Pinpointing Transient Performance Problems with SMF 98 & 99

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:

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

The SMF 98 and 99 records can provide performance data at a much finer time granularity than we normally get from the RMF and SMF data: down to 2 second intervals for some data! In this webinar, Scott Chapman will show how we can leverage these under-appreciated records to gain deeper insights into performance issues that may be relatively short-lived.

Agenda

- What do we mean by “Transient Performance Problems”
- Reminder about SMF sizes
- SMF 99 examples
- SMF 98 examples
- Preview of some new SMF 99 reports

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!

Transient Performance Problems

- Transient Performance Problems are those that are short lived: seconds to maybe a few minutes
- Can be difficult to impossible to diagnose with 15 minute RMF/SMF interval
  - Even 5 minute intervals may not catch them
- Real-time monitors can be a useful diagnostic, but you have to have history enabled and retained for a sufficient time period
- But we have some SMF records that let us gather data at much shorter intervals:
  - SMF 99 subtypes 1, 2, 3, 6: 10 seconds
  - SMF 99 subtype 12: 2 seconds
  - SMF 98: 5 to 60 seconds (site-determined, except for minutes 0, 15, 30, 45)

Example scenarios

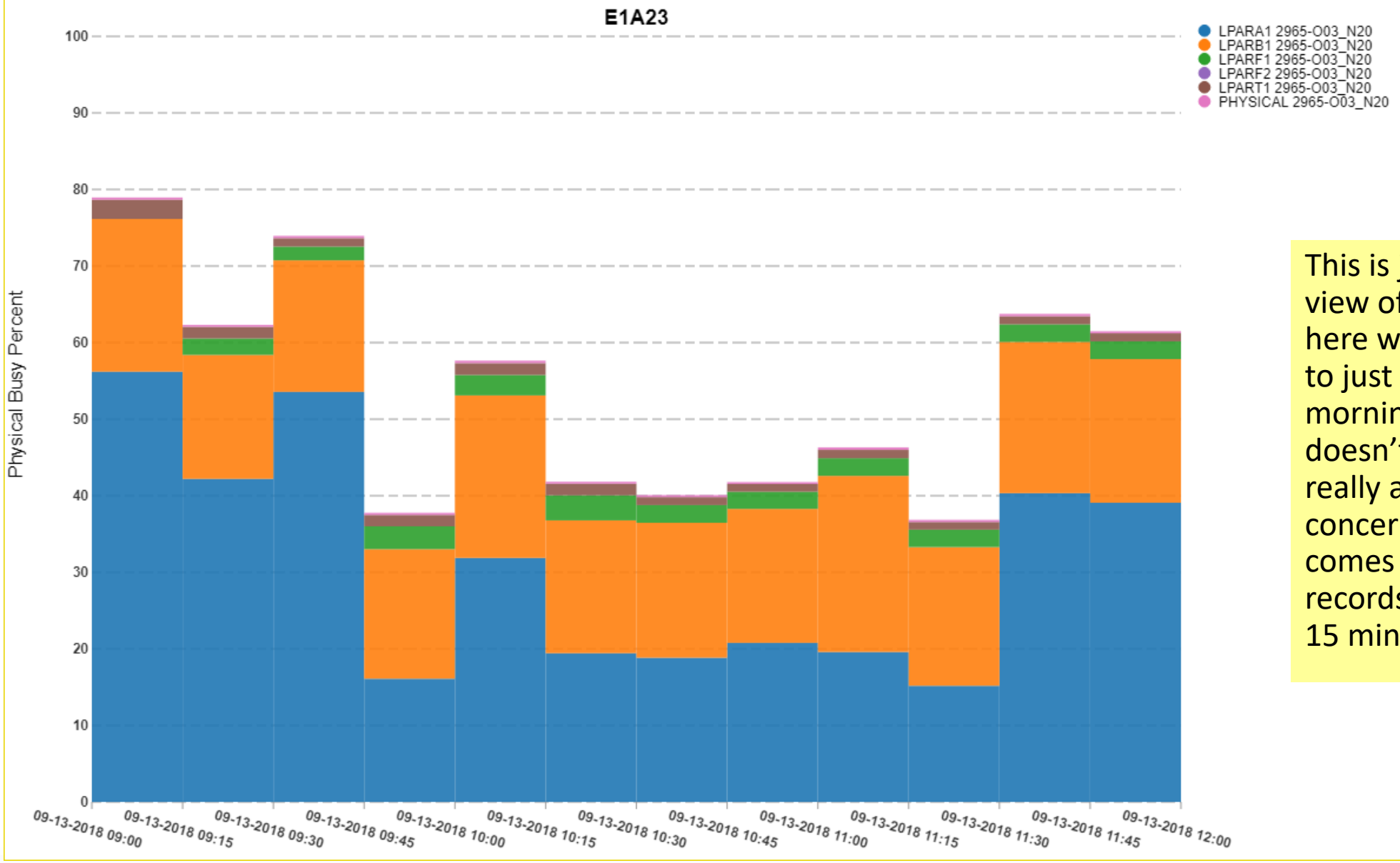
- Users are complaining of slow TSO response time, but
  - 15 minute averages RT look fine
- Application complaining about some slow DDF transactions, but
  - May have somewhat elevated zIIP or CPU delay sample counts, but the machine utilization for that interval looks fine

Reminder: volumes aren't bad

- SMF 98
  - 15 second interval = 100-150 MB / system / day
  - 5 second interval = 400-500MB / system / day
- SMF 99 minimum recommended subtypes:
  - 6, 10, 11, 12 ,14 = 50-100 MB / system / day
  - 8 = 10-30 MB / system / day
- SMF 99 additional recommended subtypes:
  - 1, 2, 3 = 1-1.5 GB / system / day
- Many customers are cutting a total of 10s to 100s of GBs / system / day
  - Even enabling the 99.2 isn't going to move the needle that much

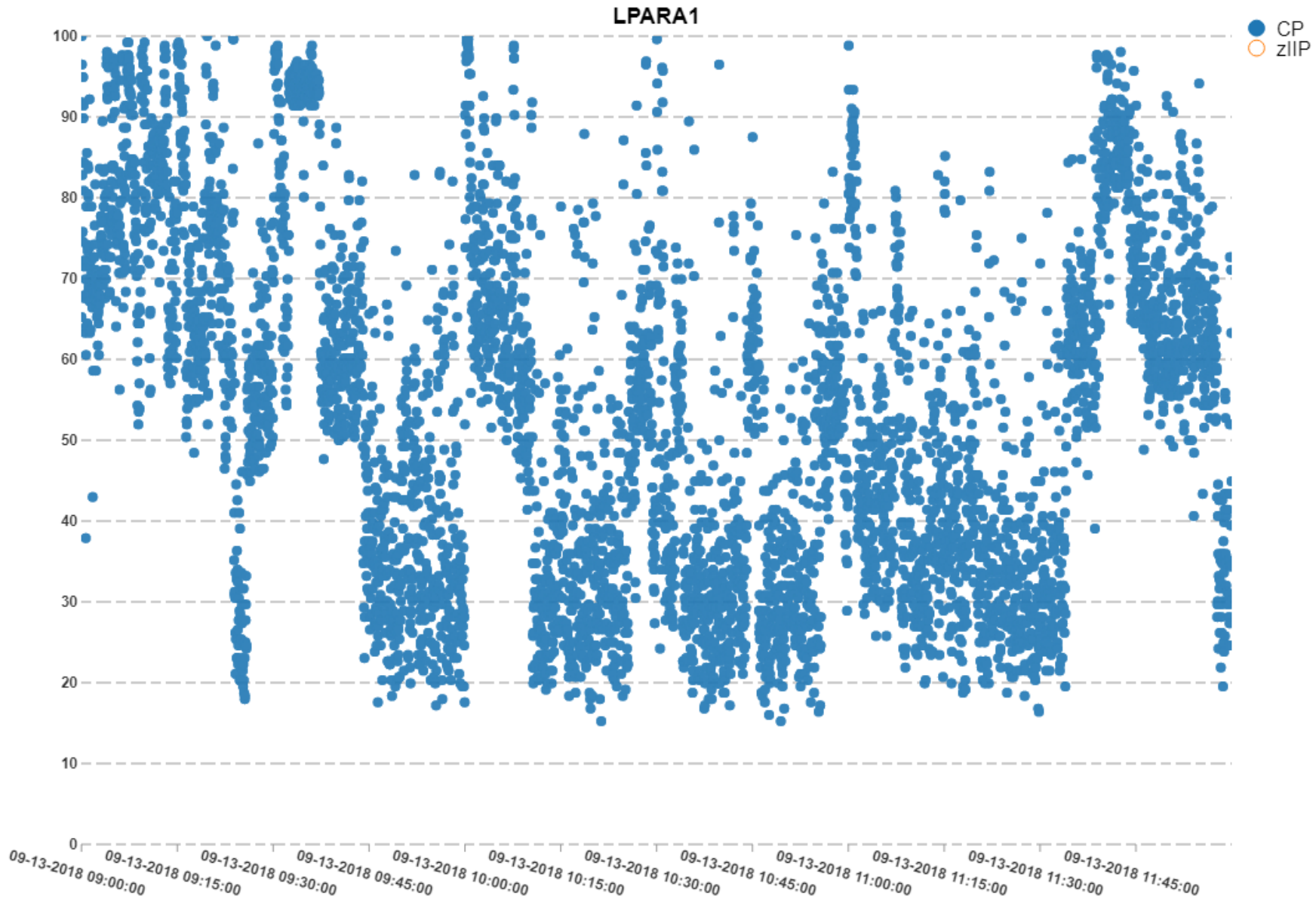
Examples of Using the 98s and 99s

CEC Physical Machine CP Busy% by CEC Serial Number

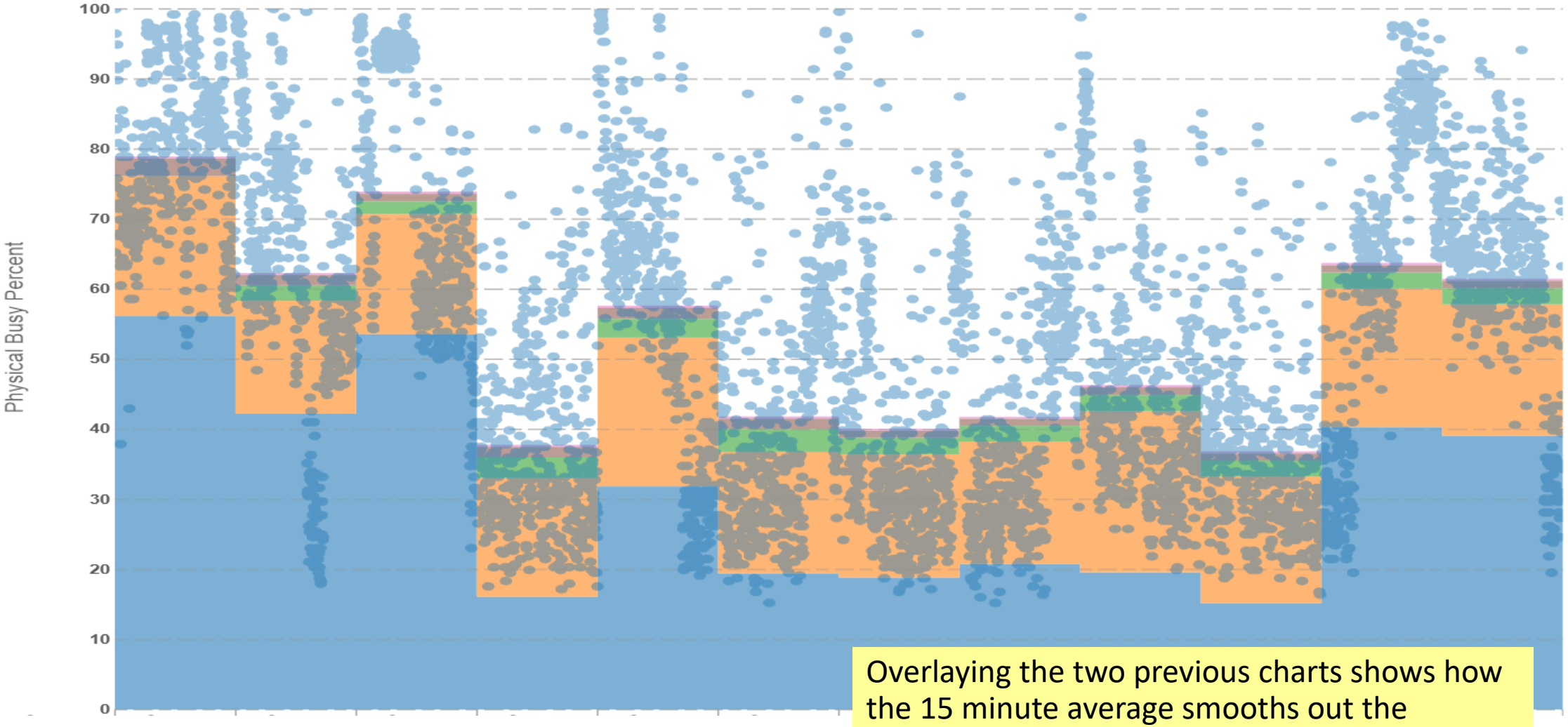


This is just a standard view of CEC Utilization, here we've narrowed in to just 3 hours in the morning, where it doesn't appear there's really any capacity concerns. This data comes from the SMF 70 records, in this case on 15 minute intervals.

HiperDispatch CEC Utilization

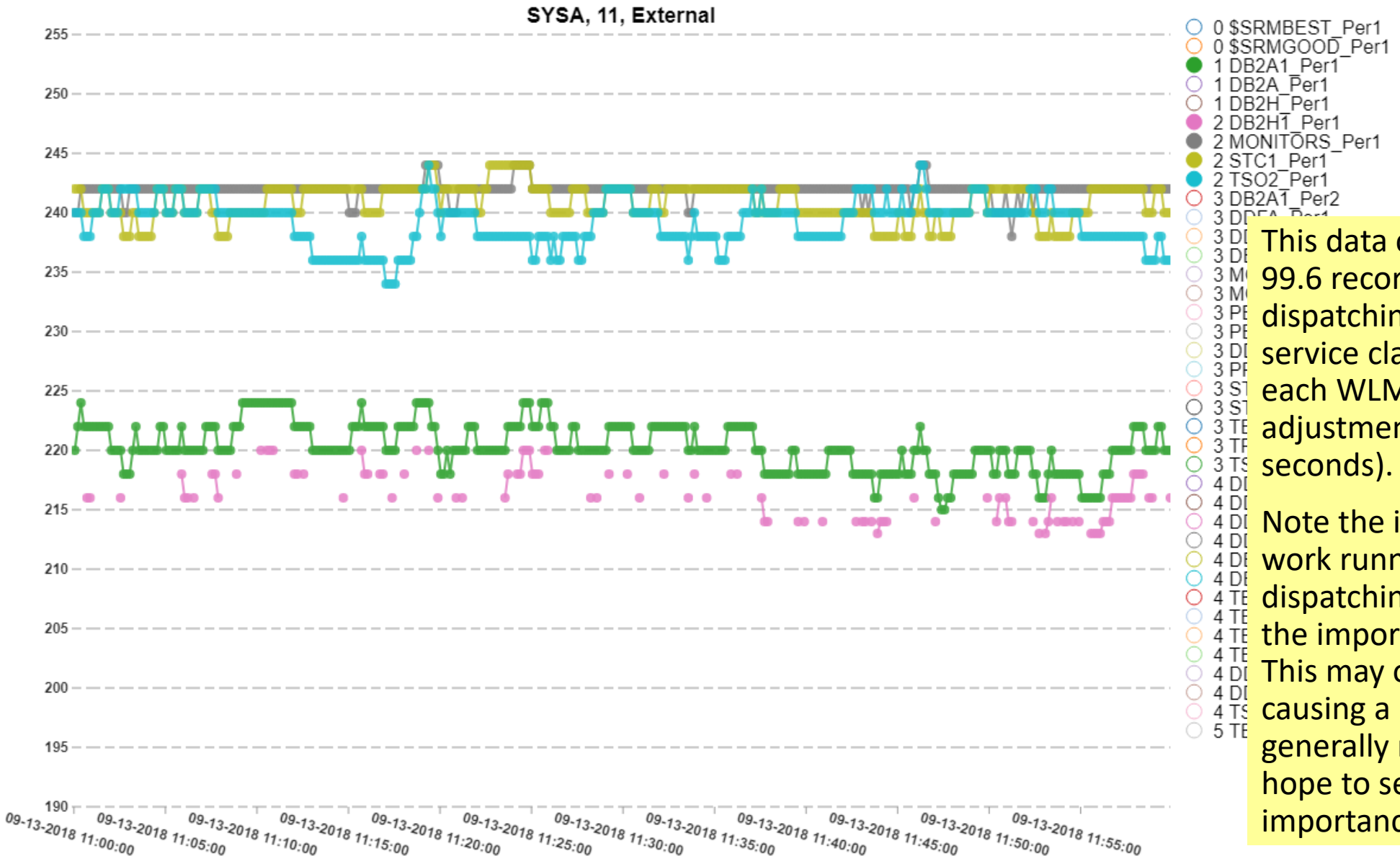


This data comes from the 99.12 HyperDispatch records and shows the CEC utilization at 2 second(!) intervals. Note that this tells a different story than the 15 minute RMF intervals.



Overlaying the two previous charts shows how the 15 minute average smooths out the utilization levels. There may have been performance issues due to running at or nearly 100% busy at times, but that's really lost in the 15 minute averages.

WLM SMF 99.6 - CPU Dispatching Priority

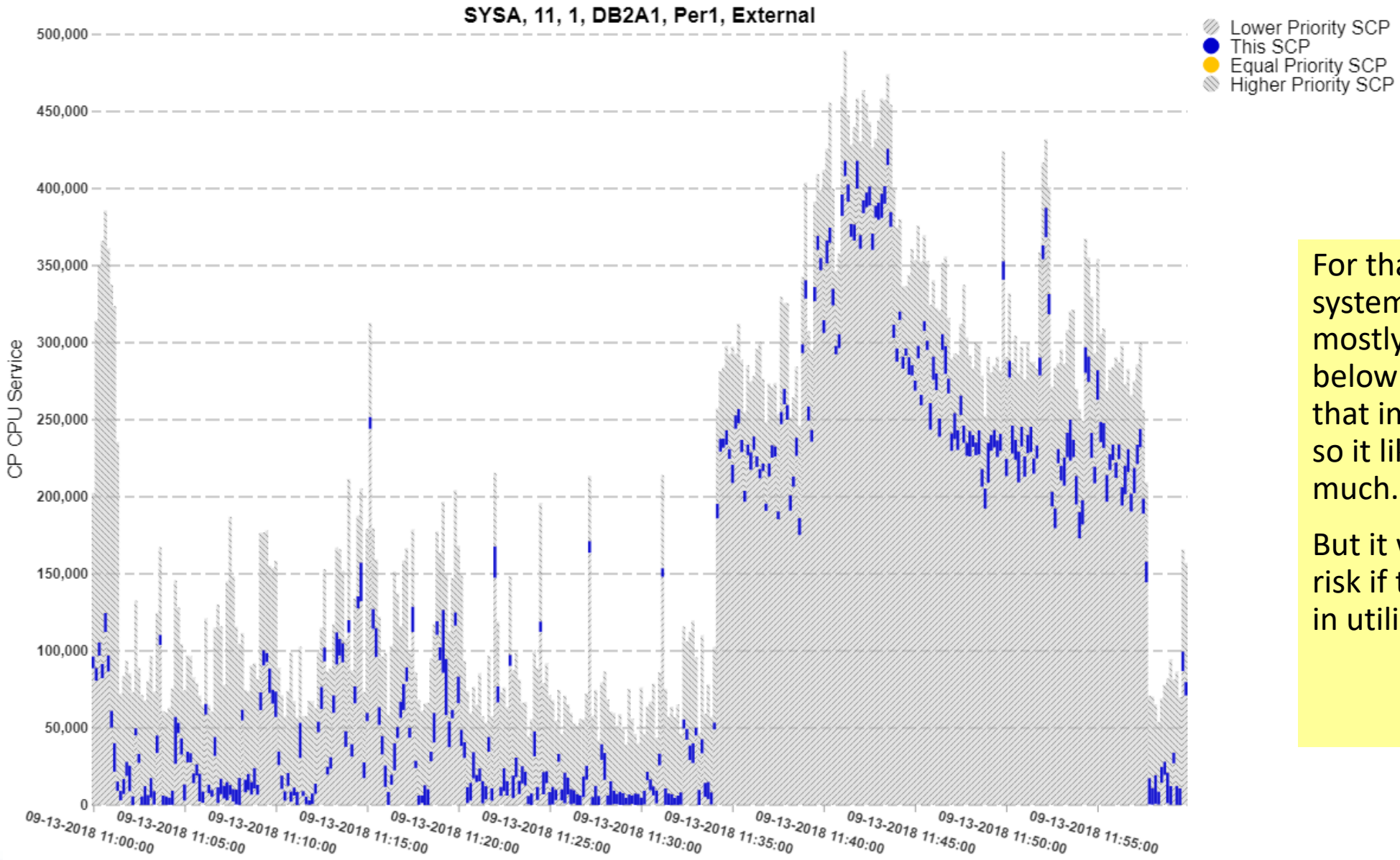


This data comes from the 99.6 records and shows the dispatching priority of each service class period for each WLM policy adjustment interval (10 seconds).

Note the importance 1 work running at a lower dispatching priority than the importance 2 work. This may or may not be causing a problem, but is generally not what we hope to see at these importance levels.

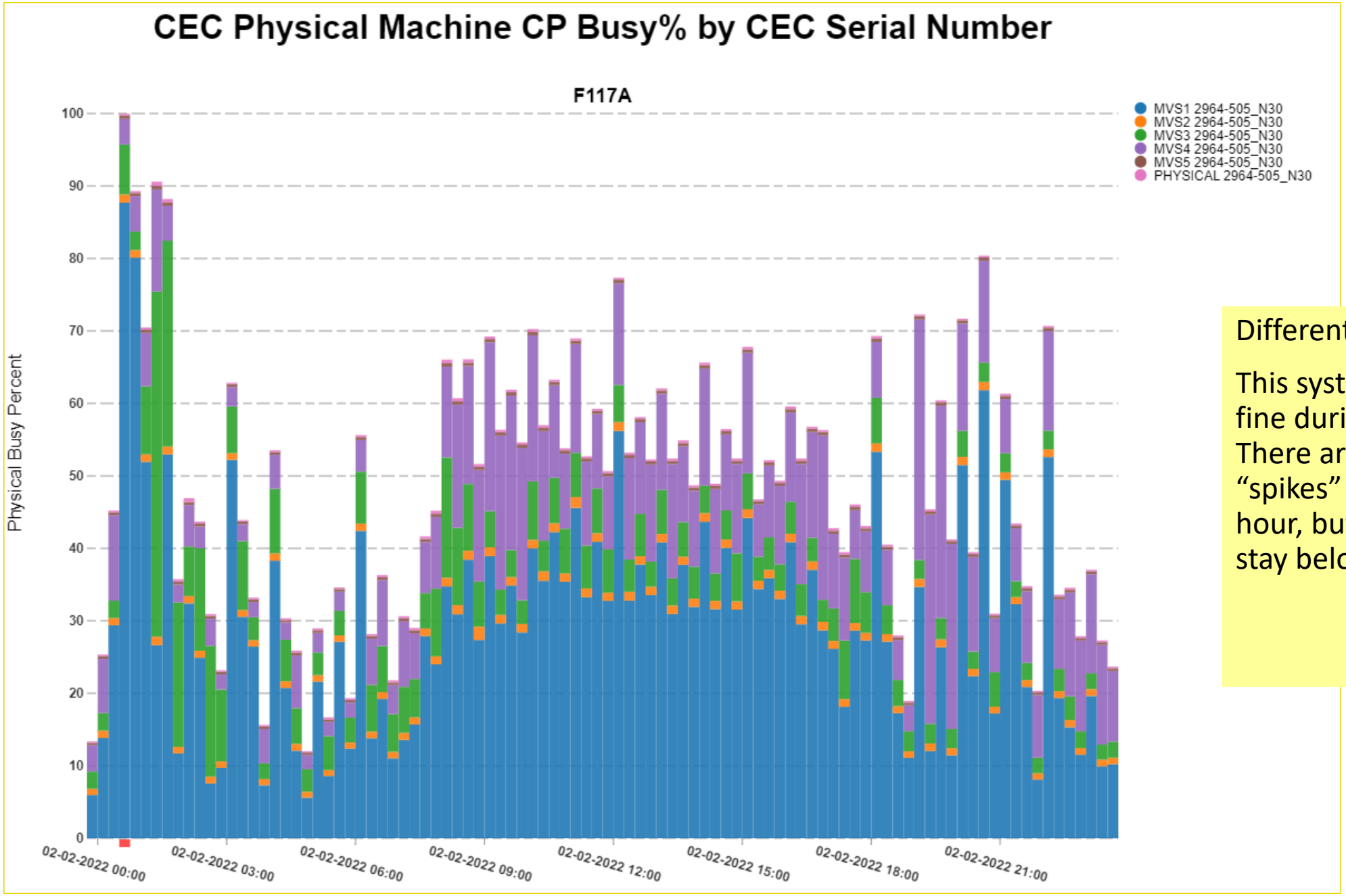
CP CPU Service Accumulated Above / Below SCP

From SMF 99.6



For that hour, when the system got busy, it was mostly work running below the priority of that importance 1 work, so it likely didn't suffer much.

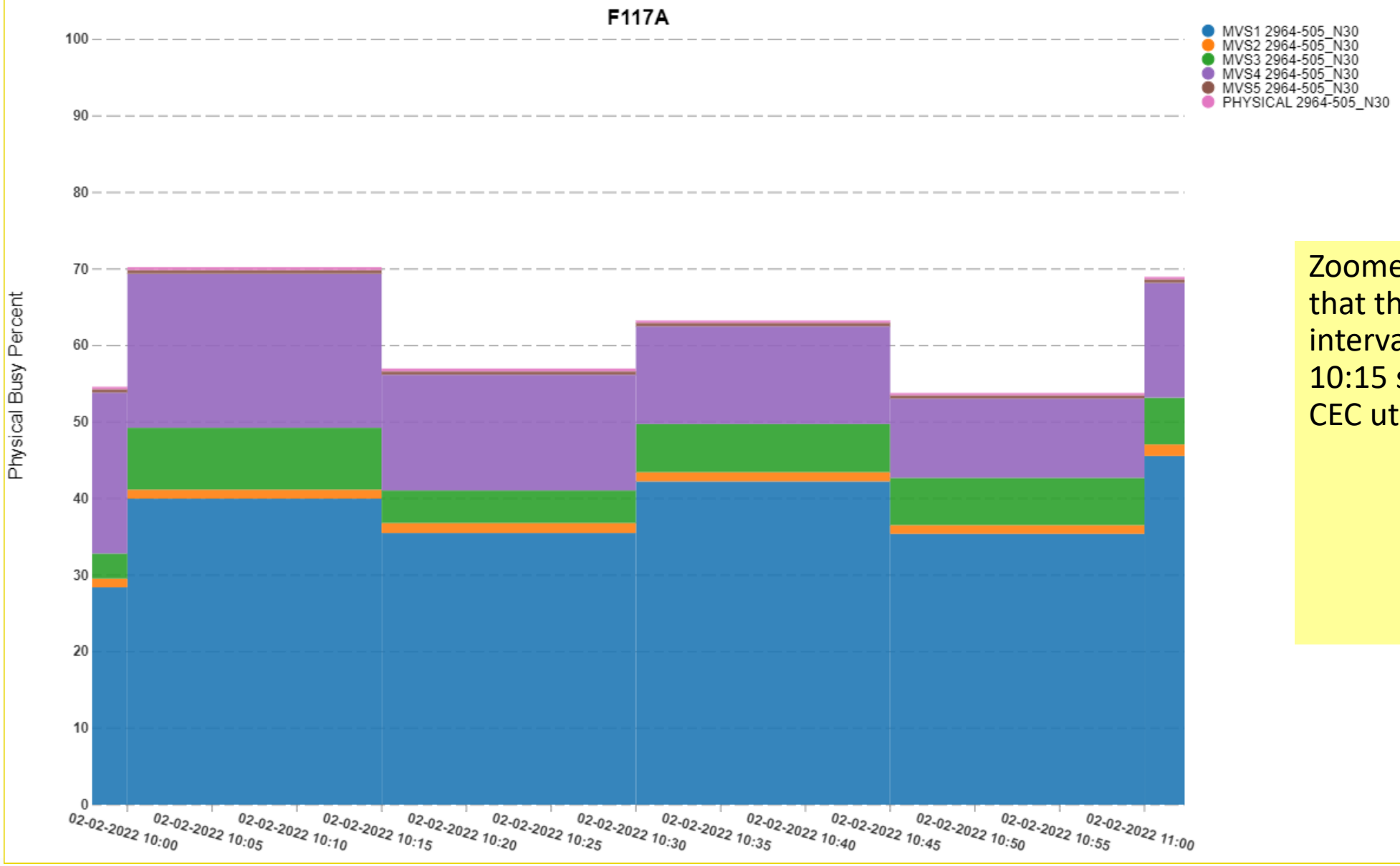
But it was definitely at risk if there was a spike in utilization.



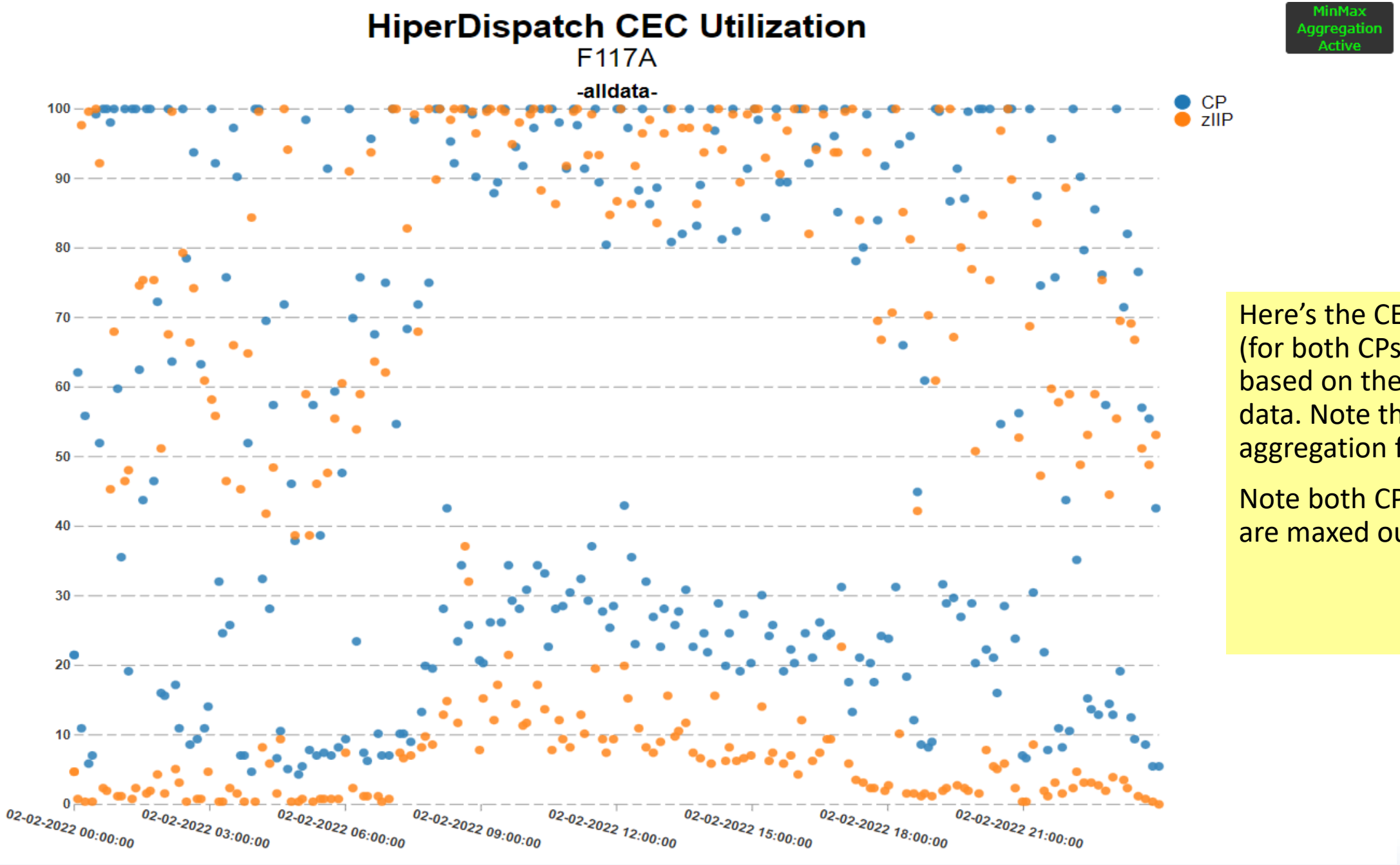
Different example:

This system looks quite fine during the day. There are regular "spikes" at the top of the hour, but they mostly stay below 70%.

CEC Physical Machine CP Busy% by CEC Serial Number

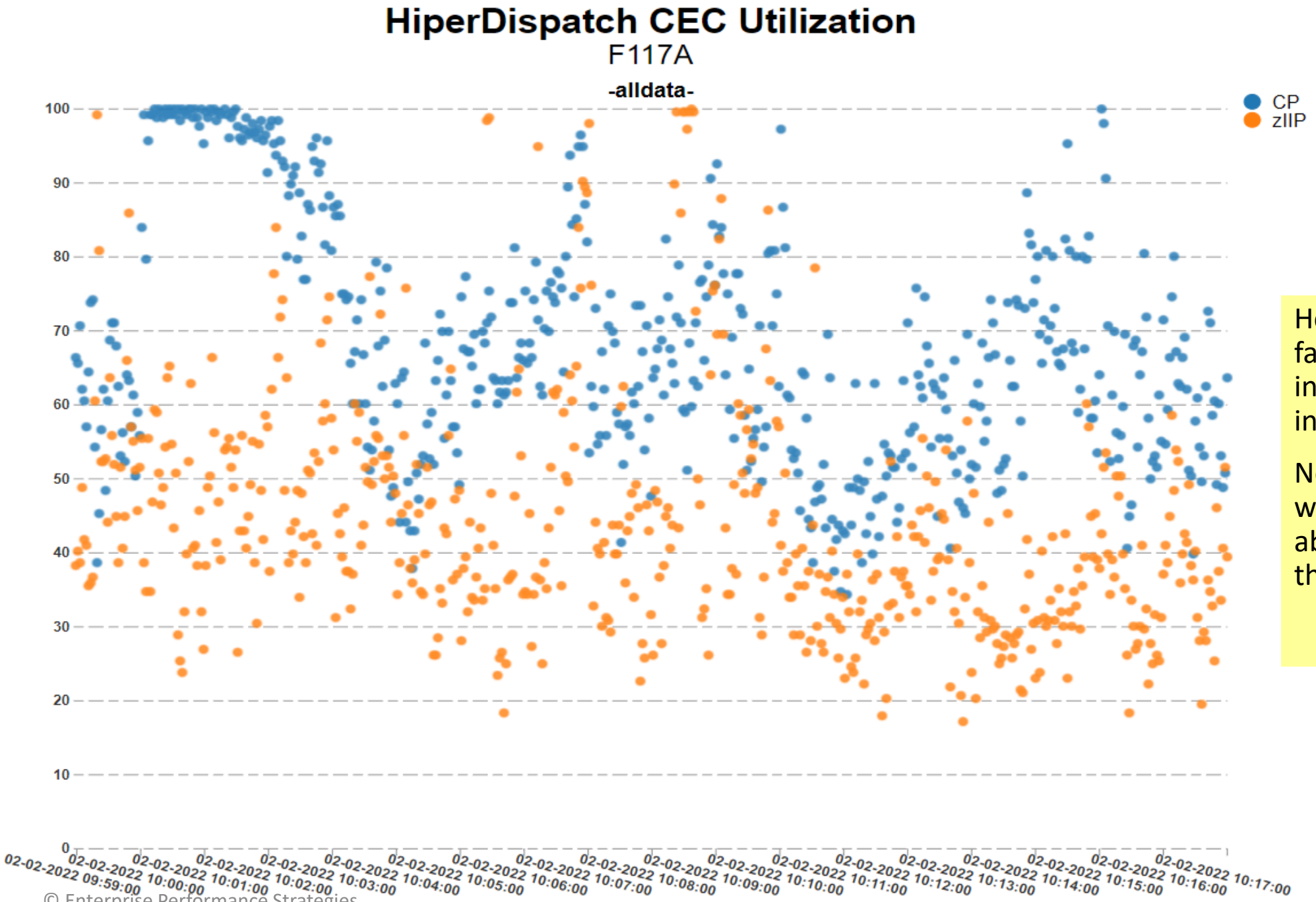


Zoomed in to better see that the 15 minute interval from 10:00 to 10:15 shows an overall CEC utilization of 70%.



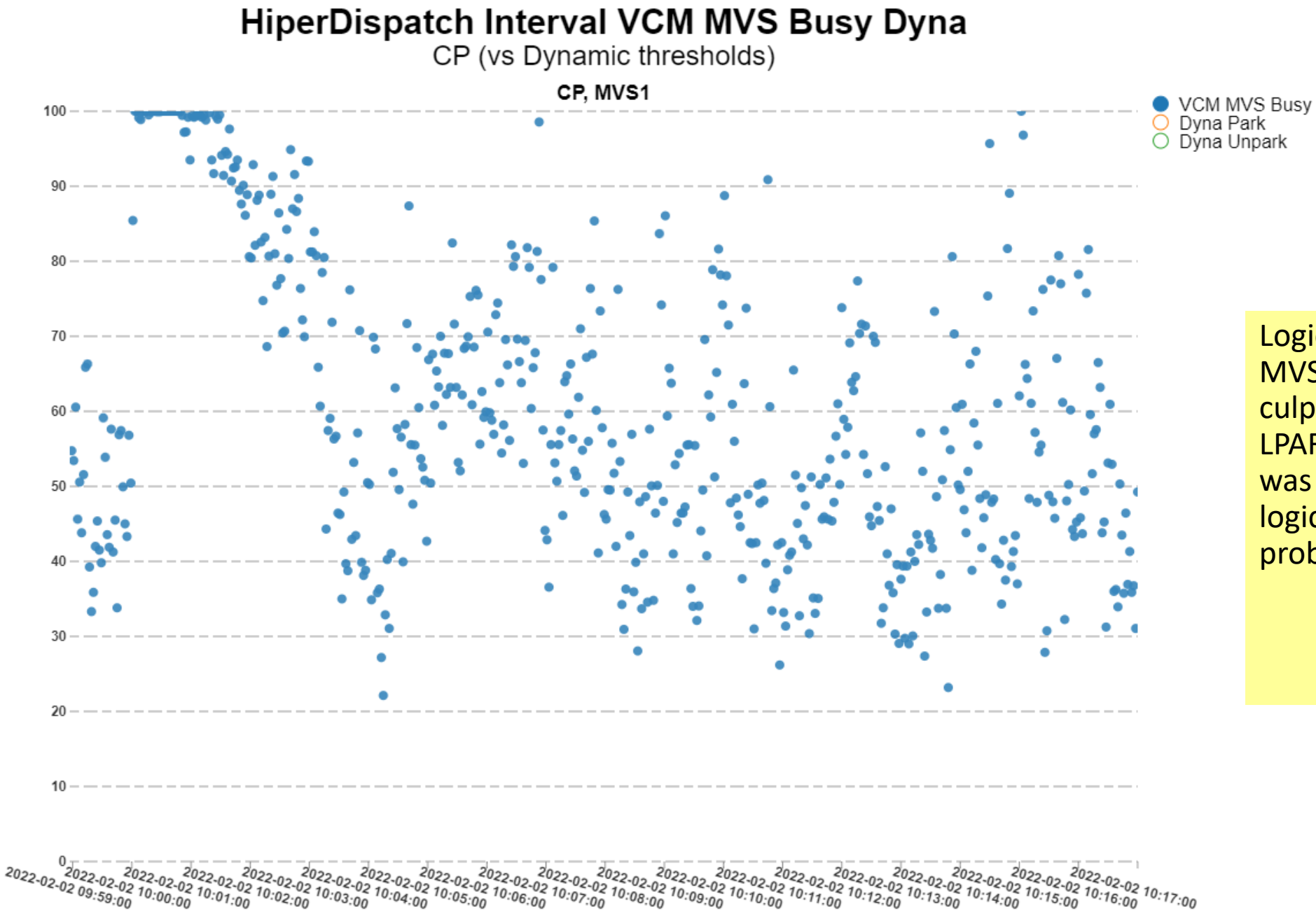
Here's the CEC utilization (for both CPs and zIIPs) based on the 2 second data. Note the MinMax aggregation flag.

Note both CPs and zIIPs are maxed out at times!

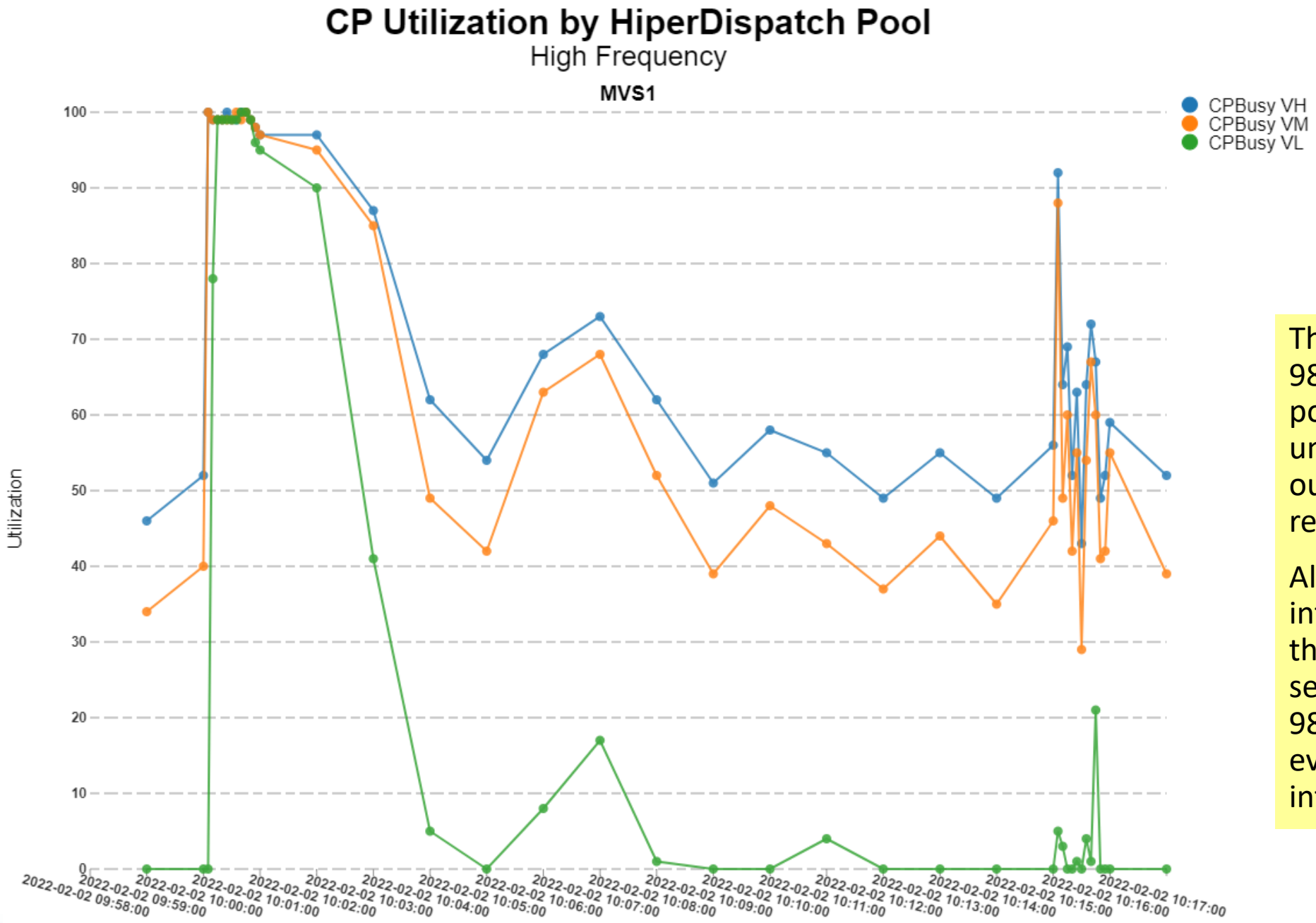


Here we've zoomed in far enough to see the individual 2 second intervals.

Note that the machine was >=95% busy for about 2 minutes right at the top of the hour!



Logical MVS busy shows MVS1 was probably our culprit: it's the major LPAR on the system and was running at 100% logical busy during the problem time.

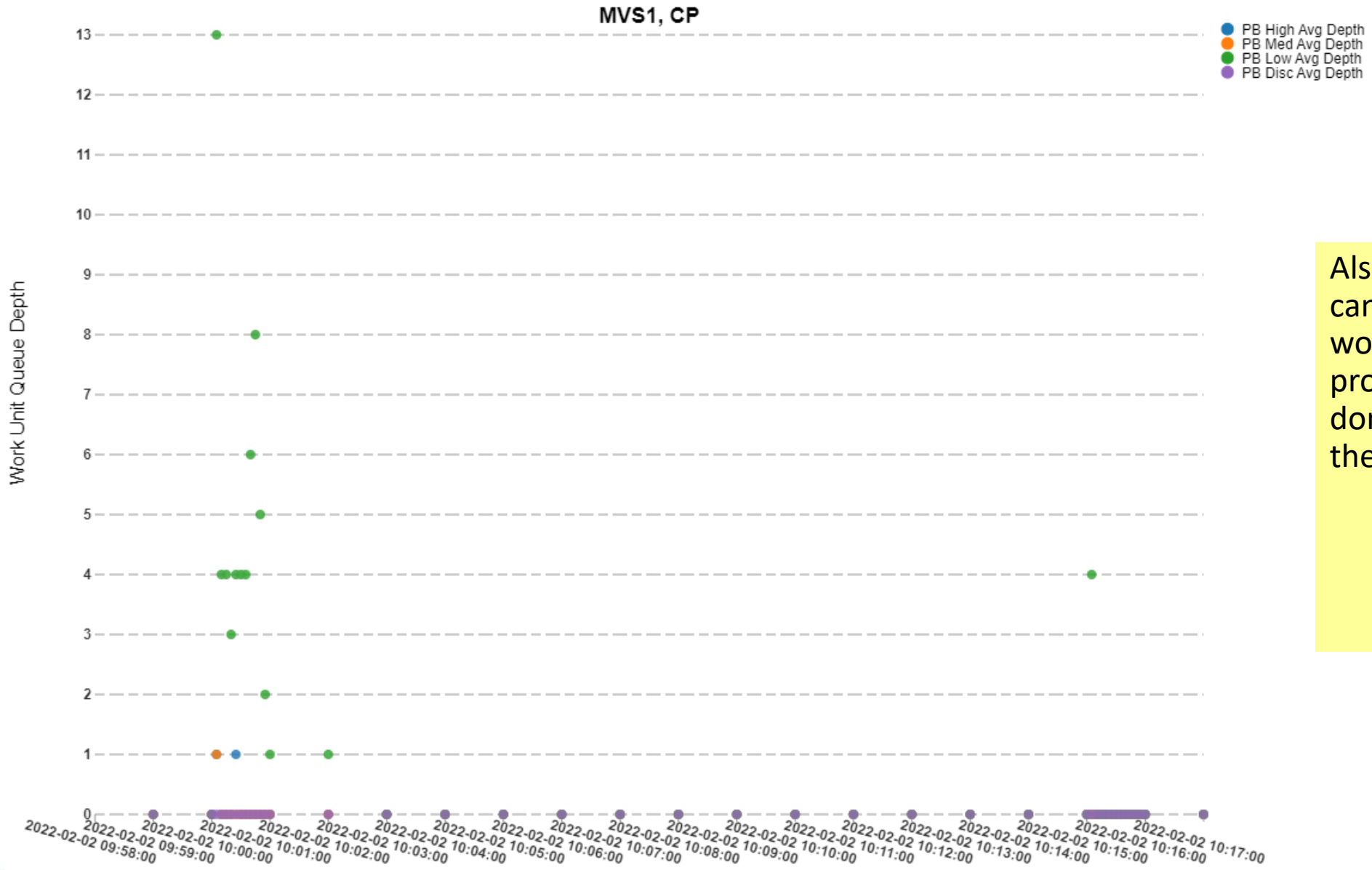


This chart from the SMF 98s shows how the low pool processors got unparked to try to help out MVS1 when it got really busy.

Also note the variable interval width because this site was using 60 second intervals for SMF 98, but 4 minutes of every hour get 5 second intervals regardless.

Avg Work Unit Queue Depth by Priority Bucket

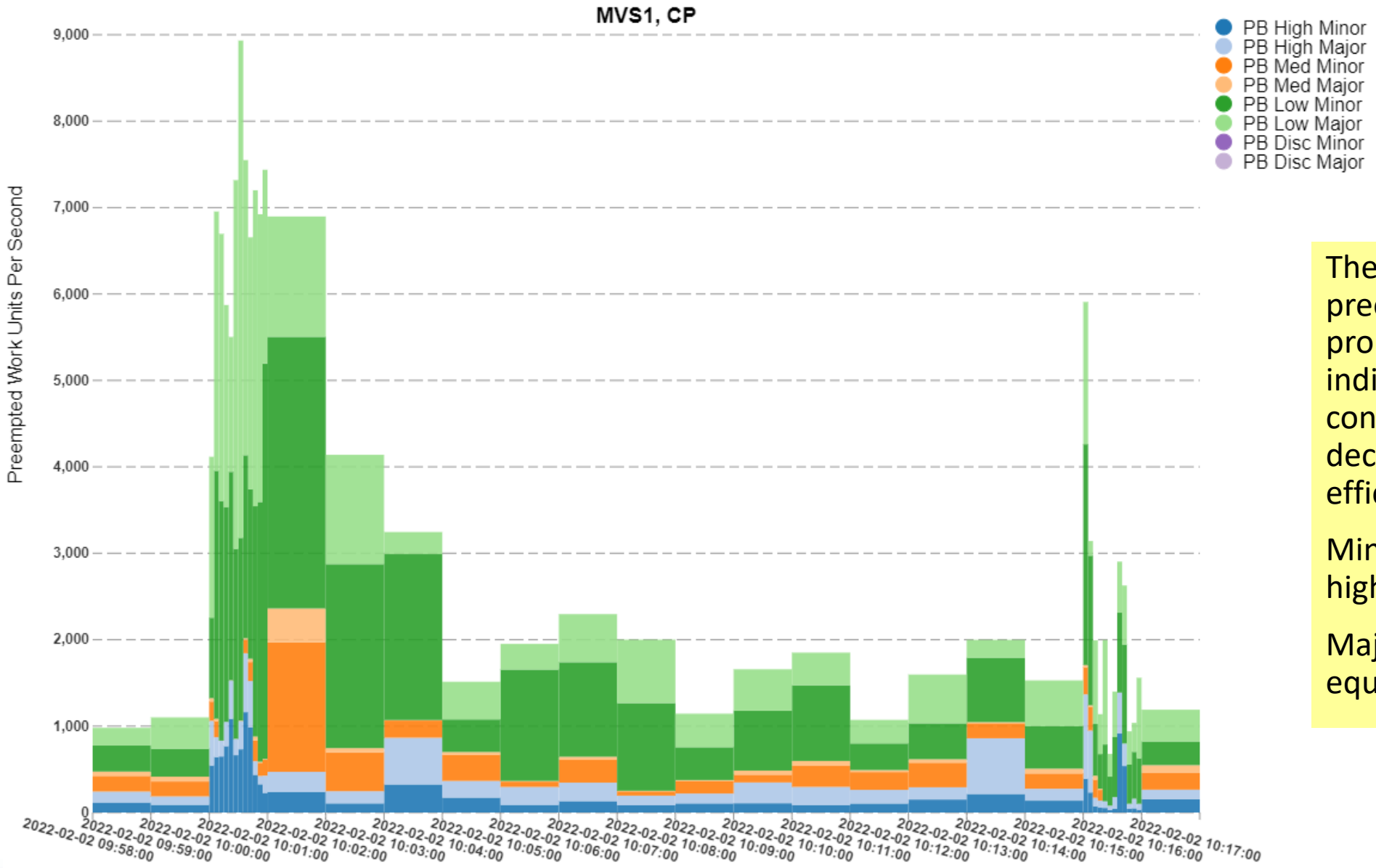
High Frequency



Also from the 98s, we can see that we have work queues during the problem period that we don't have outside of the problem period.

Preemptions per Second by Time Slice

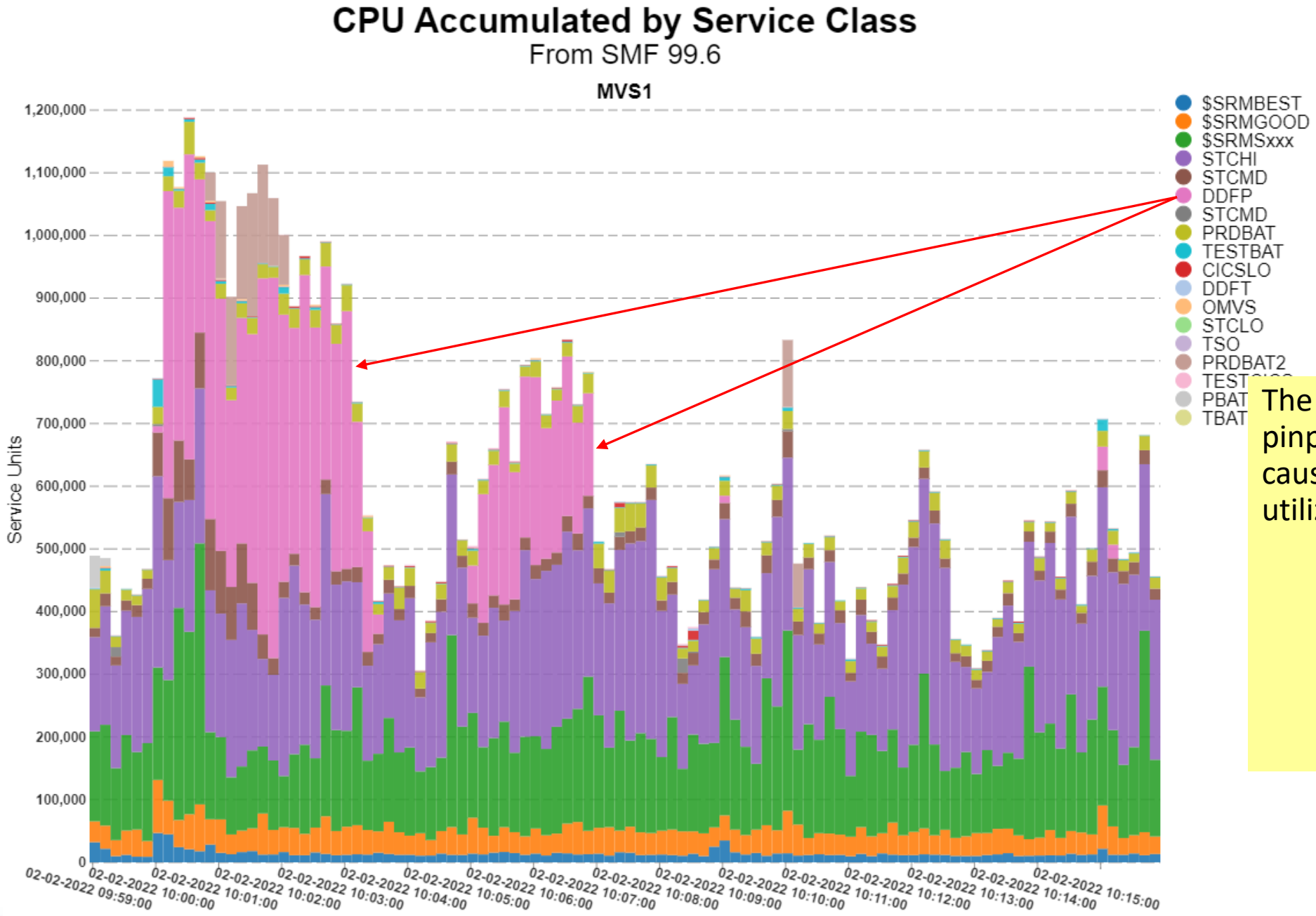
High Frequency



The increase in preemptions during the problem time is another indicator of CPU contention. (And a decrease in overall CPU efficiency.)

Minor = preempted by higher-priority work

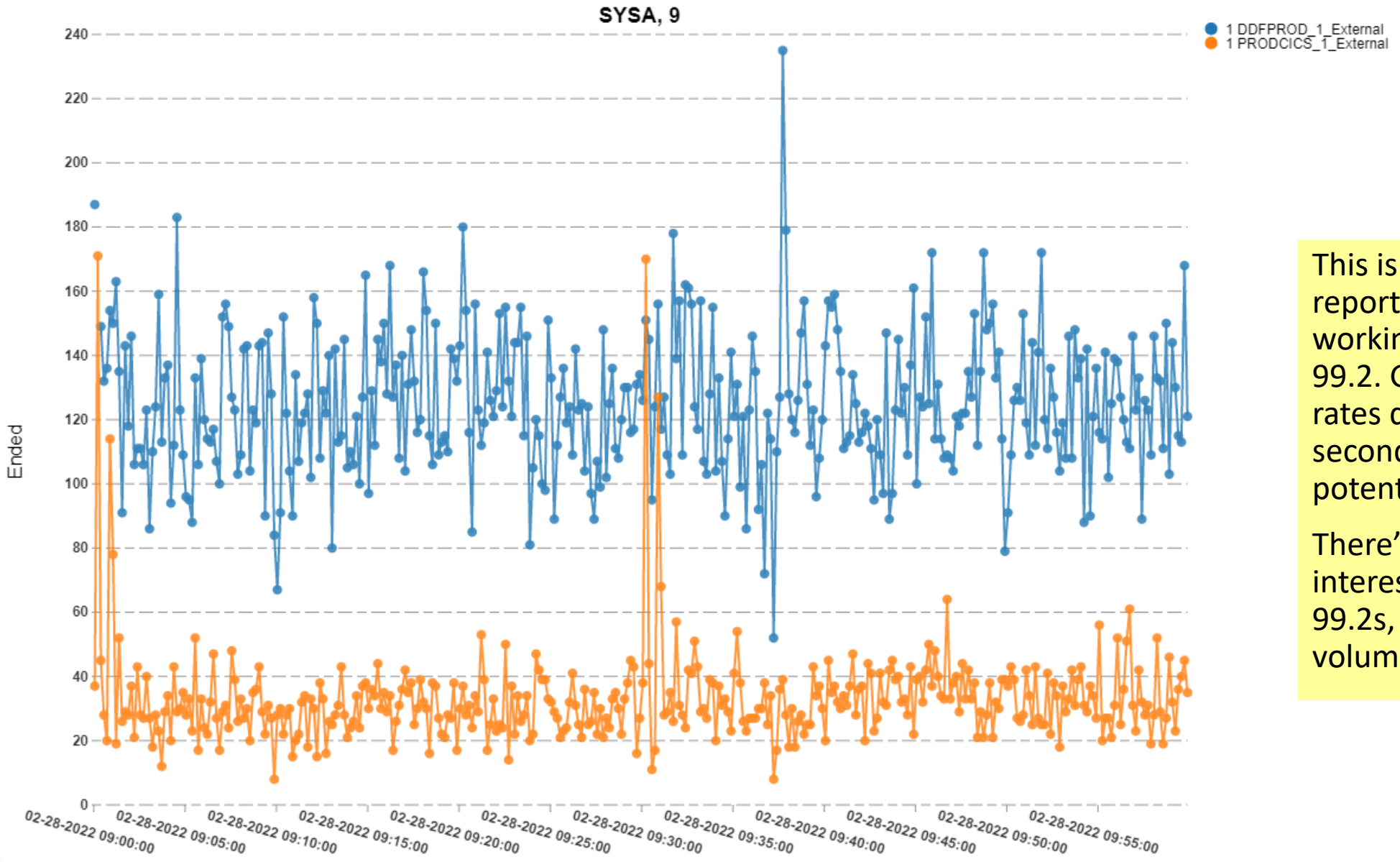
Major = preempted by equal-priority work



The 99.6 can also help us pinpoint what work was causing that increased utilization.

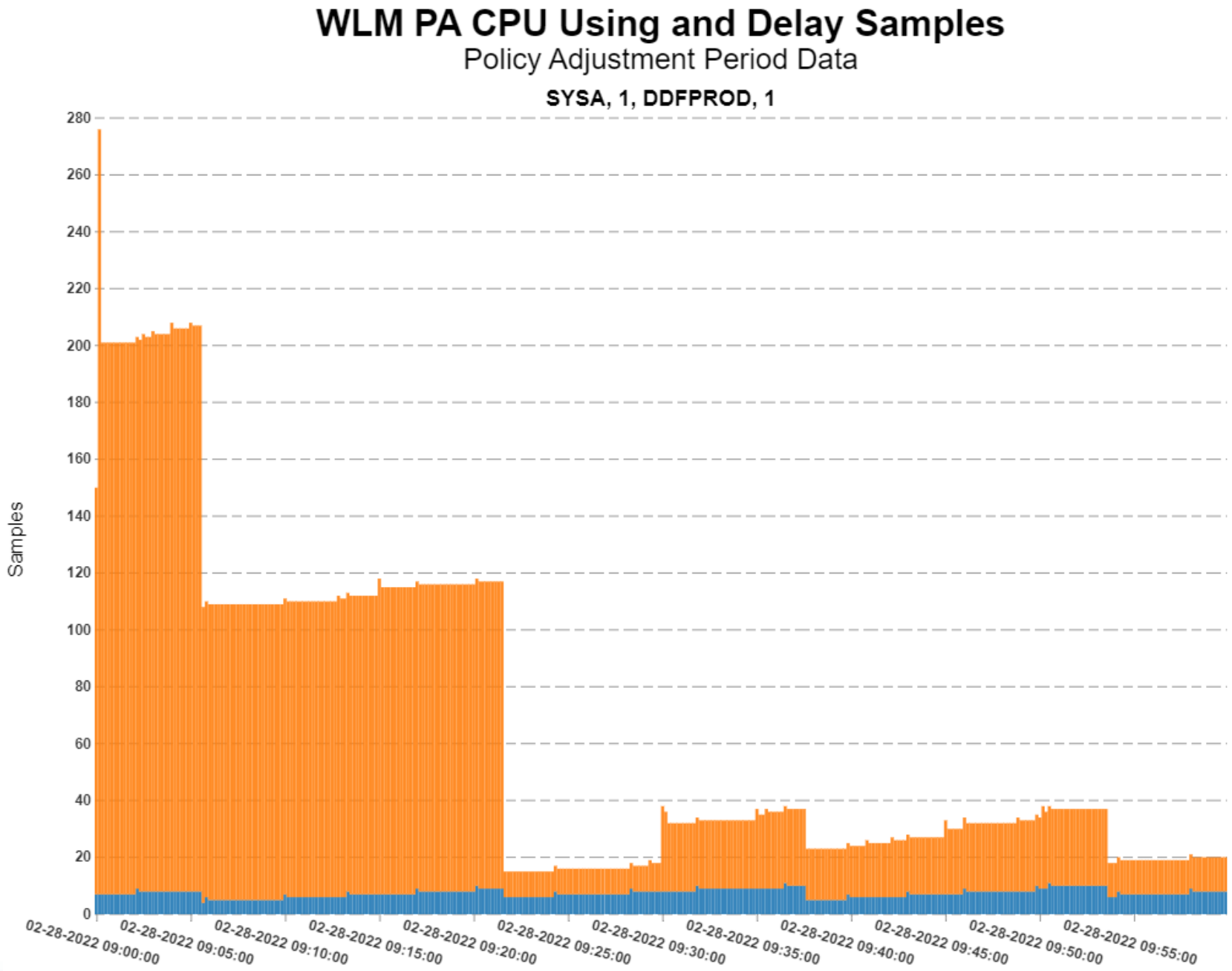
WLM PA - RT Goal Ext Per Ended Transactions Last PA Interval

Policy Adjustment Period Data



This is a preview of some reporting Peter has been working on out of the 99.2. Getting transaction rates down to a 10 second interval sounds potentially useful.

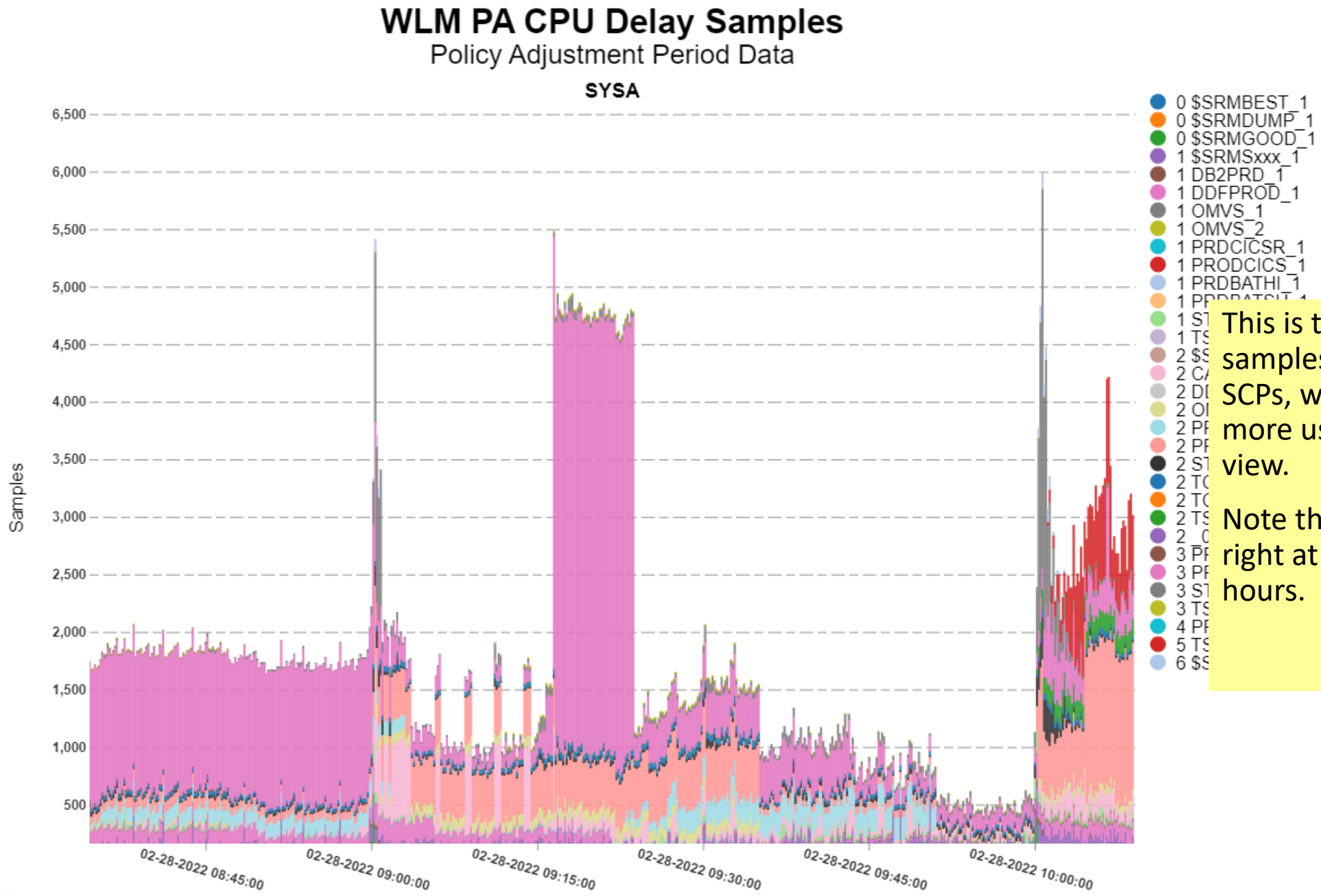
There's a lot of interesting data in the 99.2s, but they are quite voluminous.



Another preview: CPU using and delay samples at a 10 second interval basis from the 99.8.

These records are much less voluminous than the 99.2 but still have some interesting data such as this.

The large ratio of delays relative to using is interesting here and may be indicative of a flood of work.



This is the 99.8 CPU delay samples but across all SCPs, which is perhaps a more useful “at a glance” view.

Note the contention spikes right at the top of the hours.

Summary

- In today's storage abundant world, recording the 98s and 99s make sense
- There's a lot of interesting data available to intervals of just a few seconds
- If you're a Pivotor customer, be sure to send this data
  - And look for some new transient performance reports in the coming weeks
- With these records it should be rare to need to use very short RMF intervals

Thanks!
Questions?

Email: scott.chapman@epstrategies.com