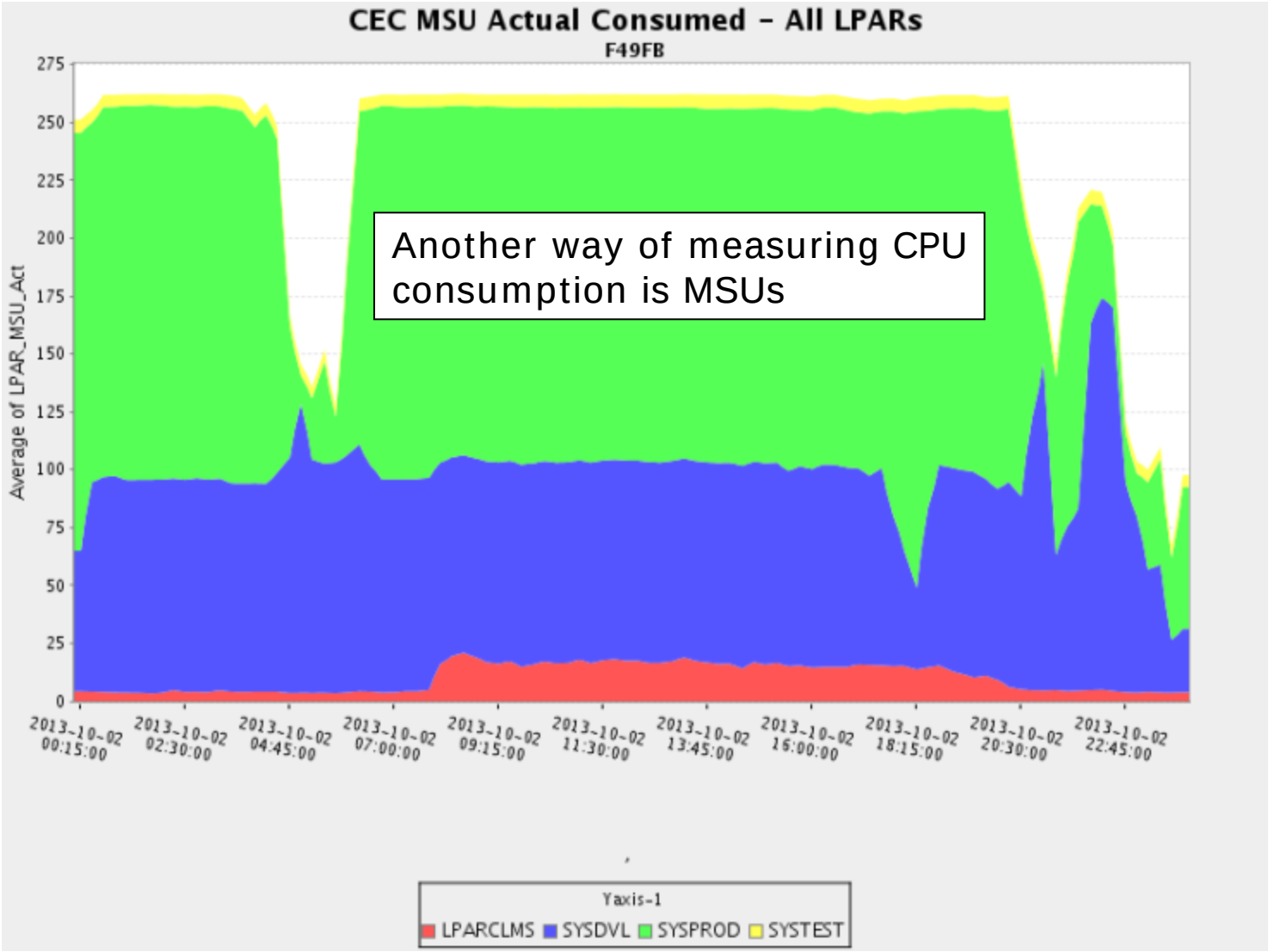
Decomposing CPU Consumption
-
Machine Level Analysis

- Which LPARs are using the physical CPUs?
  - Utilization
  - MSUs
- Look at LPAR Management Busy% to ensure it is within guidelines
- Was there LPAR weight enforcement?

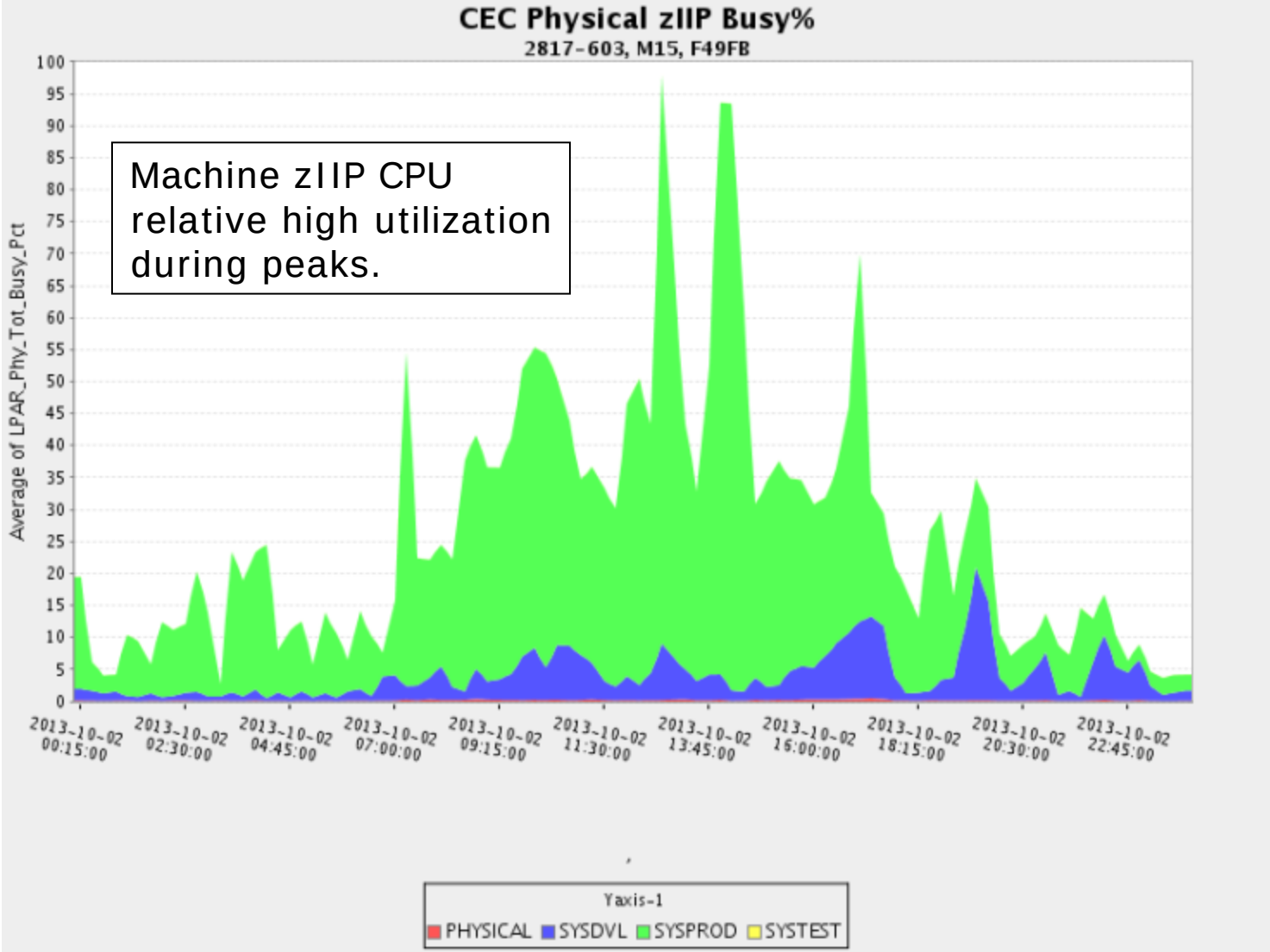
Machine Busy – CP Percent Busy



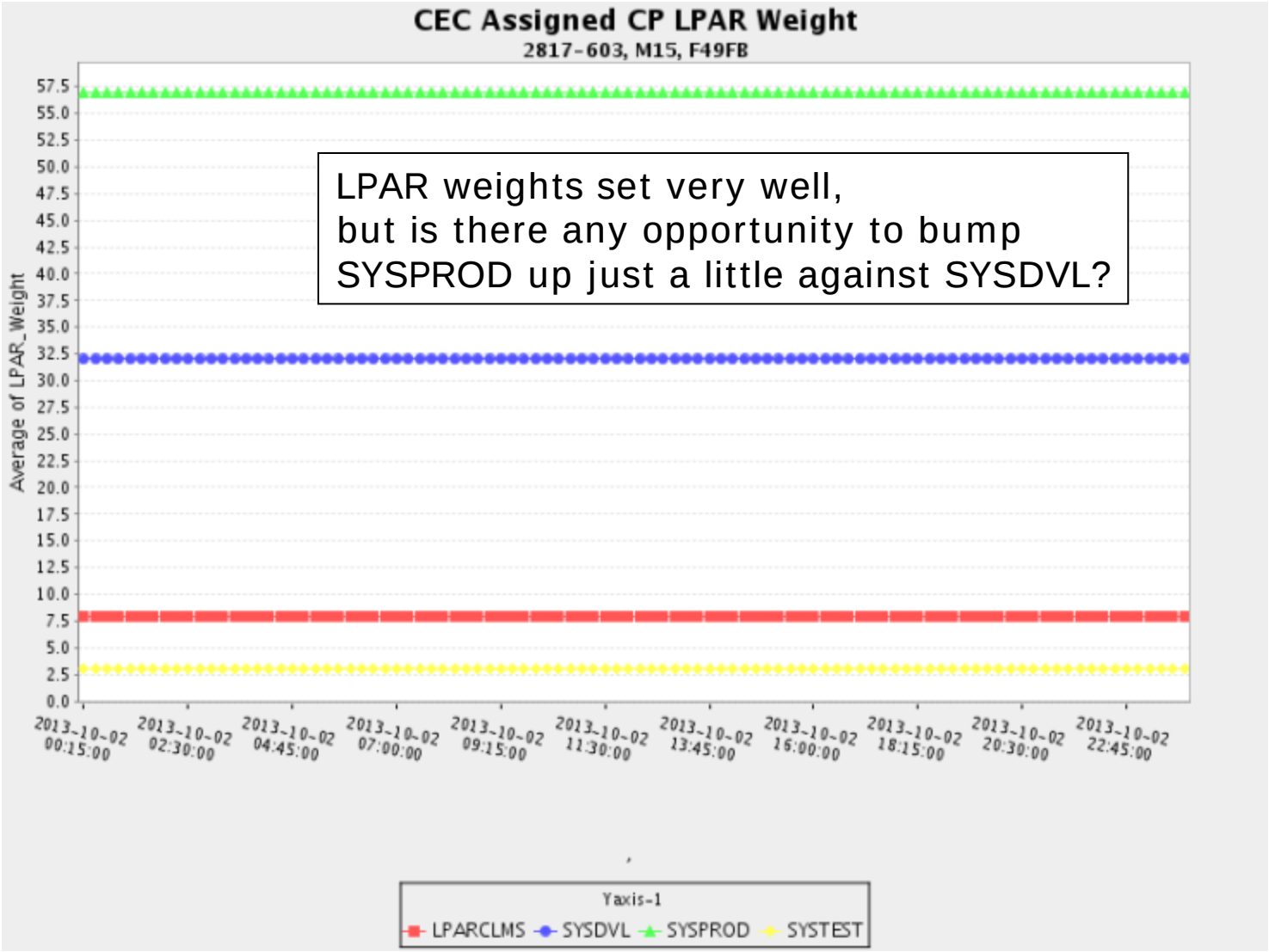
Could measure CPU Consumption in MSUs



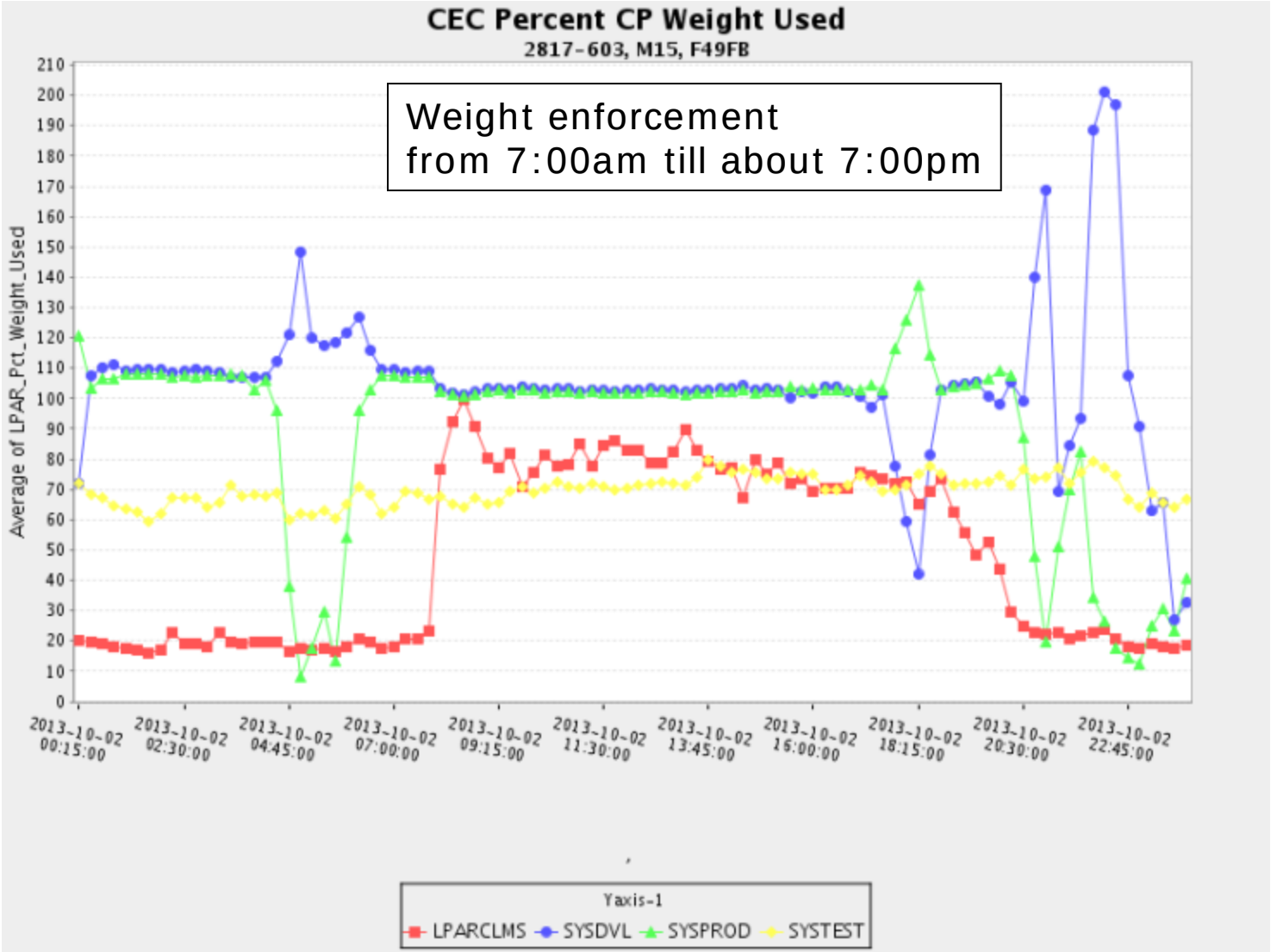
Machine Busy – zIIP Percent Busy



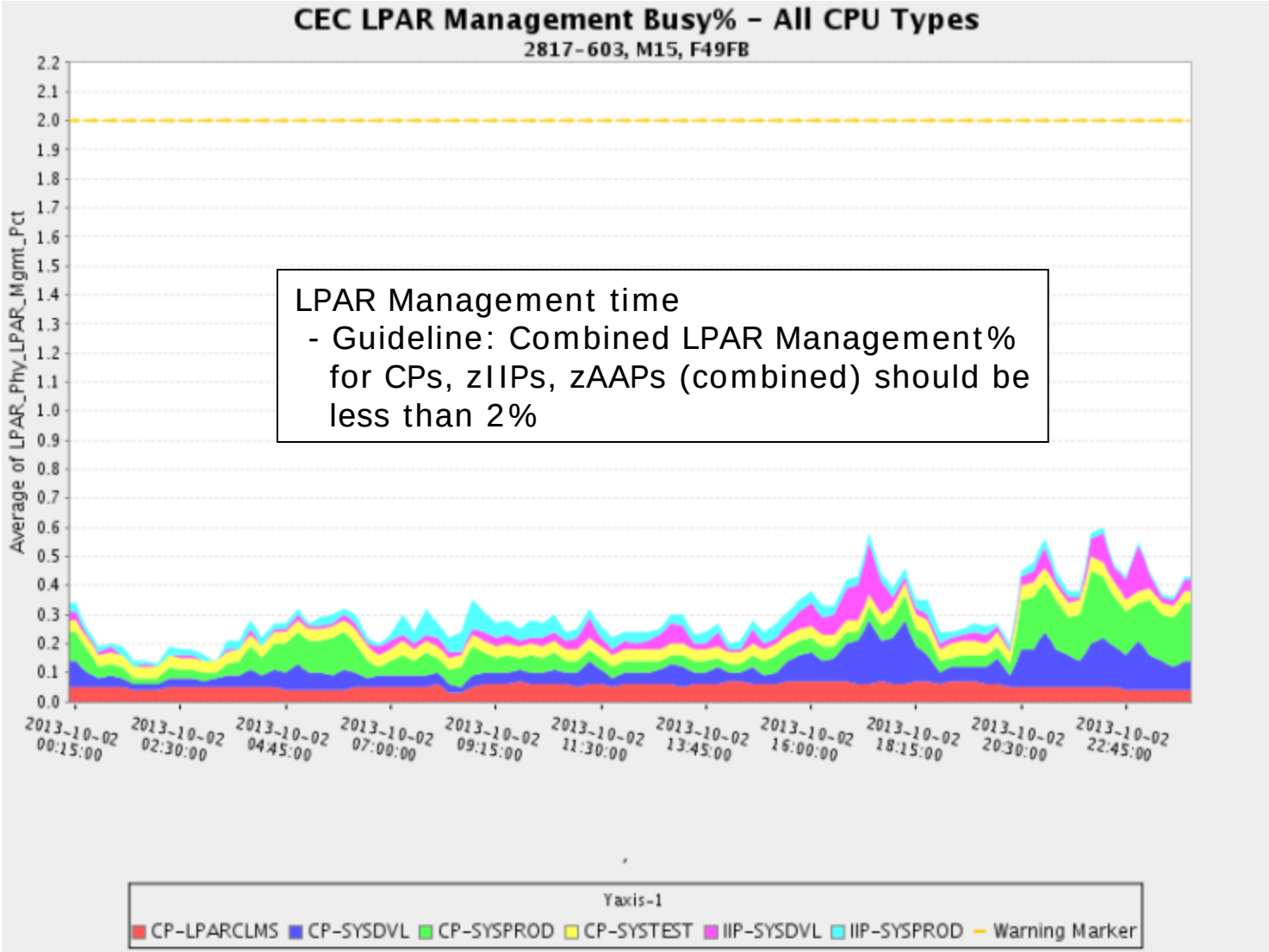
Assigned LPAR Weights for CP Engines



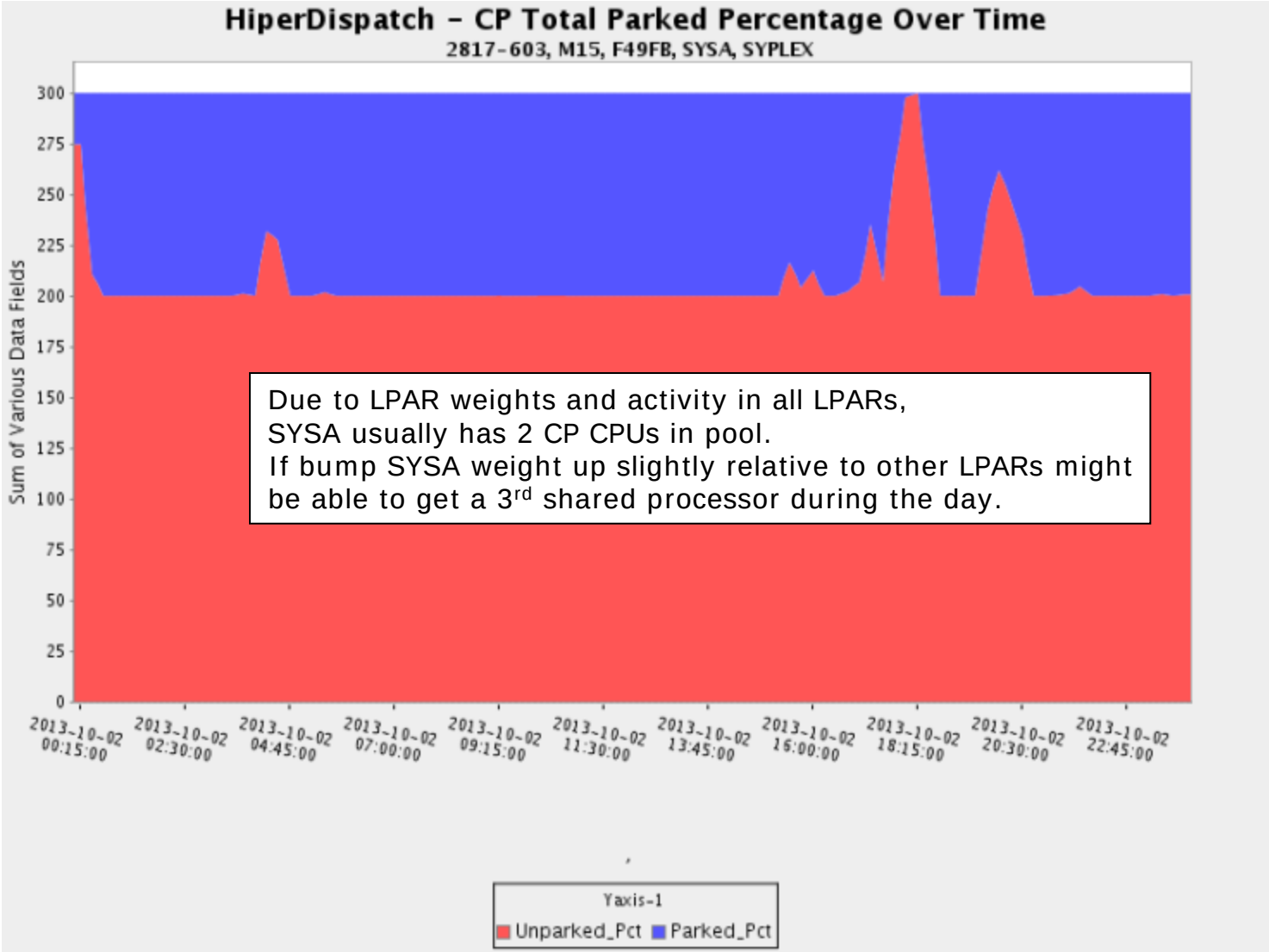
Weight Enforcement



LPAR Management Time



HiperDispatch Parked and Unparked CPs



Decomposing CPU Consumption
-
LPAR Level Analysis

- CPU Consumed by an LPAR
  - The LPAR utilization trinity (LPAR Busy%, Workload Busy%, and MVS Busy %
  - Capture ratios
  - Work Unit distribution to gain insights to latent demand
  - Host Effect CPU Consumption

z/OS CPU Times

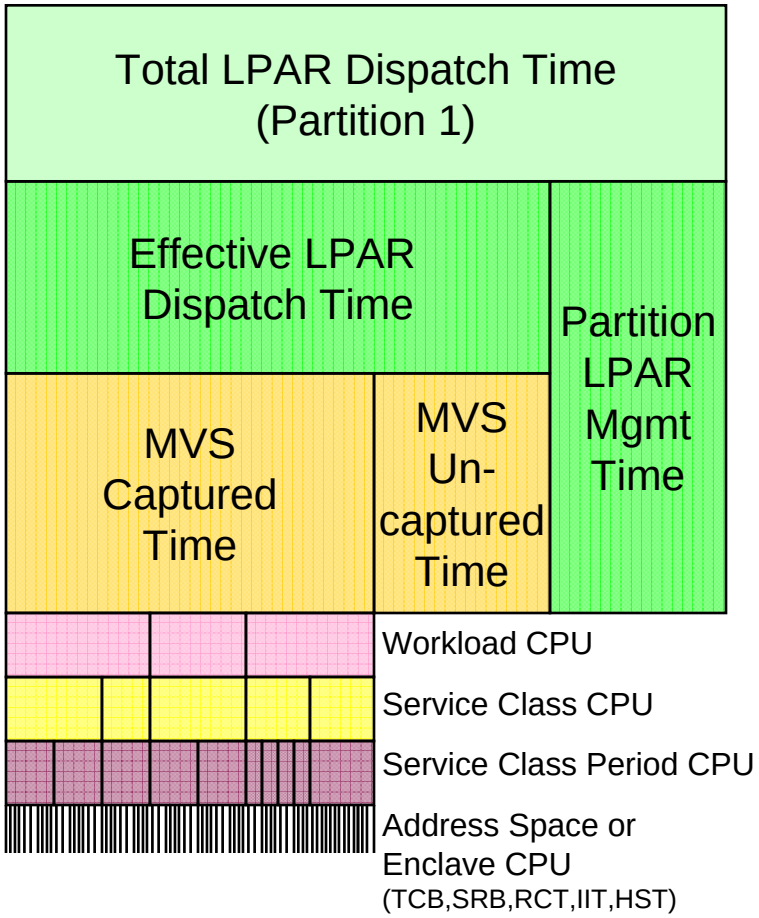
- Capture Ratios used to understand the stability and cost of system overhead

- Effective Dispatch Time
  - Time that the z/OS and the workloads were executing on the CPU

- MVS Capture Time
  - Time that can be accounted for towards specific workloads

- MVS Un-captured Time
  - System overhead

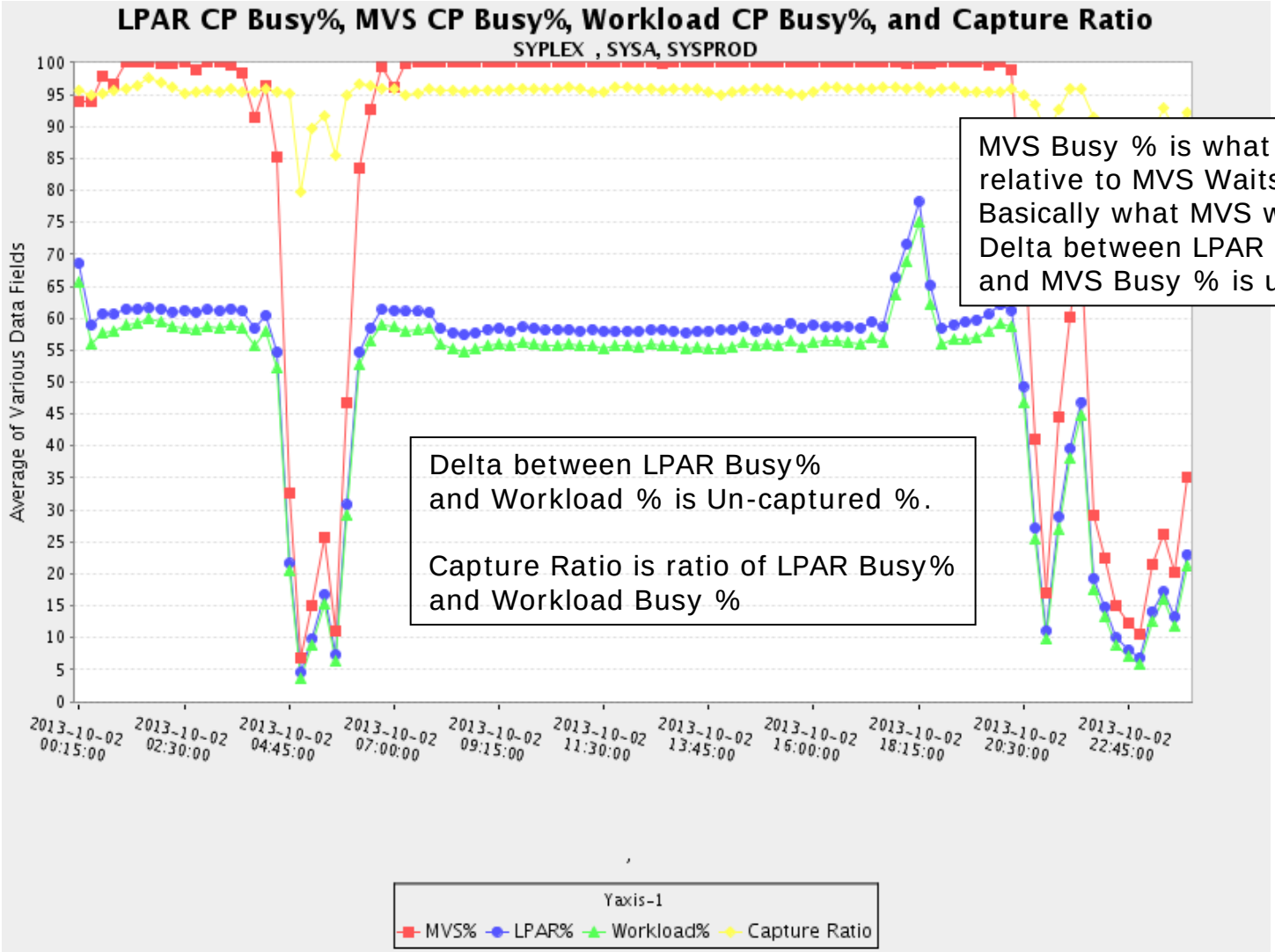
- Capture Ratio
  - Ratio of MVS Capture Time to Effective Dispatch Time



Causes for Uncaptured Time

- Many causes for uncaptured time. Common causes are as follows:
  - High page fault rates
  - Full preemption
  - Suspense lock contention
  - Spin lock contention
  - Getmain/Freemain activity (recommend cell pools)
  - SRM time-slice processing
  - Interrupts
  - SLIP processing
  - Long queues being processed in uncaptured processing
  - Affinity processing (such as need for a specific CPU or crypto facility)

LPAR Busy%, Workload%, MVS%, and Capture Ratio



MVS Busy % is what utilization relative to MVS Waits.
Basically what MVS wanted.
Delta between LPAR Busy % and MVS Busy % is usually due to capping

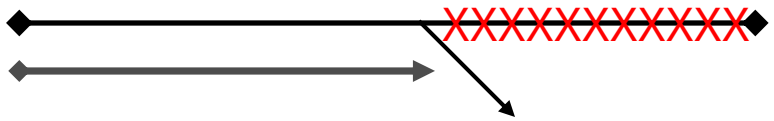
Delta between LPAR Busy% and Workload % is Un-captured %.

Capture Ratio is ratio of LPAR Busy% and Workload Busy %

Understanding Dispatching to Gain Insight to MVS Busy %

- Dispatch Time
  - Time logical processor is associated with a physical processor

- MVS Time
  - Time z/OS was busy before voluntarily giving up a processor



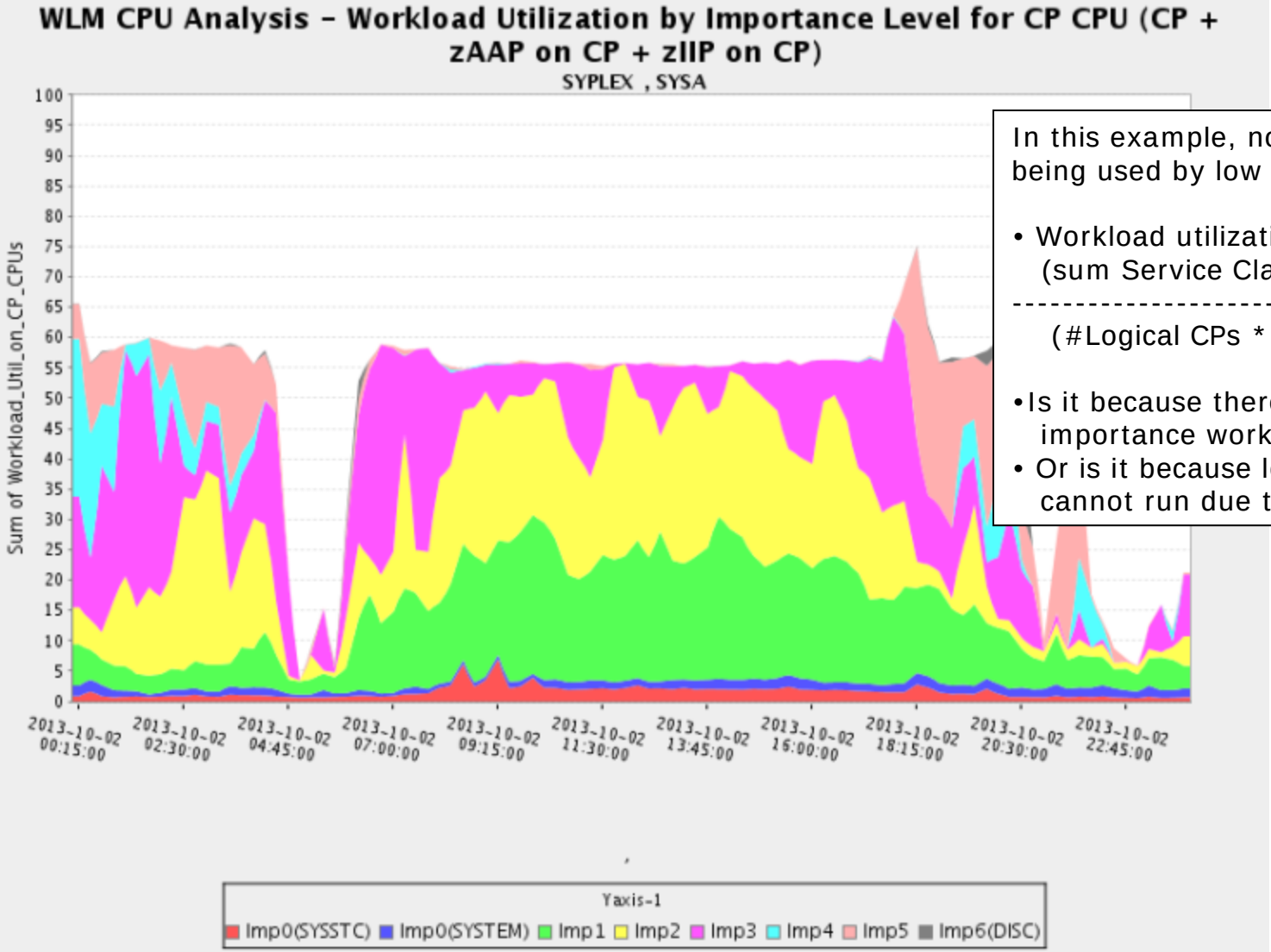
Voluntary Wait
- z/OS voluntarily gives up the processor
- MVS time equals dispatch time

Involuntary Wait (mostly on vertical mediums)
- z/OS does not give up the processor voluntarily
- Instead PR/SM un-dispatches the partition
- MVS time will be greater than dispatch time

Decomposing CPU Consumption
-
WLM Workload Level Analysis

- CPU Consumption at the importance level
- CPU Consumption at the WLM Service Class and Service Class Period Level
- Commentary about Report Classes
- Other CPU consumption measurements
  - –CPU consumed at promotion
  - –Did lower importance work not consume CPU due to lack of demand or due to lack of CPU?

Workload Utilization by Importance Level



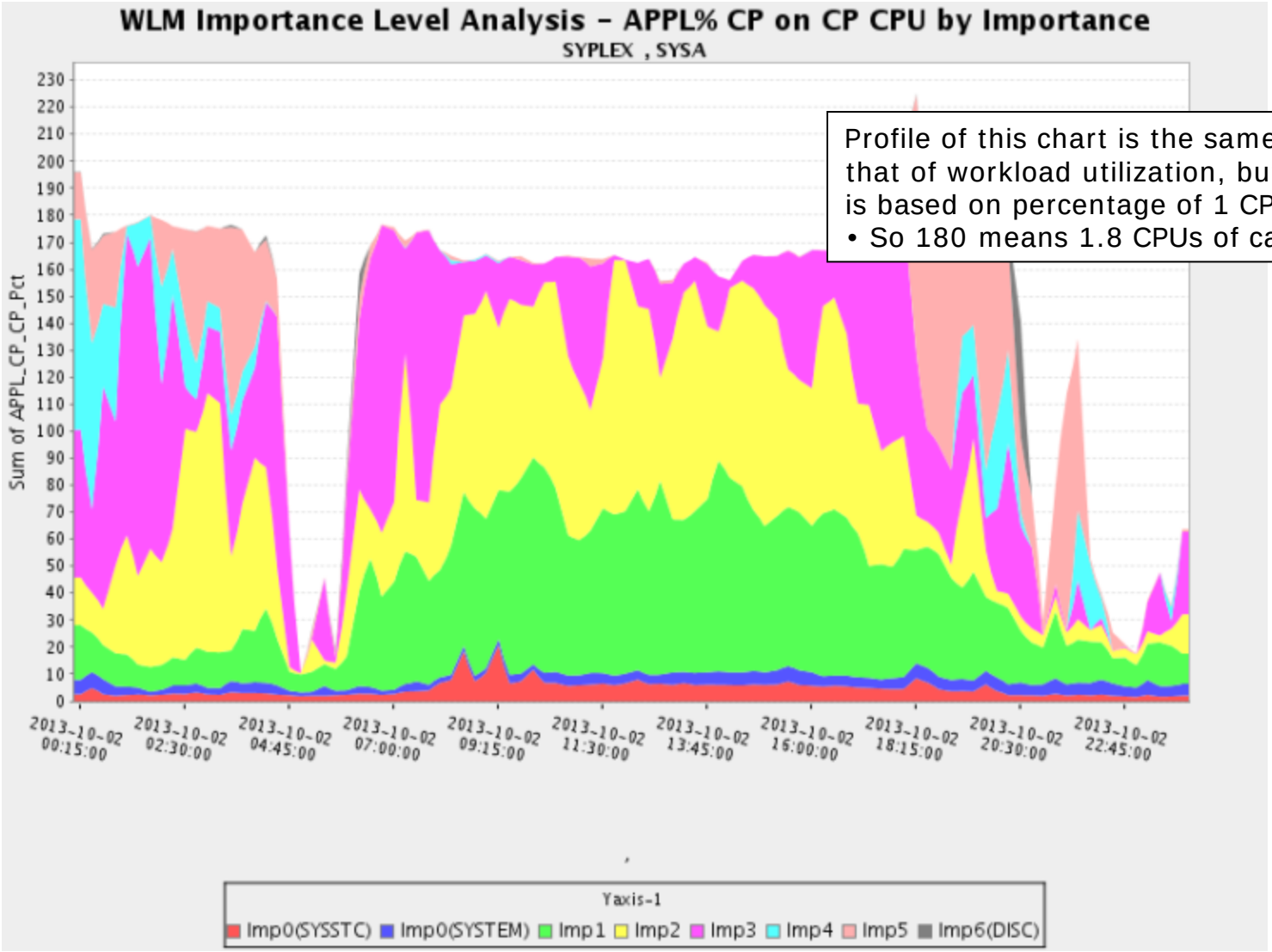
In this example, notice that little CPU is being used by low importance work.

- Workload utilization is calculated as

$$\frac{\text{(sum Service Class CPU Tim)}}{\text{(\#Logical CPs * Interval Time)}}$$

- Is it because there is little to no low importance work?
- Or is it because low importance work cannot run due to lack of capacity?

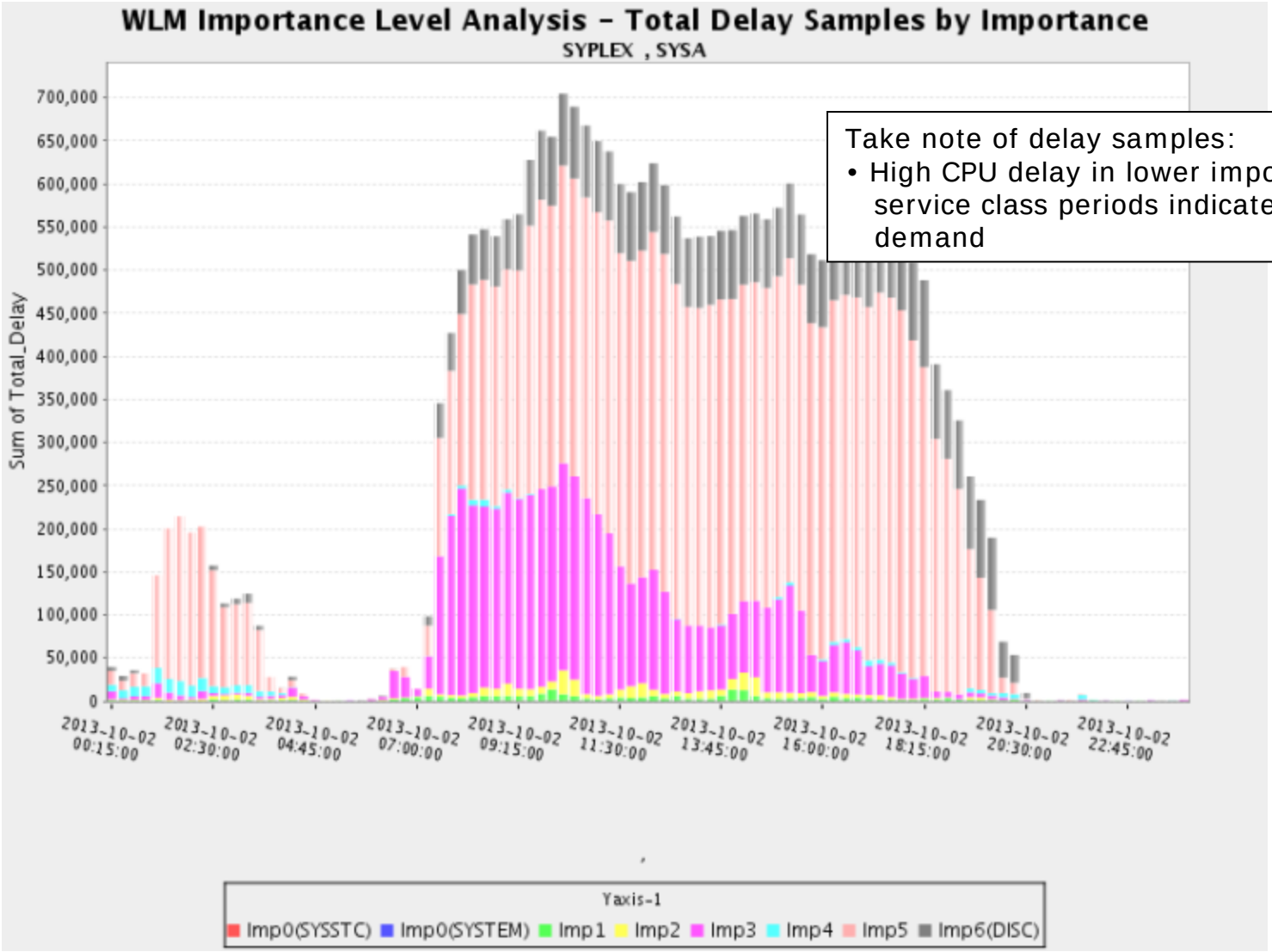
APPL% by Importance Level



Profile of this chart is the same as that of workload utilization, but scale is based on percentage of 1 CPU.

- So 180 means 1.8 CPUs of capacity.

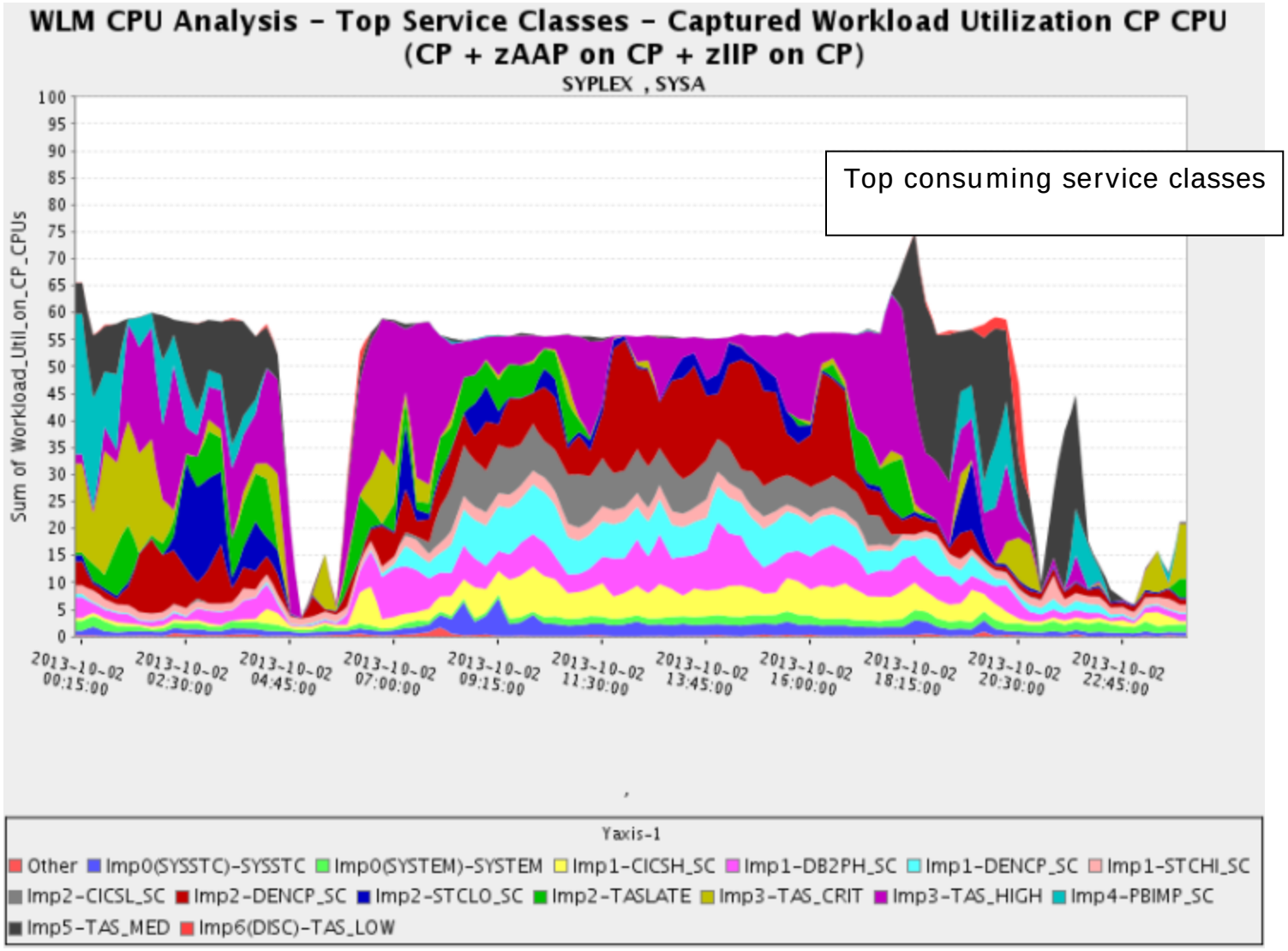
Delay Samples by Importance Level



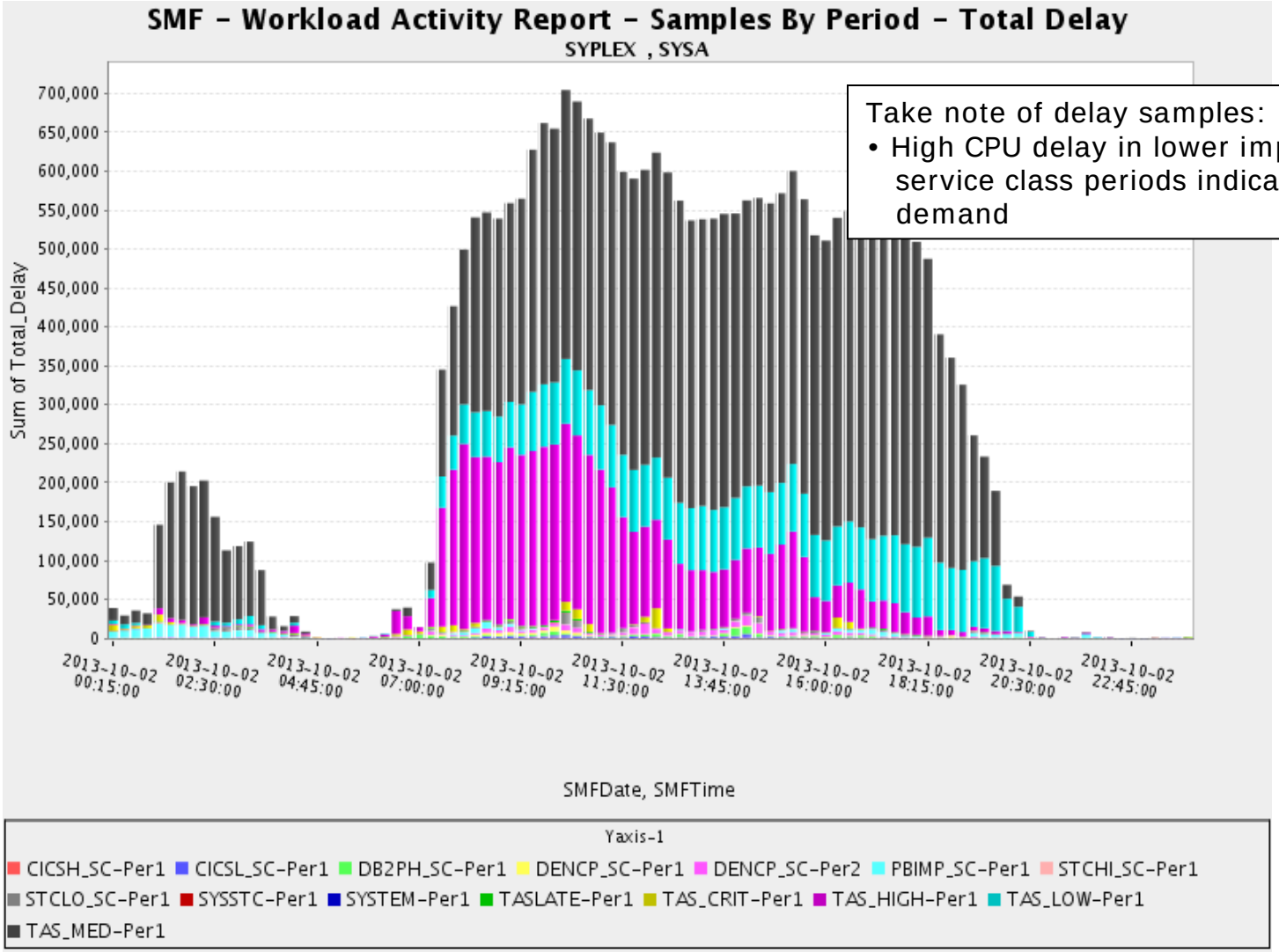
Take note of delay samples:

- High CPU delay in lower importance service class periods indicates latent demand

Workload Utilization by Service Class Period



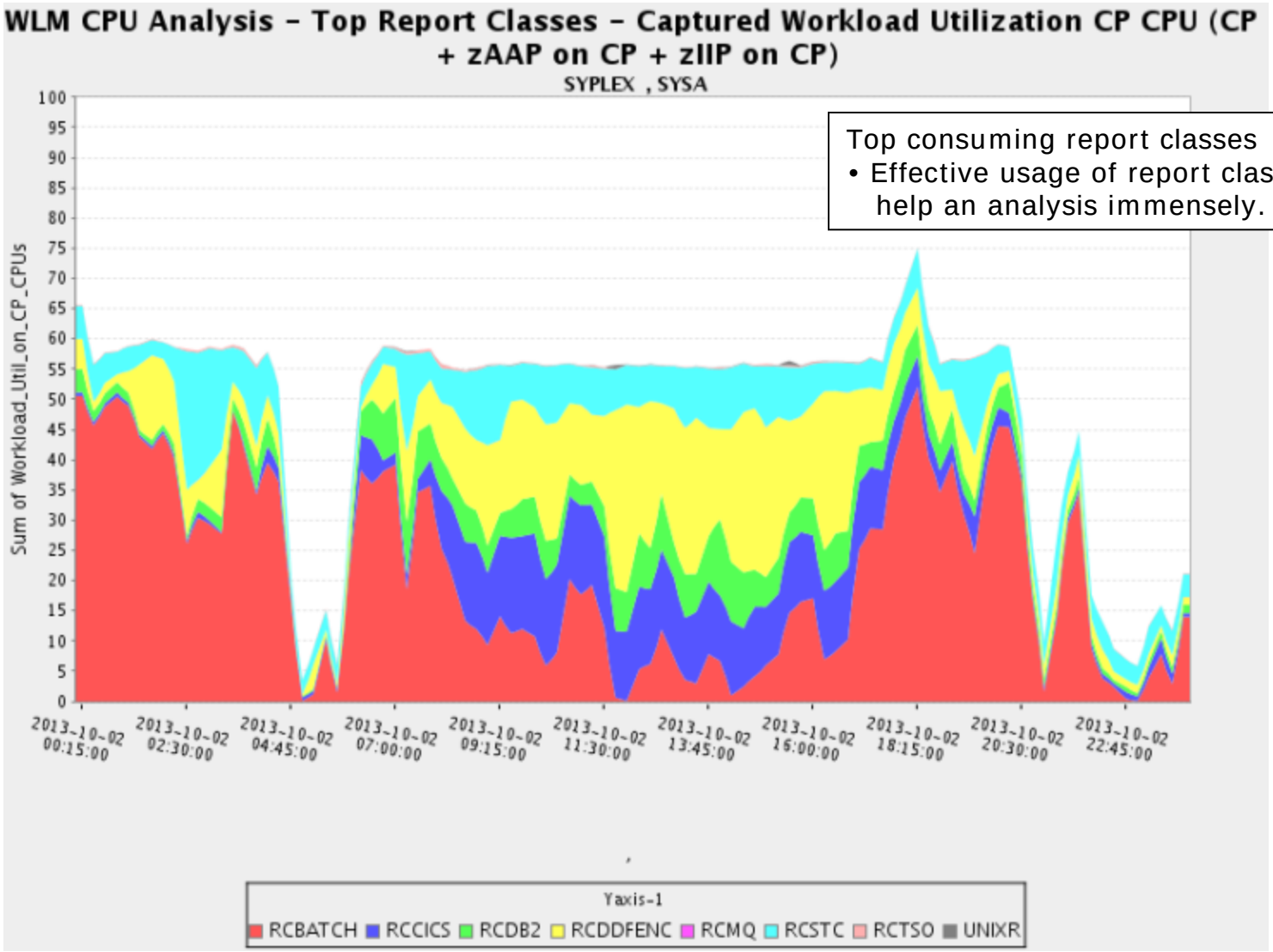
Delay Samples by Service Class Period



Take note of delay samples:

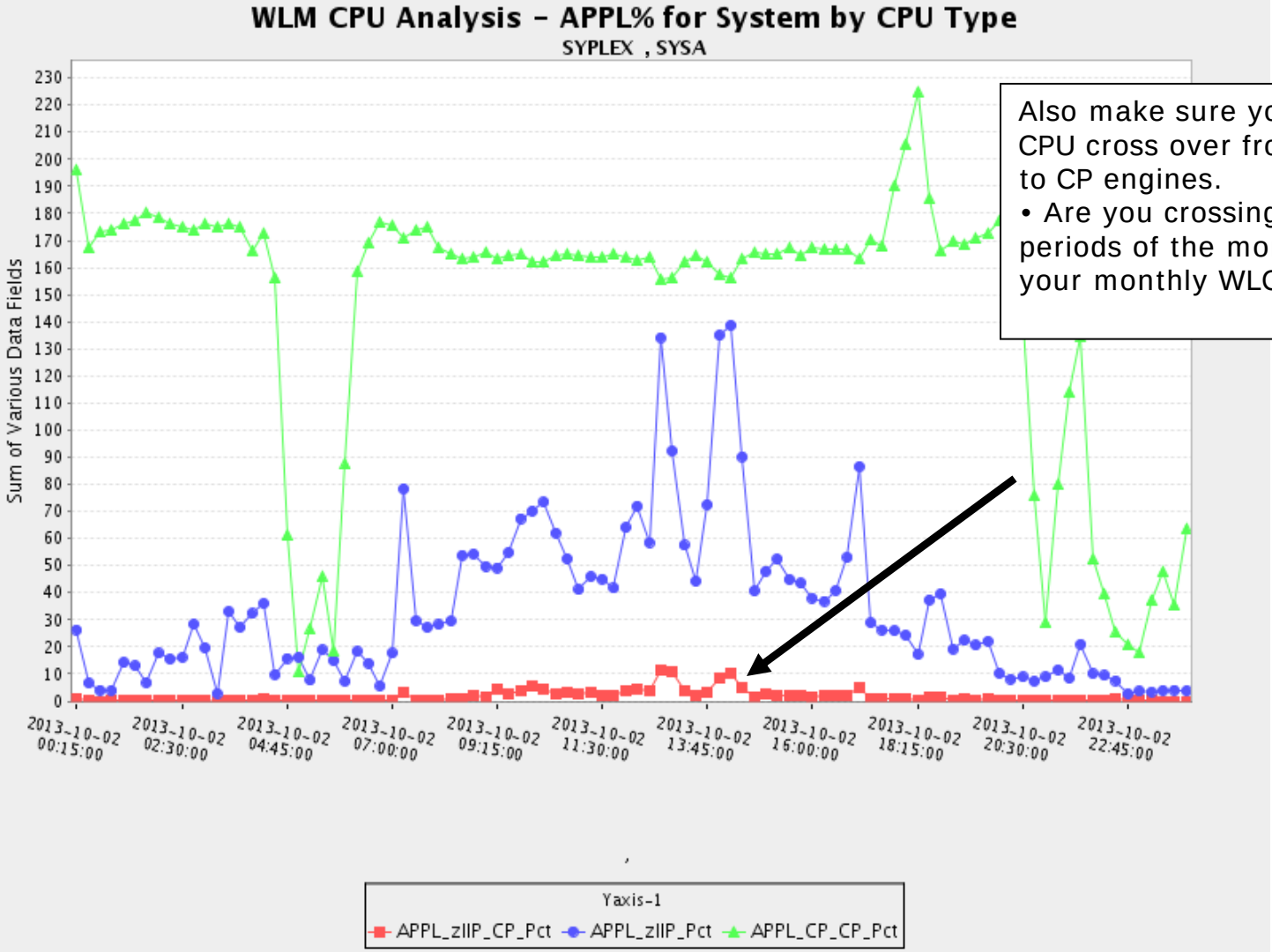
- High CPU delay in lower importance service class periods indicates latent demand

Workload Utilization for Top Report Classes



Top consuming report classes
- Effective usage of report classes can help an analysis immensely.

Workload Utilization as a Percentage of 1 CPU (APPL%) - By CPU Type



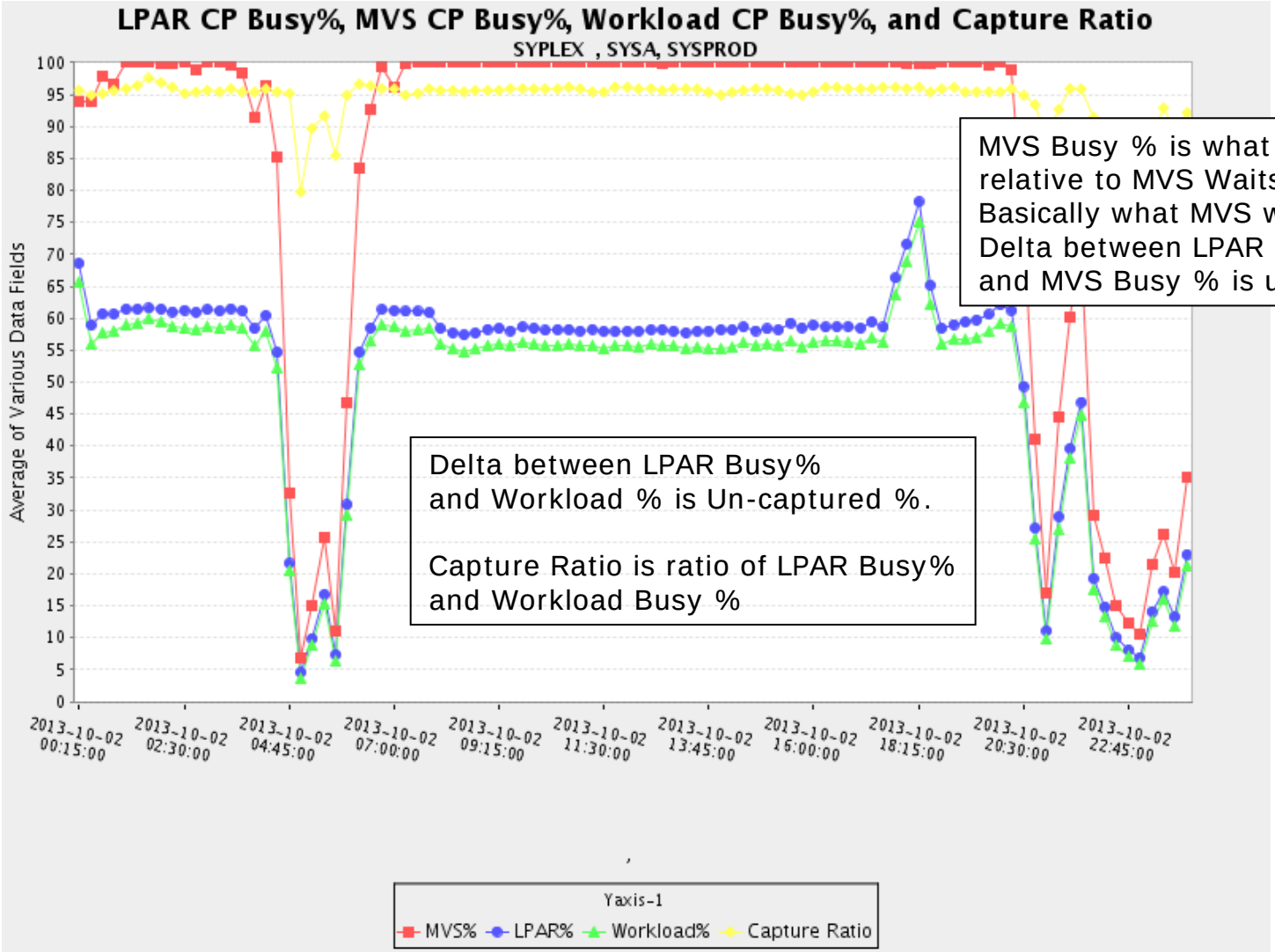
Also make sure you understand the CPU cross over from zIIP and zAAP engines to CP engines.

- Are you crossing over during peak periods of the month and possibly affecting your monthly WLC software bill?

Frustrating CPU Time Problem

Sometimes you may see more CPU time consumed by your workloads than the CPU time that PR/SM is dispatching to the LPAR.

LPAR Busy%, Workload%, MVS%, and Capture Ratio

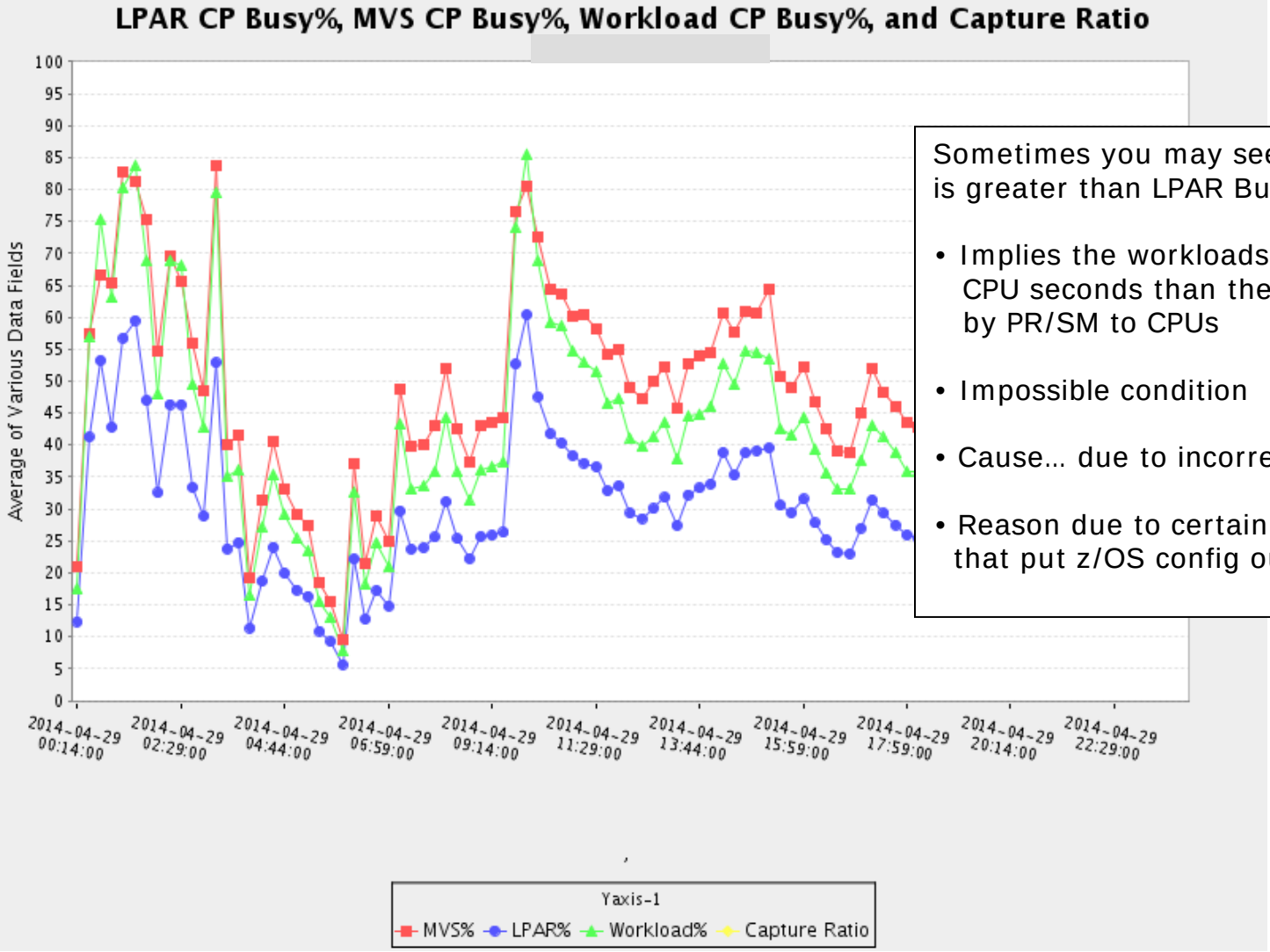


MVS Busy % is what utilization relative to MVS Waits.
Basically what MVS wanted.
Delta between LPAR Busy % and MVS Busy % is usually due to capping

Delta between LPAR Busy% and Workload % is Un-captured %.

Capture Ratio is ratio of LPAR Busy% and Workload Busy %

Beware… sometimes workload CPU can be greater than dispatch CPU. (???????)



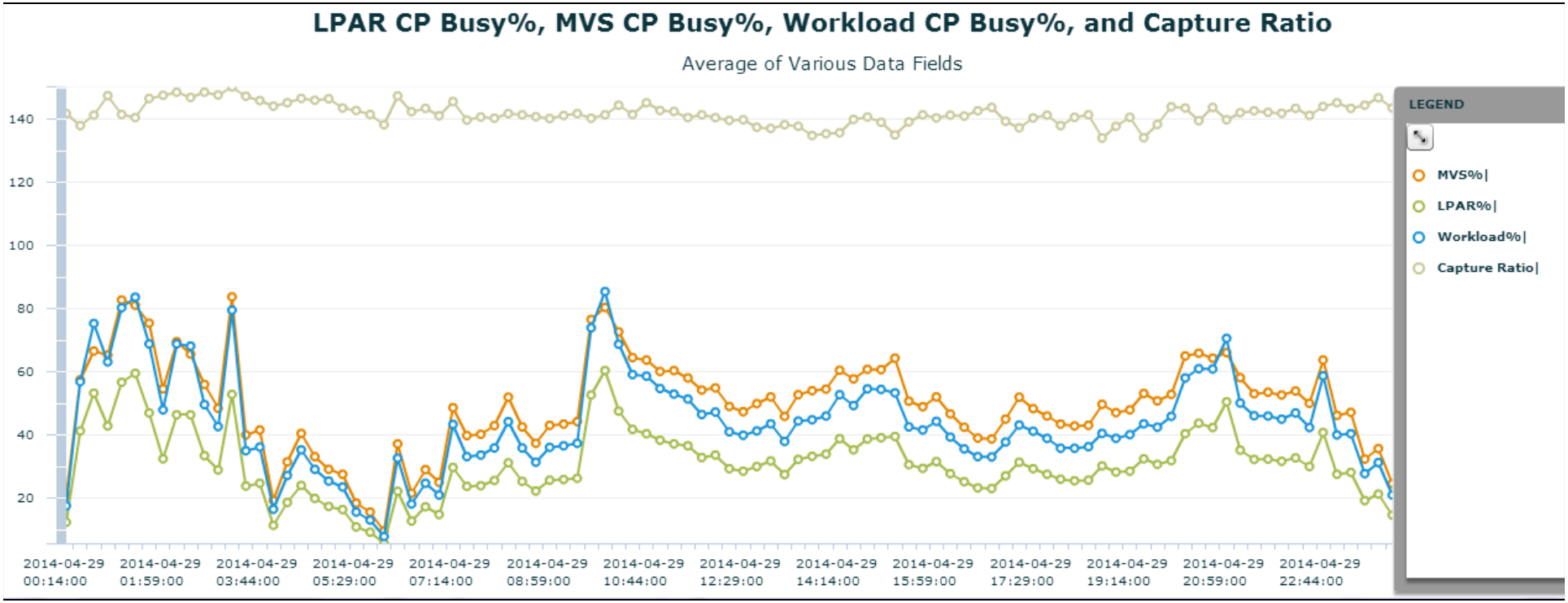
Sometimes you may see that Workload % is greater than LPAR Busy %

- Implies the workloads consumed more CPU seconds than the system was dispatched by PR/SM to CPUs
- Impossible condition
- Cause… due to incorrect SU/Sec constant
- Reason due to certain configuration changes that put z/OS config out of sync with PR/SM

Beware… sometimes workload CPU can be greater than dispatch CPU. (???????)

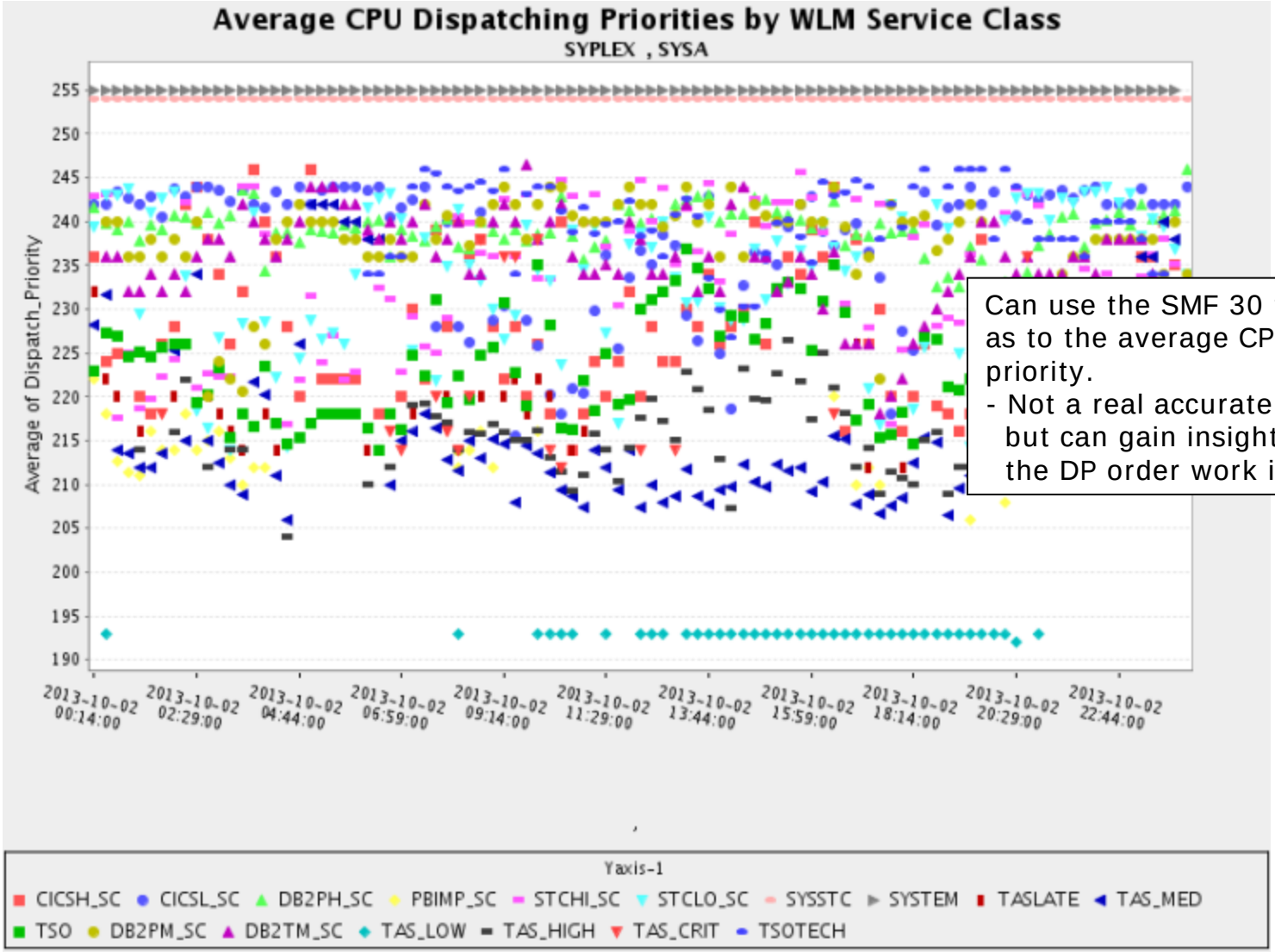
- This is the same chart as previous slide only showing the capture ratios of greater than 100% since Workload Time > LPAR Dispatch Time

- Results in higher than expected CPU time per transactions… screws up CPU



Looking at CPU Dispatching Priorities (an approximation)

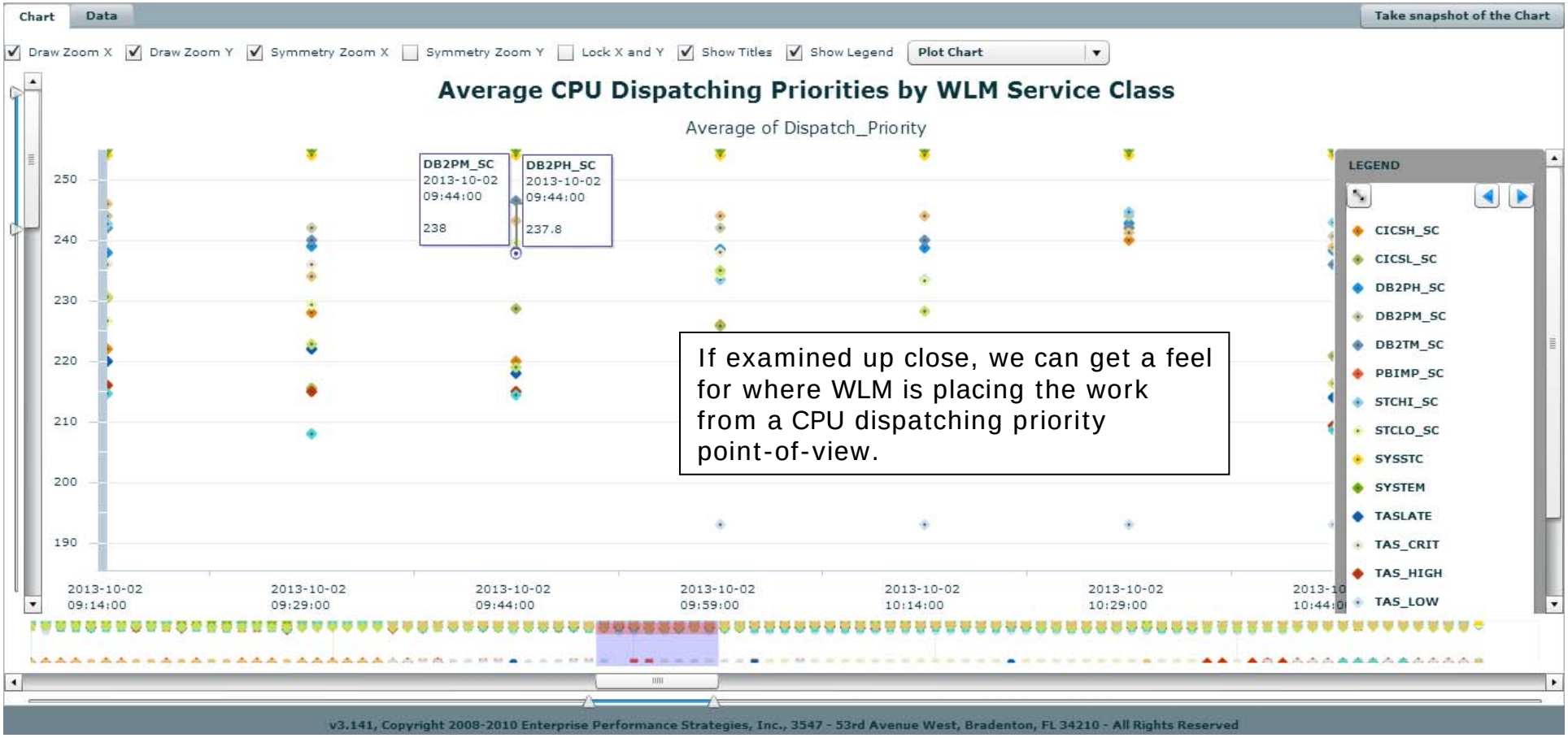
Average CPU Dispatching Priorities for Address Spaces



Can use the SMF 30 to gain insights as to the average CPU dispatching priority.

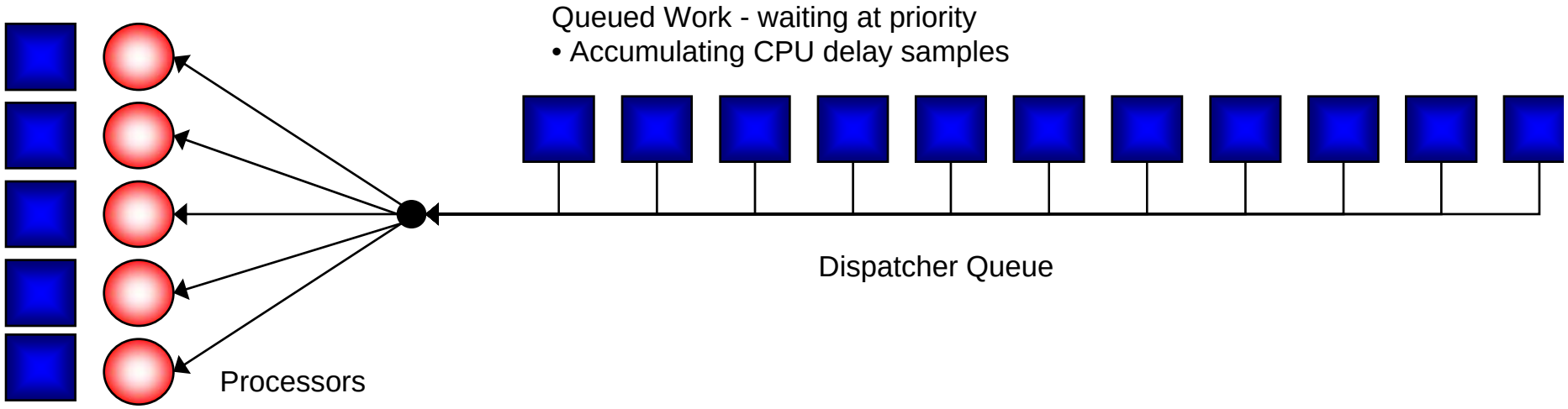
- Not a real accurate measurement, but can gain insight into where in the DP order work is.

Average CPU Dispatching Priorities for Address Spaces

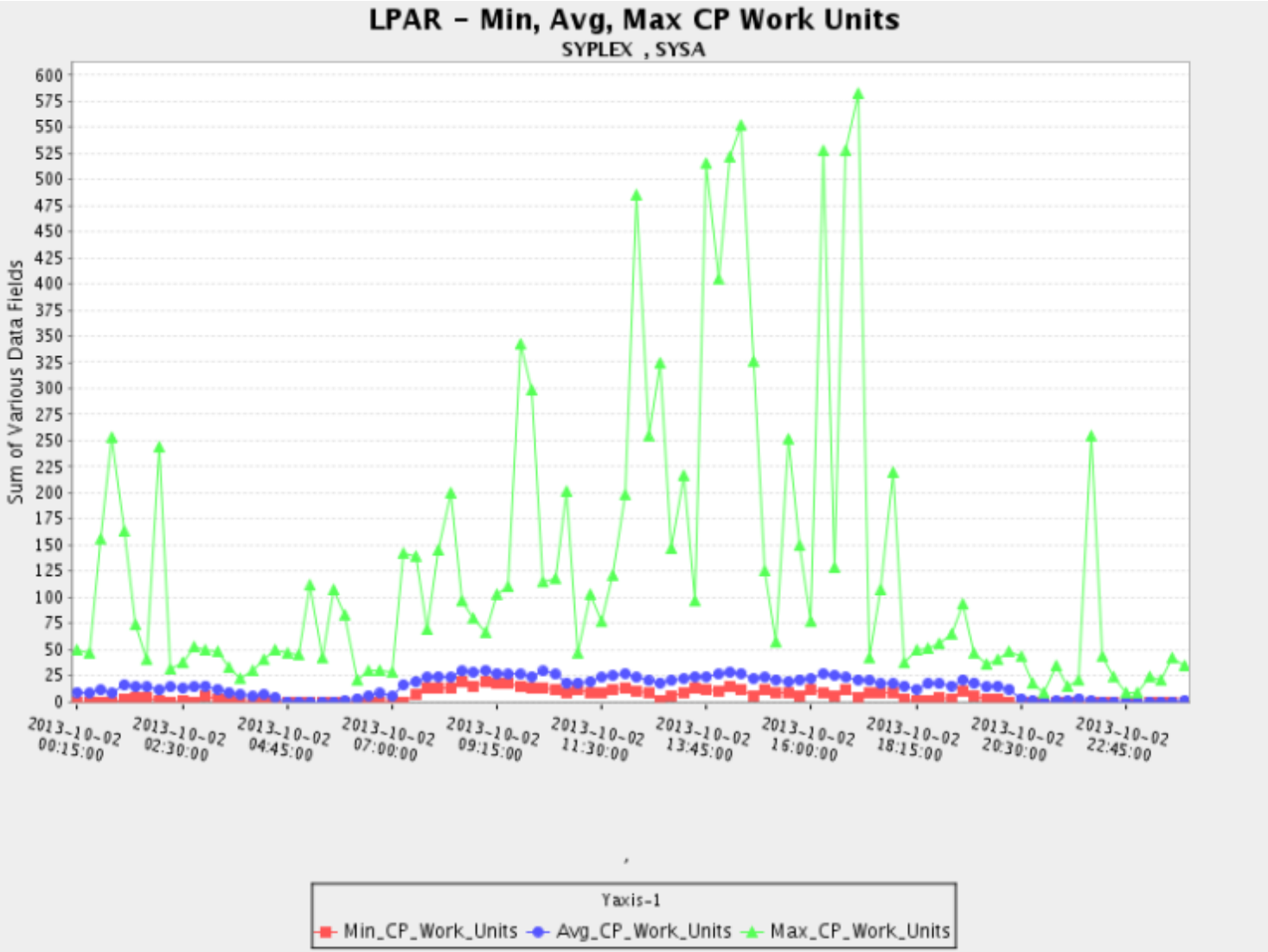


Insights into Latent Demand

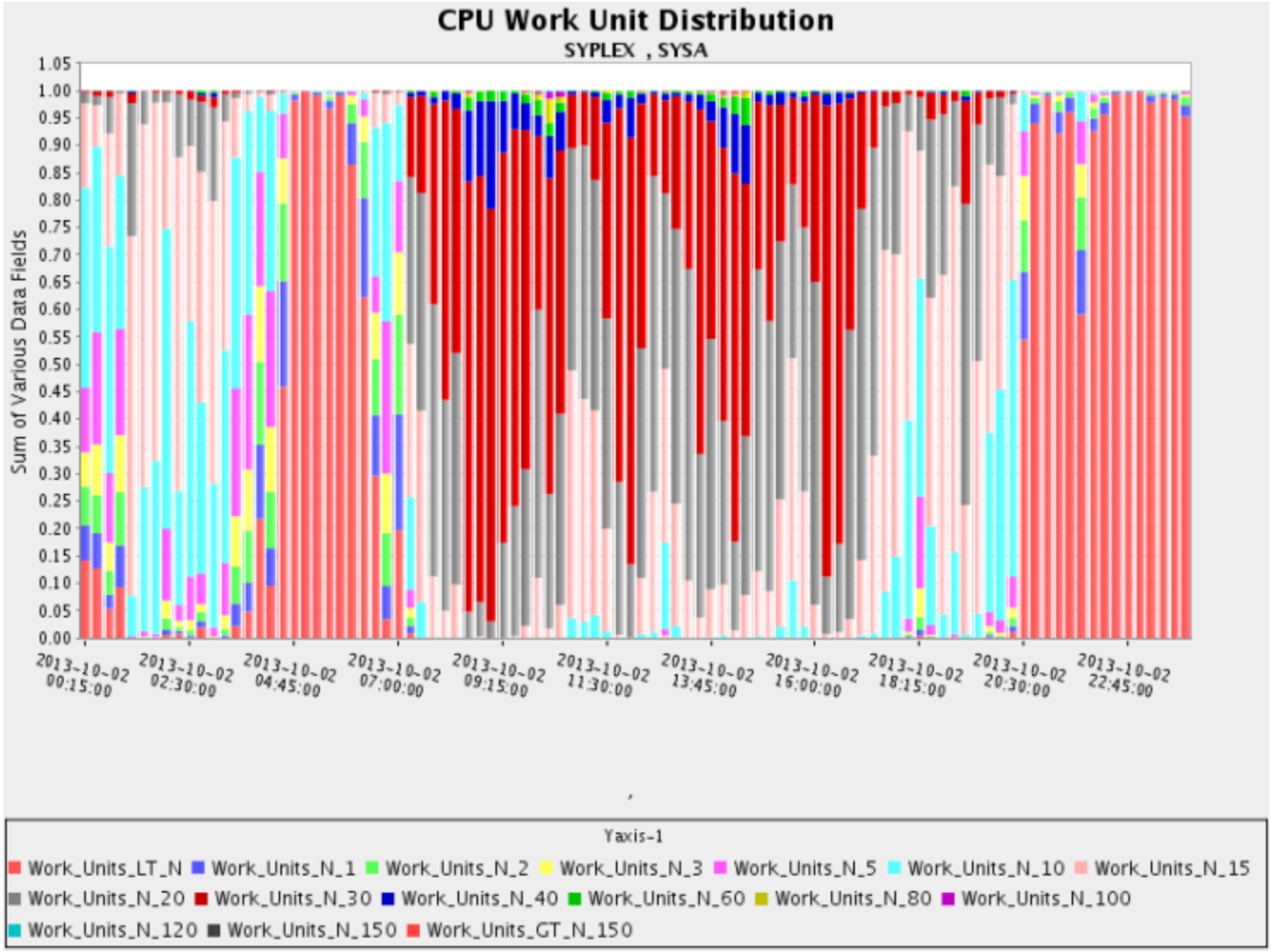
Dispatched Work
• Accumulating CPU Using Samples



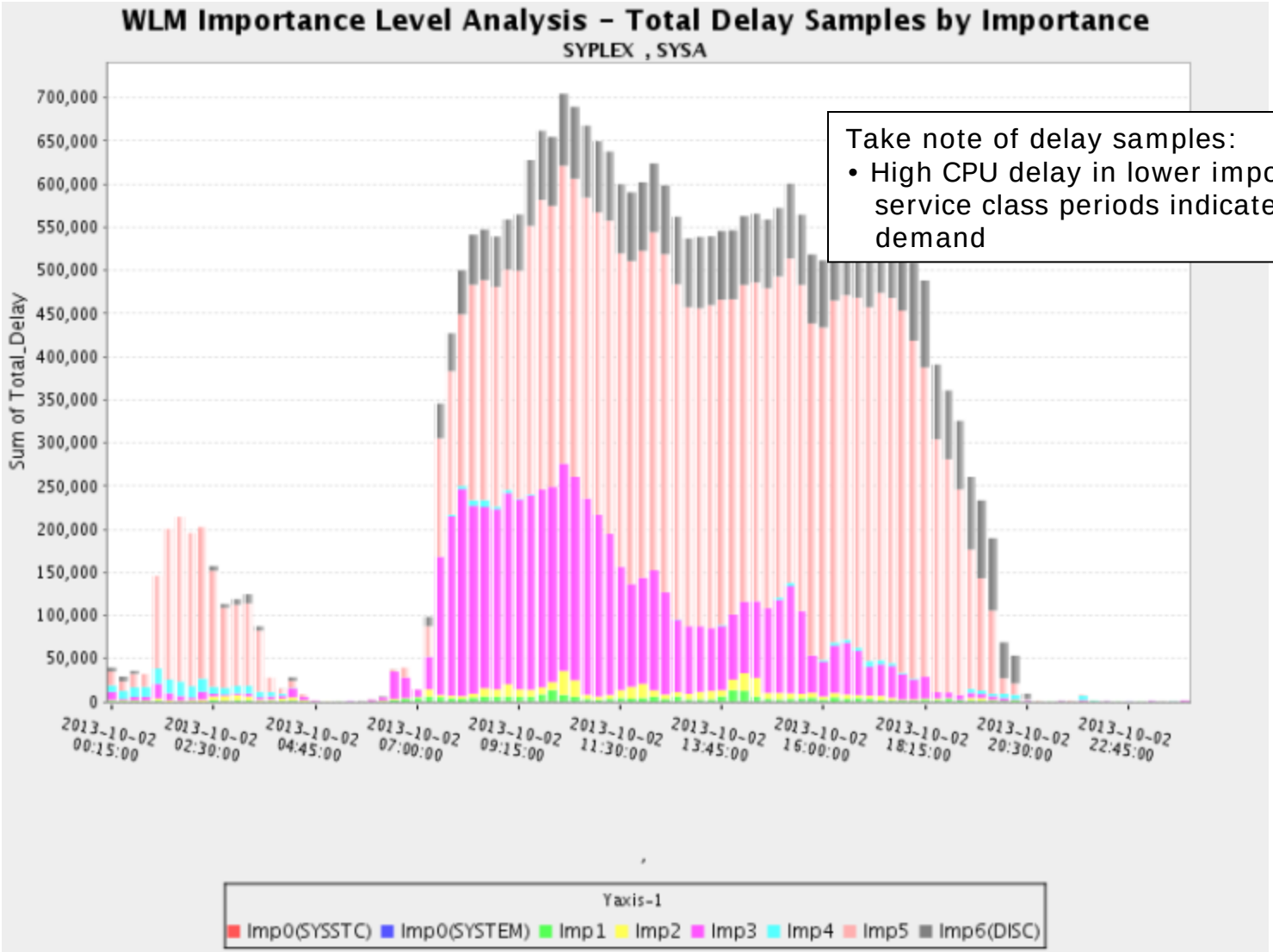
Min / Max / Avg Work Unit Queuing



Work Unit Distribution
Showing Latent Demand



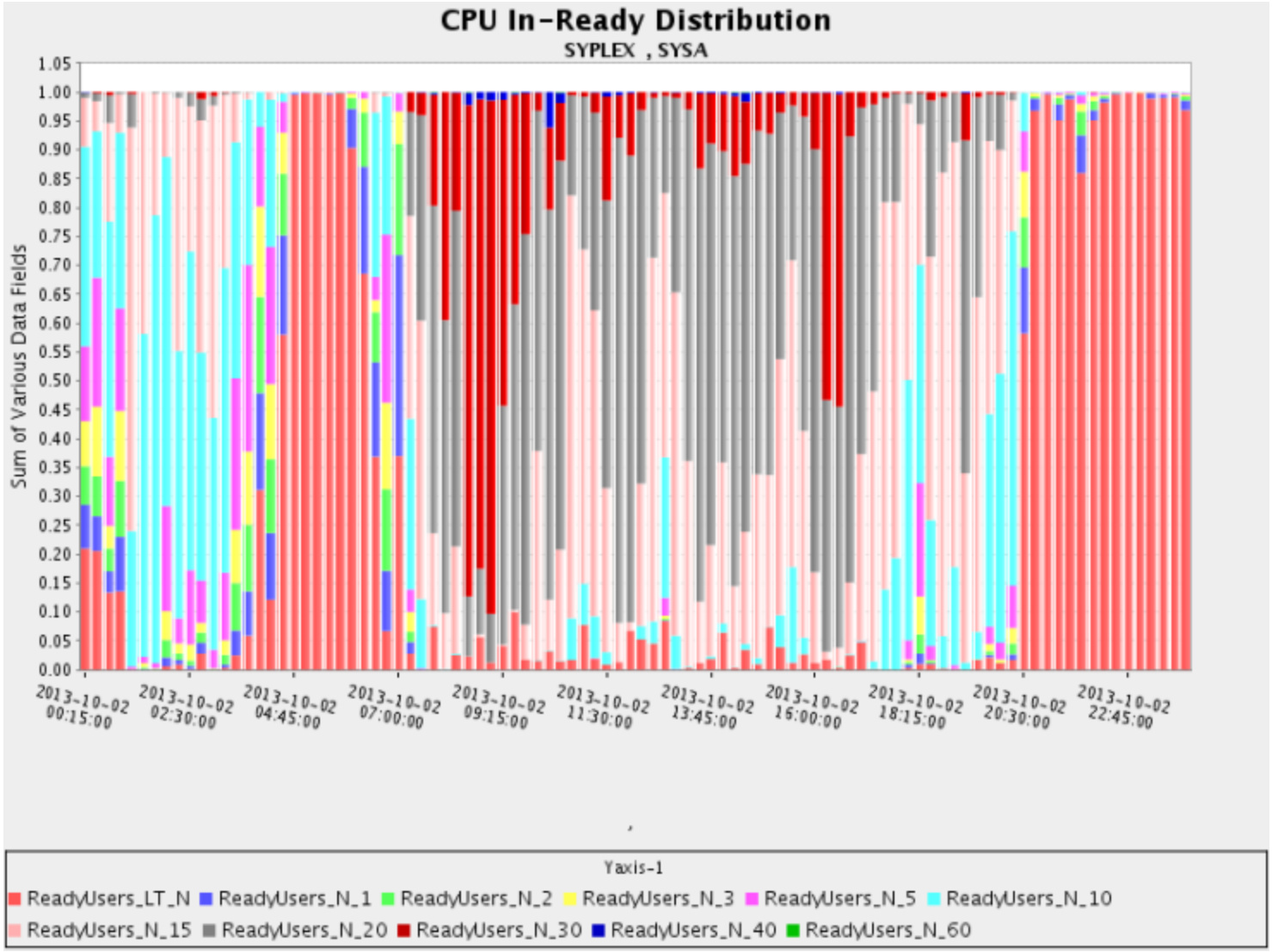
Delay Samples by Importance Level



Take note of delay samples:

- High CPU delay in lower importance service class periods indicates latent demand

Older Style In-Ready Distribution – Less Accurate Latent Demand



Top Address Spaces Consuming CPU

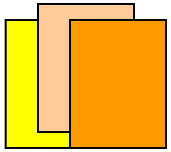
Top 20 Address Spaces Consuming Most CPU in 24 Hours

SC_Name	RC_Name	Job_Name	AS_Type	SYS1	SYS2	SYS3	SYS4	Sum	Machine%
DB2PH_SC	DB2R	DSNDIST	STC	35,883.6				35,883.6	13.8%
CICSH_SC	CICSR	CICSHADP	STC	13,921.3				13,921.3	5.4%
CICSL_SC	CICSR	CICSH81P	STC	10,527.0				10,527.0	4.1%
DB2PH_SC	DB2R	DSNDBM1	STC	10,127.9				10,127.9	3.9%
STCLO_SC	STCR	DFHSM	STC	7,964.1			214.6	8,178.7	3.2%
CICSH_SC	CICSR	CICSH11P	STC			5,797.9		5,797.9	2.2%
STCHI_SC	STCR	OMEGDSST	STC	1,622.2	1,019.7	1,146.9	1,827.2	5,616.0	2.2%
SYSTEM	STCR	WLM	SYS	535.2	342.7	211.5	1,890.2	2,979.6	1.1%
HPS_HIGH	BATCHR	HM026D03	JOB	2,376.6				2,376.6	0.9%
SYSSTC	STCR	NET	STC	1,005.0	44.8	485.3	749.8	2,285.0	0.9%
HPS_HIGH	BATCHR	IT110D01	JOB	2,145.8				2,145.8	0.8%
SYSTEM	STCR	CATALOG	SYS	1,540.7	11.3	14.7	572.2	2,138.9	0.8%
SYSSTC	STCR	TCPIP	STC	1,476.8	98.9	118.5	374.9	2,069.1	0.8%
TBATL_SC	BATCHR	DB2HRWS0	JOB				1,924.4	1,924.4	0.7%
CICSH_SC	CICSR	CICSMG1P	STC	1,735.9				1,735.9	0.7%
TBATL_SC	BATCHR	SITH085U	JOB				1,685.2	1,685.2	0.7%
DB2TH_SC	DB2R	HPDQDIST	STC				1,683.1	1,683.1	0.6%
DB2TH_SC	DB2R	HPDQDBM1	STC				1,551.3	1,551.3	0.6%
PBIMP_SC	BATCHR	HPSVSAM1	JOB	1,302.6				1,302.6	0.5%
HPS_HIGH	BATCHR	HM026D01	JOB	1,296.8				1,296.8	0.5%

Objective of WLM Management of CICS & IMS

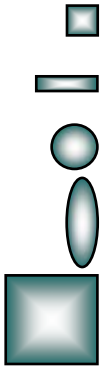
- Allow assignment of goals to the transactions and let the WLM determine which regions need the resources to meet these goals.

Region Goals



IMP 1, Velocity 60

Transaction Goals



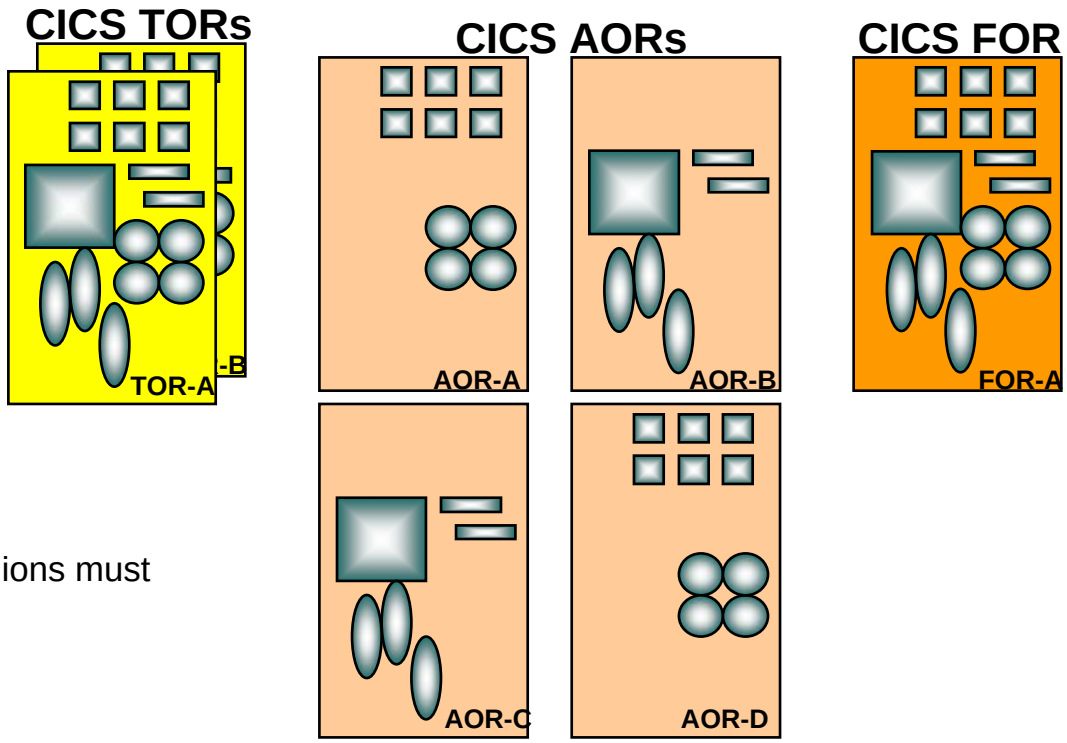
IMP 1, RT .5 sec, 90%

IMP 1, RT .75 sec, 90%

IMP3, RT 2 sec, 90%

IMP 3, Avg RT 3 sec

IMP 5, RT 20 sec, 85%

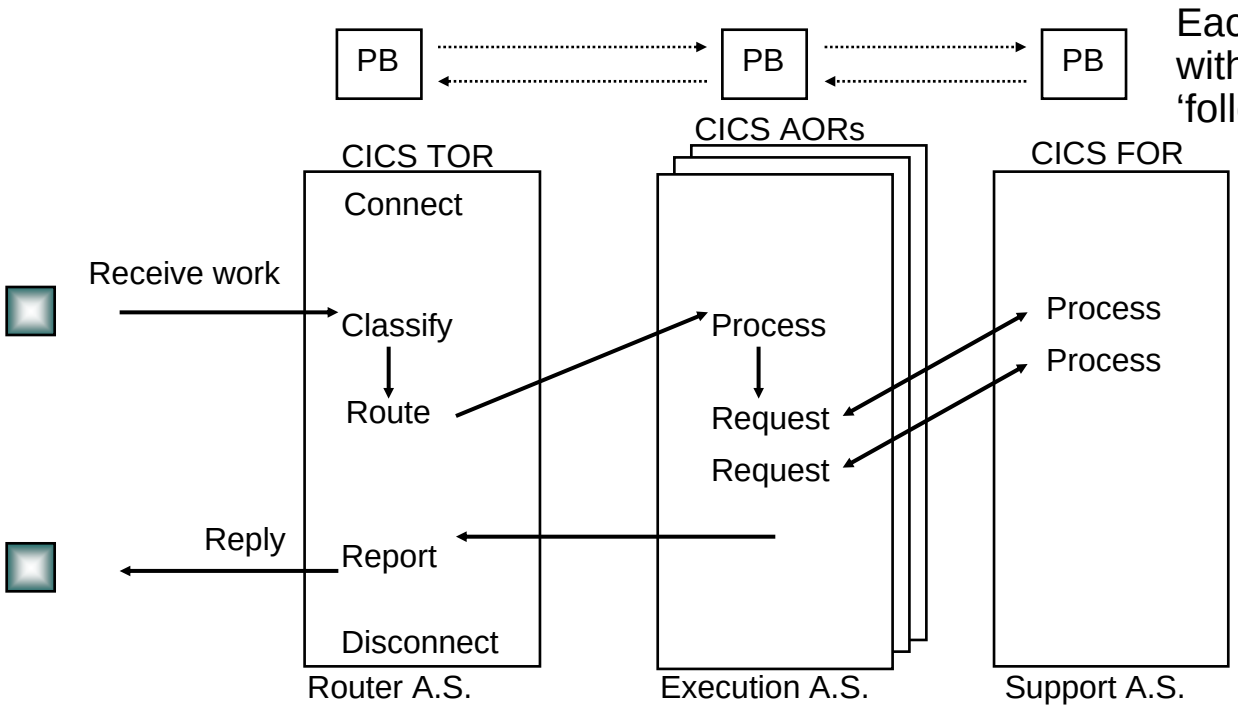


To meet the RT goals of the following regions must be managed:

- CICS TOR-A, TOR-B
- CICS AOR-A, AOR-D
- CICS FOR-A

WLM needs an awareness of which regions are processing which transactions, and how often

- CICS and IMS exploit WLM Work Manager services
  - Regions '***Connect***' (ie 'register') to WLM during startup & obtain current service policy
  - At transaction startup, region uses WLM '***Classify***' to associate incoming transaction with a service class
  - At transaction end, region uses WLM '***Report***' to signal end and report response time
  - Other important services to make this all work

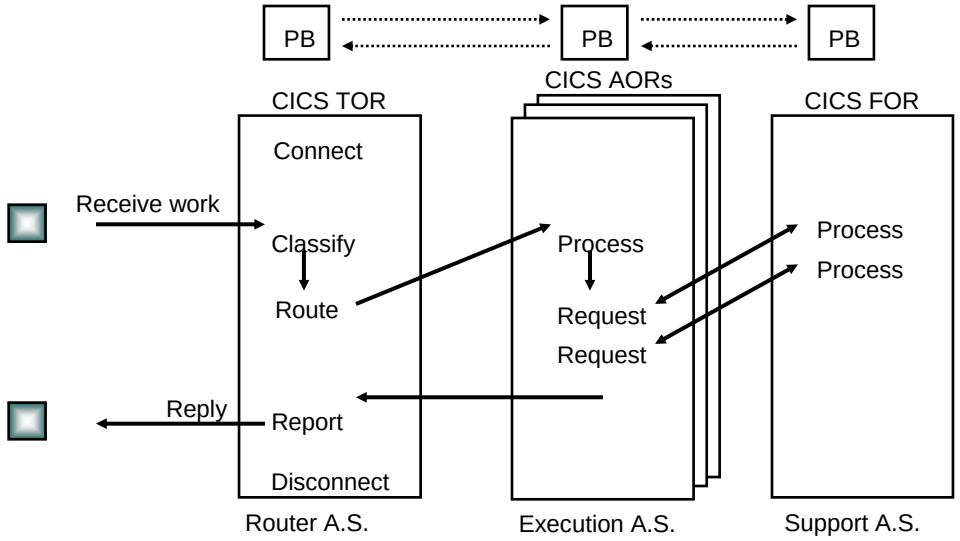


Each transaction is 'associated' with a Performance Block(s) that 'follows' the transaction.

Note: IMS looks a little different, but similar concept

WLM Sampling and CICS MAXTASK Parameter

- Beware of excess sampling overhead due to CICS MAXTASK parameter!
  - In a CICS environment, one PB is pre-allocated for each possible task as set by the CICS MAXTASK parameter
- All PBs are sampled every 1/4 second
    - Could cause lots of WLM sampling overhead!
  - Check CICS MAXTASK parameter to make sure it is not set unnecessarily high
    - Set to your system's true high water mark
  - Mostly resolved, but still watch MAXTASK



Top 20 Address Spaces Consuming Most CPU in 24 Hours

SC_Name	RC_Name	Job_Name	AS_Type	SYS1	SYS2	SYS3	SYS4	Sum	Machine%
DB2PH_SC	DB2R	DSNDIST	STC	38,655.5				38,655.5	14.9%
CICSH_SC	CICSPRHR	CICSHADP	STC	14,269.1				14,269.1	5.5%
DB2PH_SC	DB2R	DSNDBM1	STC	7,147.9				7,147.9	2.8%
CICSH_SC	CICSPRHR	CICSH81P	STC	5,032.1				5,032.1	1.9%
STCLO_SC	OMEGAMON	OMEGDSST	STC	1,403.0	807.8	928.4	1,340.4	4,479.6	1.7%
CICSH_SC	CICSPRHR	CICSH11P	STC			3,662.2		3,662.2	1.4%
STCLO_SC	DFHSMR	DFHSM	STC	2,929.9			295.1	3,225.0	1.2%
DB2TM_SC	DB2R	DB2JDIST	STC				2,839.3	2,839.3	1.1%
SYSTEM	STCR	WLM	SYS	483.3	304.8	192.9	1,314.1	2,295.1	0.9%
SYSSTC	STCR	RMFGAT	STC	414.7	644.8	376.3	858.9	2,294.7	0.9%
SYSSTC	STCR	TCPIP	STC	1,319.6	85.6	99.4	541.3	2,045.9	0.8%
PSTD_SC	BATSTDR	DB105M00	JOB	2,007.3				2,007.3	0.8%
PMED_SC	BATMEDR	HPSVSMTH	JOB	1,939.5				1,939.5	0.7%
PHIGH_SC	BATHIGHR	IT110D01	JOB	1,860.6				1,860.6	0.7%
TBATL_SC	BATTSTR	DSNLRW00	JOB				1,717.5	1,717.5	0.7%
SYSSTC	STCR	NET	STC	728.5	35.7	364.5	507.5	1,636.2	0.6%
SYSTEM	STCR	CATALOG	SYS	1,185.5	11.2	13.6	389.6	1,600.0	0.6%
CICSH_SC	CICSPRHR	CICSMG1P	STC	1,326.8				1,326.8	0.5%
PMED_SC	BATMEDR	HPSVSMA1	JOB	1,200.9				1,200.9	0.5%
PMED_SC	BATMEDR	DB2REOF1	JOB	1,120.8				1,120.8	0.4%

Session summary

- Many areas need to be examined when decomposing CPU consumption
  - This presentation just discussions some of the many areas

- Basic Processor Consumption Analysis
  - Decomposing CPU Consumption
    - By importance level
    - Displaced workloads
    - By Service Class and Report Class
  - Looking at CPU Dispatching Priorities
  - Looking at Latent Demand

Performance Workshops Available

During these workshops you will be analyzing your own data!

- WLM Performance and Re-evaluating of Goals
  - Instructor: Peter Enrico and Scott Chapman
  - September 28 – October 2, 2015 – Columbus, Ohio, USA
- Parallel Sysplex and z/OS Performance Tuning
  (Web / Internet Based!)
  - Instructor: Peter Enrico and Scott Chapman
  - November 17 – 19, 2015 – Web Based
- Essential z/OS Performance Tuning Workshop
  - Instructors: Peter Enrico, Scott Chapman, Tom Beretvas
  - October 19 - 23, 2015 – Dallas, Texas, USA
- z/OS Capacity Planning and Performance Analysis
  - Instructor: Ray Wicks

EPS Sessions at Share

Peter Enrico

Day	Time	Location	Presentation
Wed	11:15	Asia 3	SMF 113 Processor Cache Counter Measurements – Overview, Update, and Usage
Wed	1:45	Asia 3	WLM – Effective Setup and Usage of WLM Report Classes
Thu	11:15	Asia 3	zProcessor Consumption Analysis (including z13), or What is Consuming All the CPU?

Scott Chapman

Day	Time	Location	Presentation
Tue	11:15	Asia 3	Memory Management in the TB Age
Tue	3:15	Southern Hemisphere 4	Lessons Learned from implementing an IDAA
Fri	11:15	Asia 3	WLM in One Page