

Can We All Share Fairly? Detection and Remediation of inter-LPAR Performance Impacts

Scott Chapman
Enterprise Performance Strategies
Scott.Chapman@epstrategies.com



SHARE Orlando
February 2026
Session TECH_136

Session 44850

Contact, Copyright, and Trademarks



Questions?

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

Copyright Notice:

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

Trademarks:

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

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

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

Other trademarks and registered trademarks may exist in this presentation

Abstract



“We have production and development in different LPARs so they can’t impact each other’s performance. Right?” Well... that’s not entirely true: there are a number of ways that one LPAR’s activity may impact another LPAR’s performance. While some of these ways are examples of poor configuration decisions, most are a result of necessary trade-offs while running in a shared environment. If you’ve ever wondered how one LPAR may impact the performance of work in another LPAR or wondered how to best protect LPARs from such interference then this session is for you. Scott Chapman will discuss performance measures that might tip you off to a problem as well as configuration options that can help protect your most important systems.

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
 - March 30 – April 3, 2026 (4 days, excl Wednesday the 1st)
- WLM Performance and Re-evaluating Goals
 - June 22 – 26, 2026 (4 days, excl Wednesday the 24th)
- Parallel Sysplex and z/OS Performance Tuning
 - May 12-13 2026
- Also... please make sure you are signed up for our free monthly z/OS educational webinars! (email contact@epstrategies.com)

Like what you see?



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

All Charts (132 reports, 258 charts)

All charts in this reportset.

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

Charts containing more than the threshold number of exceptions

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

Charts containing any number of exceptions

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

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

EPS presentations this week



What	Who	When	Where
z/OS Performance Management If You Only Have 20 Minutes A Day	Scott Chapman	Mon 9:45	Salon 14
PSP: z/OS Performance Tuning - Some Top Things You May Not Know	Peter Enrico Scott Chapman	Tue 13:15	Salon 18
Planning Your Next Mainframe Processor Upgrade in 2026	Scott Chapman	Tue 15:45	Salon 15
Processor MSU Consumption Analysis	Peter Enrico	Wed 13:15	Salon 14
Can We All Share Fairly? Detection and Remediation of inter-LPAR Performance Impacts	Scott Chapman	Wed 14:30	Salon 14
Standard z/OS Measurements When Monitoring Transactions	Peter Enrico	Thu 14:30	Salon 19

Agenda



- LPAR Basics
- Inter-LPAR impacts: Within a CEC
- Inter-LPAR impacts: Across a sysplex



LPAR Basics

Splitting up the Mainframe



- Modern mainframes run in LPAR mode using either PR/SM or DPM
 - Allows CEC resources to be divided up or shared among multiple operating system images (z/OS, ICF, Linux, VM, TPF, VSE)
 - PR/SM: Processor Resource / Systems Manager
 - Most common configuration
 - DPM: Dynamic Partition Manager
 - Used for Linux-only machines
- Logical Partitions (LPARs): system images run under the control of PR/SM
 - z/OS
 - z/VM
 - ICF – Coupling facility partition
 - IFL – Linux partition
 - TPF

Note you could run multiple z/OS systems under z/VM too, but there's probably little reason to do that today.

LPARs provide significant separation



- PR/SM's LPARs are certified at the Common Criteria EAL5+ level
 - Since the z800 in 2003 (reportedly first server to achieve EAL5 for partitioning)
 - AWS's Nitro hypervisor apparently EAL4+

EAL4: Methodically designed, tested and reviewed

EAL5: Semi-formally designed and tested

+ = "augmented" with assurances beyond the minimum required

- But ... just because LPARs provide a high degree of security doesn't mean that an LPAR might not feel a *performance* impact from another LPAR
 - LPARs are running in a shared environment
 - LPARs share things between them
 - Therefore, what one LPAR's activity can impact the performance of another

Defining LPARs

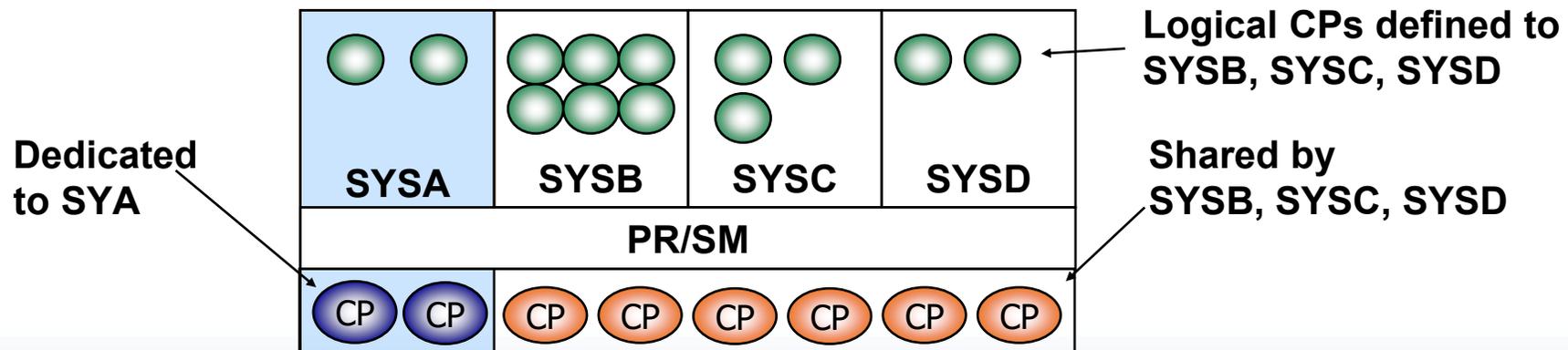


- On the HMC, you define for each LPAR a number of configuration values such as:
 - Number of processors (for each type of processor assigned to the LPAR)
 - Reserved processors can be defined to allow for future non-disruptive addition of processors
 - LPAR weight settings (for each type of processor assigned to the LPAR)
 - Capping setting
 - Central Storage (memory)
 - Various other controls that we'll be less concerned about for this presentation
 - Security controls can impact RMF's ability to see other LPAR's configuration data as well as the collection of Hardware Instrumentation Services samples

Sharing Processors



- Each system image has some number of logical processors assigned
- Dedicated processors
 - Physical processor dedicated to a partition 100% of the time
 - Used primarily for CF and z/VM Linux partitions, rarely for z/OS
- Shared logical processors
 - Physical processor that can be shared among one or more partitions
 - Not dedicated to a particular partition

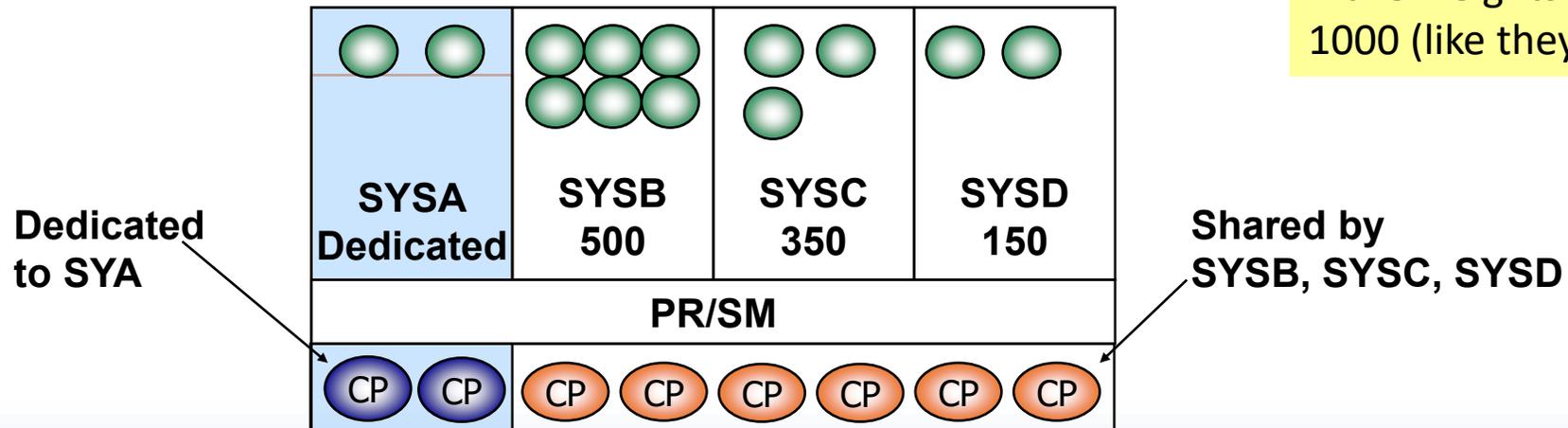


LPAR Weight Enforcement



- When there is contention for the shared physical processors, PR/SM enforces the weights assigned to each partition
- Each LPAR with shared CPs is guaranteed to get at least its share of the shared CPs
 - $LPAR\ Share = 100 * \frac{LPAR\ Weight}{\sum Weight\ of\ activated\ LPARS}$
- In below example:
 - SYSB – guaranteed 50% of shared physical processors
 - SYSC – guaranteed 35% of shared physical processors
 - SYSD – guaranteed 15% of shared physical processors

For ease of use, try to make weights add up to 1000 (like they do here).



Reminder About Physical Processors



- PR/SM dispatches cores to LPARs
- Therefore, a core (aka a CP) can only be serving a single LPAR at any one moment in time
 - Even if a zIIP has SMT enabled—the core as a whole is dispatched between LPARs
- And for a GP, only 1 task can be using the CP at a time
 - zIIPs with SMT enabled could have 2 tasks from the LPAR running simultaneously
- **If your mainframe has 3 CPs enabled, at any one moment in time, only 3 tasks are running across *all the LPARs* on the machine**
 - This is one reason why more/slower CPs are often better than fewer/faster
 - Both a 3932-O04 and 3932-T02 are 210 MSUs: do you want to be able to have 2 concurrently executing tasks or 4?

Vertical CP Management



- HiperDispatch manages CPs “vertically”, meaning it endeavors to make the logical CPs a larger percentage of a physical
- Logical processors classified as:
 - High – The processor is essentially dedicated to the LPAR (100% share)
 - Medium – Share between 0% and 100% (typically 50% and 100%)
 - Low – Unneeded to satisfy LPAR’s weight
- This processor classification is sometimes referred to as “vertical” or “polarity” or “pool”
 - E.G. Vertical High = VH = High Polarity = High Pool = HP
- Parked / Unparked
 - Initially, VL processors are “parked”: work is not dispatched to them
 - VL processors may become unparked (eligible for work) if there is demand and available capacity



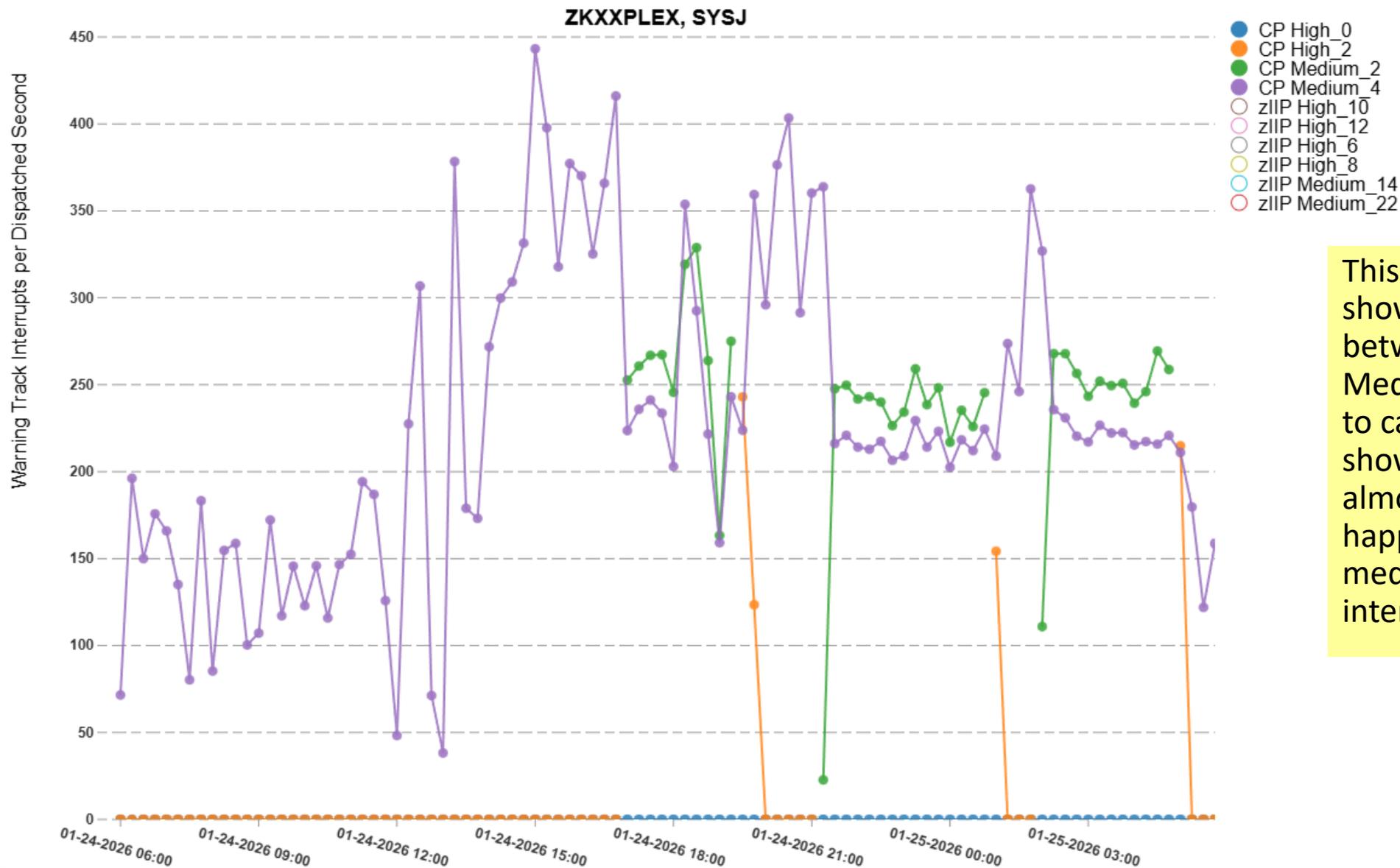
Impact from LPARs on the same CEC

Shared CPs



- Shared CPs will not always be dispatched to the LPAR
 - Will at times have to wait for a CP to become available
- PR/SM time slice documented as 12.5-25ms, but seems much shorter today
 - Low single-digit ms? Seems to vary quite a bit
- Impact mostly limited to medium & low polarity CPs
 - High-pool CPs enjoy a 100ms time slice from PR/SM
 - Won't typically see PR/SM taking a high pool CP away from an LPAR
- Hard to quantify this impact since it's happening at the millisecond scale
 - Warning Track Interrupt numbers can show how often PR/SM is asking to take the CP away from the LPAR

Warning Track Interrupts per Dispatched Second



This interesting example shows CP #2 switching between High and Medium (probably due to capping). The WTIs shown when it's "high" almost certainly actually happened when it was medium during the interval.

Shared CPs... what to do?



- Buy more CPs
 - More/slower CPs are almost always better than fewer faster
 - More things can run at once
 - More L1/L2 cache
 - Often better ratio of effective capacity to rated MSUs
 - If you only have medium CPs, they're not really fast CPs, they're short CPs
 - See also my session from Tuesday
- Try to ensure important production LPARs have at least 1 high pool CP
 - Can be hard to do, especially on machines with only a couple of physical CPs

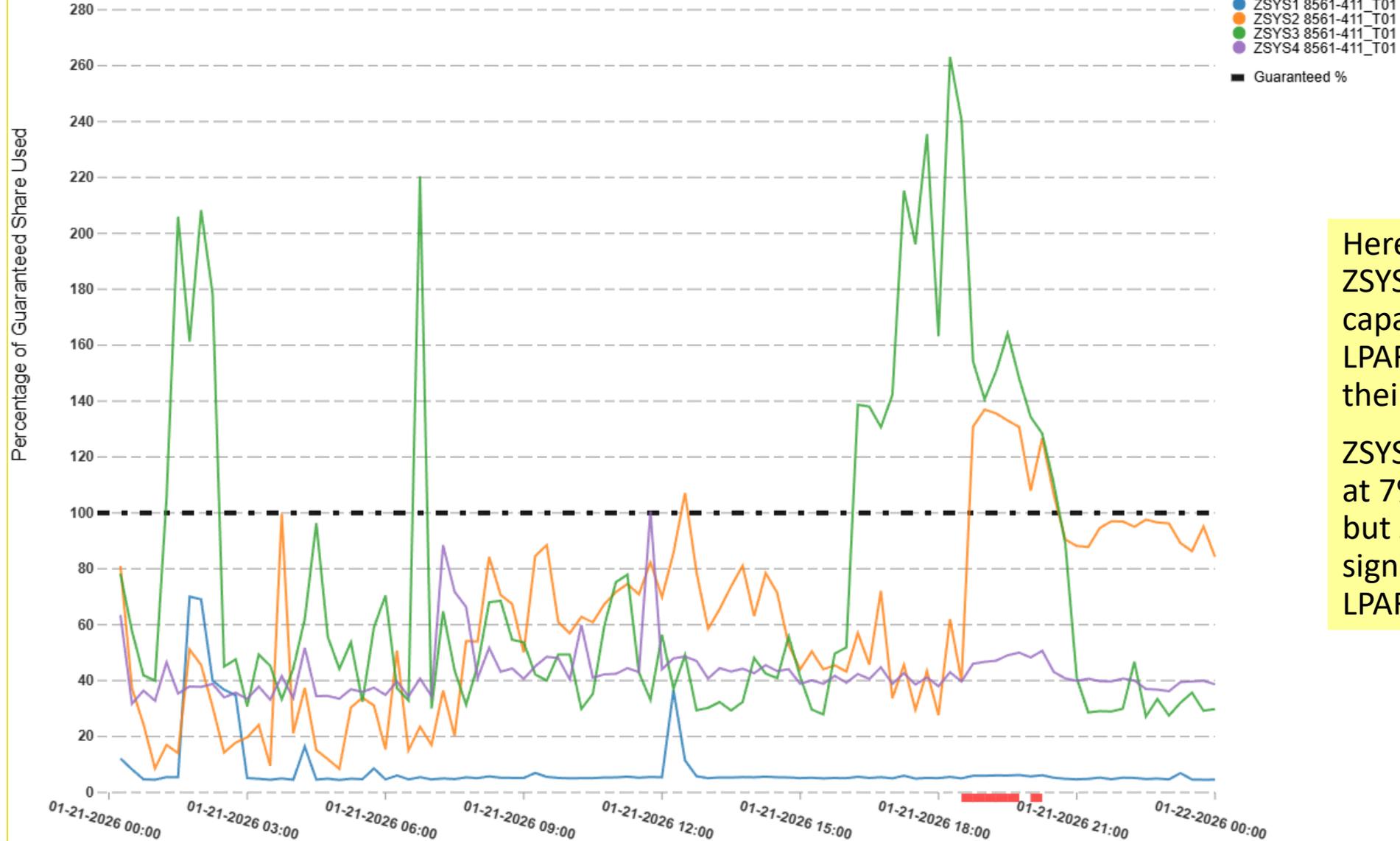
Borrowing Weight



- Most customers have their LPARs configured such that an LPAR is allowed to use more than its “fair share” defined by its weight if the other LPARs aren’t using full share
- LPARs regularly “borrowing” capacity are at risk if the other LPARs have a sudden demand for their capacity
 - E.G. you let production borrow from test overnight, but if one night test gets busy... you may find that you don’t meet your batch window because the production LPAR couldn’t use its full weight as expected
- Especially problematic when you have multiple production LPARs on the CEC with different peaks to manage
- An LPAR running above its weight is running on low pool CPs
 - Which will be inefficient at least initially
 - Can be very inefficient if on a different drawer

CEC Percent CP Weight Used

A480B



Here, LPARs ZSYS3 and ZSYS2 are “borrowing” capacity from the other LPARs that aren’t using their full share.

ZSYS3 is relatively small at 7% of the machine, but ZSYS1 and ZSYS2 are significant production LPARs.

Borrowed Weight: What to do?



- Set your weights appropriately to what the LPARs' need
- If workload varies during the day use automation to change weights to match expectations
 - E.G. shift weight for overnight vs daytime processing or for unusual events
 - Can be done via automation: see BCPii
 - <https://www.ibm.com/docs/en/zos/2.5.0?topic=bcp-ii-setup-installation>
 - <https://github.com/IBM/zOS-BCPii>

Group Capping

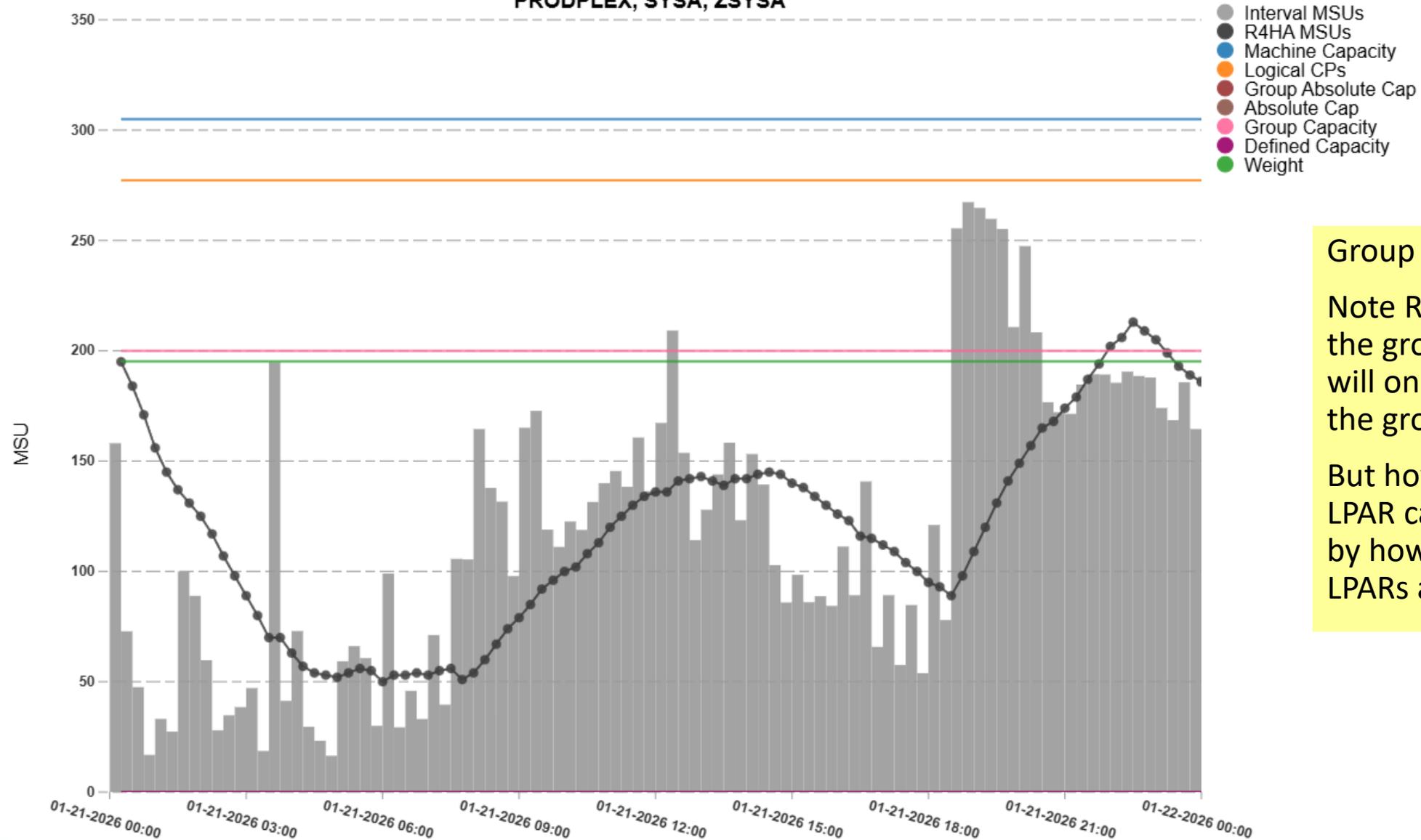


- Group caps are very handy for controlling your R4HA to limit your MLC costs
 - Only pay for the group cap limit, not the actual peak R4HA
- When the R4HA of the group exceeds the limit, entire group is capped
 - Essentially via phantom weight, so LPARs get access to less machine capacity
- All LPARs in the group potentially impacted
 - Of course if the LPAR doesn't want to use the capacity, moot point
- Having all LPARs in a CEC-wide group cap provides best cost control
 - But then if those dev/test LPARs get busy they can trigger the group cap

LPAR Limits and Utilization

Expressed as MSUs

PRODPLEX, SYSA, ZSYSA



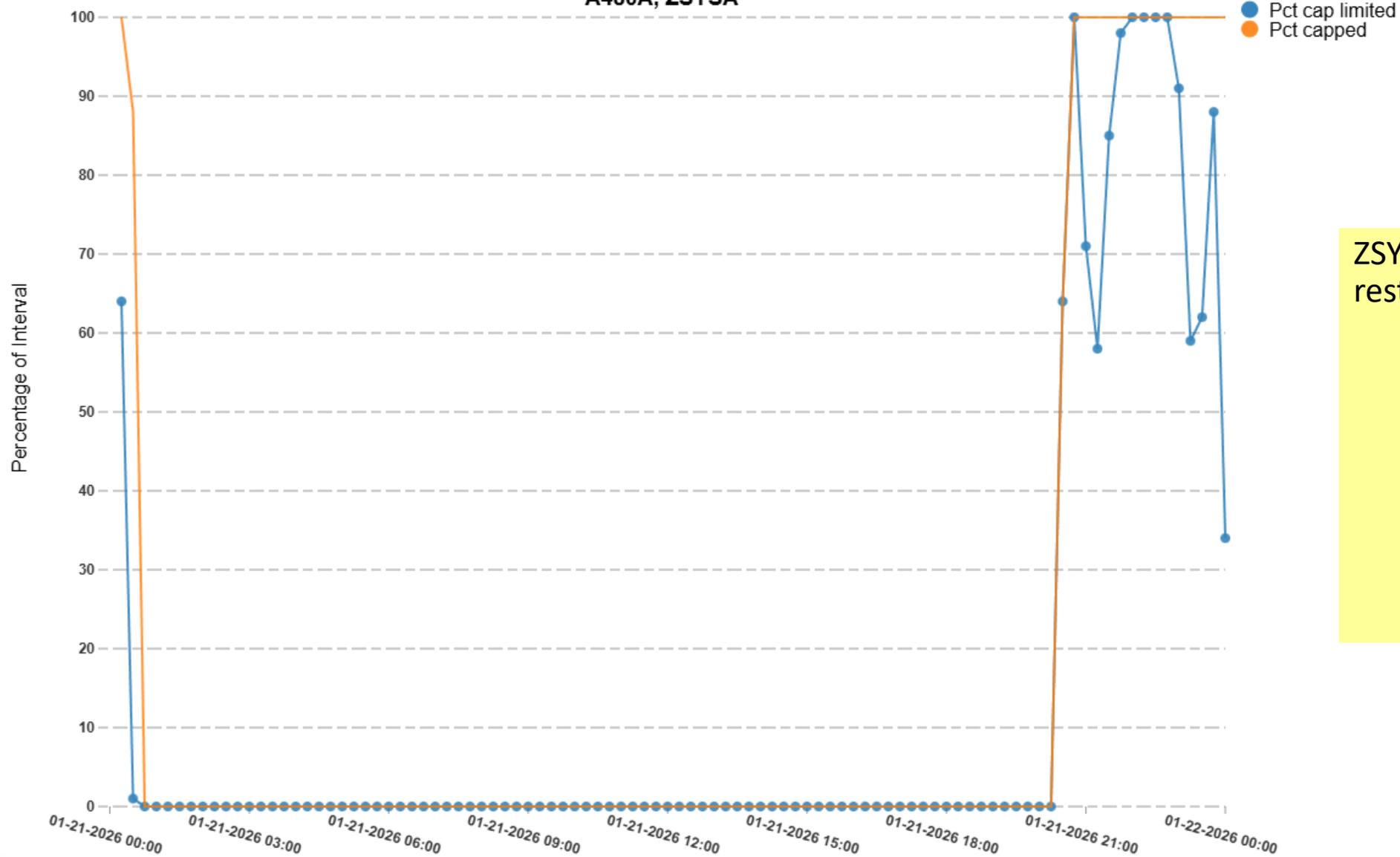
Group cap is pink line.

Note R4HA goes above the group cap line, but will only be charged for the group cap limit.

But how much CPU this LPAR can get to is limited by how much the other LPARs are using too.

CEC Capping Actually Limited Percentage vs. Percent Considered Capped

A480A, ZSYSA

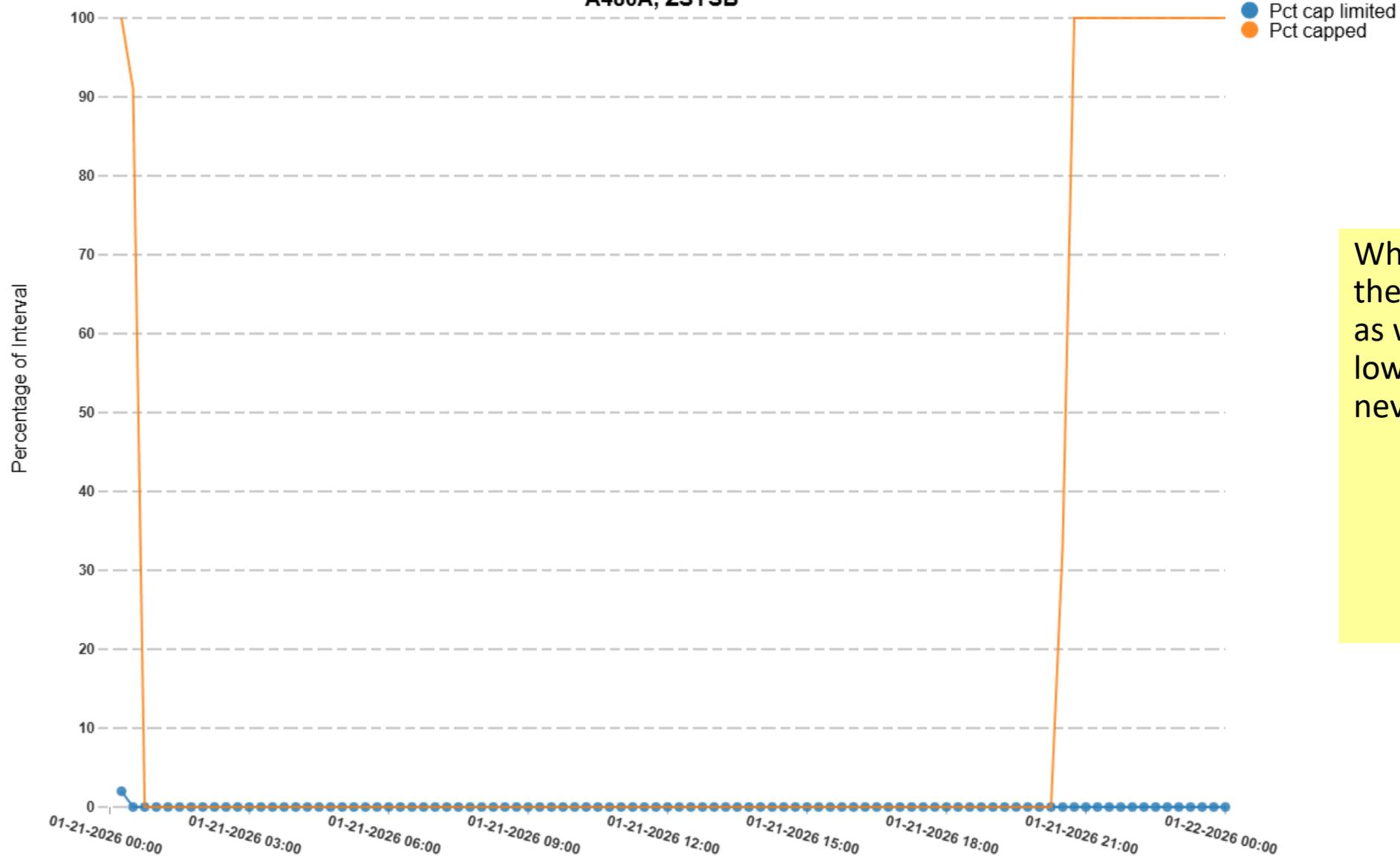


ZSYSA is actually being restricted by the capping



CEC Capping Actually Limited Percentage vs. Percent Considered Capped

A480A, ZSYSB



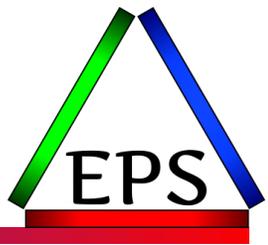
While ZSYSB was part of the group and so capped as well, its demand was low enough that the cap never impacted it.

Group Capping... what to do?



- Raise the cap!
 - Of course you put the cap in place to save money, so that may not be palatable
- If you went to Tailored Fit Pricing... do you need the cap anymore?
 - Maybe, if you have ISVs still holding you to R4HA
 - Otherwise: did you remember to remove the caps?
- If you still need the group cap
 - Consider defined capacity limits for individual LPARs, especially dev/test
 - Don't let a runaway test job drive up the cap unnecessarily
 - May consider other controls too

Shared Channels



- Rarely *should* be a problem but usually all the DASD/Tape channels are shared by all the LPARs, especially in small to medium sized configurations
- FICON channels today are fast and can handle multiple concurrent I/Os
 - But they are shared
- Occasionally we see overly busy channels
 - Usually that's where they've got 2 or maybe 4 channels to the DASD
- What to do? Don't cheap out on channels!
 - But 4 is like twice what we think we need!
 - But at some point you'll want to migrate, or will have a failure, or...
 - And channels are not expensive... just buy 8 / controller-CEC combo
 - To be fair... don't buy channels you're not going to connect
 - Did have a customer get into trouble with that once for obscure reasons



Impact from LPARs in same Sysplex

Global Enqueues

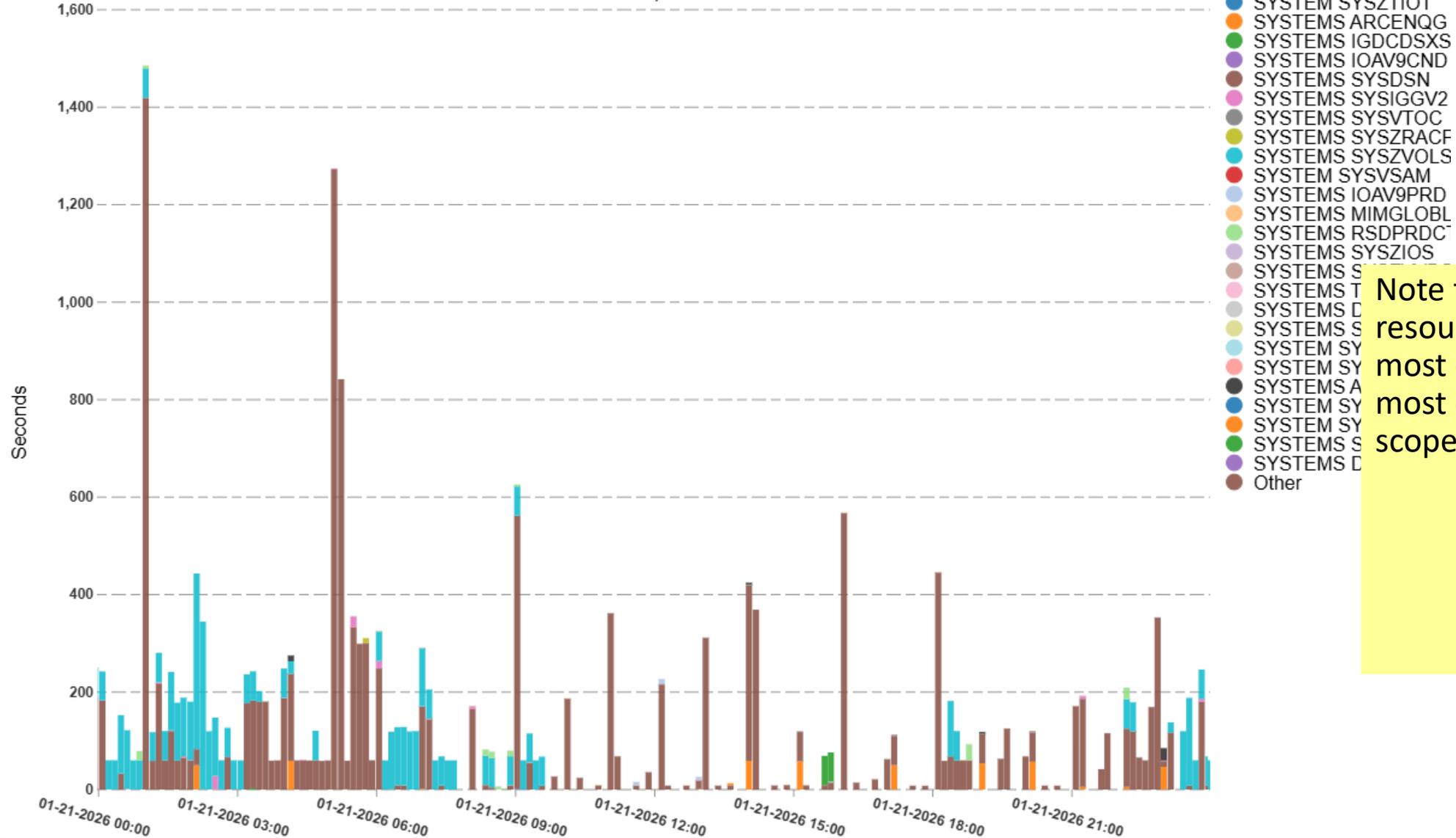


- One of z/OS's most under-appreciated features is GRS
 - Global Resource Serialization improves system and data integrity, usually with little effort on the part of the application programmer
- GRS Locks (enqueues) managed in scopes:
 - STEP: single address space
 - SYSTEM: single system
 - SYSTEMS: across all the systems in the GRSplex (likely sysplex)
- So... for those SYSTEMS scope enqueues, activity on one system can impact access to those resources on another system
 - We've probably all seen jobs waiting for a dataset that another job has

Enq - Total Cont Time Over Time

by Major Name

PLEXB, SYSB



Note that for the resource types with the most contention time, most of those have a scope of SYSTEMS

Enqueues: what to do?



- First, be aware: know how to look for the enqueue delays in your monitoring and reporting
- As with all locking mechanisms: deadlocks can occur
 - Can require scheduling or design changes
- Sometimes applications request unnecessary level of locking
 - E.G. DISP=OLD (exclusive) vs DISP=SHR (read)
- Make GRS is well configured (rarely a problem today)
 - E.G. GRS-Ring will be less efficient than GRS-Star
- Make sure common services are performing well
 - E.G. Optimize Catalog and HSM performance to avoid waits
 - Both Catalog and HSM can take advantage of RLS to improve performance

Coupling Facility Performance

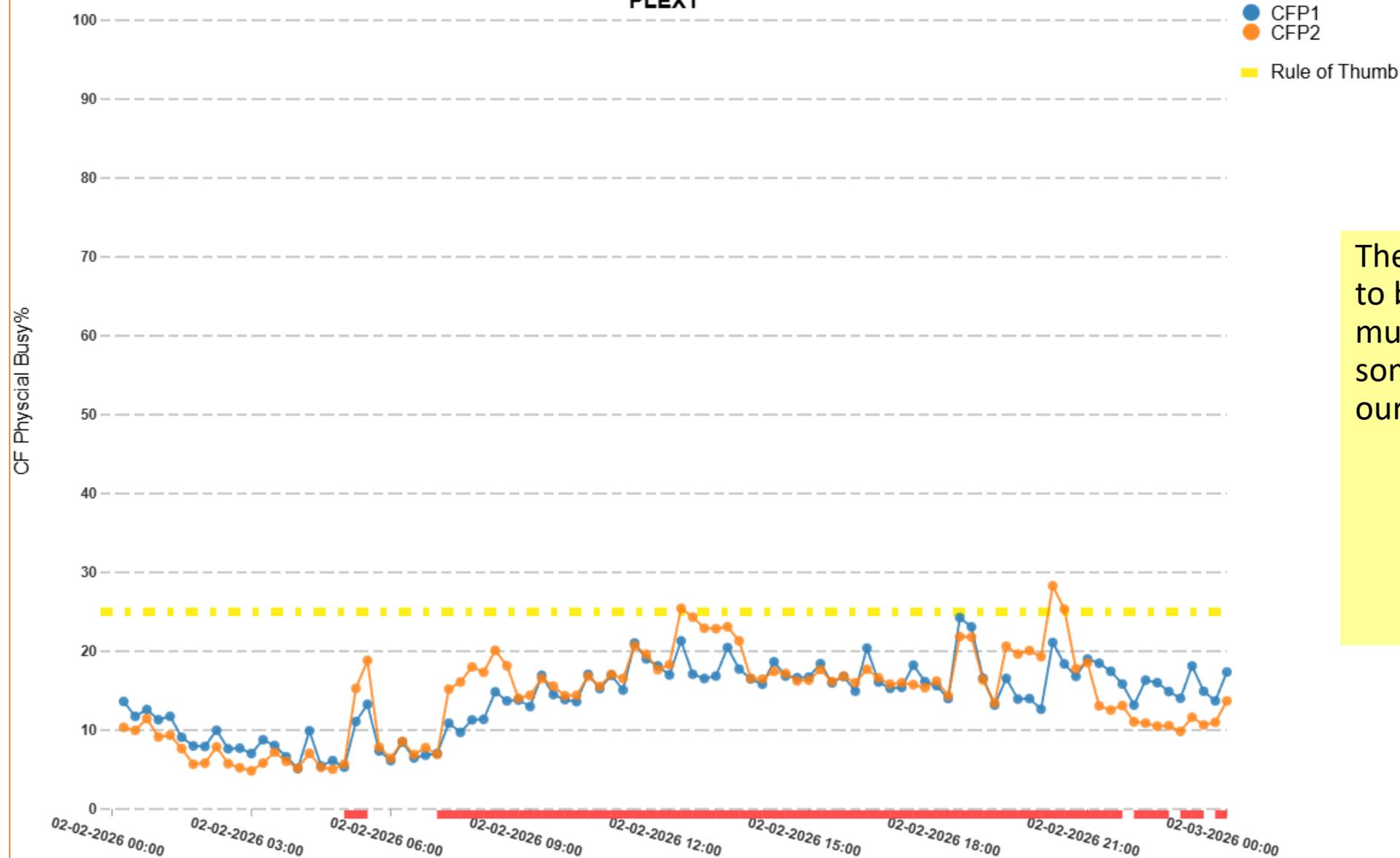


- In a parallel sysplex, Coupling Facility (CF) performance can impact performance across multiple systems
 - CF performance impacted by volume of requests coming from the different systems
- Today, this is less of a problem than it used to be
- Opportunities for improvements still sometimes are found though
 - CF engines should be run at quite low utilization levels
 - “Traditional” lock duplexing can greatly impact performance

CF CPU - CF Processor Busy Utilization

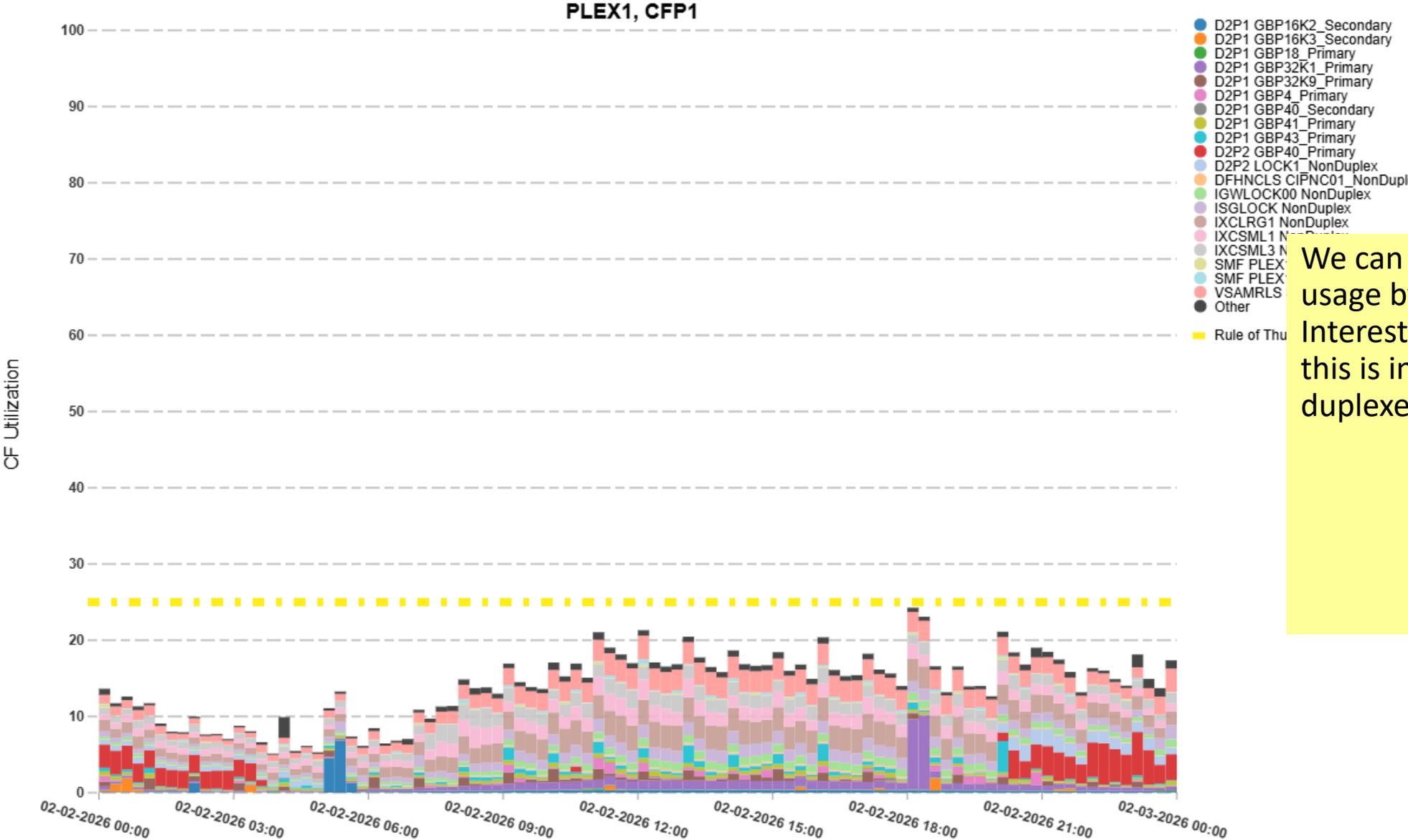
As a Percentage of Physical Processor

PLEX1



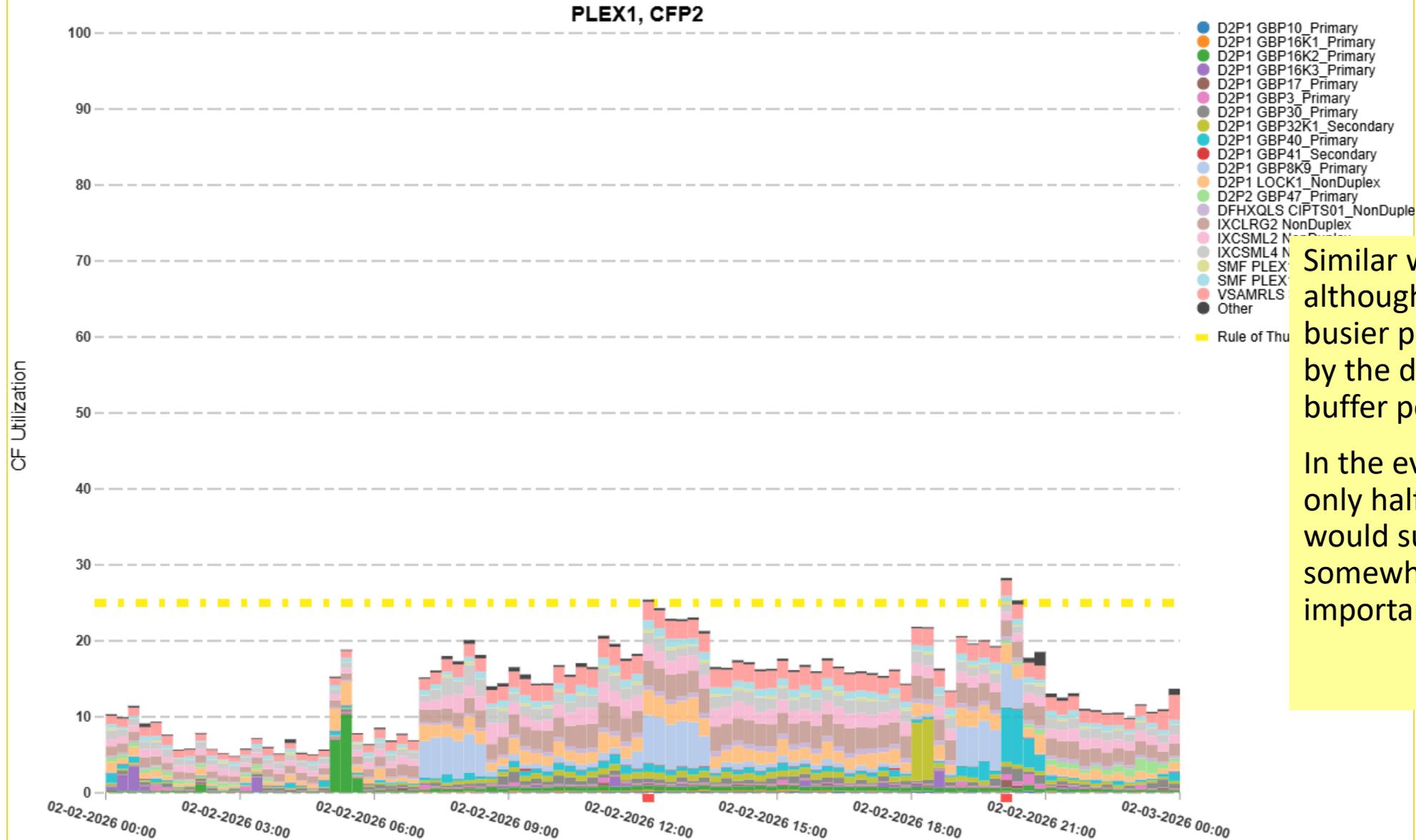
These two CFs do seem to be relatively busy for much of the day, even sometimes breaching our 25% mark.

CF CPU - Utilization by Top Structures



We can attribute the usage by structure. Interesting that a lot of this is in fact non-duplexed structures.

CF CPU - Utilization by Top Structures



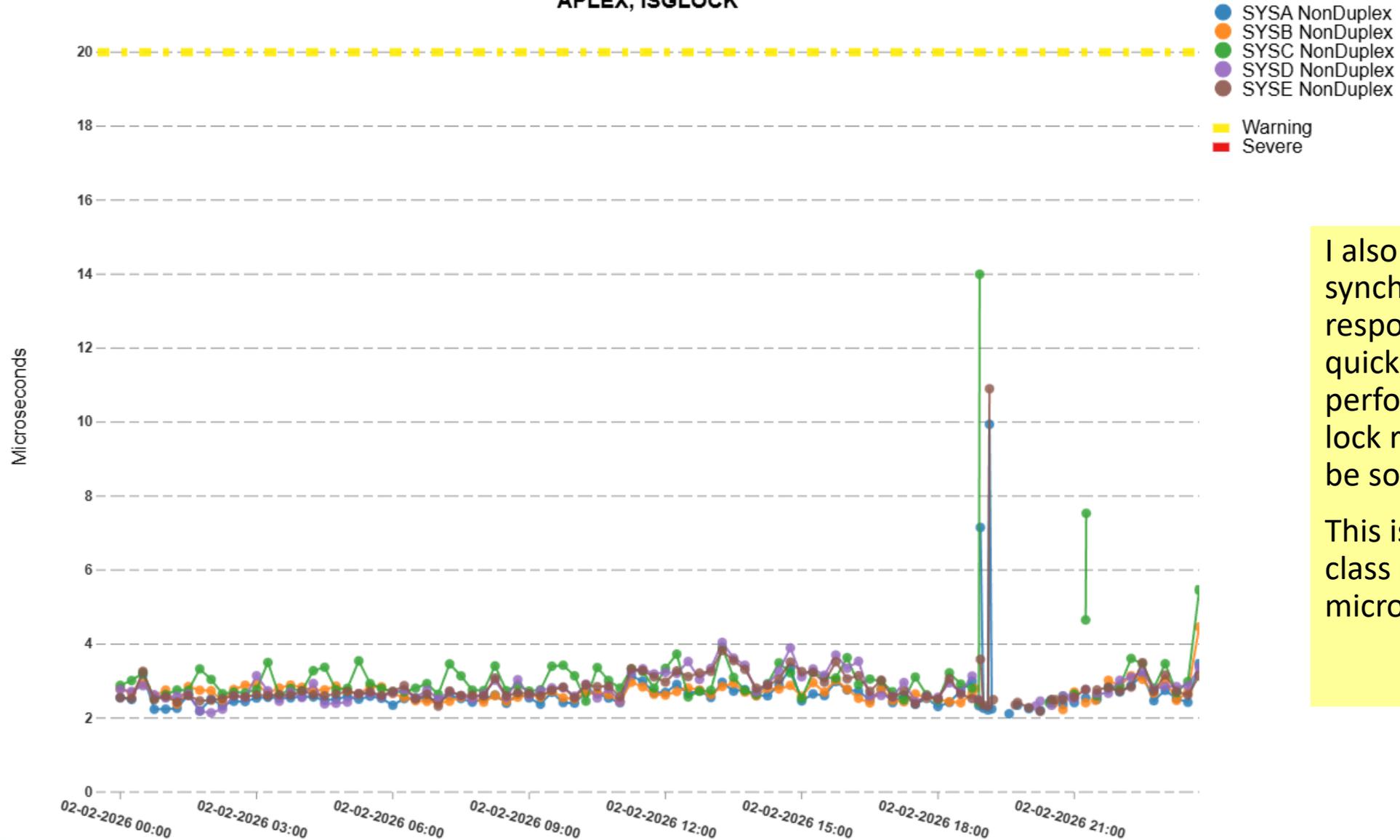
Similar with CFP2, although some of those busier periods are driven by the duplexed group buffer pools.

In the event of a failure, only half of that activity would survive, so that's somewhat less important to us.

CF Lock - Synchronous Response Time

(All systems)

APLEX, ISGLOCK



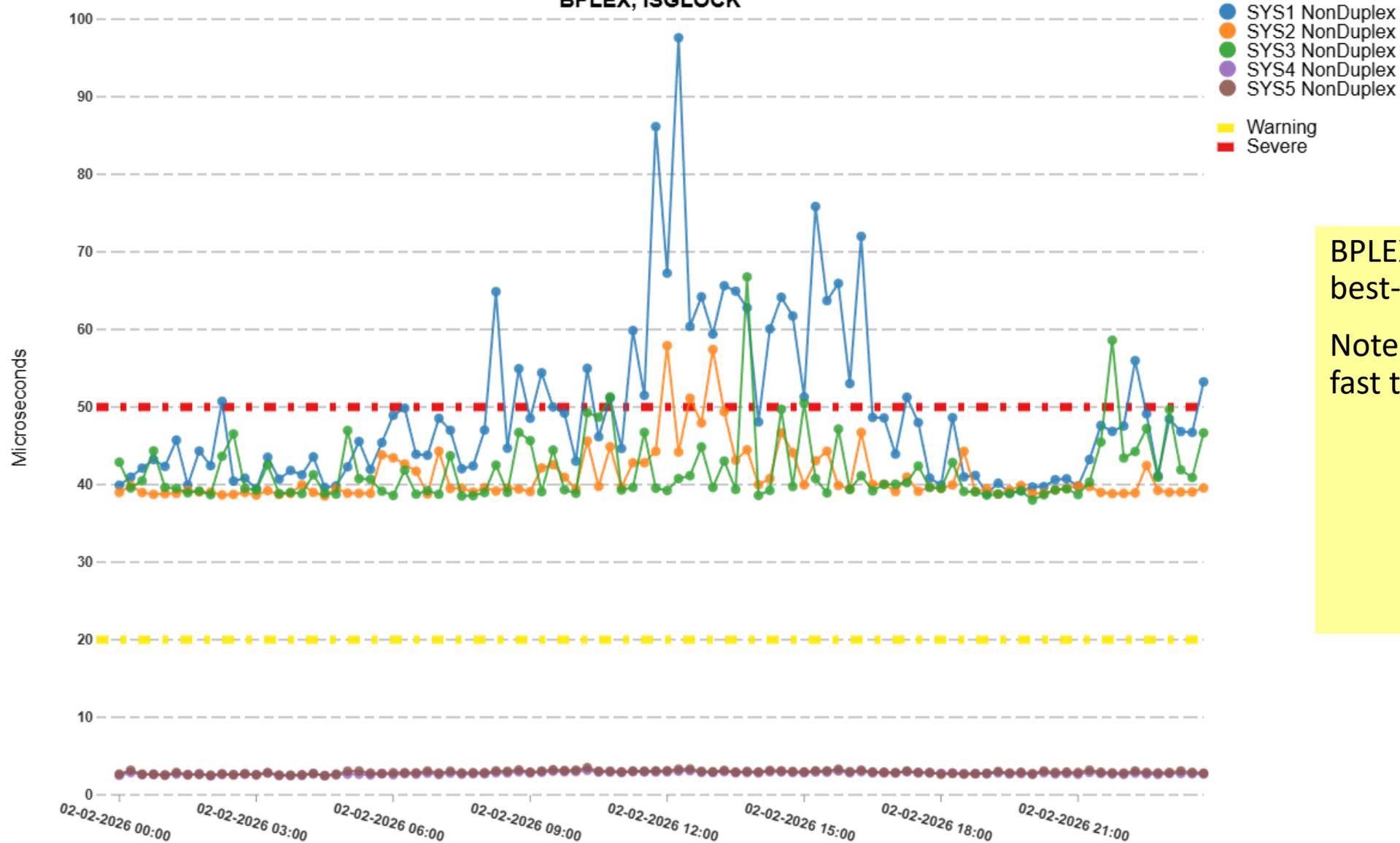
I also like looking at synchronous lock response time to as a quick sysplex performance check since lock response times can be so important.

This is currently best-in-class performance: 2-4 microseconds

CF Lock - Synchronous Response Time

(All systems)

BPLEX, ISGLOCK



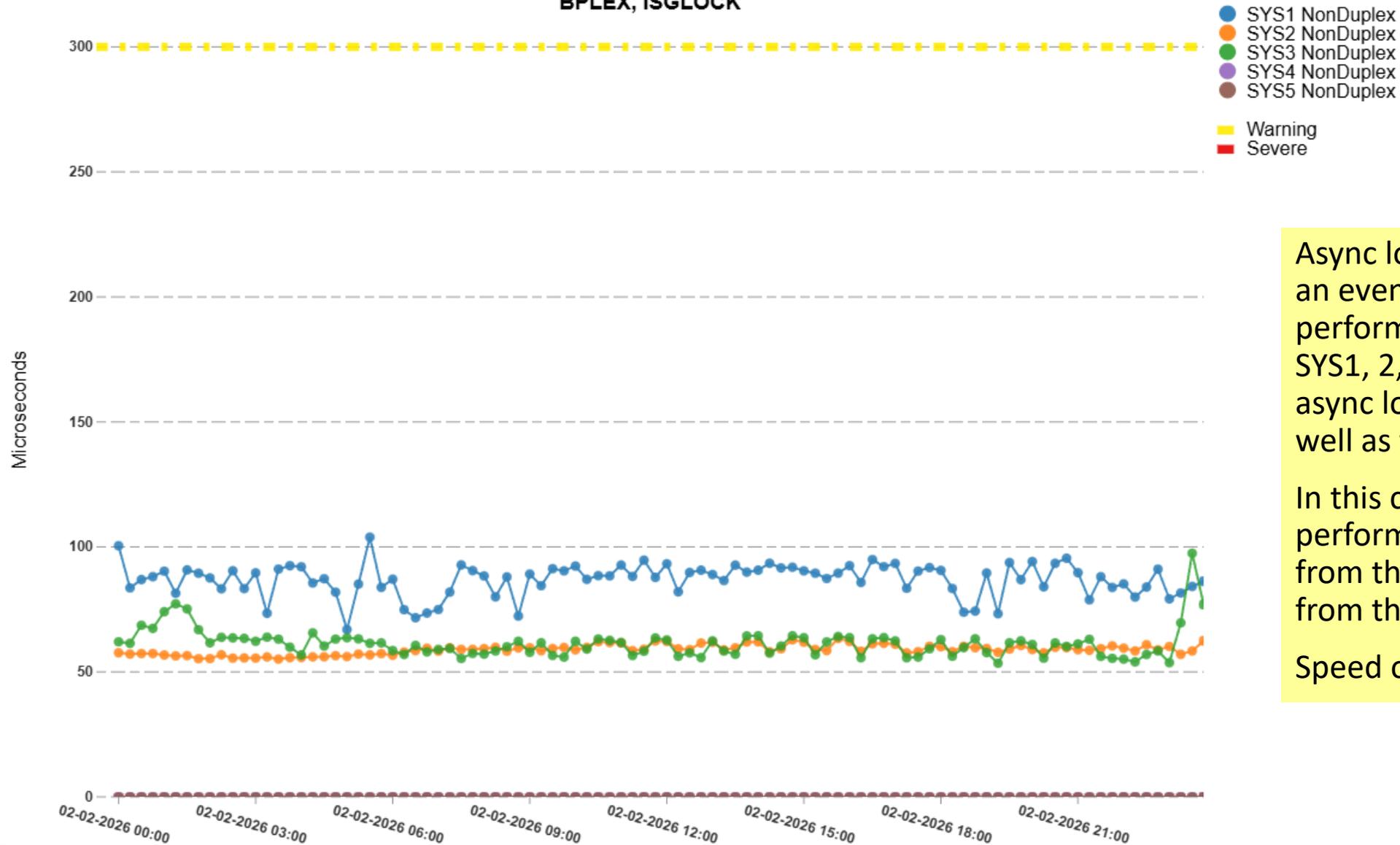
BPLEX is (mostly) not best-in-class at 40+ mics.

Note SYS5 seems to be fast though.

CF Lock - Asynchronous Response Time

(All systems)

BPLEX, ISGLOCK



Async locks are of course an even bigger performance hit. Note SYS1, 2, and 3 are doing async locks requests as well as the sync requests.

In this case the performance issue comes from the CF being 2-3km from the SYS1,2,3 LPARs.

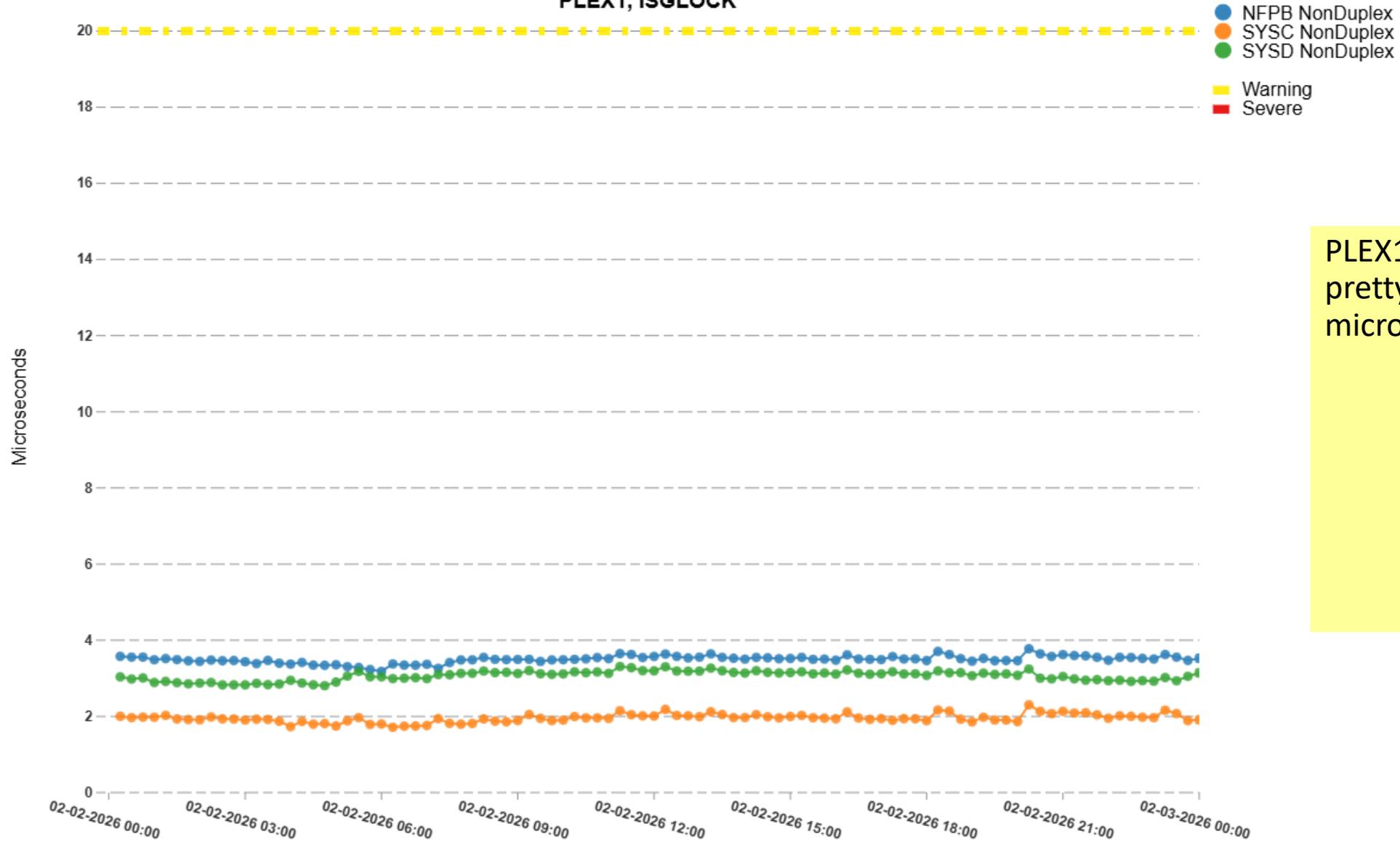
Speed of light matters!



CF Lock - Synchronous Response Time

(All systems)

PLEX1, ISGLOCK

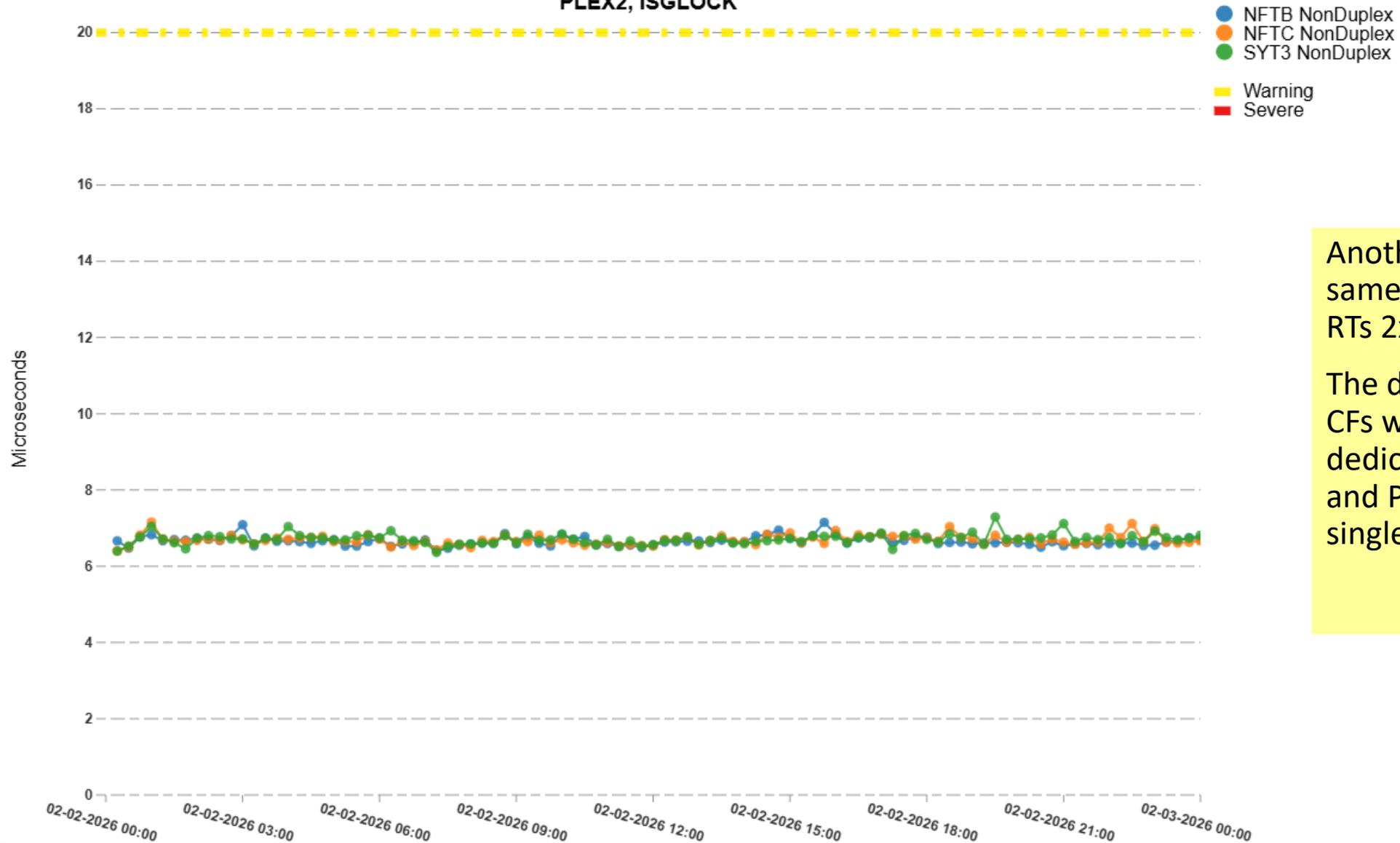


PLEX1's sync lock RTs are pretty good in that 2-4 microsecond range.

CF Lock - Synchronous Response Time

(All systems)

PLEX2, ISGLOCK



Another plex on the same hardware, but sync RTs 2x of PLEX1.

The difference? PLEX1's CFs were on multiple dedicated CF engines, and PLEX2's CFs had single shared engines.

CF Performance: What to do?



- CF Performance usually isn't a huge problem
 - Even in these examples
- Sometimes sub-optimal performance is the cost of doing business
 - I.E. distance will result in longer sync response times, but if you need the distance for resiliency reasons, that's what you have to live with
- CF engine configuration choices can make a difference though
 - Multiple dedicated engines is goodness
 - Sometimes the choice between single dedicated vs. multiple shared is hazier
- Traditional lock duplexing is also a big performance hit
 - Since z14: use async lock duplexing to significantly reduce the cost of lock duplexing
 - Or think about whether lock duplexing is really necessary for your business

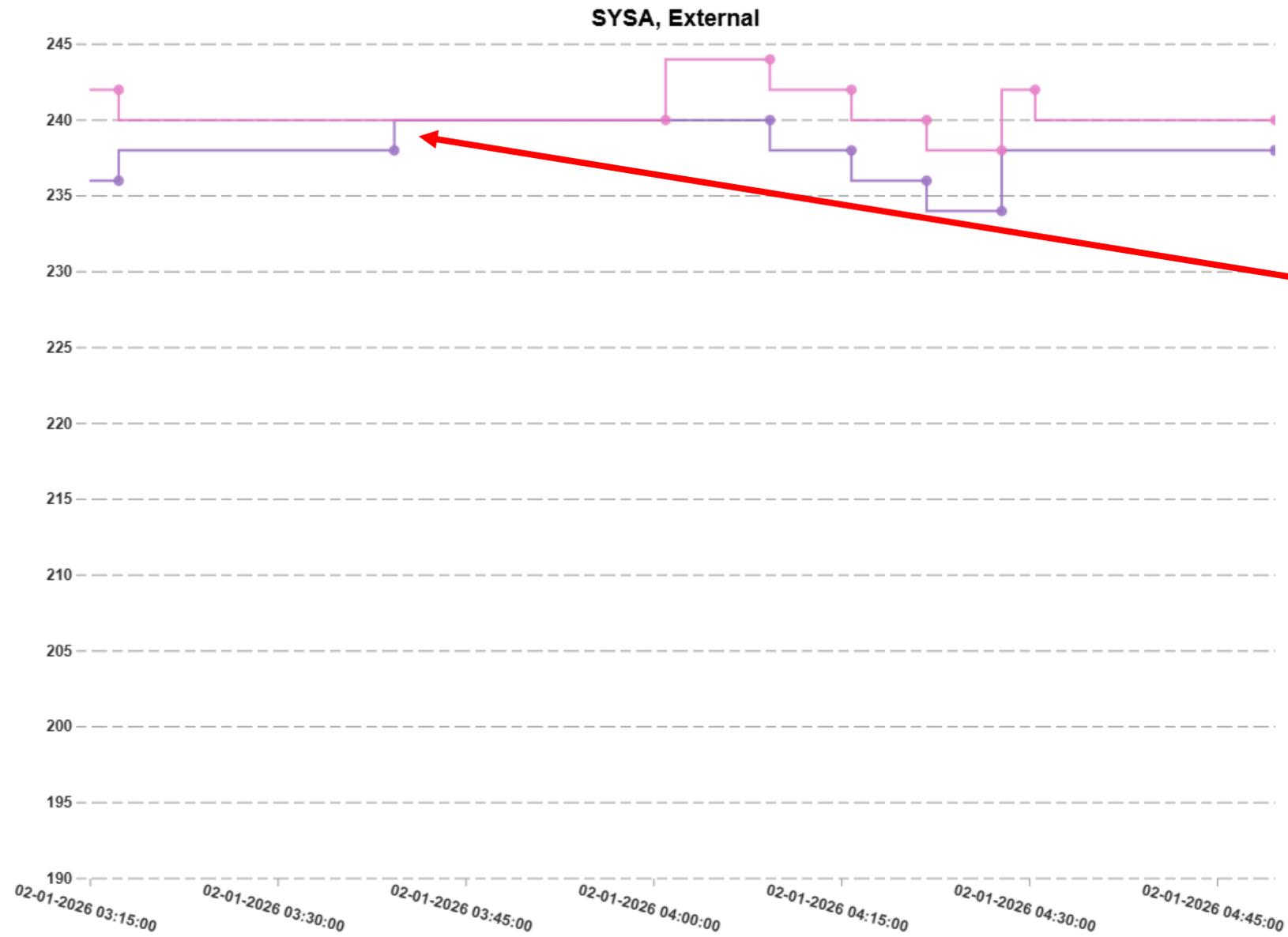
WLM: Sysplex PI



- WLM evaluates sysplex PI before system PI
 - Intent is to prioritize overall performance across the sysplex
 - The decision to help or not help work on one system can be influenced by how the work is performing on other systems
- This was a good intention in the mid 90s when WLM came out
 - Vision was for sysplexes composed of many small homogenous systems
 - With work dynamically spread across those systems without affinities
- Today reality for many sysplexes is:
 - Composed of heterogenous systems with different work and/or capacity
 - E.G. dev, test, and production may all be in the same sysplex
 - Different systems may have access to different capacity
 - Not all workloads distribute all their work dynamically across the sysplex



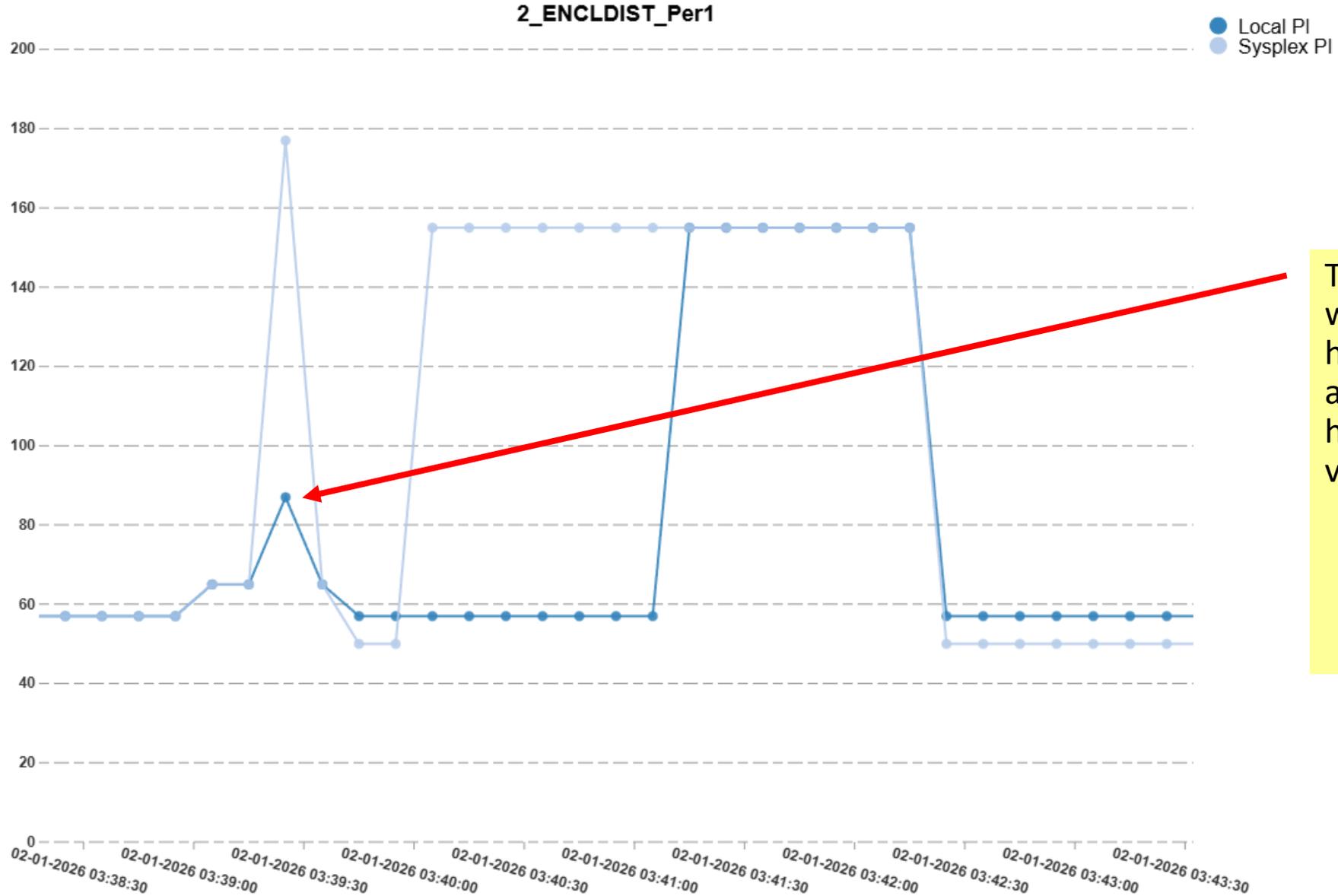
WLM SMF 99.6 - CPU Dispatching Priority



- 0 \$SRMBEST_Per1
- 0 \$SRMG00D_Per1
- 1 PSERVE_Per1
- 2 ENCLDIST_Per1
- 2 ONLPRD2_Per1
- 2 ONLPRDT1_Per1
- 2 ONLPRD_Per1
- 2 TSOPRD_Per1
- 3 ENCLBTCH_Per1
- 3 ENCLDIST_Per2
- 3 PRDBATHI_Per1
- 3 TSOPRD_Per2
- 4 ENCL
- 4 PRDB
- 4 STCM
- 4 TSOP
- 4 TSTB
- 5 ENCL
- 5 TSTB
- 5 USSP
- 6 \$SRM

Note the increase in dispatching priority here for ENCLDIST to allow it to compete with ONLPRDT1. What was its PI?

WLM SMF 99.6 Data Explorer

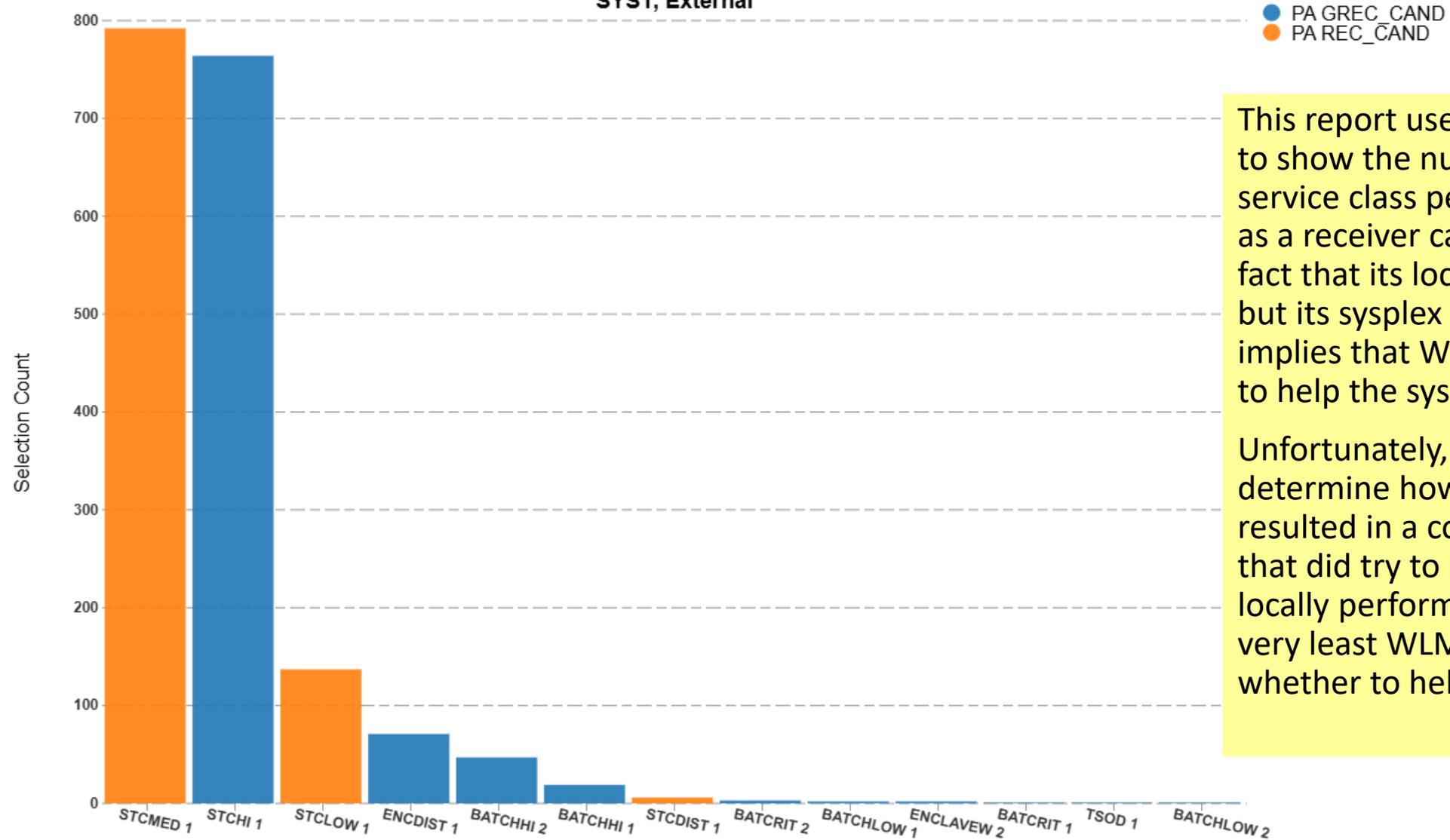


Turns out that the issue was the sysplex PI was high, but the local PI was actually under 1. (Shown here as the raw internal value of 100.)

WLM PA - Top Sysplex PI Receiver Candidates

(Policy Adjustment Decisions)

SYS1, External



This report uses the 99.1 records to show the number of times a service class period was selected as a receiver candidate despite the fact that its local PI was 1 or less, but its sysplex PI was over 1. This implies that WLM selected it to try to help the sysplex PI.

Unfortunately, it's not easy to determine how many of these resulted in a committed change that did try to help that already locally performant SCP. But at the very least WLM was investigating whether to help them.

Sysplex PI: What to do?

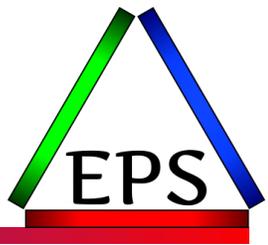


- First: this is usually not a big issue!
 - Except sometimes Sysplex PI does drive decisions more than I expected
 - More likely to be a concern when:
 - The sysplex includes both prod and non-prod systems
 - The systems in the sysplex have very different access to processor capacity
 - Where the work in a service class is not evenly spread distributed across the systems
- Maybe don't include prod and non/prod in the same sysplex
 - This is of course not trivial to change
- Use different service classes on different systems
 - This is easier to do for some work types than others
 - Can use System qualifier for STC, OMVS, and TSO
 - Can use Subsystem instance for CICS, DB2, MQ, etc.
 - May end up with more complicated classification rules
- Ask IBM for an option to look at local PIs first (or even maybe only!)
 - **Vote for the idea:** <https://ibm-z-hardware-and-operating-systems.ideas.ibm.com/ideas/ZOS-I-4771>



Wrap-up

In summary...



- LPARs provide secure separation, but it is still a shared environment
- More CPs to share is generally more better
- Make sure your weights reflect the correct “fair” shares
- Group caps can be good, but also look to protect prod from non-prod
- Don’t cheap out on channels
- Locking is there to help you, even if it doesn’t always feel like it
- CF configuration can impact performance across the sysplex
- Sysplex PI can result in suboptimal local WLM decisions

Your feedback is important!

Submit a session evaluation for each session you attend:

www.share.org/evaluation

Session 44850



SHARE Orlando
February 2026
Session TECH_135



Thanks!
Questions?

Email: scott.chapman@epstrategies.com