How Busy Is That Mainframe, Really?

Scott Chapman

Enterprise Performance Strategies, Inc.

Scott.chapman@EPStrategies.com

SHARE Webinar
November 2025

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:

Abstract (why you're here!)

It's 9:03 AM, do you know how busy your mainframe is? No points for know how busy it was on average for the 15 minutes starting at 9:00! Your 15-minute performance intervals may be sufficient for many use cases, but are likely insufficient for discovering and diagnosing performance problems that last for just a few minutes or less. In this session, Scott Chapman will discuss the SMF 98 and 99 data that lets you see CPU usage on a sub-minute basis as well as the practical utilization of such data.

Agenda

- Introduction
- Do we have a problem?
- Exploring transient performance measures
- Do we care?
- Best practice recommendations

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
    https://www.pivotor.com/content.html
    https://www.pivotor.com/webinar.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!

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

Like what you hear?

- Free z/OS Performance Educational webinars!
  - Examples from our Summer / Fall 2025 webinars:
    - *Overseeing z/OS Performance Management With Your Outsourcer*
    - *Back to basics - Processor Consumption Analysis*
    - *Pivotor Pointers*
    - *WLM and CPU Critical Control*
    - *Back to Basics - Evaluating Latent Demand*
    - *Understanding SMF 98 Locking Measurements*
    - *Standard Measurements when Monitoring Transactions*
    - *Processor Comparison Discussion*
    - *z/OS Performance Management in an AI World*
    - *Understanding z/Architecture Processor Topologies*
    - *SMF 99 WLM Decision Making Traces*
    - *Understanding SMF 98 Address Space Consumption Measurements*

email: contact@epstrategies.com

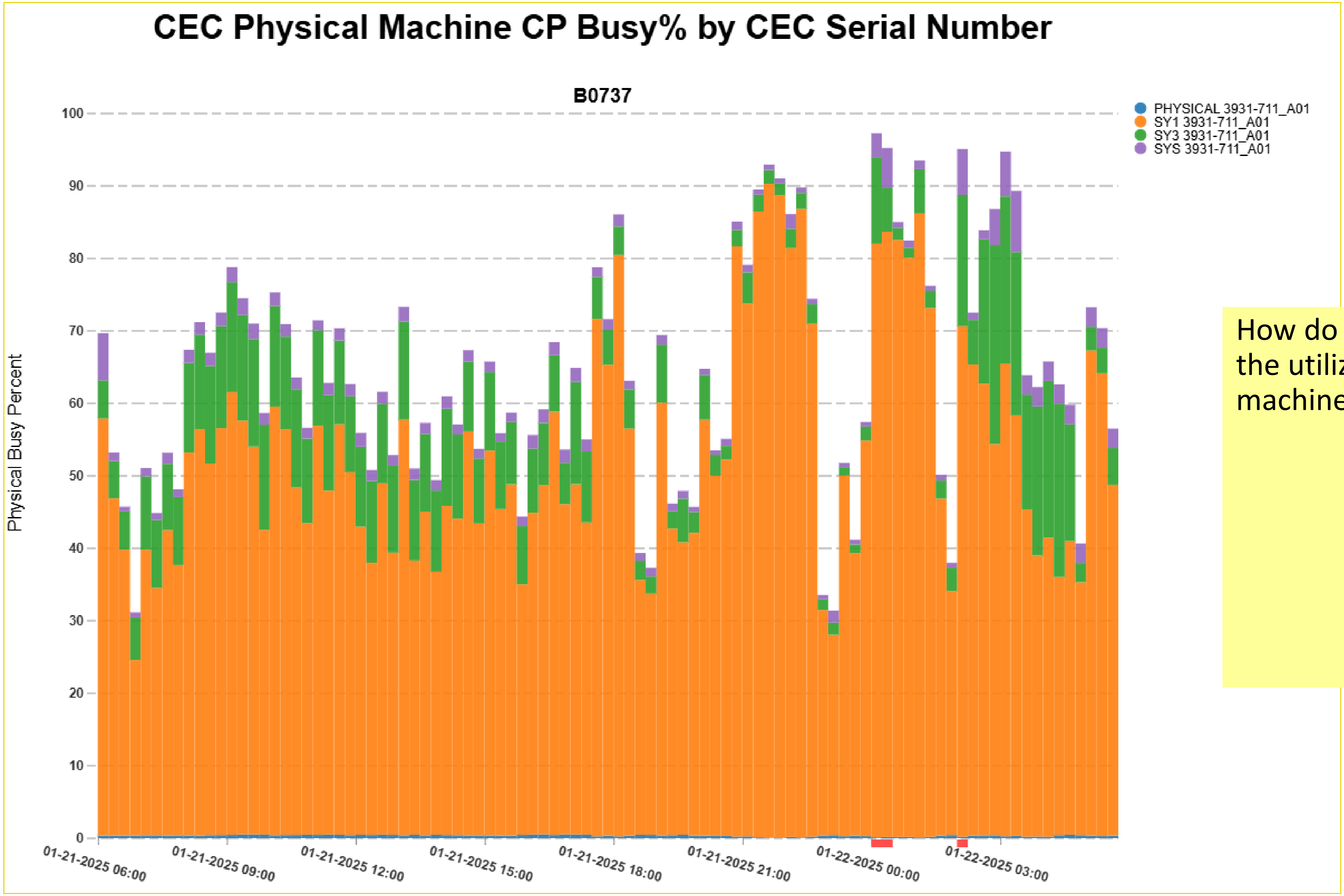
- If you want a free cursory review of your environment, let us know!
  - We’re always happy to process a day’s worth of data and show you the results
  - See also: http://pivotor.com/cursoryReview.html

z/OS Performance workshops available

During these workshops you will be analyzing your own data!

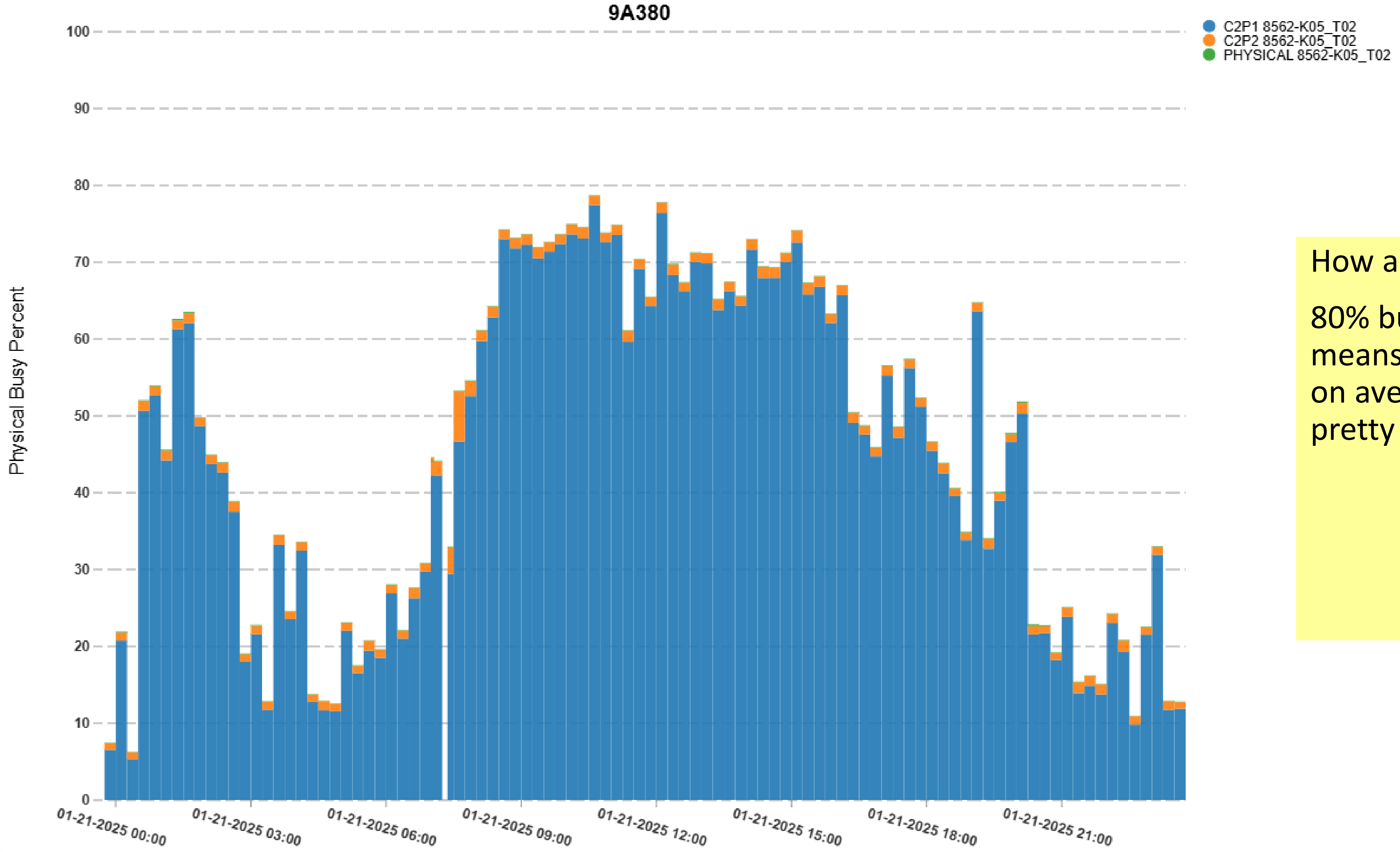
- WLM Performance and Re-evaluating Goals
  - TBD, 2026 (4 days, M, Tu, Th, F)
- Parallel Sysplex and z/OS Performance Tuning
  - July 15-16, 2026 (2 days)
- Essential z/OS Performance Tuning
  - March 30 – April 3, 2026 (4 days, M, Tu, Th, F)
- Also… please make sure you are signed up for our free monthly z/OS educational webinars! (email contact@epstrategies.com)

Do we have a problem?



How do you feel about the utilization of this machine?

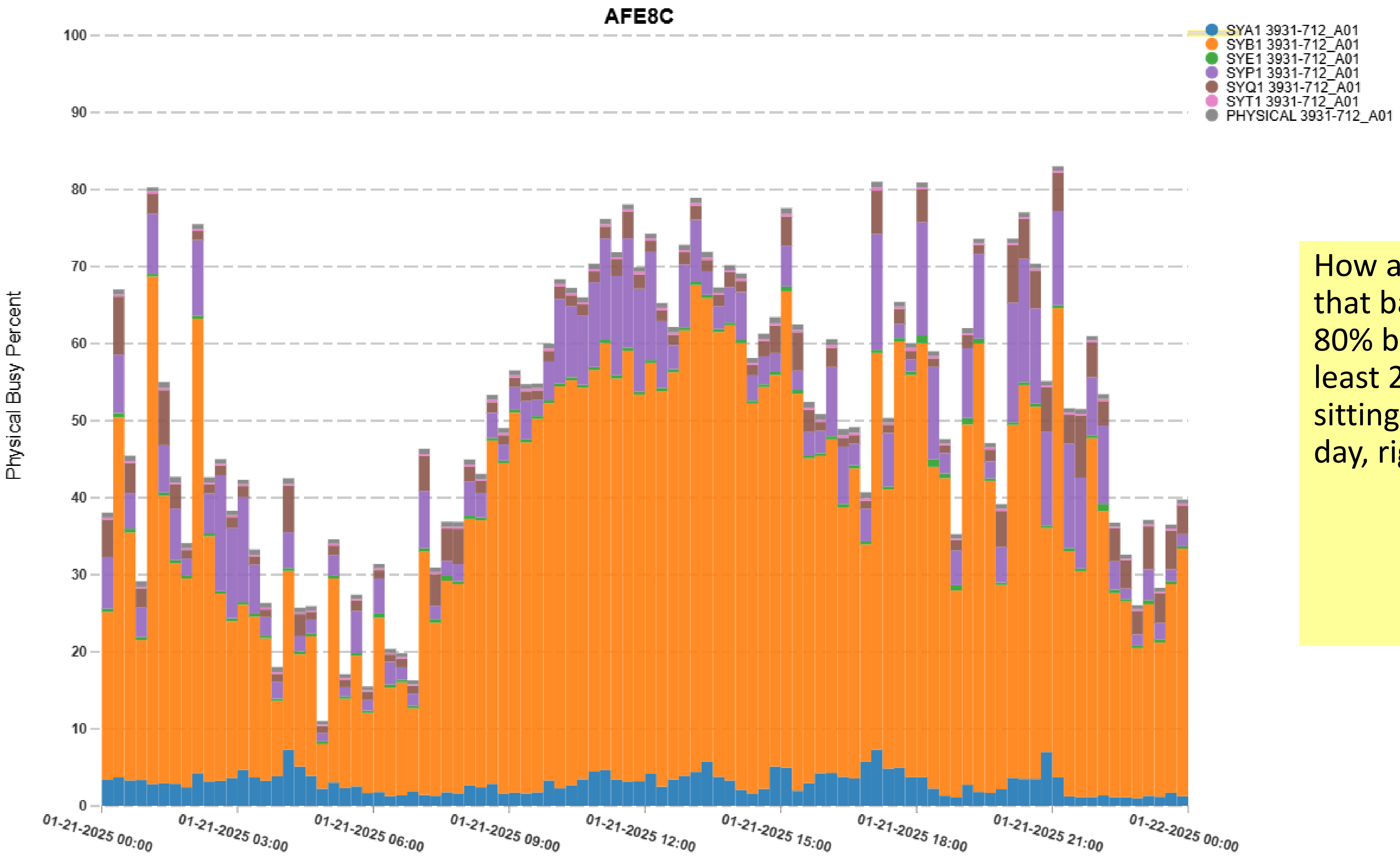
CEC Physical Machine CP Busy% by CEC Serial Number



How about this one?

80% busy of a 5 way means 1 CPU is unused on average, so surely is pretty good.

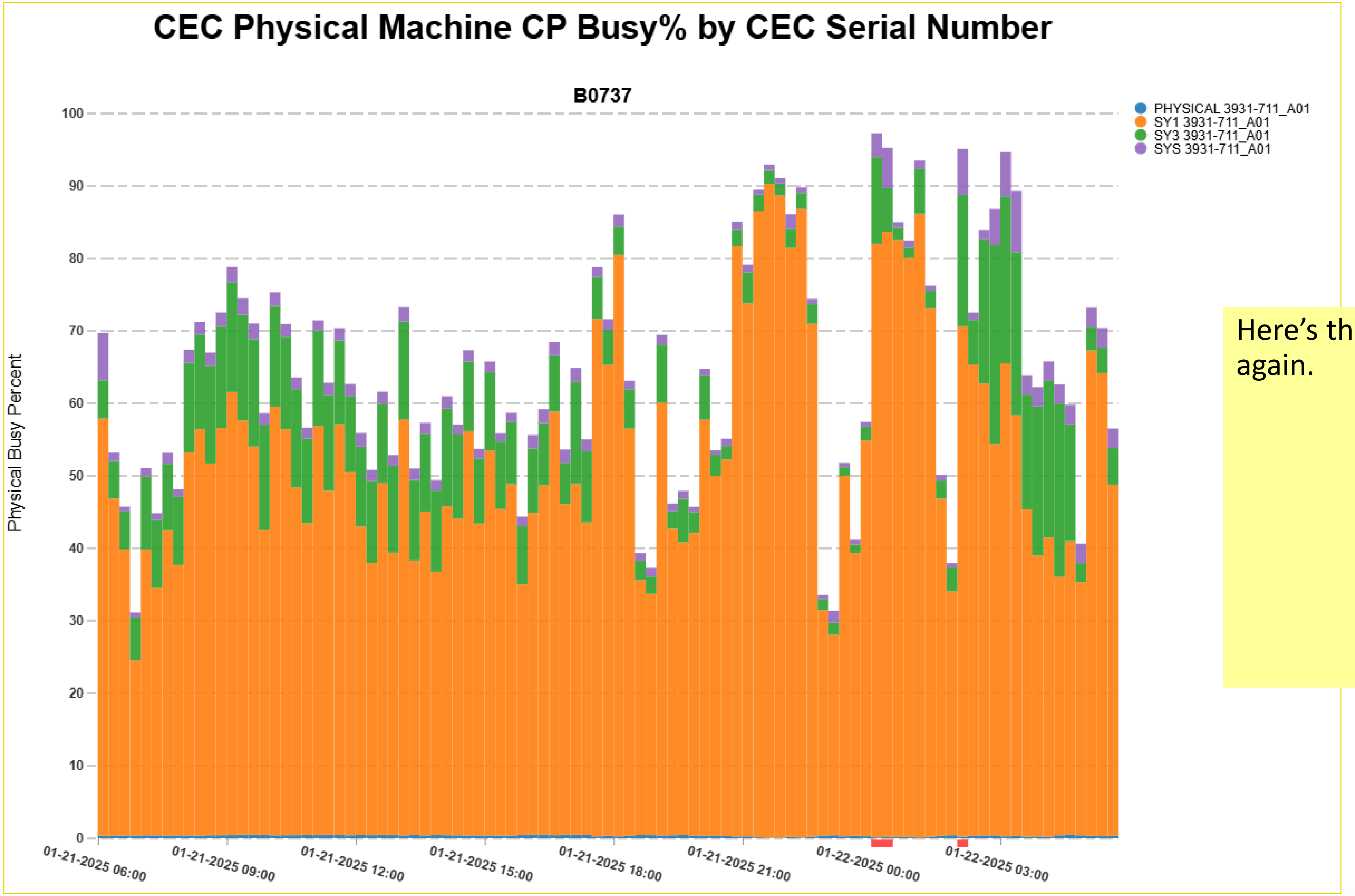
CEC Physical Machine CP Busy% by CEC Serial Number



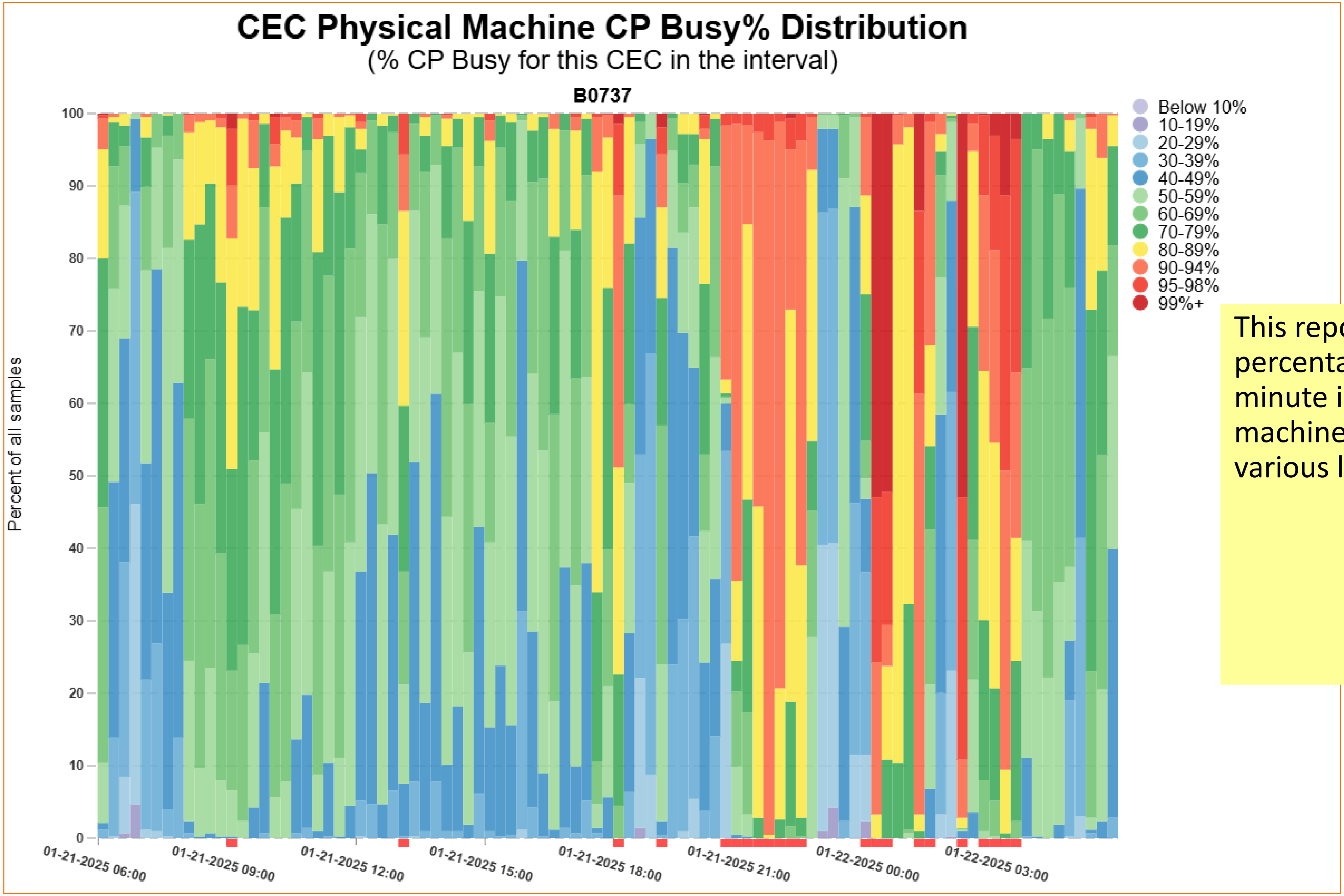
How about a 12-way that barely breaks past 80% busy? It's like at least 2 engines are sitting there unused all day, right?

How busy is that machine?

- Probably most z/OS performance/capacity people would say those machines are fine from a capacity perspective
  - And they probably (mostly) are
- Long-standing mantra of “you can run the mainframe 100% busy unlike those other servers that run at 10% busy”
  - Usually leaving off the part about assuming you have a mixed workload including some delayable work
  - Not everything that can be done should be done
  - LSPR testing done at around 90% busy
  - (But yes, in general, MF can run at much higher utilizations compared to others)
- Can you have performance problems despite only being 80% busy?



Here's that first machine again.

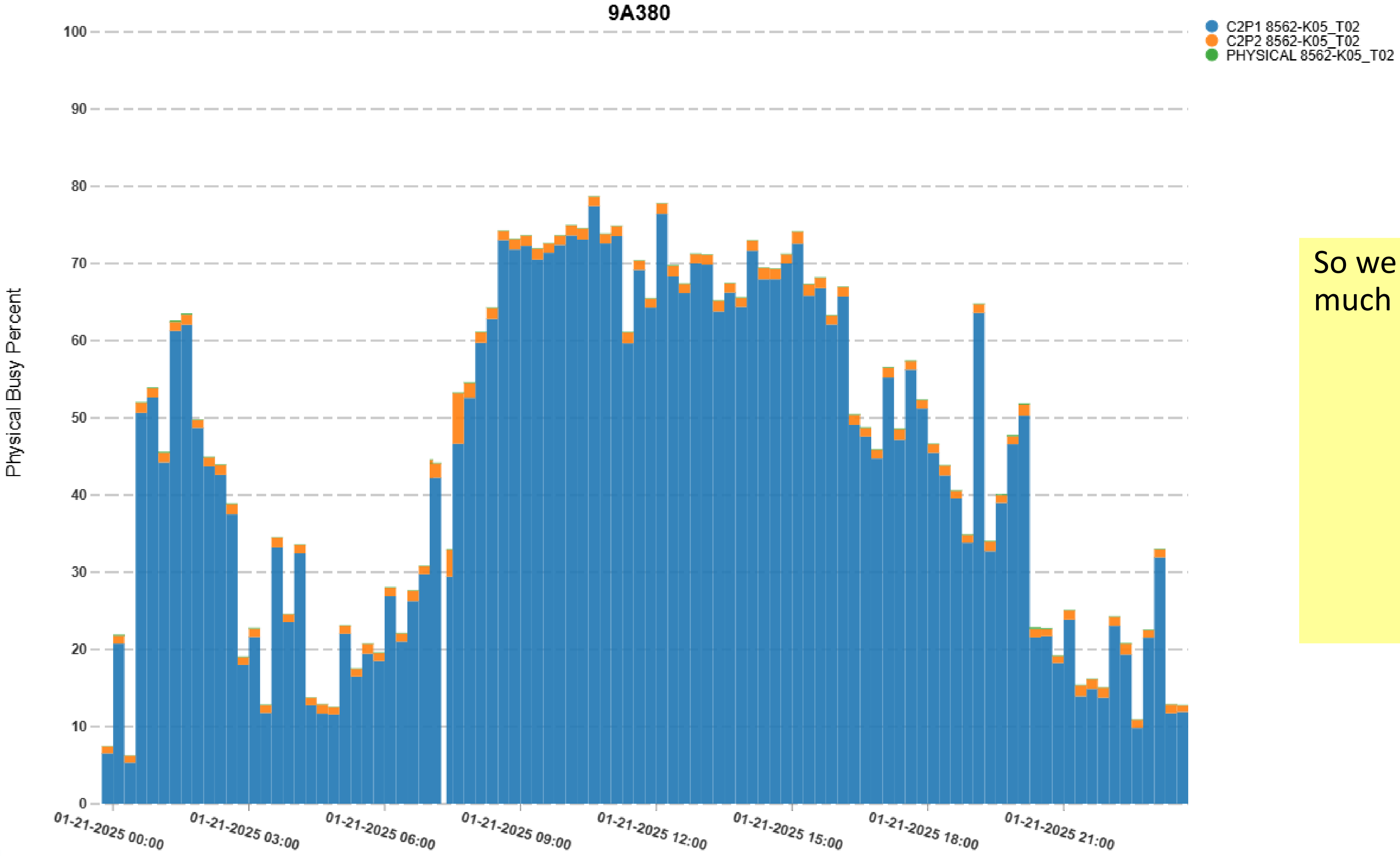


This report shows what percentage of the 15-minute interval the machine was running at various levels of busy.

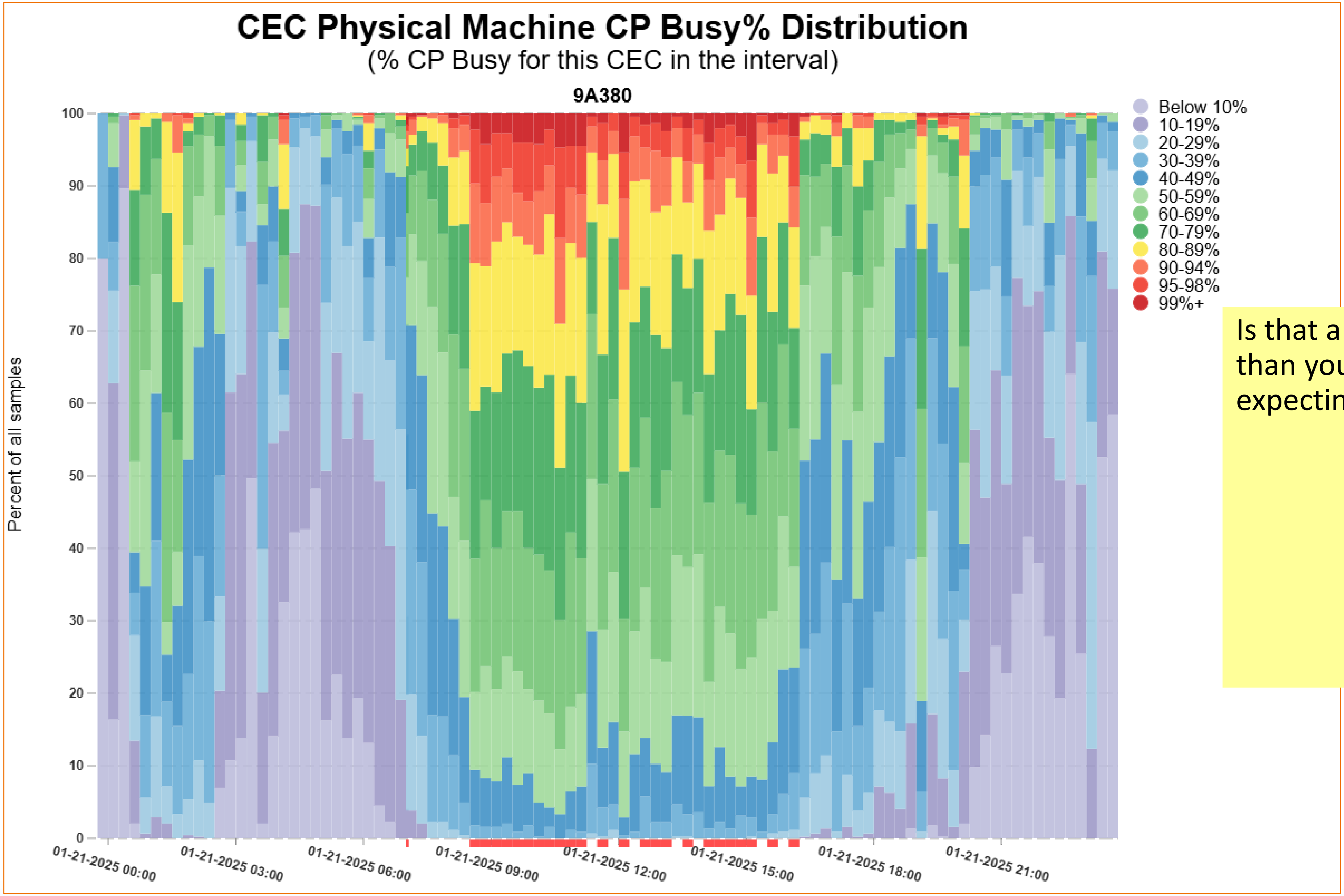
Where did that data come from??

- That previous report comes from the SMF 99.12 records
- HiperDispatch details, every 2 seconds
  - Including CEC utilization
  - Interestingly: 2 second intervals, but only written every 10 seconds
- Much less room to hid spikes within 2s intervals than 900s intervals!
- The distribution report basically summarizes the 2-second measurements that occurred within the 15 minute RMF interval

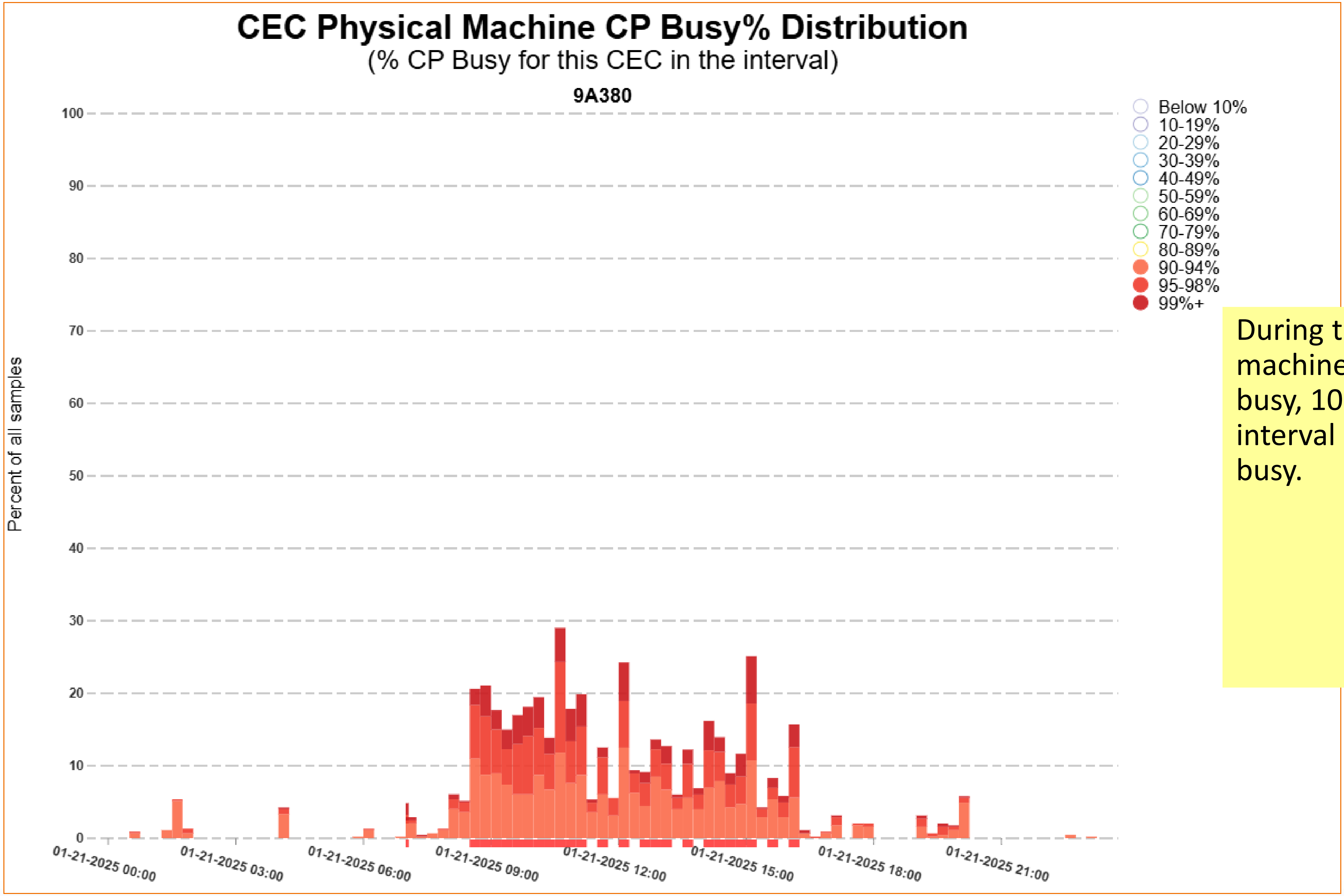
CEC Physical Machine CP Busy% by CEC Serial Number



So we shouldn't have much red with this one…

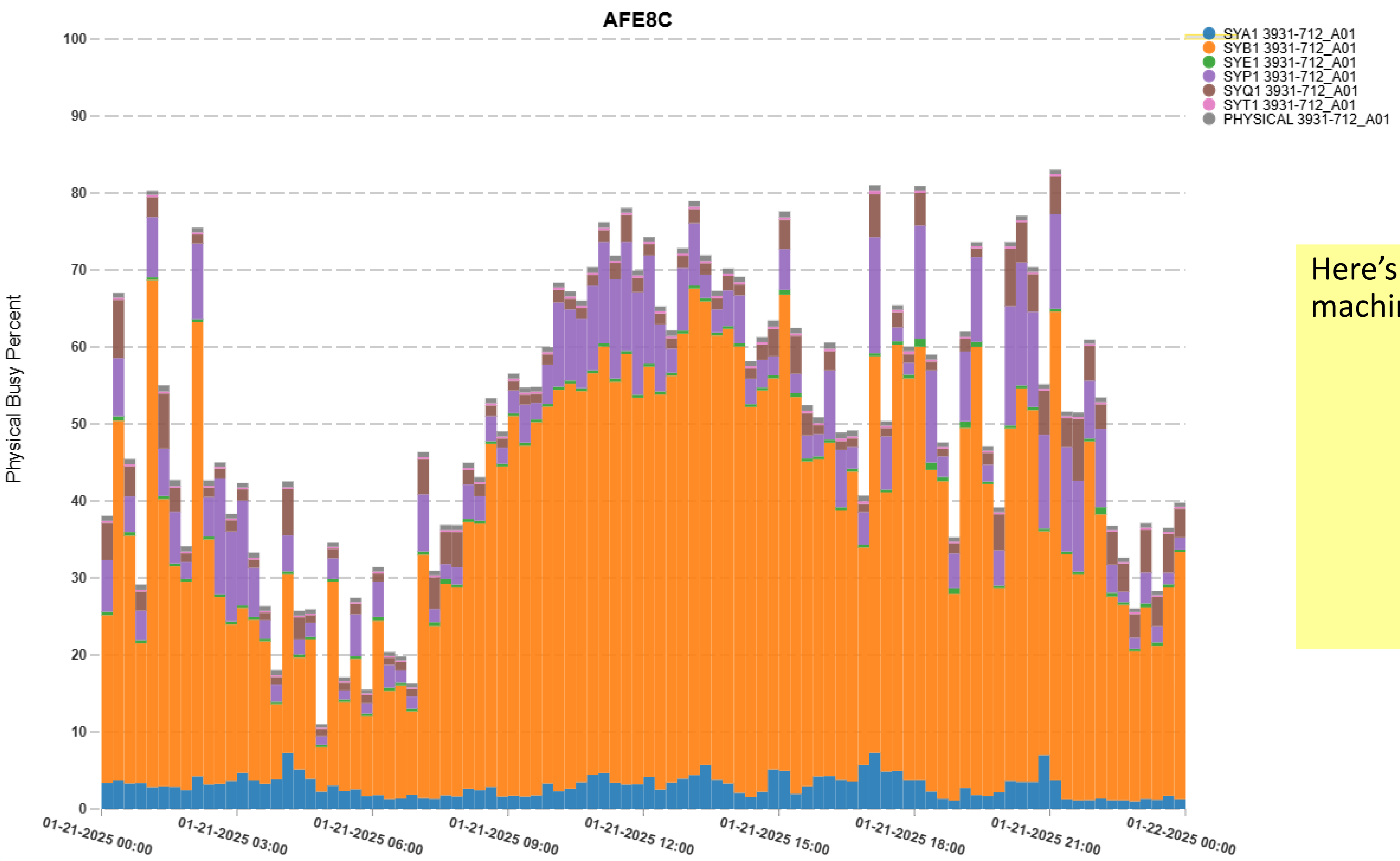


Is that a different picture than you were expecting?

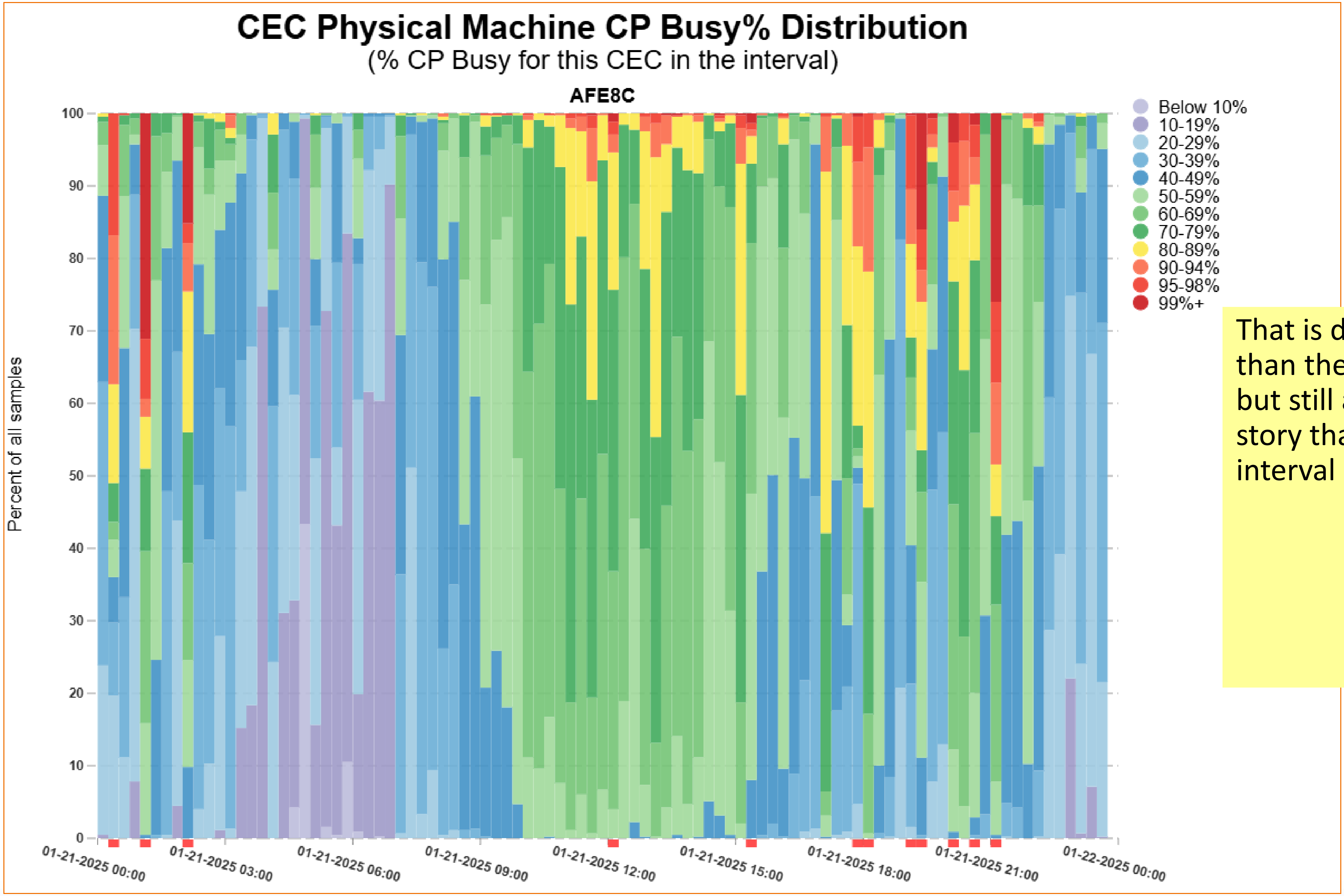


During the day when the machine was under 80% busy, 10-20% of the interval it was still >90% busy.

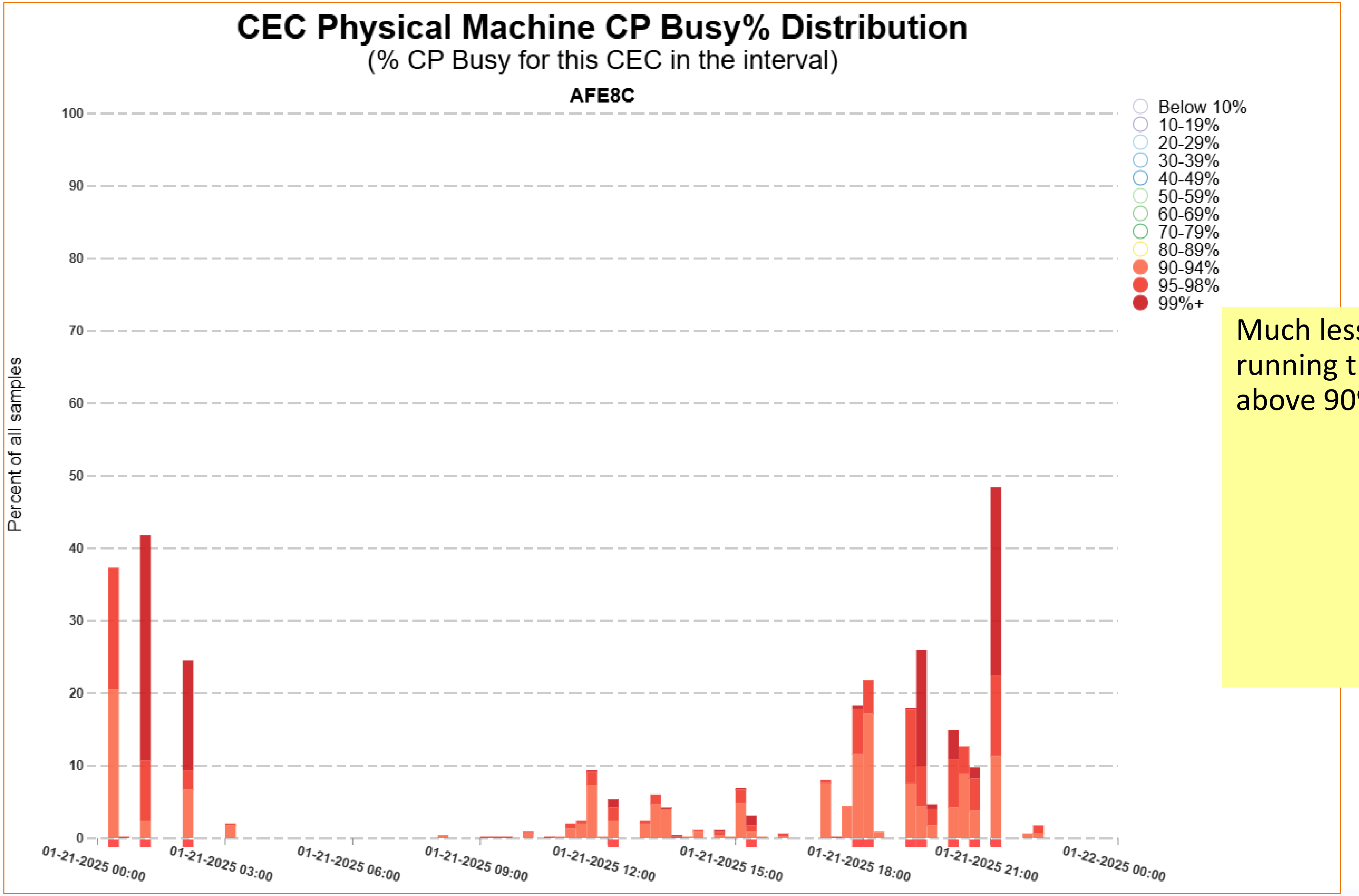
CEC Physical Machine CP Busy% by CEC Serial Number



Here's the 12-way machine…



That is definitely better than the 5-way machine, but still a bit different story than the 15-minute interval data implies.



Much less evidence of running those 12 CPs above 90% busy.

Why would we care?

- In many cases, you might not
  - Sometimes it's better not to look for problems
- But sometimes there are brief (minutes to seconds) mystery problems
  - 15-minute intervals can hide a lot of transient problems
  - Sometimes transient backups can cause a broader cascade of problems
- Commonly heard: "Well the WLM policy really only matters when the system is busy, and we're not that busy"
  - Mostly true, but you're probably busier more often than you think
- Do not assume you have no CPU problems because you're only 80% busy
  - Within that 15 minutes there could be multiple minutes at 100% busy!

Exploring transient performance

SMF for Transient Performance Analysis

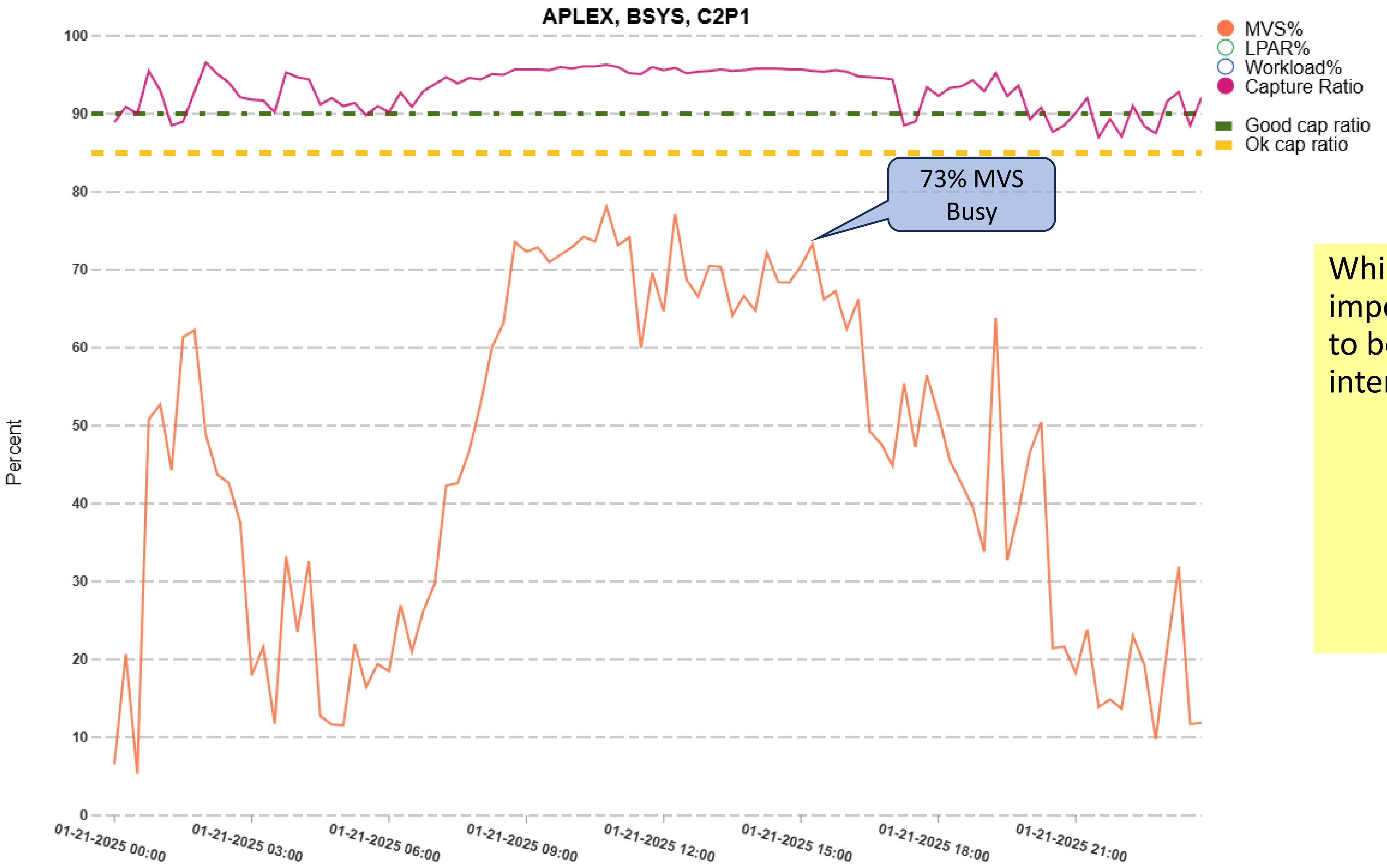
- Transactional Records
  - Records for CICS, DB2, Websphere, etc. like SMF 110, 101, 120
  - Can be extremely voluminous—100s of GBs/day quite possible
  - Not going to be discussed here
- SMF 98
  - High Frequency Throughput Statistics
  - Also, specific records for Workload Interaction Correlator (IBM Product)
  - Customer-defined interval (mostly) of 5-60 seconds
- SMF 99
  - SRM/WLM Decisions
  - Primarily 10 seconds, subtype 12s at 2 seconds

Data volumes

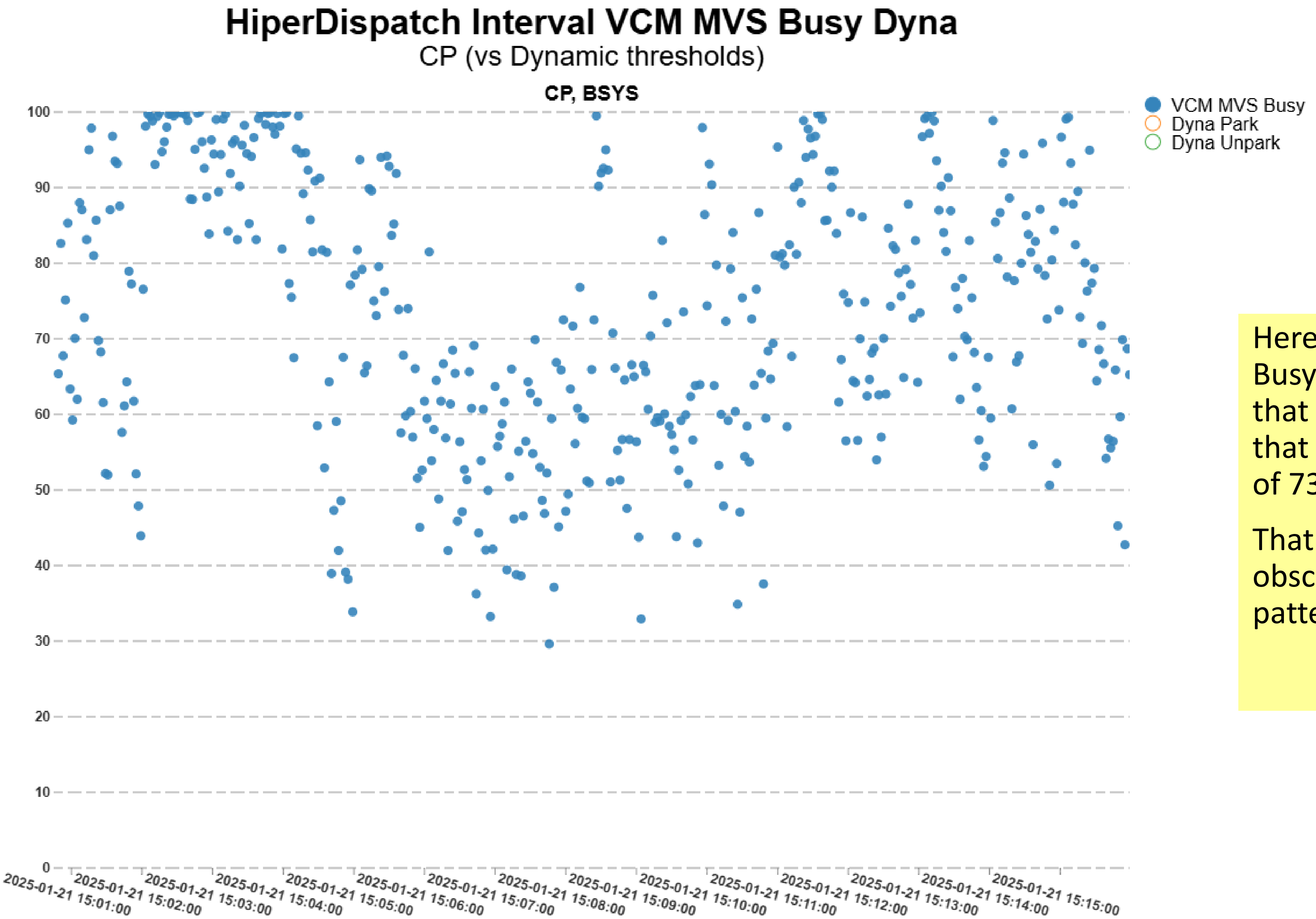
- Transactional records
  - Based on transaction volume but can easily be 10s to 100s of GBs/day
  - Most customers are in fact already cutting (at least most of) these
    - Many don't cut the below despite those records being only MBs/day
- SMF 98
  - 15 second interval = 100-150 MB / system / day
  - 5 second interval = 400-500 MB / system / day
- SMF 99 minimum recommended subtypes:
  - 6, 8, 10, 11, 12 ,14 = 100-150 MB / system / day
- SMF 99 additional recommended subtypes:
  - 1, 2, 3 = 850-1500 MB / system / day

Subtype 8 is new(ish) to our minimum recommendation

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



While MVS Busy is less important than it used to be, sometimes it's still interesting to look at.

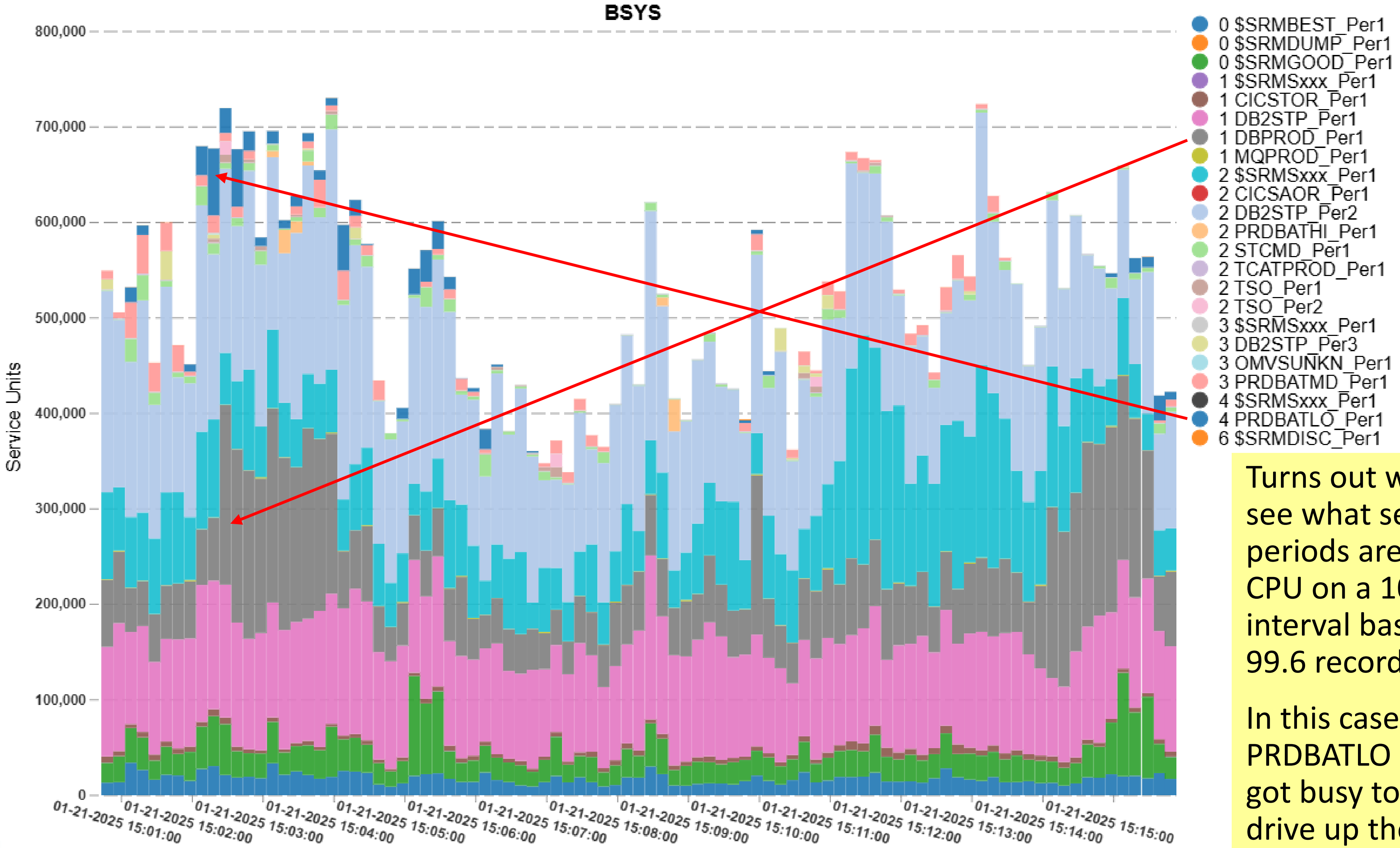


Here's the 2-second MVS Busy measurements for that 15-minute interval that showed an average of 73% busy.

That 73% average obscures the actual pattern of utilization.

CPU Accumulated by Service Class Period

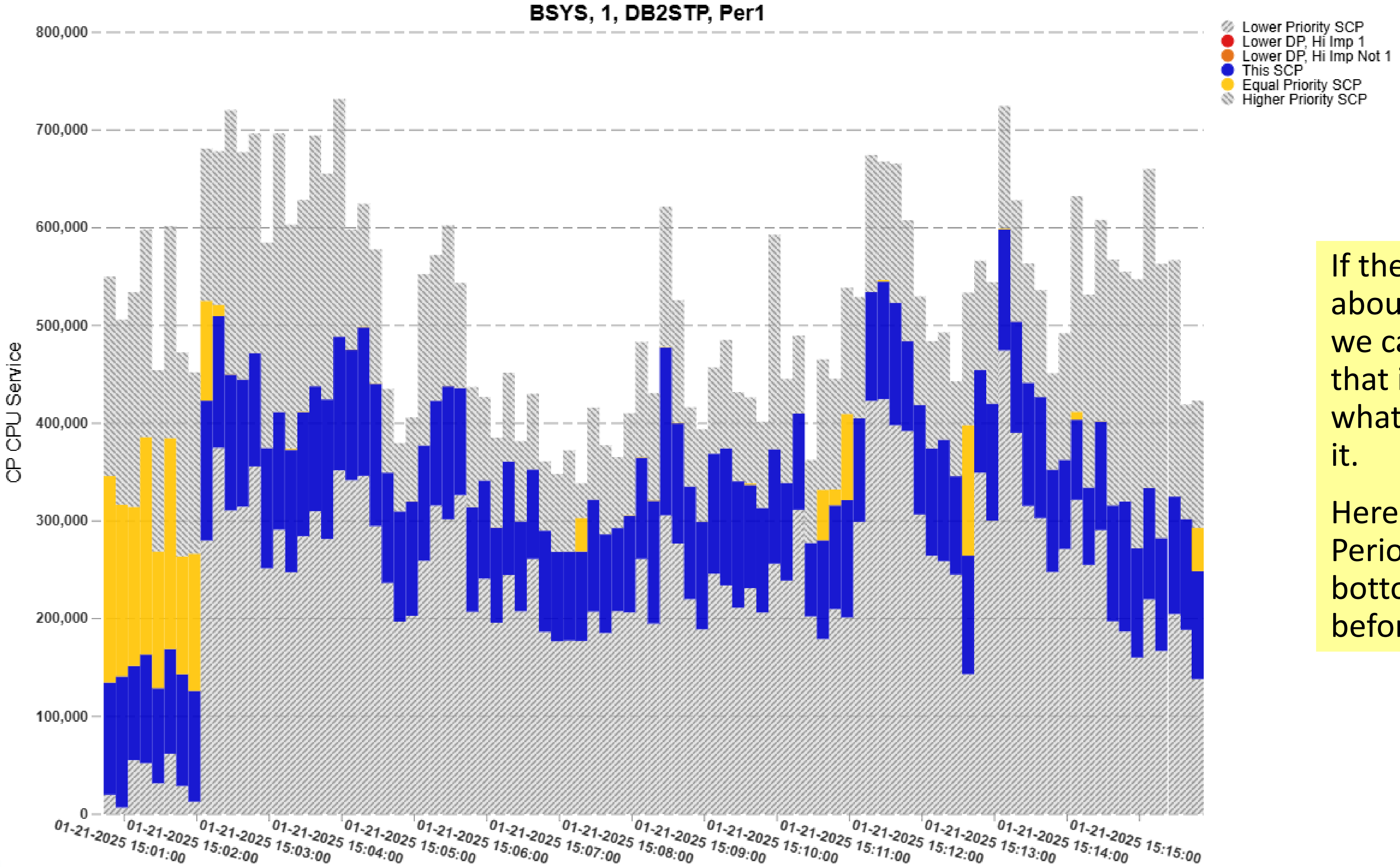
From SMF 99.6



Turns out we can also see what service class periods are using the CPU on a 10-second interval basis from the 99.6 records.

In this case looks like PRDBATLO and DBPROD got busy together to drive up the utilization.

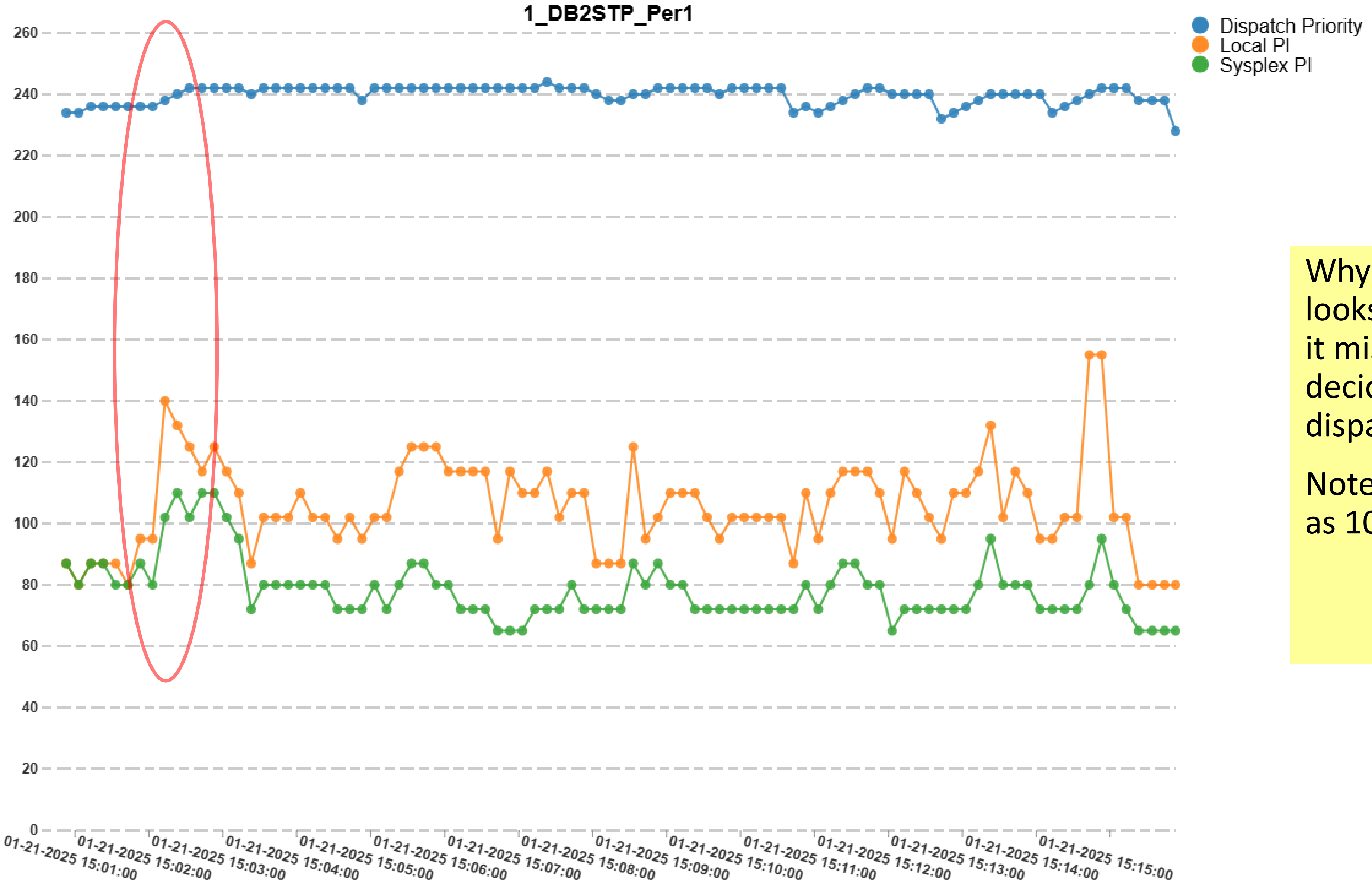
CP CPU Service Accumulated Above / Below SCP



If there's a question about a particular SCP we can dive deeper on that in the 99s to see what was going on with it.

Here we see DB2STP Period 1 running at the bottom of the heap before being moved up.

WLM SMF 99.6 Data Explorer

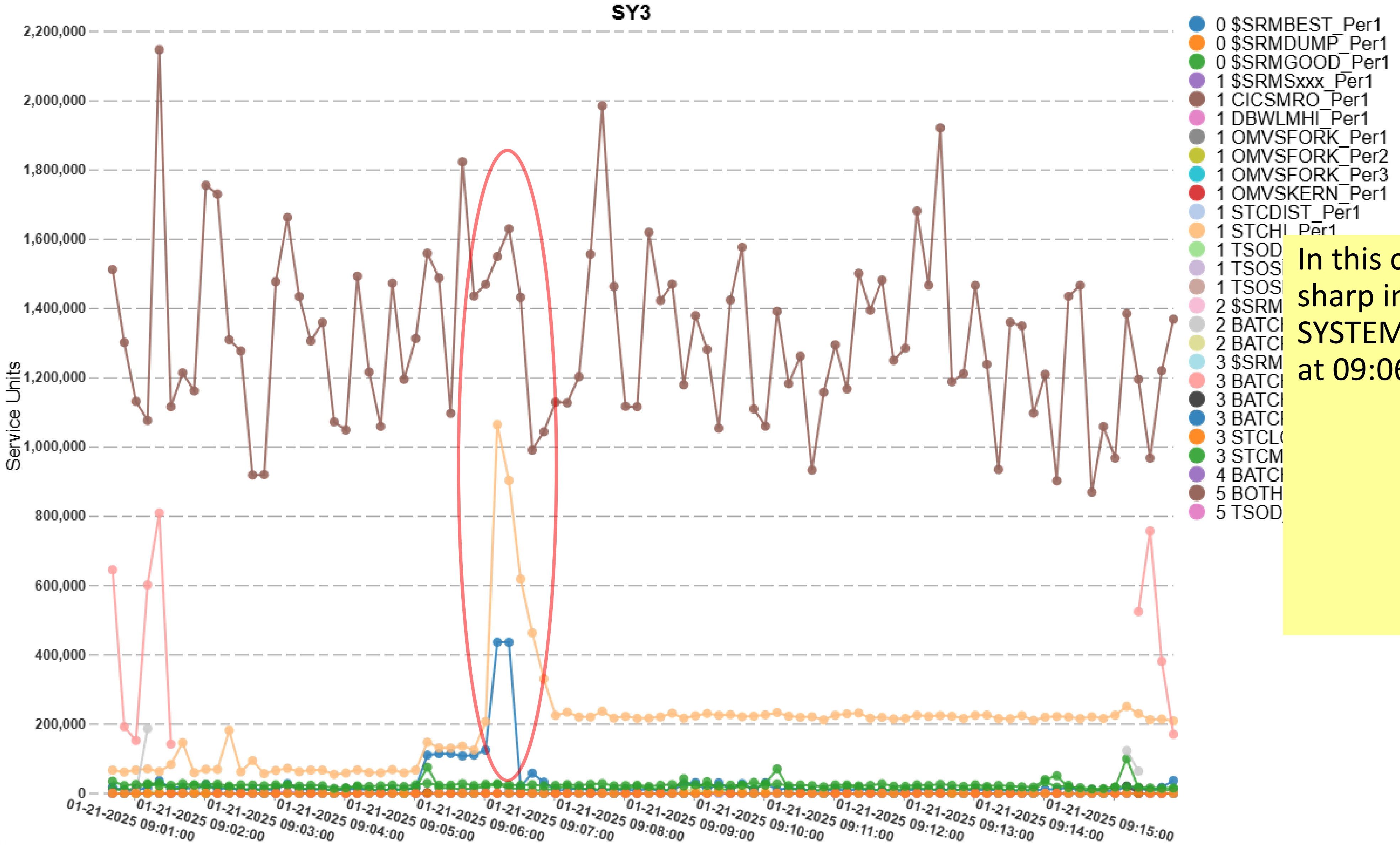


Why did that happen? It looks like WLM noticed it missing its goal and decided to bump up its dispatching priority a bit.

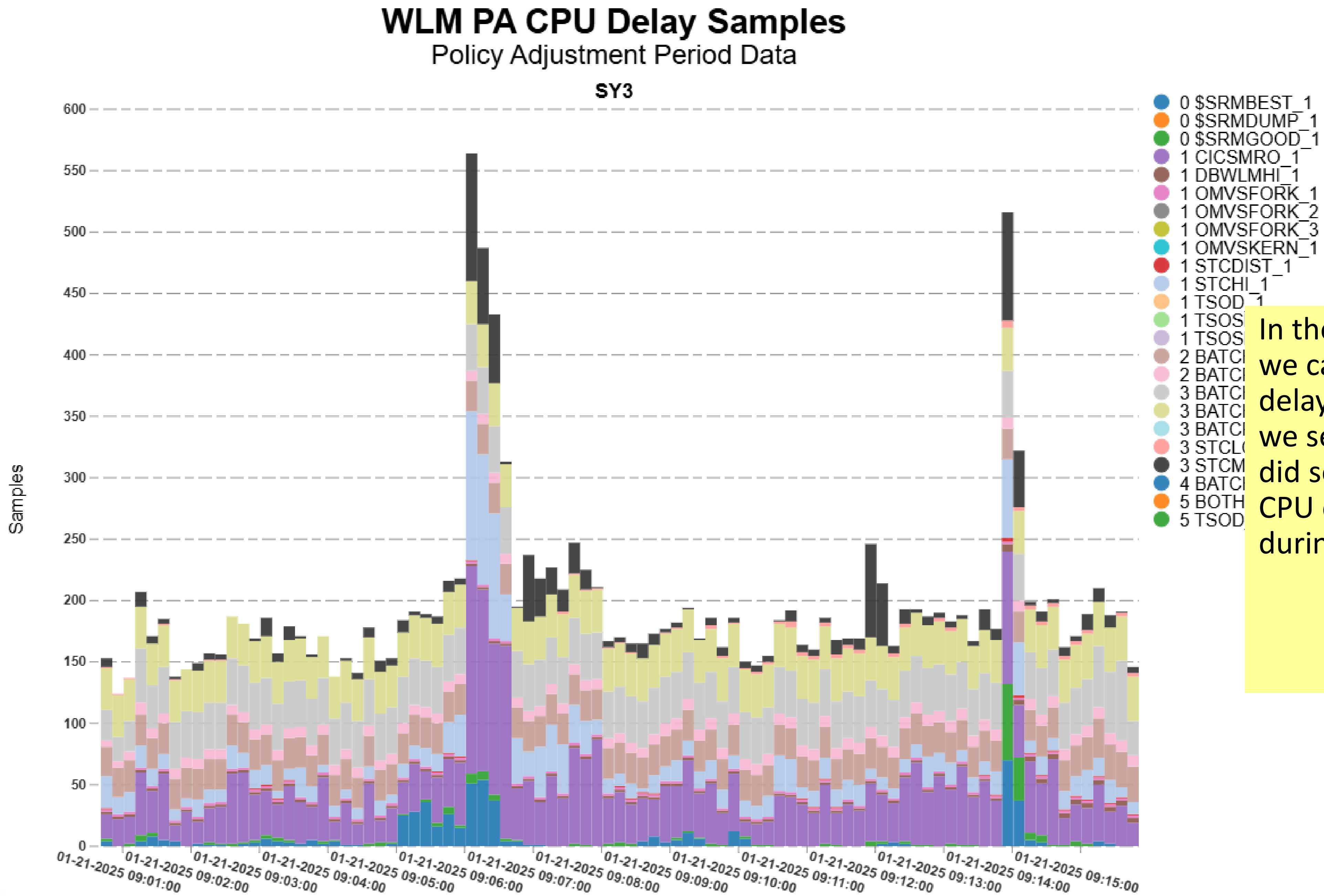
Note PI of 1 is recorded as 100 in this 99 data.

CPU Accumulated by Service Class Period

From SMF 99.6



In this case we can see a sharp increase in both SYSTEM and STCHI work at 09:06:11



In the SMF 99.2 and .8 we can see the CPU delay sample counts and we see that CICSMRO did see an increase in CPU delay samples during that 40 seconds.

WLM - SMF 99 Decision Trace

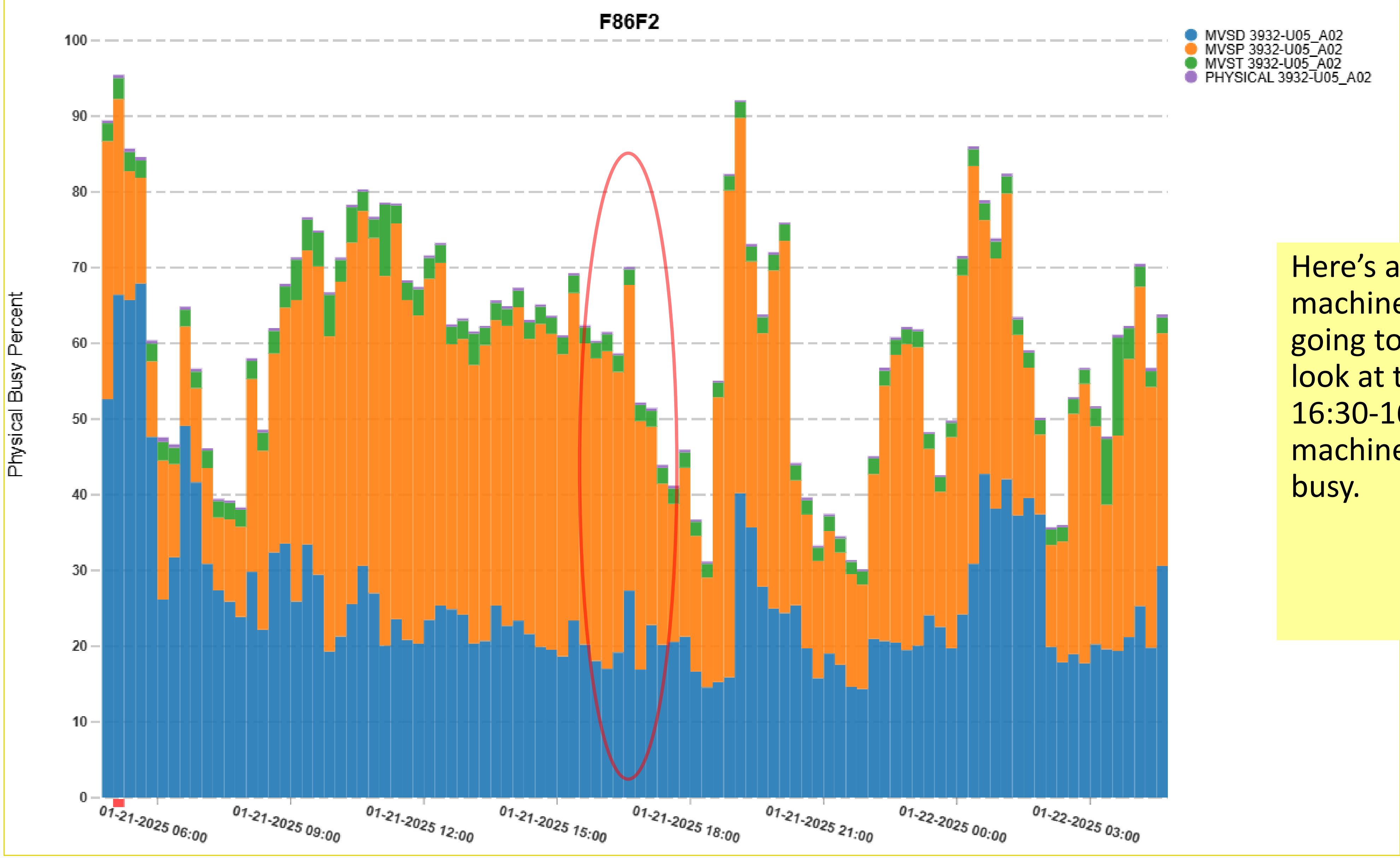
SY3, 9

Freeze 4 Filter: system = Value Apply Clear All Toggle Sums 31251 Initial rows; 2 filters leaving 5 rows (0%)

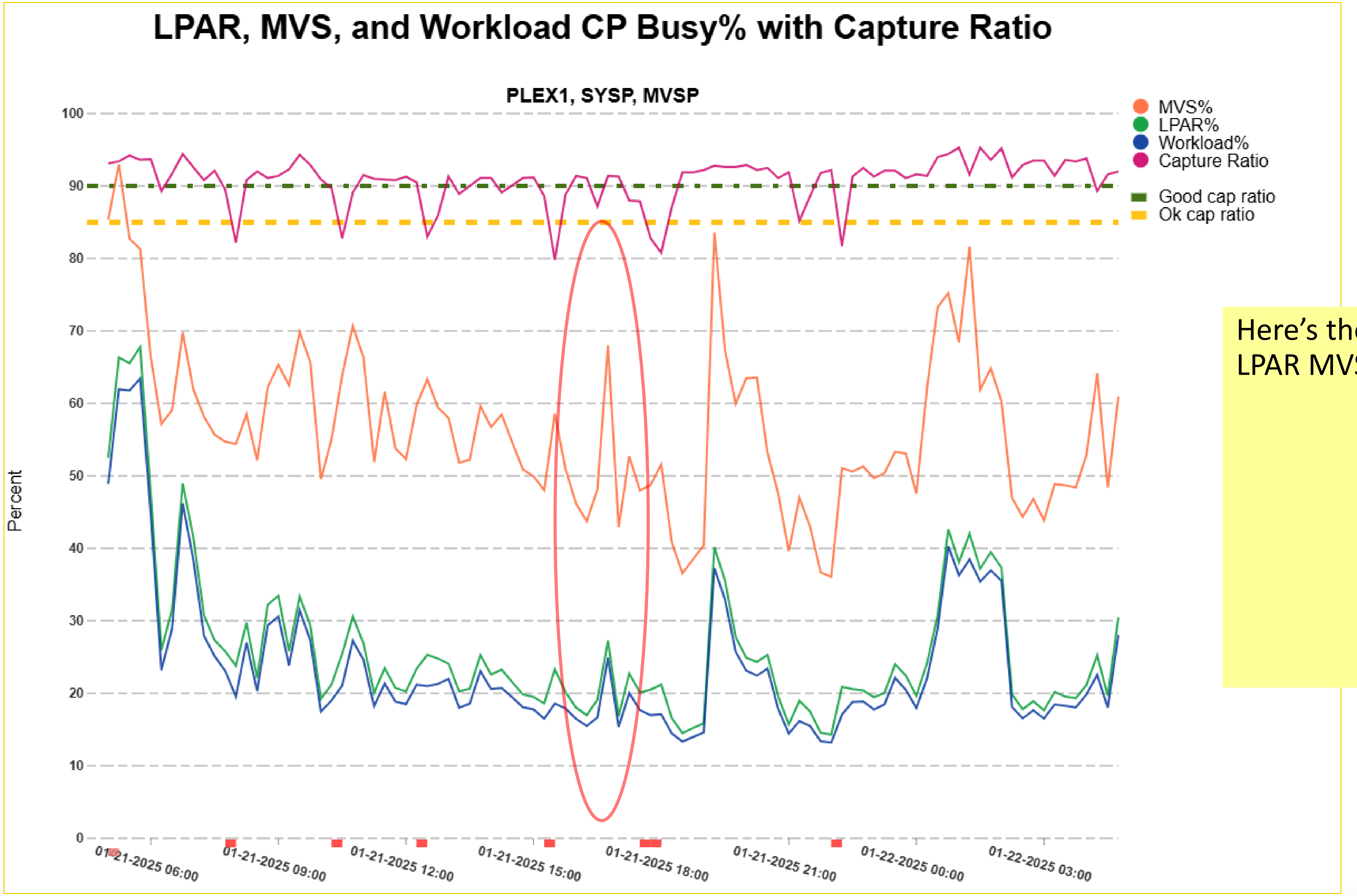
S...	H...	Date	Time	Event	P... I...	R... I...	C...	SC Name	Period	Local PI	Sysplex PI	A... S...	Symbol	Explanation
								CICSMR(					PA_GREC_CAND	
5	45			281.0	3...	3...	120...		5.0	5.3	5.3			
SY3	9	2025-01-21	09:06:11	22.000	36...	45...	240....	CICSMRO	1.000	1.140	1.140		PA_GREC_CAND	Policy adjustment; goal receiver candidate selected.
SY3	9	2025-01-21	09:06:21	25.000	37...	50...	240....	CICSMRO	1.000	1.160	1.160		PA_GREC_CAND	Policy adjustment; goal receiver candidate selected.
SY3	9	2025-01-21	09:06:31	25.000	38...	55...	240....	CICSMRO	1.000	1.010	1.010		PA_GREC_CAND	Policy adjustment; goal receiver candidate selected.
SY3	9	2025-01-21	09:29:01	61.000	17...	21...	240....	CICSMRO	1.000	1.030	1.030		PA_GREC_CAND	Policy adjustment; goal receiver candidate selected.
SY3	9	2025-01-21	09:47:31	148.000	28...	4....	240....	CICSMRO	1.000	0.910	0.910		PA_GREC_CAND	Policy adjustment; goal receiver candidate selected.

Looking at the 99.1 trace records we can see that in fact WLM at least considered helping CICSMRO during 3 policy adjustment intervals when the other work got busy.

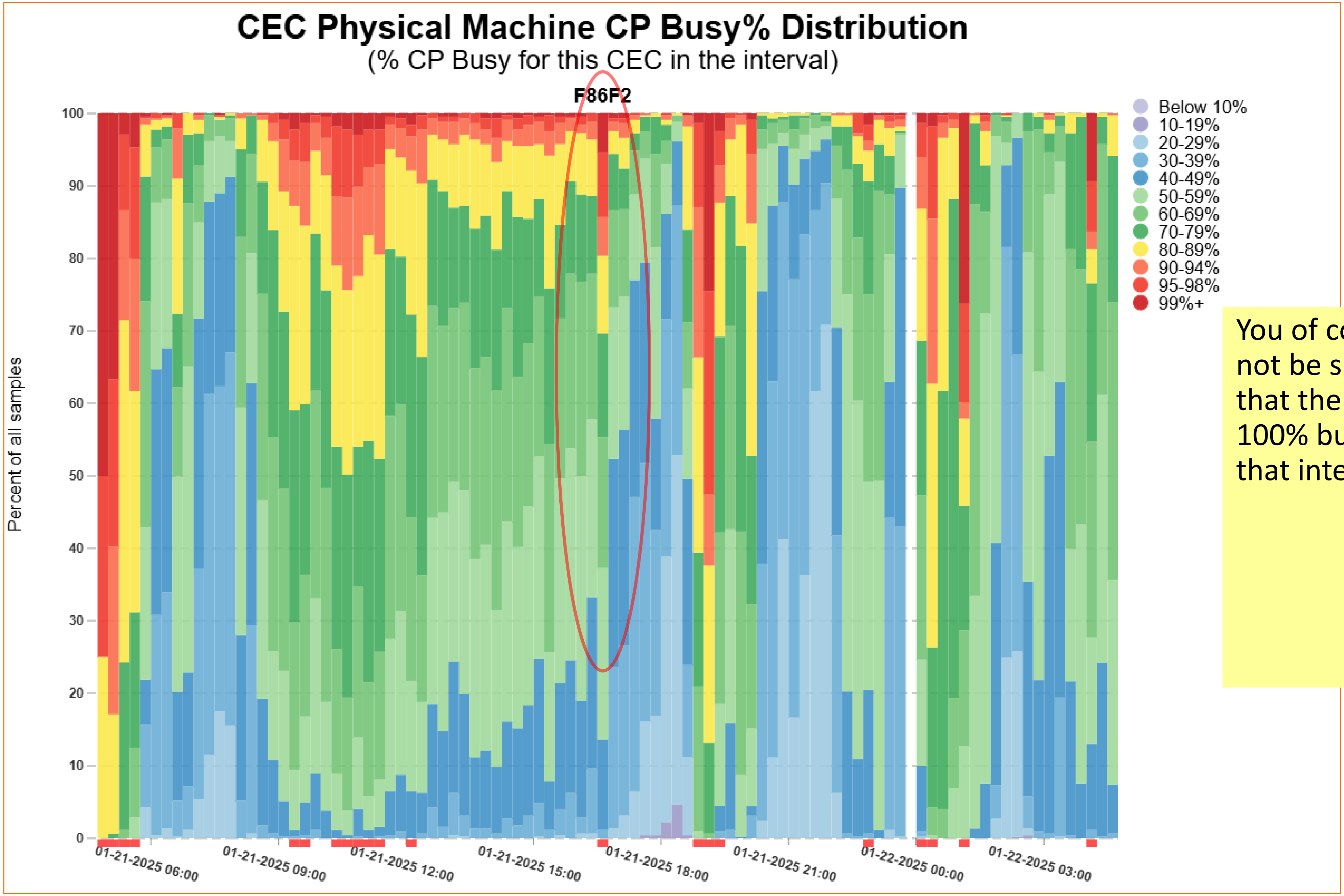
CEC Physical Machine CP Busy% by CEC Serial Number



Here's a different machine that we're going to take a closer look at that interval from 16:30-16:45 where the machine appears 70% busy.

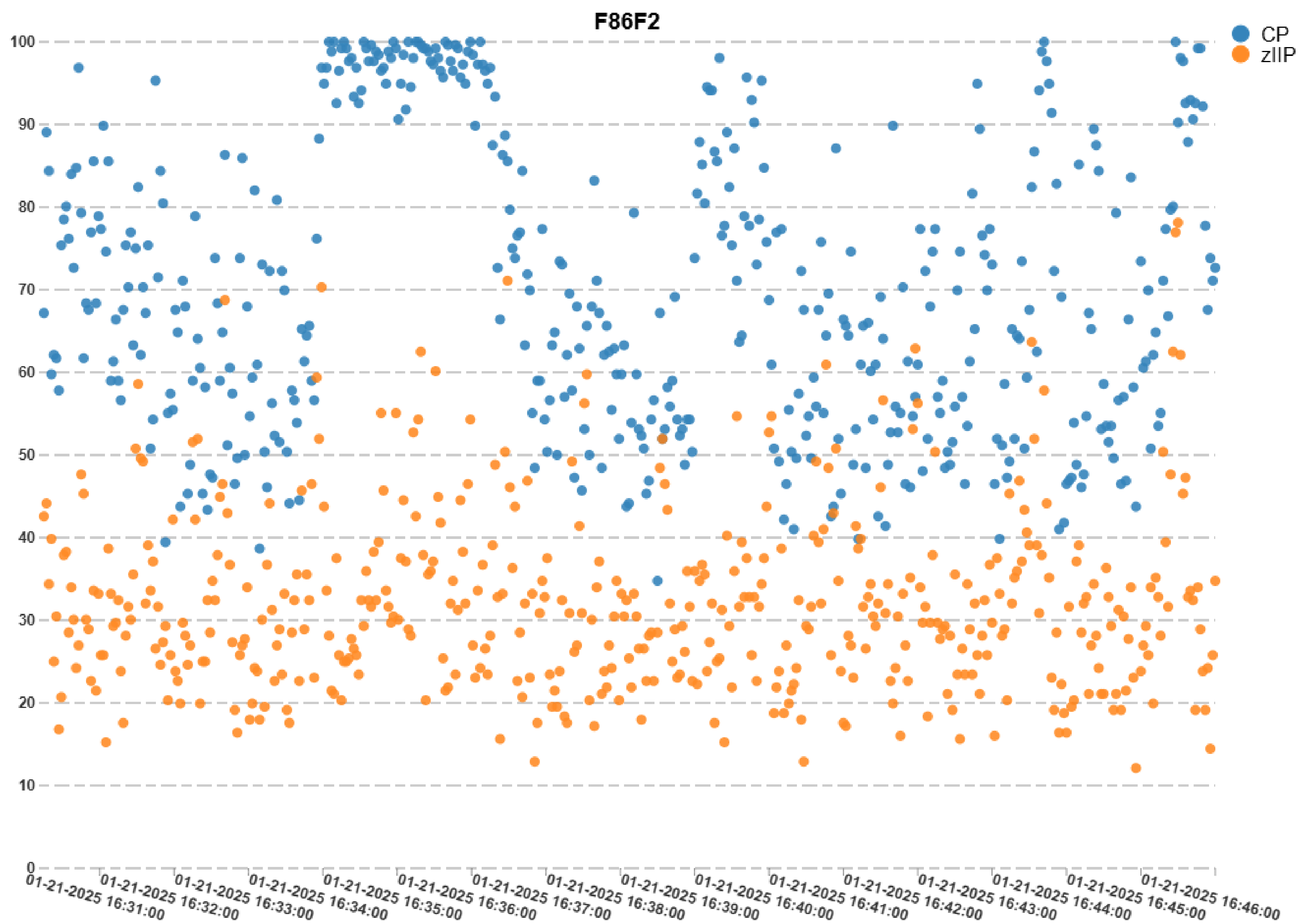


Here's the Production LPAR MVS Busy

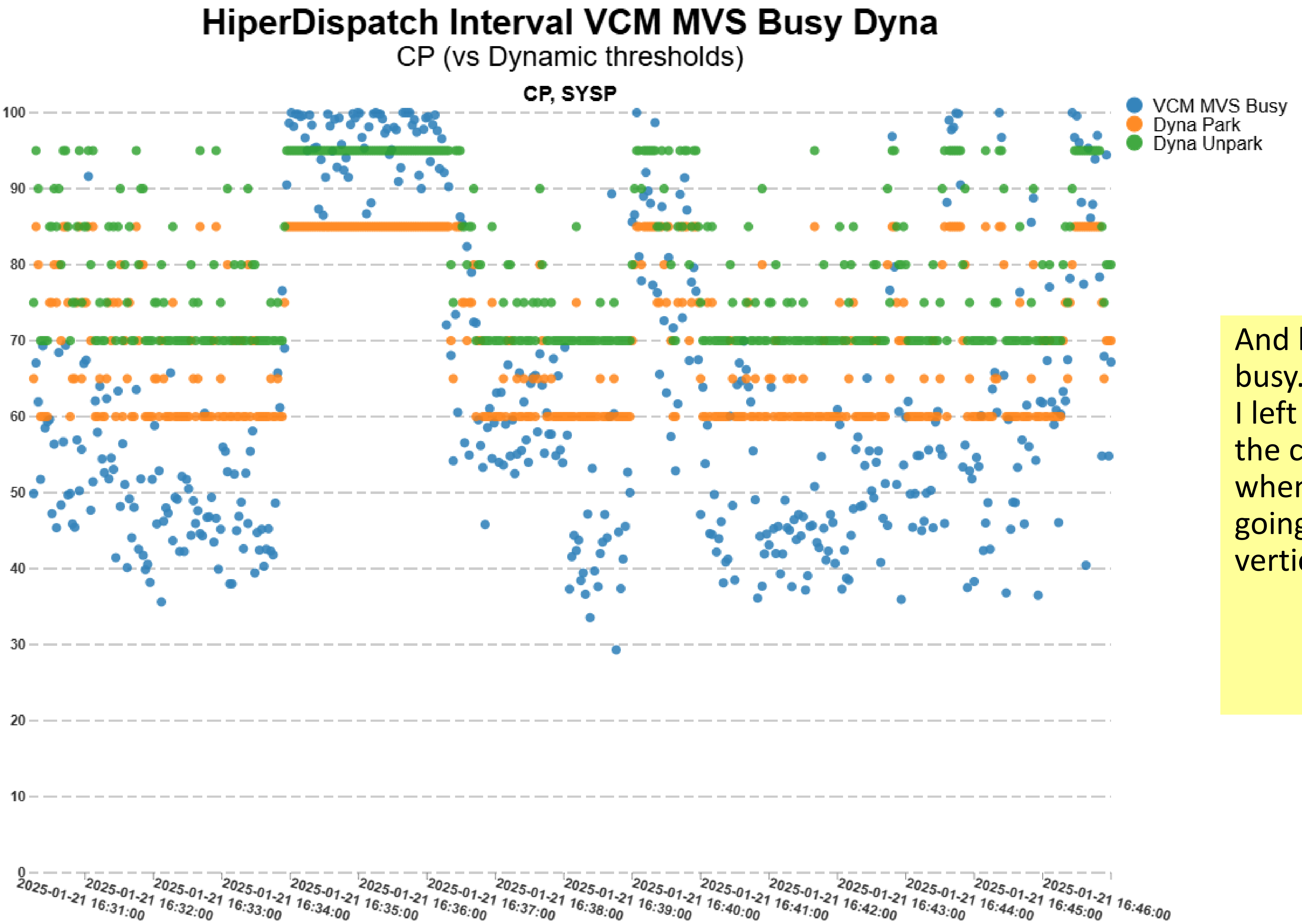


You of course will now not be surprised to hear that the machine was 100% busy for part of that interval.

HiperDispatch CEC Utilization

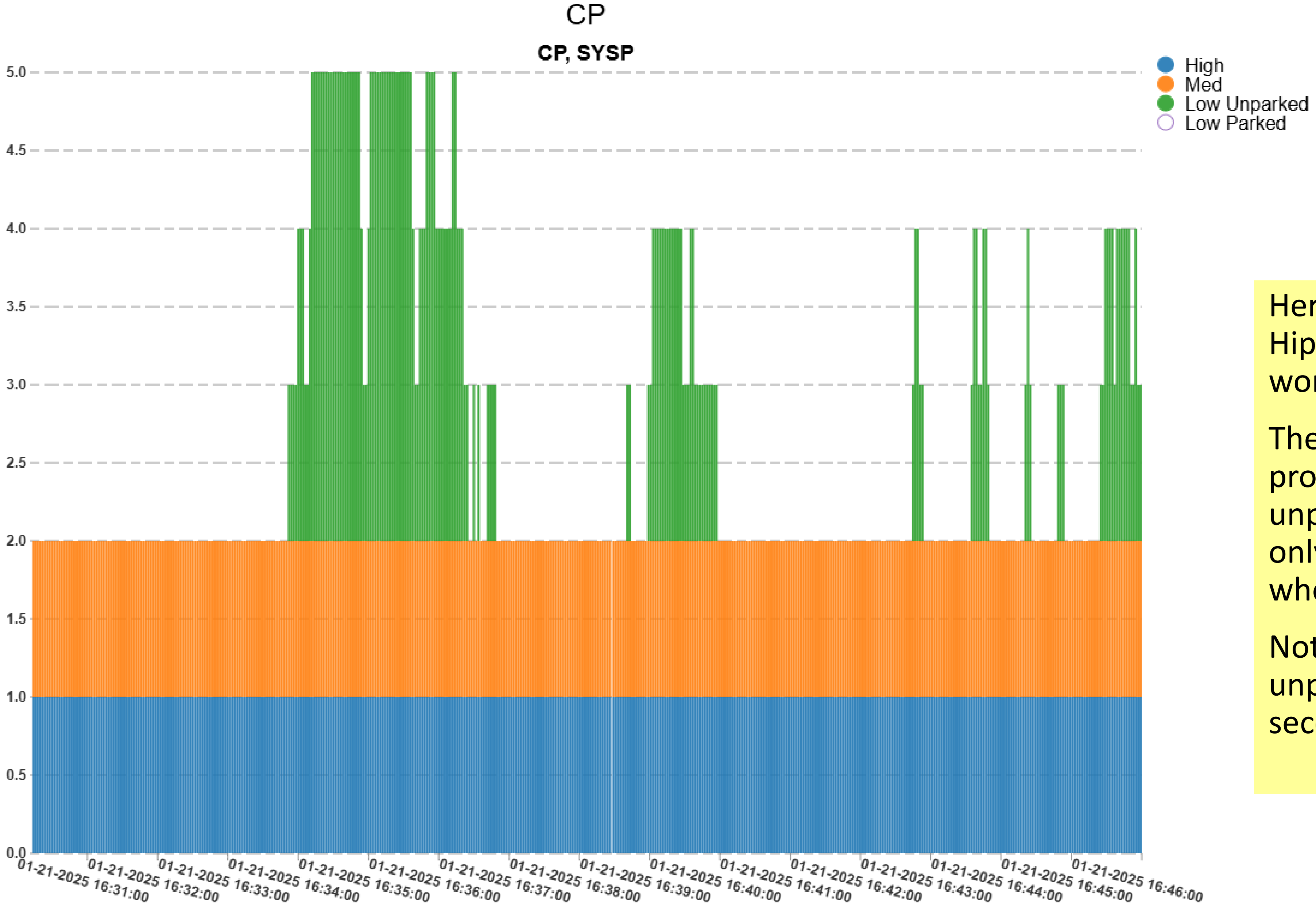


Here's the 2 second CEC utilization: got quite busy for a couple of minutes around 16:35.



And here's the logical busy. Note that this time I left the thresholds on the chart that show when HiperDispatch is going to unpark/park vertical low CPs.

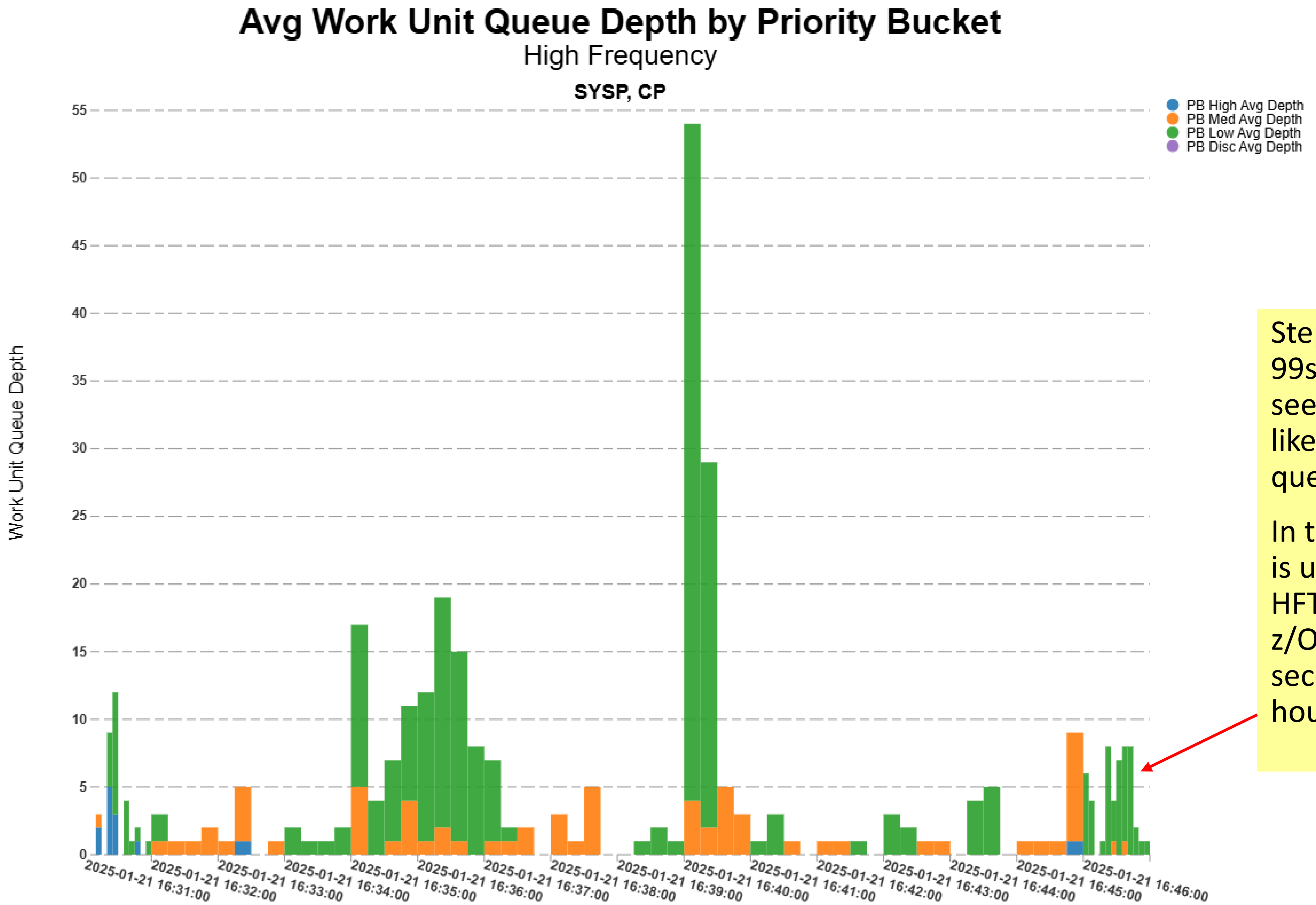
HiperDispatch Interval Processor Counts



Here we can see HiperDispatch doing its work.

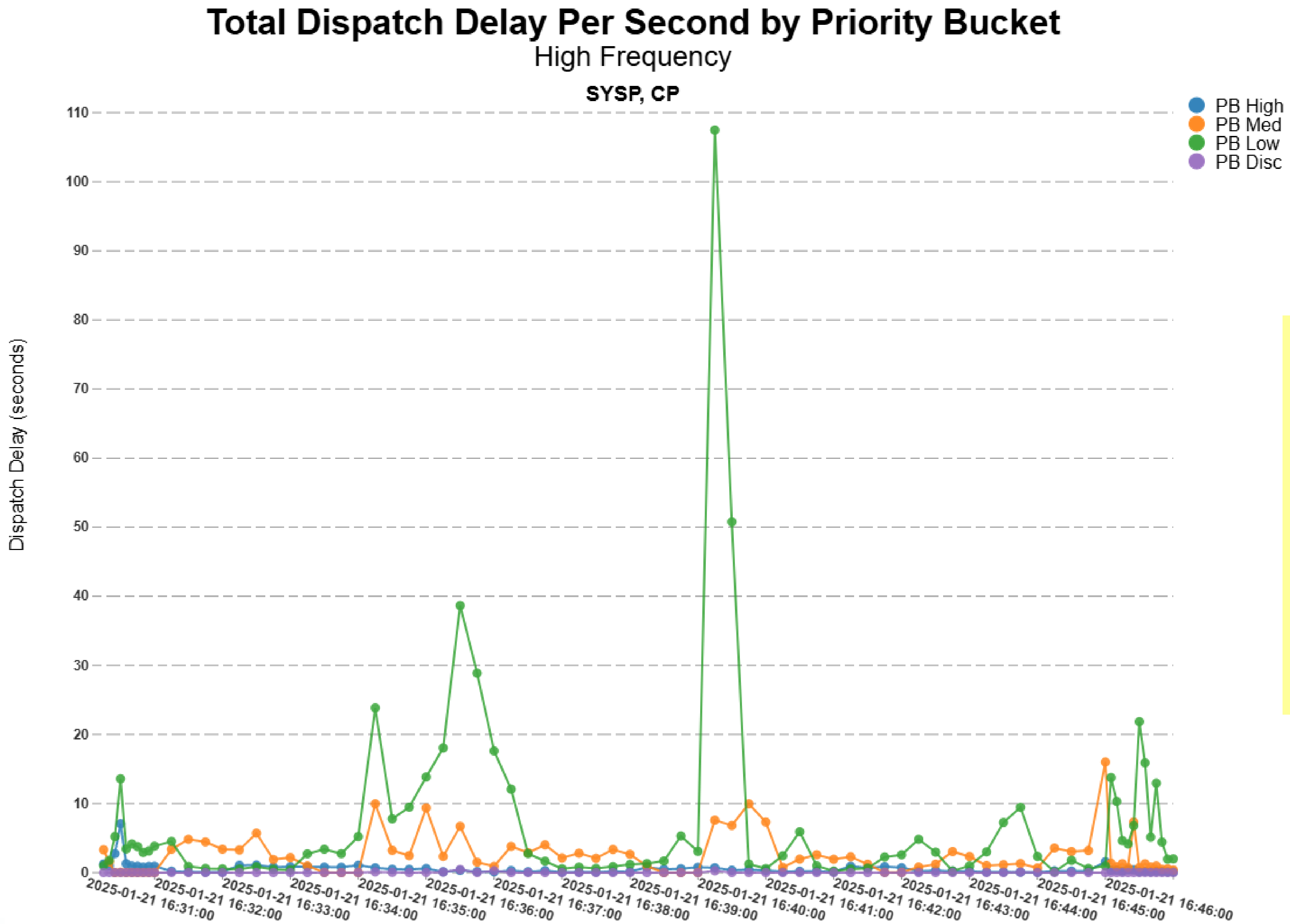
The VH and VM processors are always unparked, but the VLs only come into play when needed.

Note that it only parks or unparks 1 CP every 2 seconds.



Stepping away from the 99s, in the 98s we can see interesting things like how long of a work queue did we have?

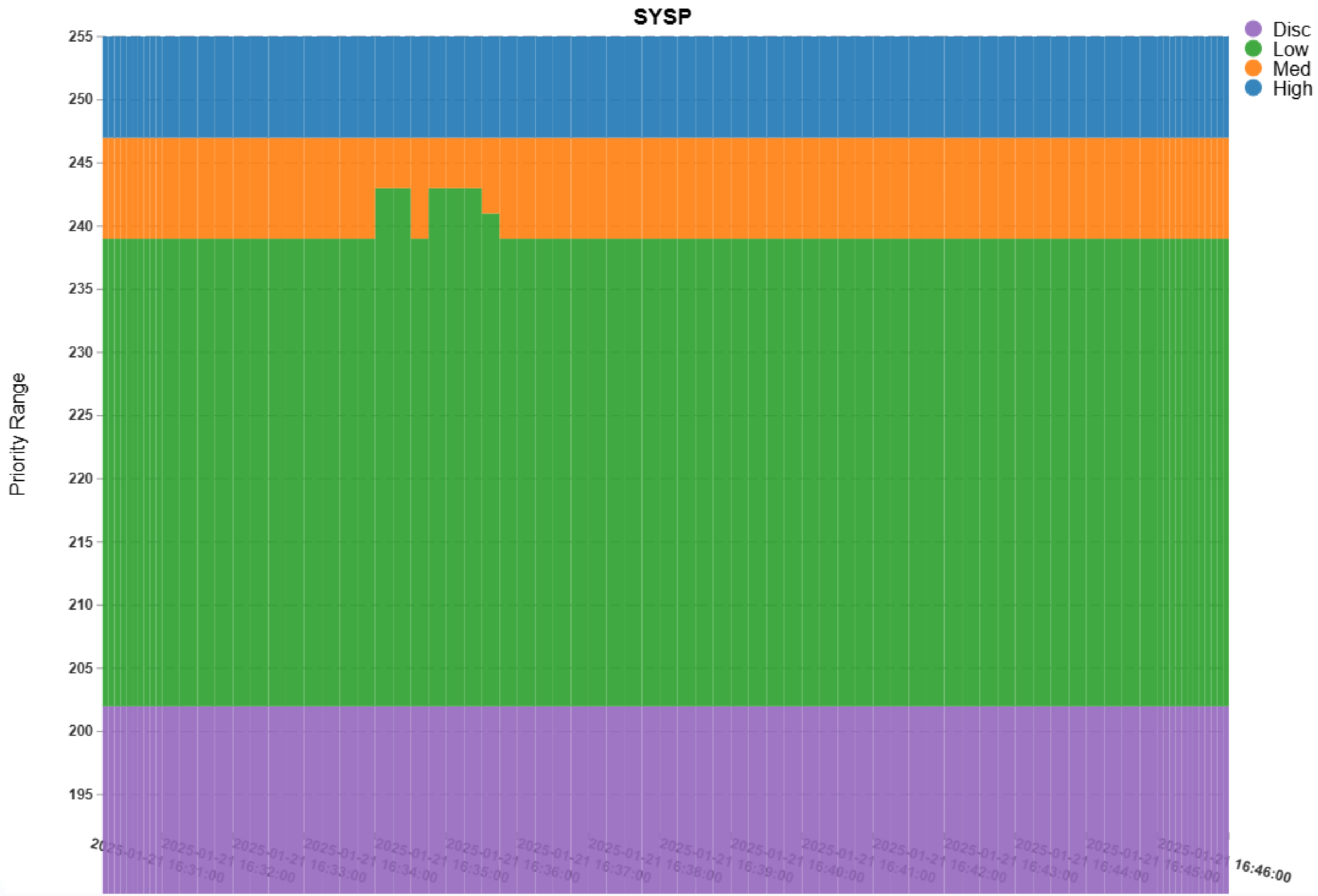
In this case the customer is using a 15 second HFTS interval, but note z/OS forces it to 5 seconds on the quarter hour marks.



We also can see how much delay those queues are adding to the work.

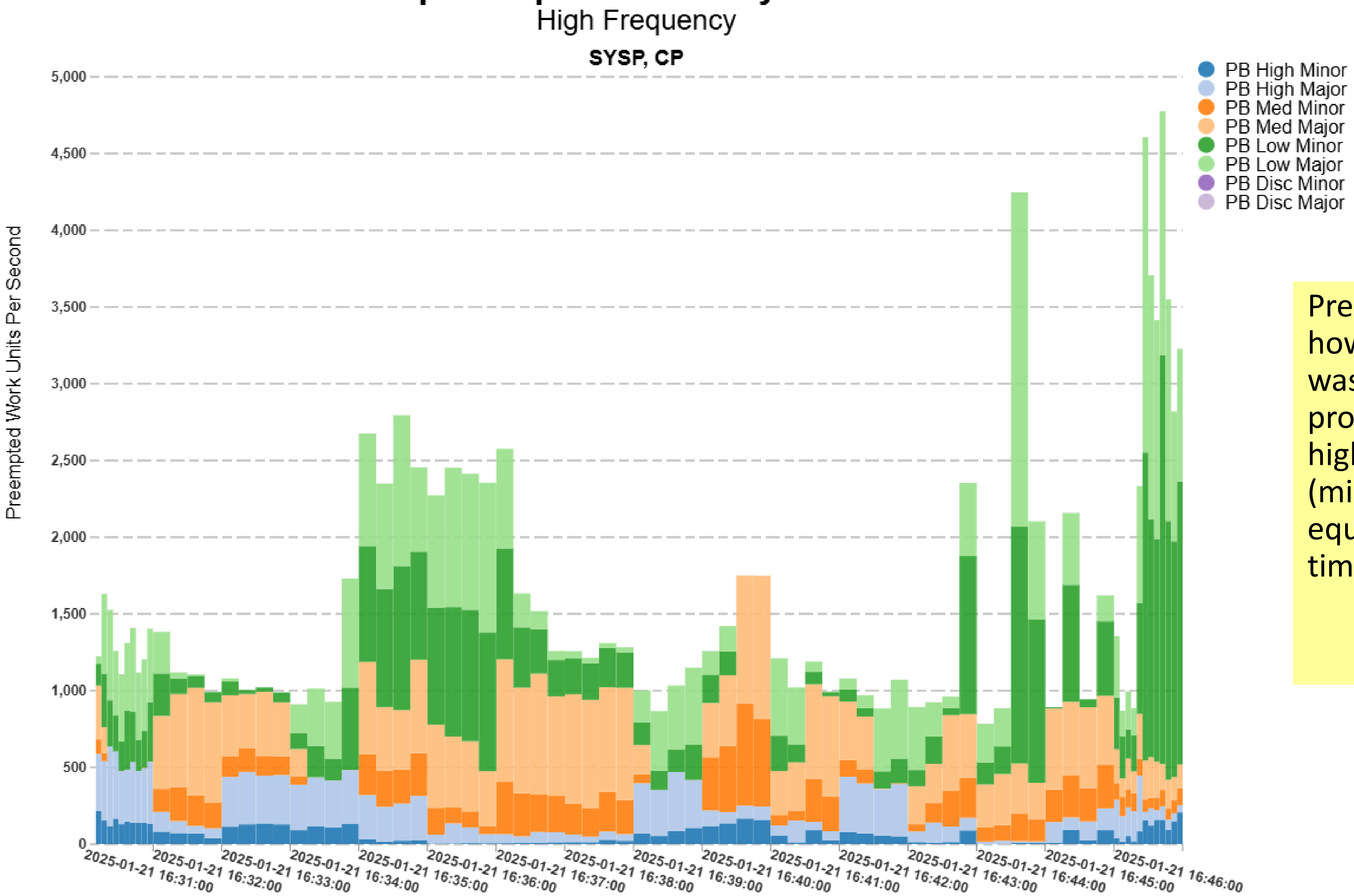
Both of these are by dispatch priority "bucket". Unfortunately, the exact bucket thresholds can vary.

HFTS Priority Bucket Ranges

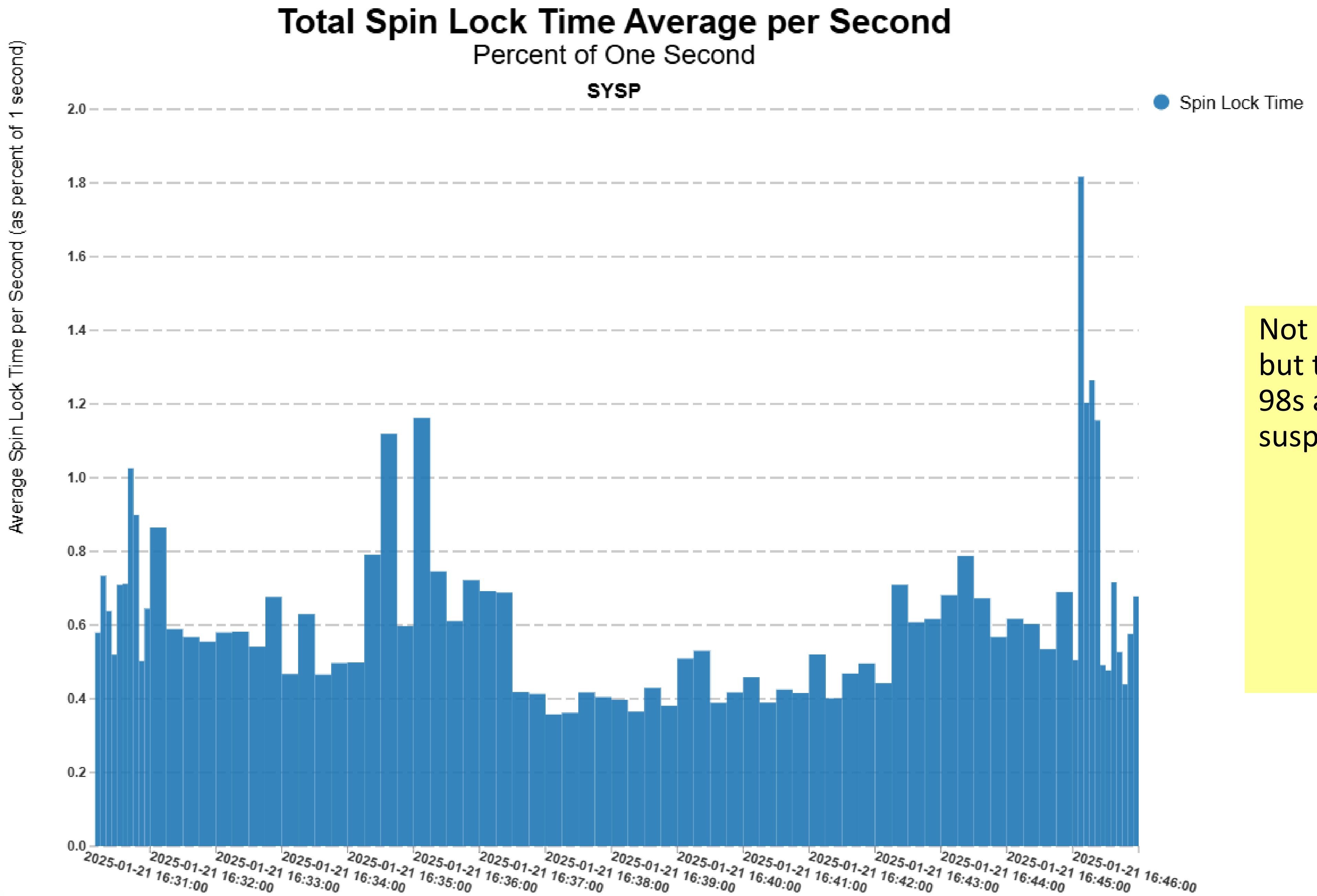


Not the greatest visualization, but at least we can see what the ranges were.

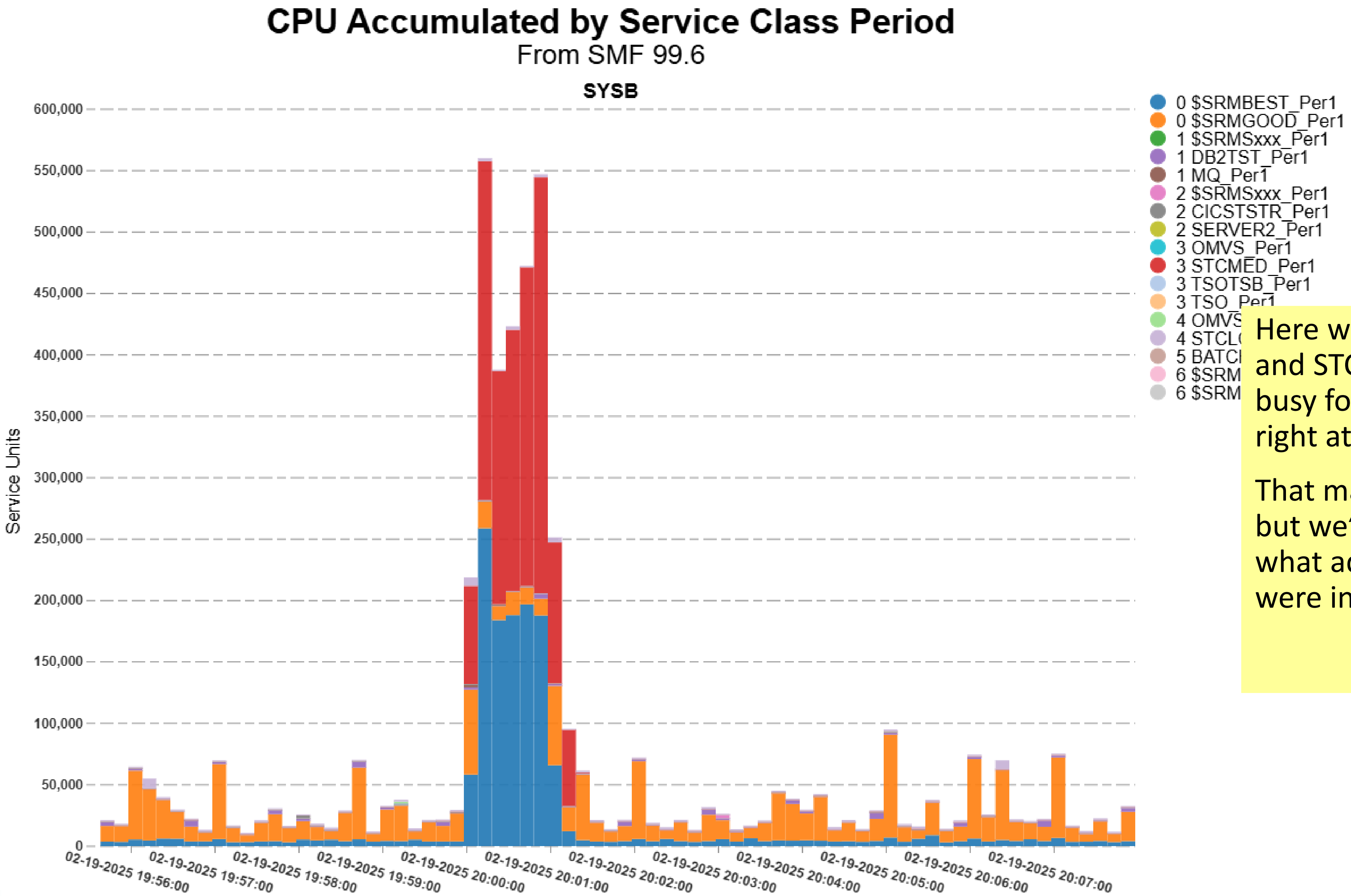
Preemptions per Second by Time Slice



Preemptions goes to how often something was kicked off the processor due to a higher priority work unit (minor time slice) or an equal priority WU (major time slice).

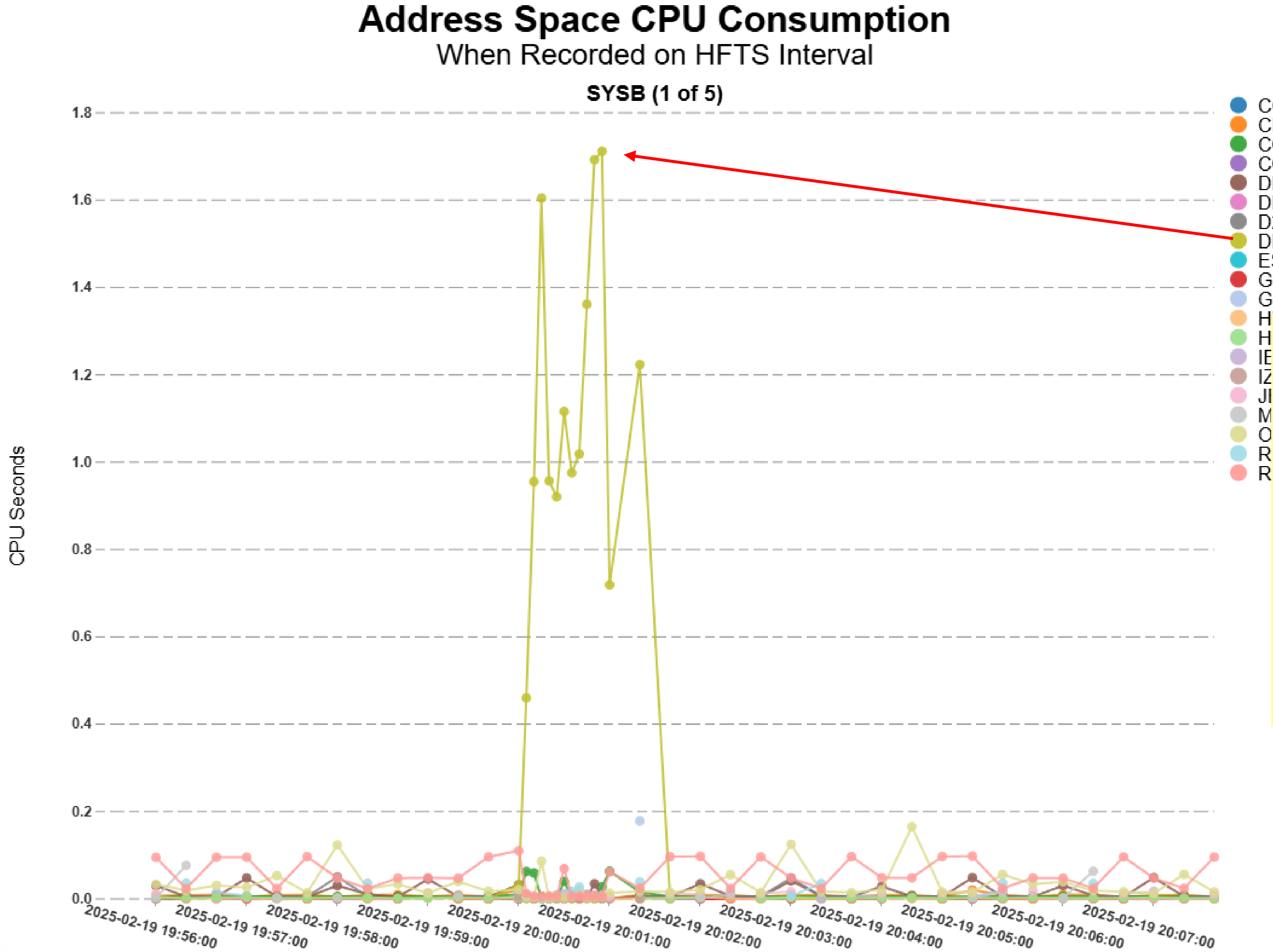


Not (usually) interesting, but there’s details in the 98s about spin and suspend locks too.



Here we have SYSTEM and STCMED getting busy for about 1 minute right at 20:00.

That may be interesting, but we'd like to know what address spaces were involved in that.

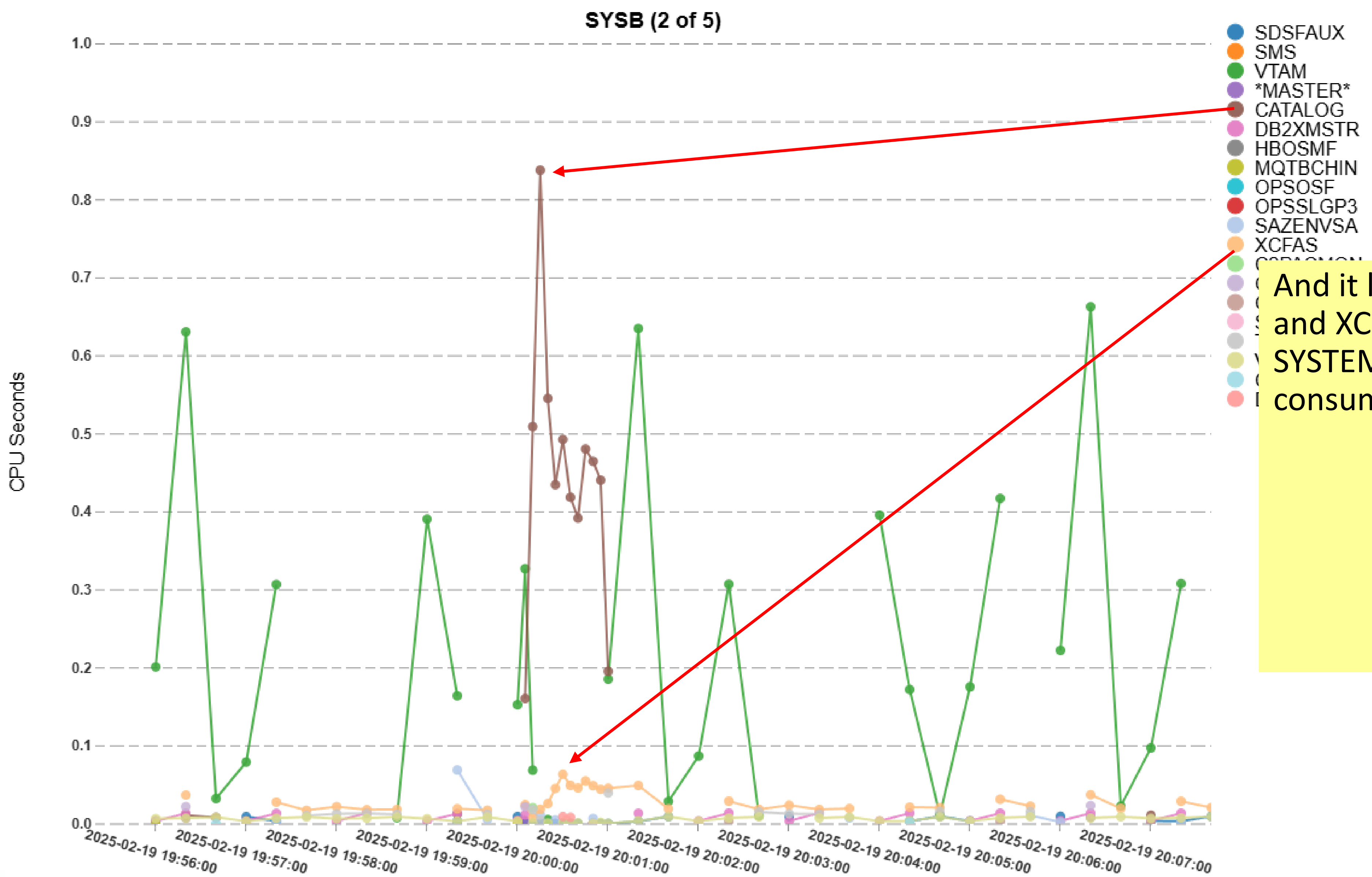


We can turn to the SMF 98 records to see if we can see what might have been involved.

In this case it appears that DFHSM was involved. That's the STCMED.

Address Space CPU Consumption

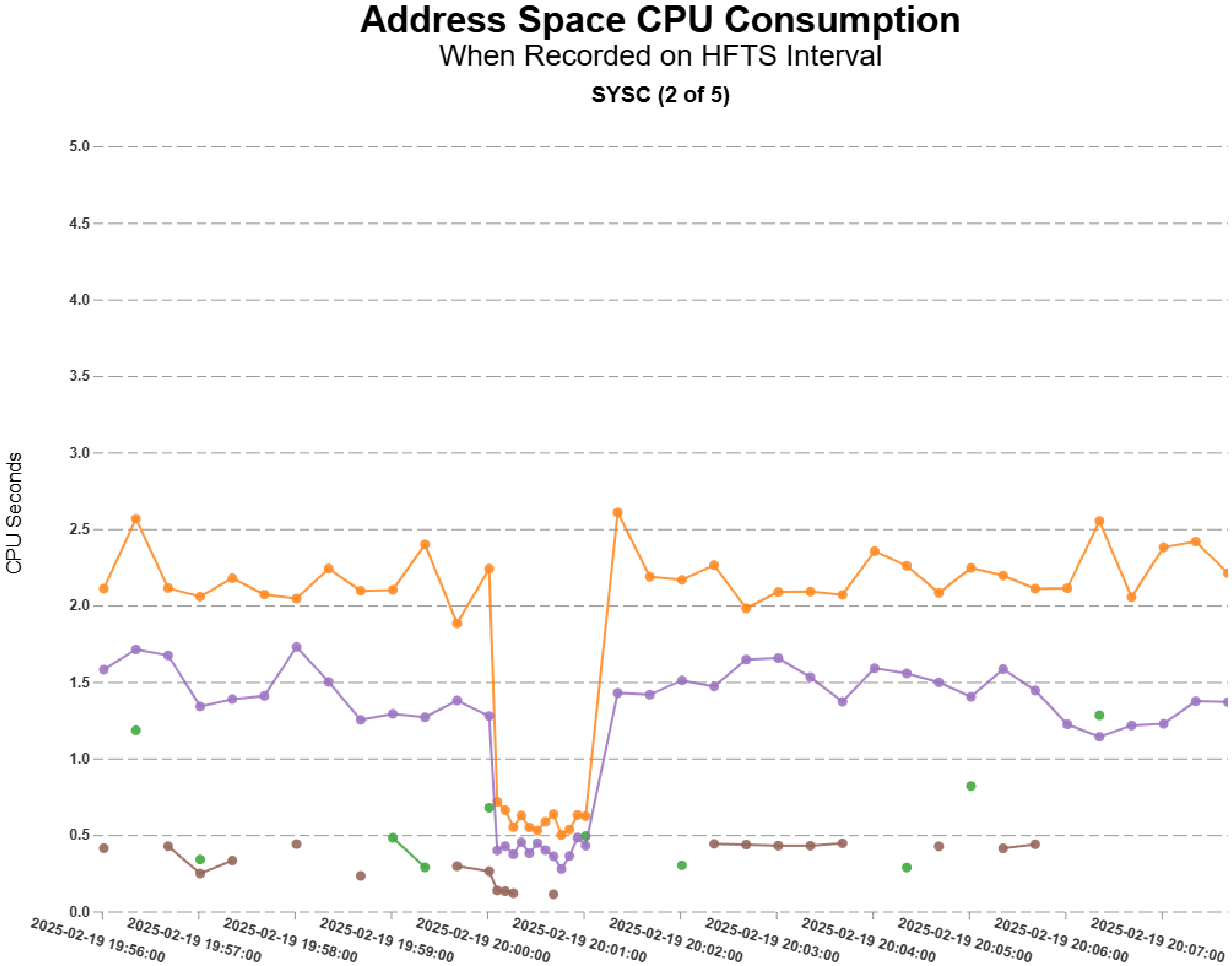
When Recorded on HFTS Interval



And it looks like Catalog and XCFAS were the SYSTEM things consuming more CPU

SMF 98 Address Space Data Notes

- Only some address spaces recorded each interval
  - The top address space according to some measure in some bucket/sub-bucket
  - Yes, this is confusing, not going into more detail here
- Lack of data does not mean that the address space was not using CPU!
  - Just means something else was the "top" address space
- CPU seconds recorded
- HFTS interval is user selected 5 to 60 seconds
  - But on the quarter hour it becomes 5 seconds
  - If using > 5 second intervals be aware of this as that drop in CPU is not what it seems



Note how the dots are closer together at the top of the hour.

That drop in in CPU used by TCP/IP and XCFAS by about 75% is due to the short intervals.

Also note VTAM and *MASTER* probably weren't idle when not shown… they just weren't recorded.

Do we really care?

- Most days you don’t care about this level of detail
- But if you’re investigating a problem, this can be helpful
- Understanding when low pool processors get parked/unparked goes to how many things can be running in the LPAR simultaneously
- Queue depths and dispatch delays help us understand why things might be delayed
  - Also are indicative of latent demand
- Sub-minute CPU measurements of Service Class Periods and Address Spaces can help:
  - Determine when an issue started/stopped
  - Determine what was consuming CPU, possibly at the expense of something else

Summary & Best Practices

- Your machine is probably busier than you think more often than you think
- Consider whether 15 minute RMF interval is still good for your environment
  - There may be benefit to going to a 10 or even 5 minute RMF interval
- SMF 98 and 99 data can be quite useful for diagnosing transient performance problems
  - Can be summarized to help understand how busy your system really is
- The SMF 98 and 99 data is not that large relative to modern systems

- **Record the SMF 98 and 99 data!**
  - If you're a Pivotor customer, add them to your daily transmission
  - If you're not, make sure you have reporting for the data

Real example record sizes

Small Environment
1 System
5s SMF 98

Rec Type	Recs	Bytes
type 098	17,281	400,665,728
type 099	267,523	882,010,880
subtype 099_001	8,766	94,304,416
subtype 099_002	198,189	681,960,832
subtype 099_006	8,634	24,409,416
subtype 099_008	8,634	37,919,432
subtype 099_012	43,014	43,272,084
subtype 099_014	286	144,716
type 030	75,242	127,542,360
type 072	8,352	12,758,400
type 074	3,840	109,555,584
type 119	263,496	91,755,080

Medium Environment
1 System
15s SMF 98

Rec Type	Recs	Bytes
type 098	6,528	155,556,832
type 099	699,570	1,250,404,352
subtype 099_001	9,456	132,003,656
subtype 099_002	315,780	807,837,248
subtype 099_003	140,600	44,485,216
subtype 099_006	8,640	34,461,172
subtype 099_008	8,640	35,088,776
subtype 099_010	2	16,320
subtype 099_011	288	586,656
subtype 099_012	43,175	47,578,848
subtype 099_013	172,699	148,151,024
subtype 099_014	290	195,460
type 030	160,380	315,278,400
type 042	1,375,384	846,715,968
type 072	9,397	17,668,548
type 074	10,976	306,955,648
type 100	5,764	10,221,576
type 101	2,730,869	1,539,210,112
type 102	3,612	7,482,643
type 110	33,683	521,378,528
type 116	570,456	3,054,344,960

Large Environment
9 Systems
20s SMF 98

Rec Type	Recs	Bytes	Bytes/system
type 098	50,079	1,287,791,488	143,087,943
type 099	6,184,194	12,406,321,152	1,378,480,128
subtype 099_001	87,294	1,237,456,896	137,495,211
subtype 099_002	2,424,340	7,715,728,384	857,303,154
subtype 099_003	1,416,581	565,265,280	62,807,253
subtype 099_005	5,437	989,534	109,948
subtype 099_006	82,912	258,727,040	28,747,449
subtype 099_008	82,912	250,740,240	27,860,027
subtype 099_010	4	32,640	3,627
subtype 099_011	1,796	13,252,116	1,472,457
subtype 099_012	414,453	474,274,272	52,697,141
subtype 099_013	1,665,662	1,886,882,560	209,653,618
subtype 099_014	2,803	2,972,085	330,232
type 030	1,067,142	2,495,804,672	277,311,630
type 042	6,755,265	4,795,271,168	532,807,908
type 070	3,419	71,175,576	7,908,397
type 072	152,332	271,338,848	30,148,761
type 074	58,185	1,114,929,792	123,881,088
type 100	404,310	702,351,808	
type 101	307,838,528	172,787,793,920	
type 102	33,389,874	13,666,774,016	
type 110	53,572,512	465,821,106,176	
type 119	47,320,428	26,800,586,752	