System Recovery Boost: Hitting the Turbo Button on z/OS

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

SRB is one of the more interesting things that IBM has introduced with the z15 and certainly can help certain installations shut down, start up, and recover faster. But what are the practical implications of using SRB? How does enabling the different flavors of SRB influence both the performance and measurement of the systems, even potentially the systems that aren't being boosted? Join Scott Chapman in this session as he explores SRB and discusses why System Recovery Boost is like Solid Rocket Boosters: exciting and getting you where you're going quickly, but needs to be handled with some care.

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
  - Via Zoom, June 21-25, 2021
- WLM Performance and Re-evaluating Goals
  - Via Zoom, September 20-24, 2021
- Parallel Sysplex and z/OS Performance Tuning
  - Via Zoom, November 16-17, 2021
- Also… please make sure you are signed up for our free monthly z/OS educational webinars!

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

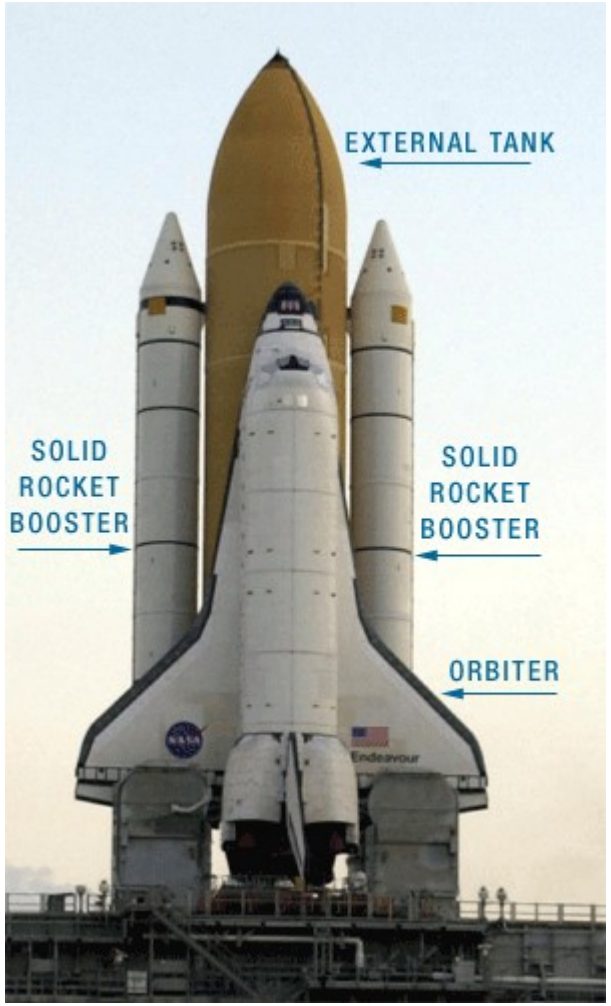
- We also have a free Pivotor offering available as well
  - 1 System, SMF 70-72 only, 7 Day retention
  - That still encompasses over 100 reports!

Agenda = Questions to be answered

- What is SRB and what types of boosts are available?
- When will boosts be triggered?
- Will this impact our software bill?
- Which LPARs will be impacted (directly or indirectly)?
- How is the SMF performance data impacted?

- Summary/Recommendations

What is SRB?
What are the types of boosts?



System Recover Boost

- Overall goal: provide additional capacity and performance to better recover from certain planned or unplanned events
  - Shutdown faster
  - Startup faster
  - Faster DR site switch
  - Faster sysplex recovery
- Requires z15 machine
- No additional charge for basic System Recovery Boost
  - Extra cost item discussed in just a moment

Types of boost

- Speed Boost: *GP engines run as full speed engines*
  - Applicable only to sub-capacity systems
  - I.E. this is of no use to 8561-7xx and 8562-Z0x systems
  - Only boosted LPARs run as full speed!

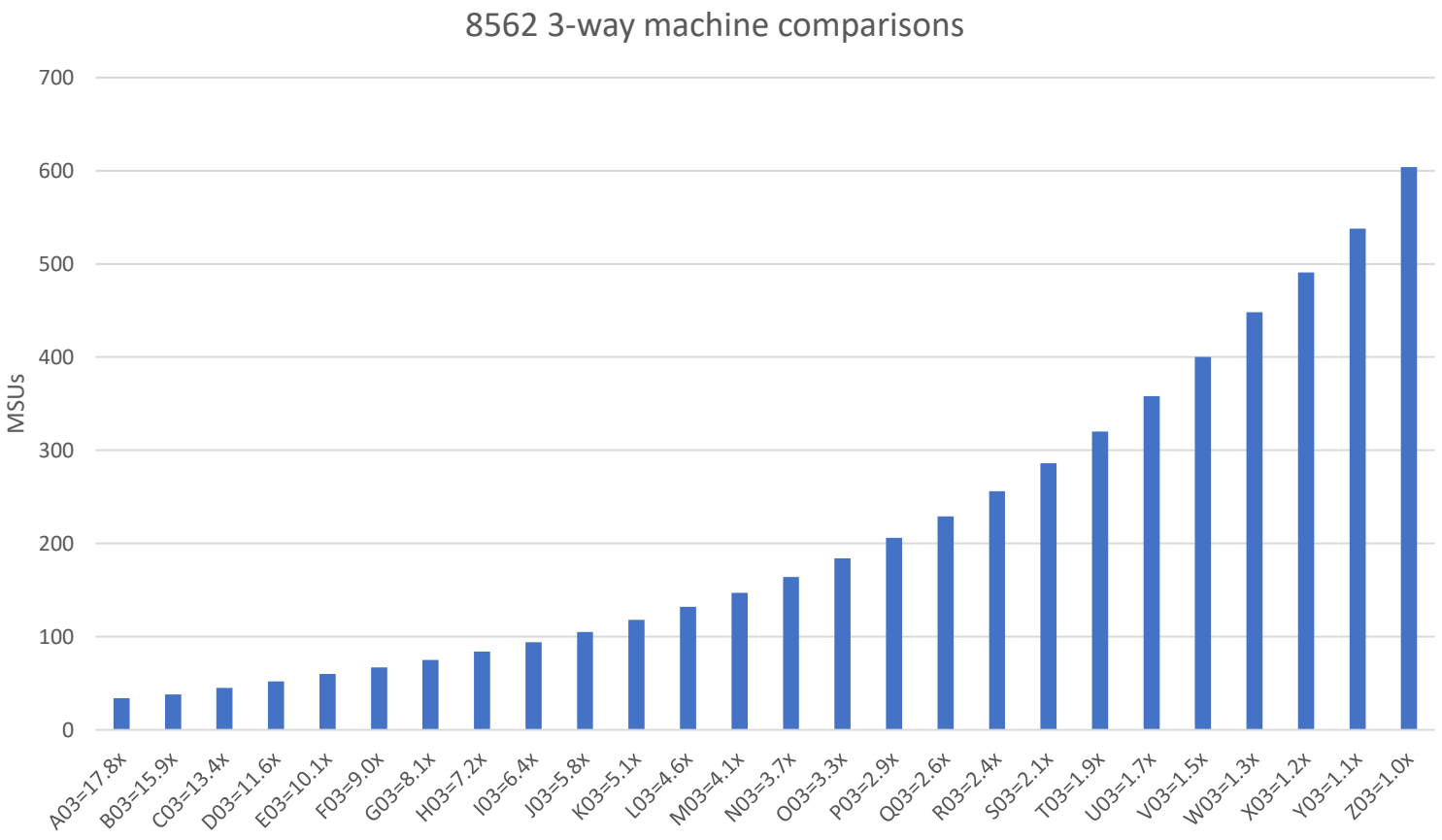
- zIIP Boost: *zIIPs will run work not eligible to run on zIIPs*
  - I.E. ***any*** workload may be dispatched to a zIIP
  - Reserved (but physically available) zIIPs brought online to the boosted LPAR
  - Note that zIIPs always run full speed
  - What if you want more zIIPs for zIIP boost?

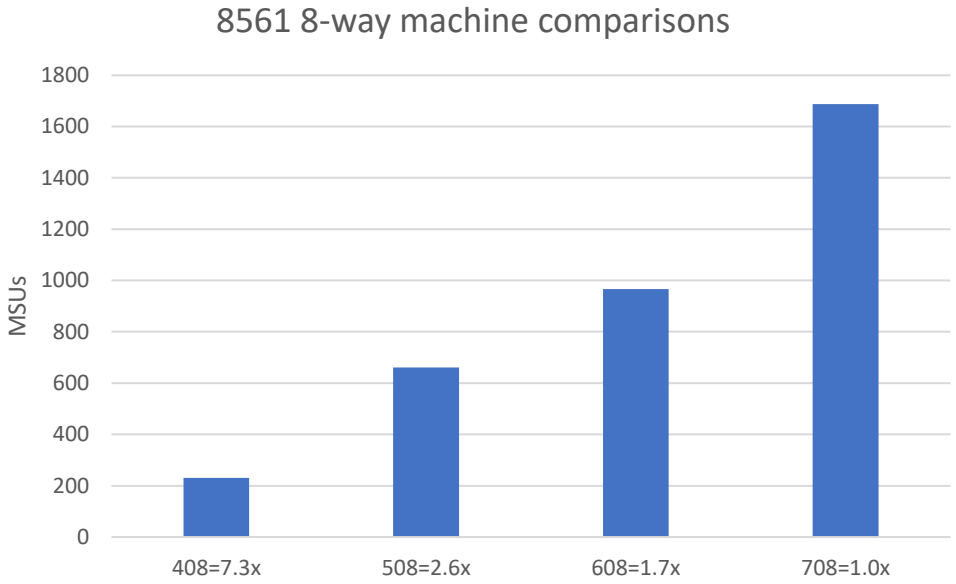
System Recovery Boost Upgrade Record

- Pay IBM some money and you can have more zIIPs for boost periods!
- Subscription feature that allows for activation of uncharacterized PUs as zIIPs during boost periods
  - Can add up to 20 zIIPs
  - Does allow for violating the 2 to 1 zIIP to GCP ratio
- Handled much like a capacity upgrade record
  - Boost Upgrade record is activated in 6 hour increments for the machine
  - But it appears boosts are still limited to their normal time frames
- Primarily expected to be useful for larger 7xx systems with multiple LPARs that may need to be IPLed in short succession

How big of a speed boost?

- Customers on slower sub-capacity machines could see a huge speed boost!





Plus zIIP boost could add even more capacity!

When will boosts be triggered?

Boost classes = triggering events

Class	Where	Duration	WLM Work Routing
Shutdown	Single system	30 minutes	Routes work away from boosted system
IPL	Single system	60 minutes	Routes work to boosted system
Recovery Process	Multiple systems	1-3 minutes	No change

- Recovery Processes
  - CF data sharing member recovery
    - Triggered by disconnection from lock structure while locks were held
  - Sysplex partitioning
    - System removed from a sysplex
  - CF structure recovery
    - Structure rebuild, duplex failover, or reduplexing
  - HyperSwap
    - Recovery from storage controller failover

IPL Boost

- Enabled by default
  - Can be controlled in IEASYSxx with BOOST=SYSTEM|ZIIP|SPEED|NONE
    - Selecting ZIIP or SPEED means to only use that type of boost
  - For most customers the default (both) is probably appropriate
- Lasts for 60 minutes
  - Very early IPL only benefits from speed boost because it's single-threaded
  - Hour apparently starts when WLM comes up and zIIP boost is (possibly) enabled

```
13:32:22.40  *MASTER*  IEE252I MEMBER IEASYM00 FOUND IN …
13:32:29.18  *MASTER*  IEA681I IPL speed boost is active
13:33:08.33  *MASTER*  IEA675I IPL zIIP boost is active with 1 transient zIIP cores
13:33:08.59  WLM       IWM064I BOOST ACTIVATED.
14:33:08.35  *MASTER*  IEA678I All IPL boosts have ended
14:33:08.62  WLM       IWM064I BOOST ENDED.
```

- WLM Routing Services will report boosted capacity
  - May result in more than normal work being routed to boosted LPAR

Shutdown Boost

- Initiated by running proc IEASDBS
  - Obviously need to change your procedures to do this
  - Shutdown boost can not be invoked again until the LPAR is IPLed
- WLM starts avoiding the LPAR for new work
- Lasts for 30 minutes, until LPAR comes down, or IEABE is run
  - At end of 30 minutes, lose boost advantage but WLM will continue to avoid directing work to the LPAR
  - Must figure out when it makes most sense to run IEASDBS to improve shutdown time: probably just before the timeframe of the most CPU consumption even if it means final shutdown steps run unboosted
  - But don't want to run it too far in advance because of the first point
- Respects BOOST= in IEASYSxx

Recovery Process

- Automatically initiated by the system
  - Again, respects BOOST= in IEASYSxx
  - Won't be initiated if the LPAR in IPL or Shutdown boost
- Short duration (<5 minutes)
  - Max of 30 minutes Recovery Process Boost time within any 24 hour period
  - Sysplex partitioning boost seem to be 2 minutes
  - Multiple reasons can overlap, extending the boost (see next slide)
- zIIP Boost limited to bringing online 2 reserved zIIPs
  - Probably doesn't make much sense for more given the short duration

Recovery Boost messages


```
SYSA 13:29:02.45            V XCF,SYSC,OFFLINE
SYSA 13:29:02.46  *MASTER* *66 IXC371D CONFIRM REQUEST TO VARY SYSTEM SYSC OFF
SYSA 13:29:11.06            R 66,SYSNAME=SYSC
SYSA 13:29:11.18  *MASTER*  IEA681I Recovery process speed boost is active
SYSA 13:29:11.18  *MASTER*  IEA675I Recovery process zIIP boost is active with 0 transient zIIP cores
SYSA 13:29:11.18  *MASTER*  IEA687I Recovery process boost requestor: CF Datasharing Member Recovery
SYSB 13:29:11.18  *MASTER*  IEA681I Recovery process speed boost is active
SYSB 13:29:11.18  *MASTER*  IEA675I Recovery process zIIP boost is active with 0 transient zIIP cores
SYSB 13:29:11.18  *MASTER*  IEA687I Recovery process boost requestor: CF Datasharing Member Recovery
SYSA 13:29:11.45  WLM       IWM064I BOOST ACTIVATED.
SYSB 13:29:11.46  WLM       IWM064I BOOST ACTIVATED.

SYSA 13:29:31.74  *MASTER*  IEA686I Recovery process boost(s) have been extended
SYSA 13:29:31.74  *MASTER*  IEA687I Recovery process boost requestor: Sysplex Partitioning
SYSB 13:29:31.75  *MASTER*  IEA686I Recovery process boost(s) have been extended
SYSB 13:29:31.75  *MASTER*  IEA687I Recovery process boost requestor: Sysplex Partitioning

SYSA 13:31:31.74  *MASTER*  IEA678I All recovery process boosts have ended
SYSB 13:31:31.75  *MASTER*  IEA678I All recovery process boosts have ended
SYSA 13:31:32.01  WLM       IWM064I BOOST ENDED.
```

- Interesting case of Member Recovery being replaced by Partitioning boost

Will this impact our software bill?

Software Cost Impact

- By design: none
- MSU capacity rating of the machine not changed during boost periods
  - E.G. running on a 8562-N04, boosted LPAR consumes 100% of 2 CPs + 100% of 1 zIIP
  - 8562-N04 = 214 MSUs, 2 engines = half of that capacity = 107 MSUs consumed
    - Even though those 2 engines, while boosted, did *395* MSUs worth of work
    - Actually, maybe you saved some
- Theoretically speaking: shifting work around in time might impact the timing of peaks
  - Practically speaking: most people aren't IPLing in/near their peaks
- For TFP customers: all CPU time counts, as always
- SMF70 and SMF89 do contain boost indicators so… IBM or ISVs that read those records could also make adjustments

Which LPARs will be impacted, and how?

Be ready for the turbo to kick in!

LPARs being boosted

- Seems obvious: they get more work done!
- But may need changes to prepare them to get more work done
  - Maybe: define reserved (offline) zIIPs to be brought online during boost period
  - May need to add zIIP weight to be able to access newly online zIIPs
- Increased zIIP usage could cause some zIIP work to cross over to the GPs
  - Work with HONORPRIORITY=NO prevented from crossing over
  - So *may* want to consider changing from HONORPRIORITY=NO during IPL boost
    - Probably unnecessary in most cases
    - Recovery process boosts likely short enough to not be a major concern
    - Work should be routed away from the LPAR during shutdown boost anyways
- Boosted LPARs over-achieving work won't be capped to help discretionary
- New SMF interval started when boost starts/ends

LPARs *not* being boosted

- These LPARs *could* be negatively impacted
- CPU cache effectiveness may be impacted by boosted LPARs (at least theoretically)
  - Speed boosted LPARs do more work per unit of time on the GPs
  - zIIP boosted LPARs may drive more work to zIIPs
  - zIIP caps ignored during boost periods
- Higher physical zIIP utilization may impact non-boosted LPAR's ability to get work dispatched on zIIPs
  - Could potentially lead to more crossover
  - zIIP caps ignored but weight enforcement still applies
  - May need to change relative weights during boost periods, if trying to protect unboosted LPARs
- Resource Group caps with sysplex scope don't count work running on boosted LPARs
  - May allow more than expected work to run on unboosted systems
- Conversely:
  - If boosted systems consume less of GP capacity, might help non-boosted systems

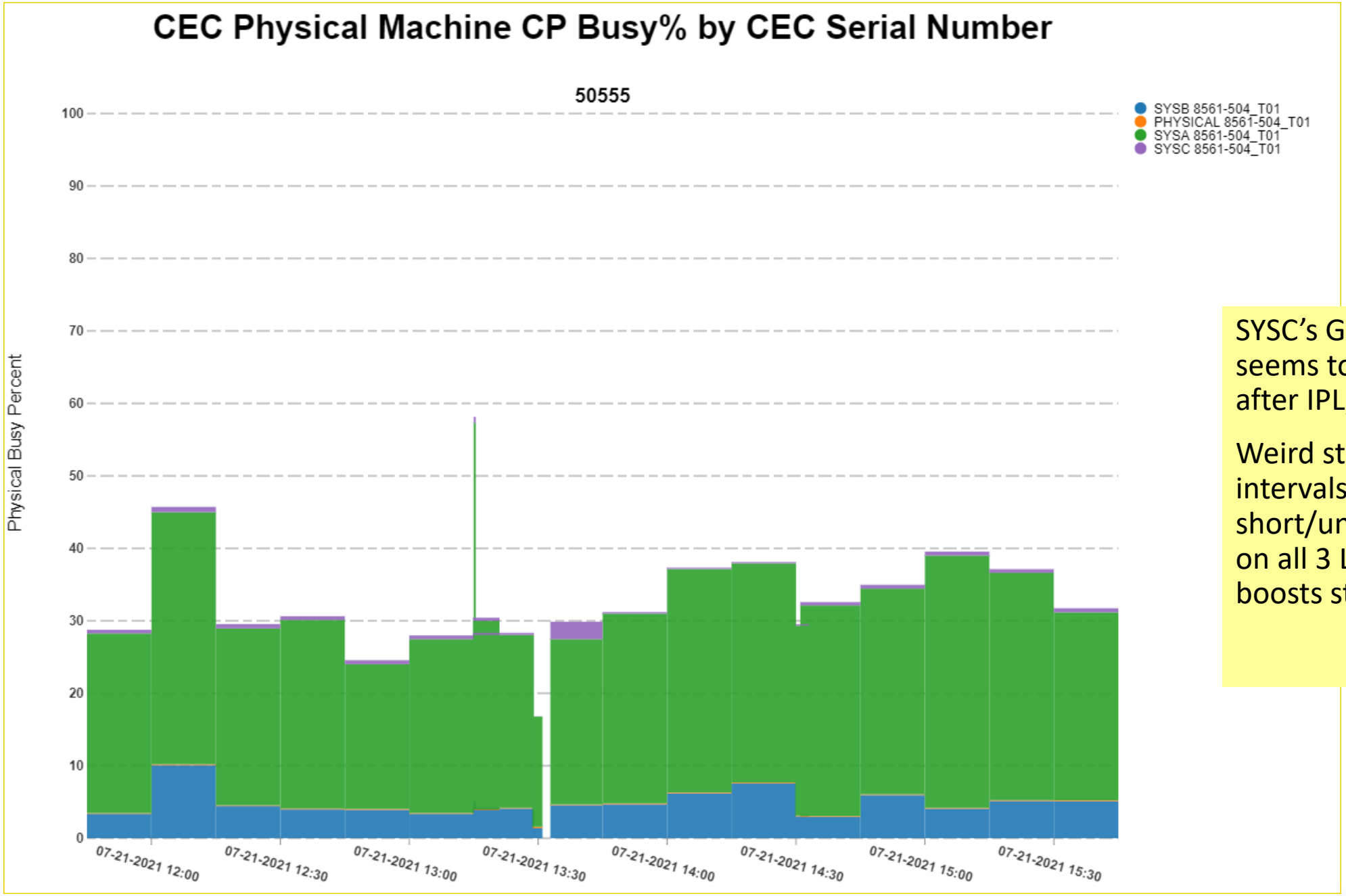
Case Study: IPL SYSC

- SYSA = Prod, SYSB = Dev/Test, SYSC = SysProg Test System
- All part of same sysplex
- z15-504 (~2.6x speed boost)
- 1 Reserved zIIP on SYSC
- Not much running on SYSC, so easier to see boost impacts on system background

```
13:15:28 - Activate Shutdown Boost on SYSC
13:29:11 - Vary SYSC out of the plex
13:29:11 - Recovery Boost started on SYSA, SYSB for Datasharing Member Recovery
13:29:31 - Recovery Boost extended for partitioning
13:31:31 - Recovery Boost ended on SYSA, SYSB
13:32:11 - First SYSC IPL messages
13:32:29 - SYSC IPL Speed Boost active message
13:33:08 - SYSC IPL zIIP bost active message
14:33:08 - SYSC IPL Boost ended message
```

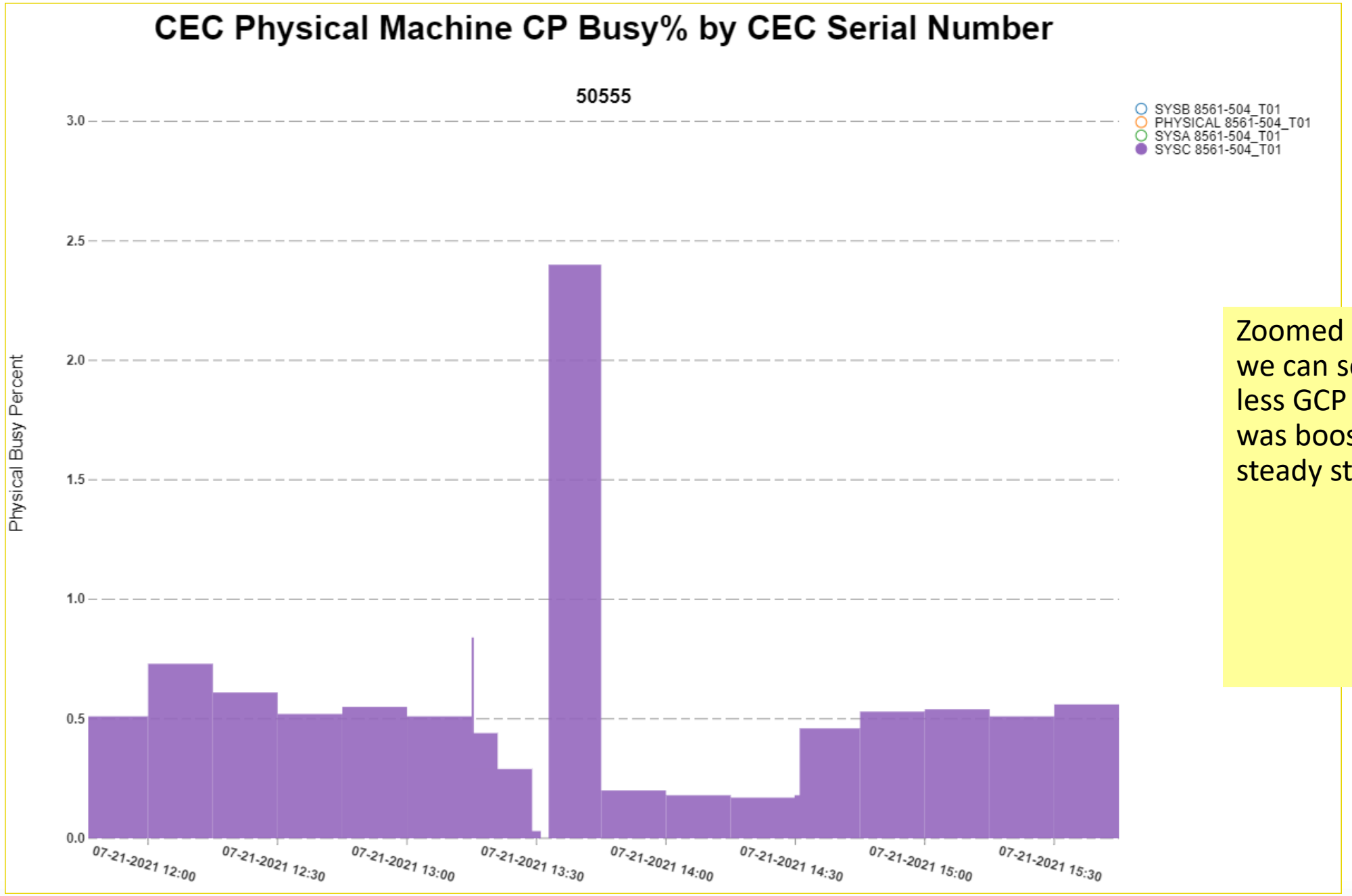
GRS, XCF, etc still initializing

Most base system components up

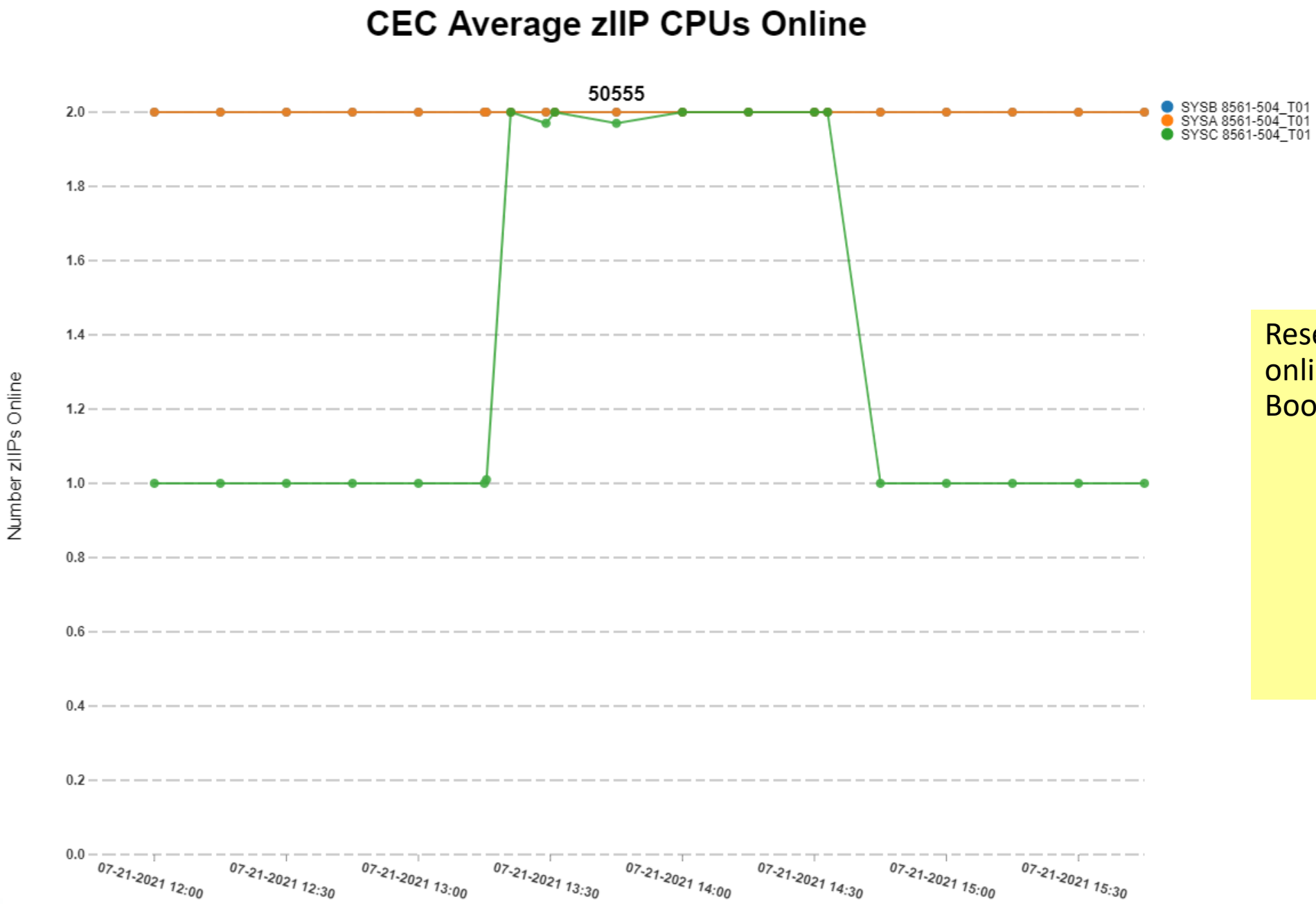


SYSC's GP utilization seems to go up right after IPL, then down.

Weird stacking of intervals is due to short/unsynced intervals on all 3 LPARs due to boosts starting/stopping.

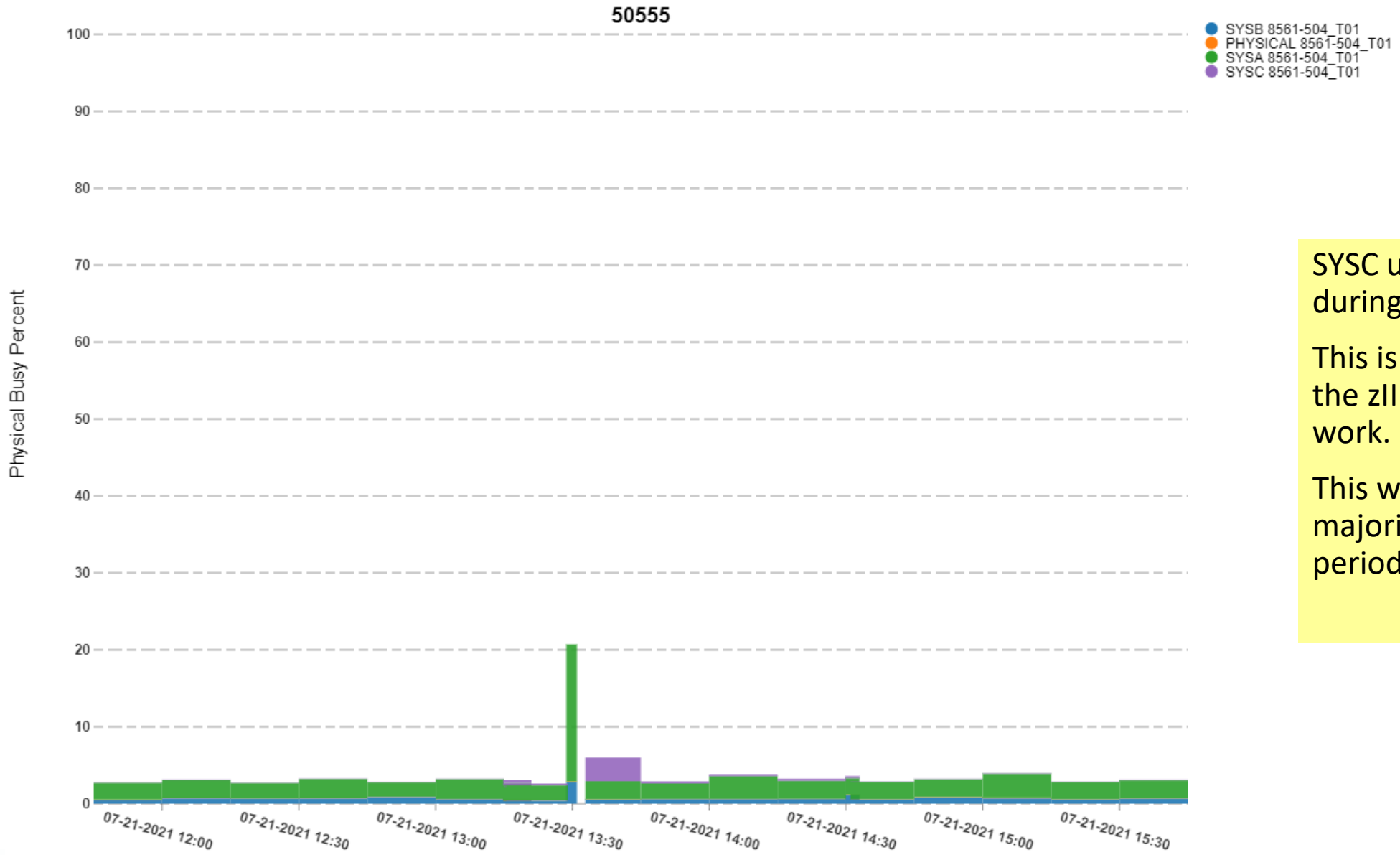


Zoomed into just SYSC we can see it consumed less GCP capacity while it was boosted vs. normal steady state.



Reserved zIIP brought online to SYSC for zIIP Boost.

CEC Physical Machine zIIP Busy%

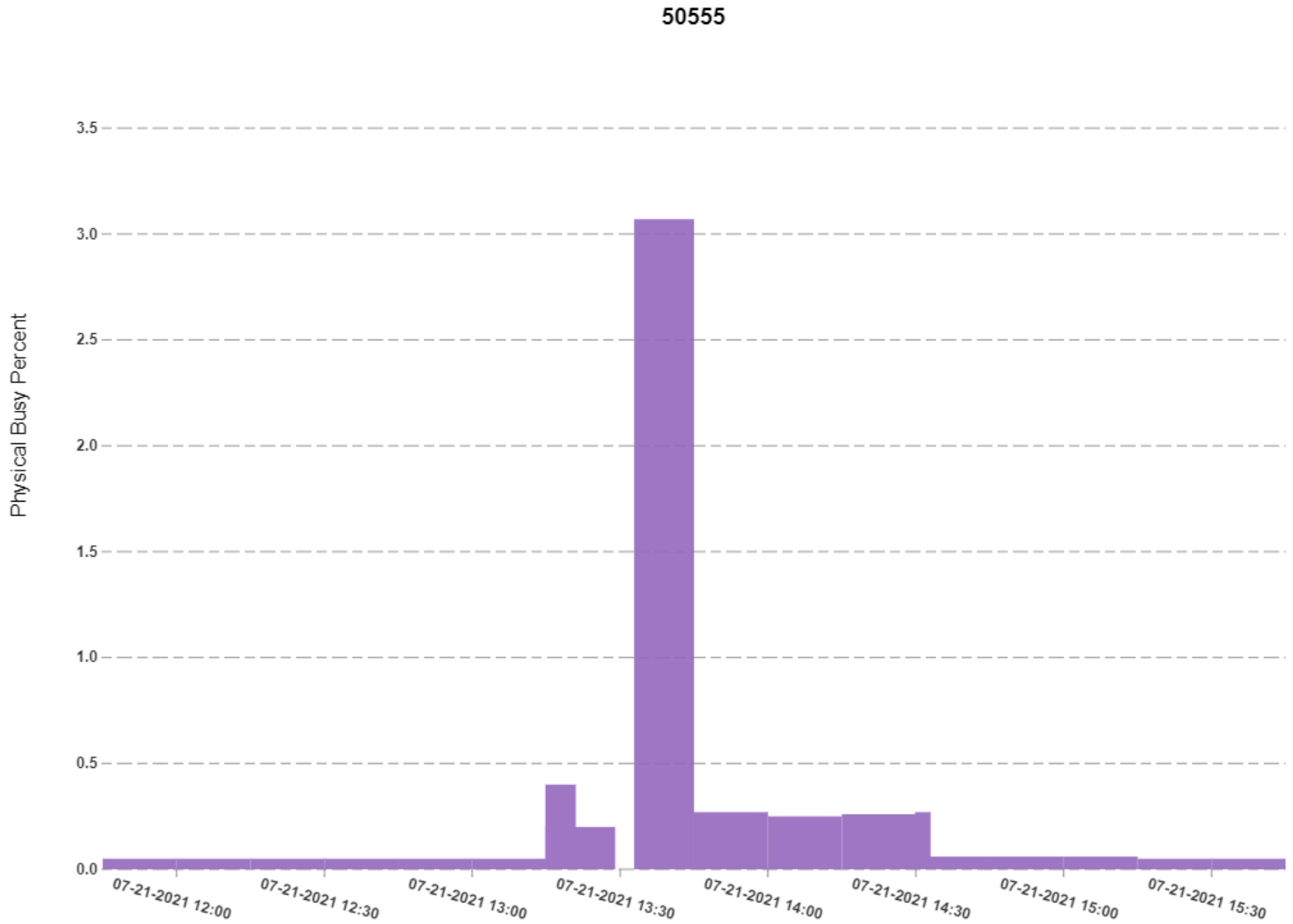


SYSC used more zIIP during it's boost period.

This is as we'd expect: the zIIPs are running GP work.

This was true in the vast majority of boost periods that I examined.

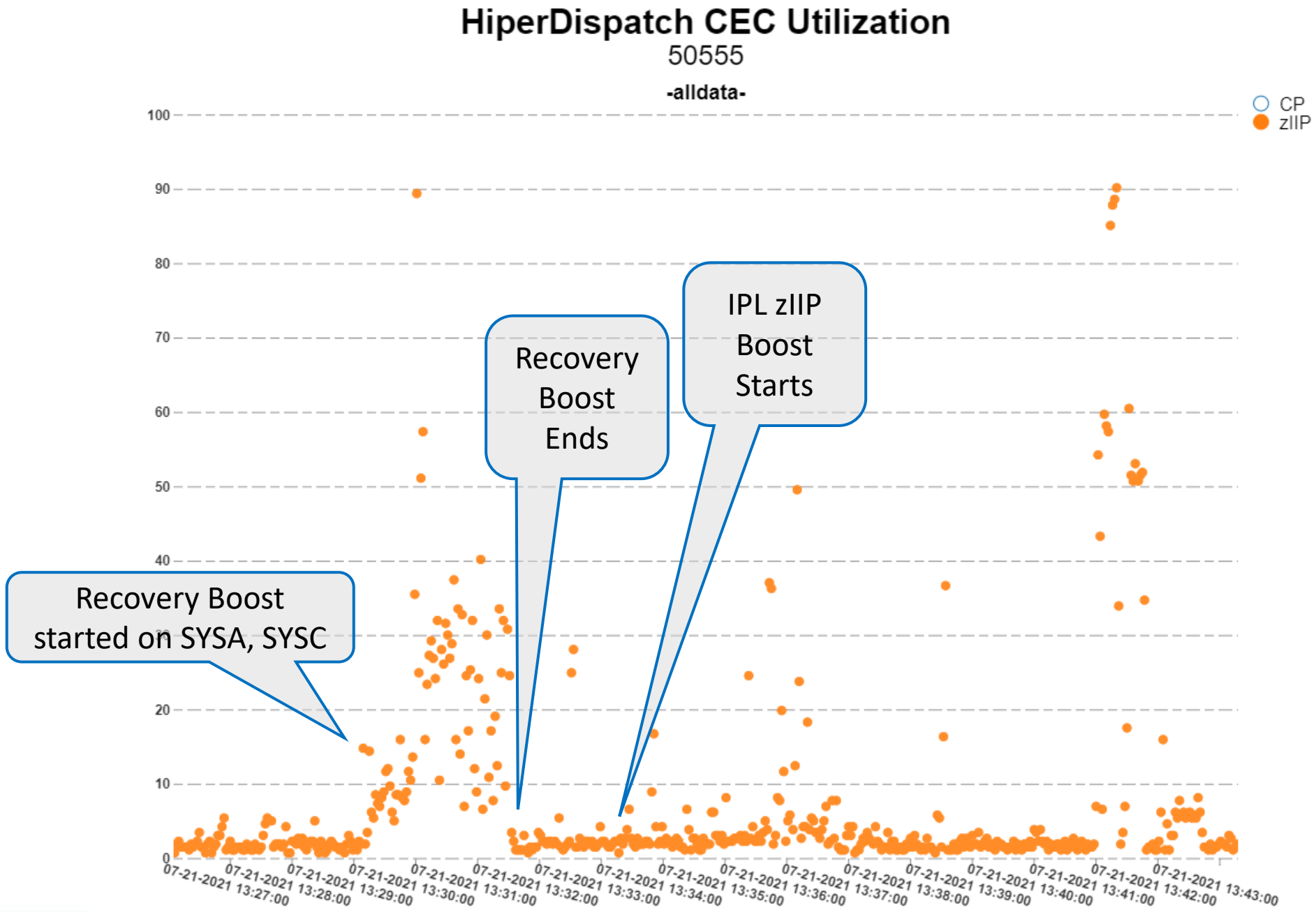
CEC Physical Machine zIIP Busy%



Here we see SYSC using more zIIP because some of that GP work ran on the zIIPs.

Note this is a small system the relative change might not be as dramatic on larger systems. Or it might be larger if more zIIPs were enabled relative to the GPs.

In other words: YMMV

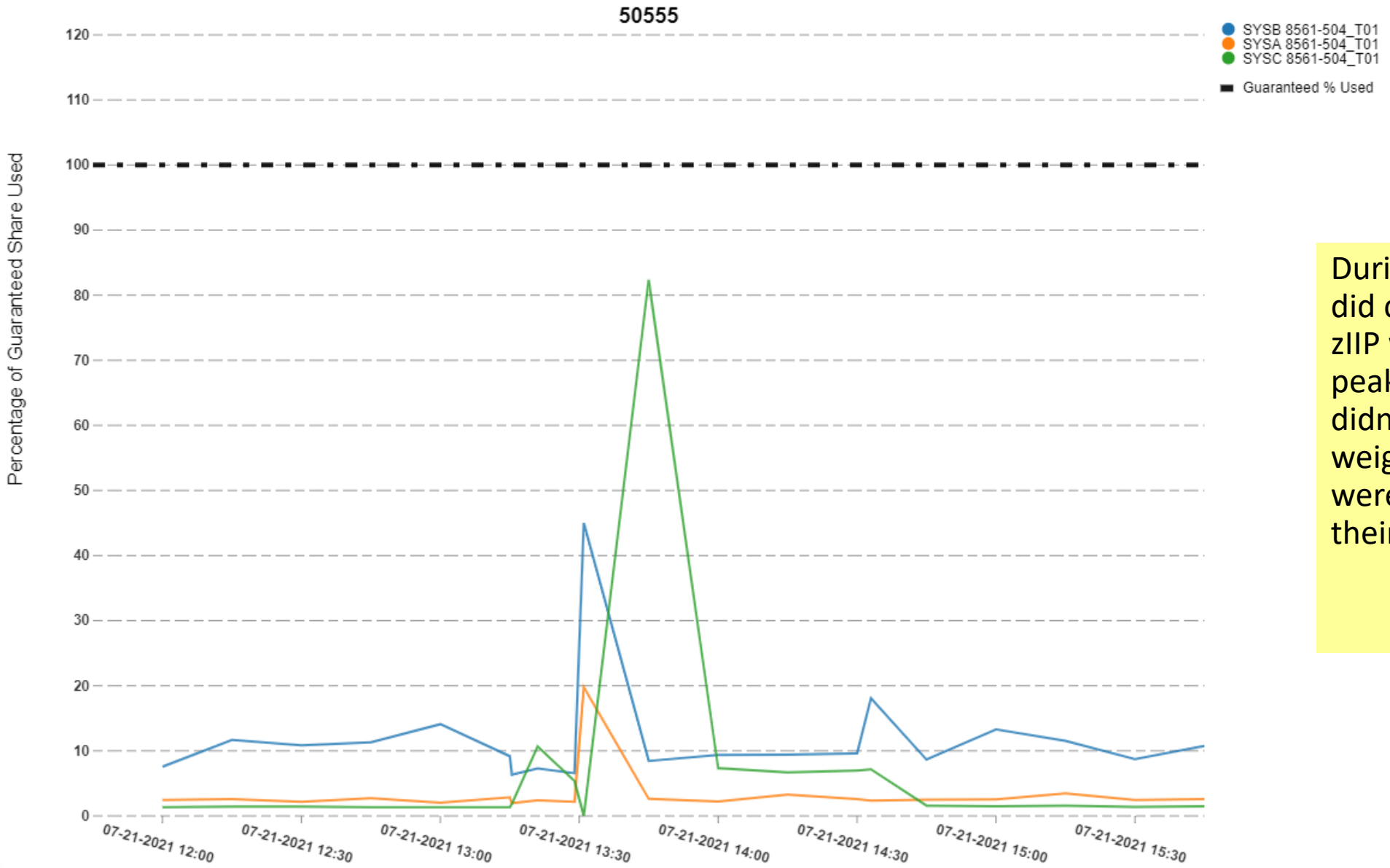


2-second zIIP utilization levels.

Weighty issues

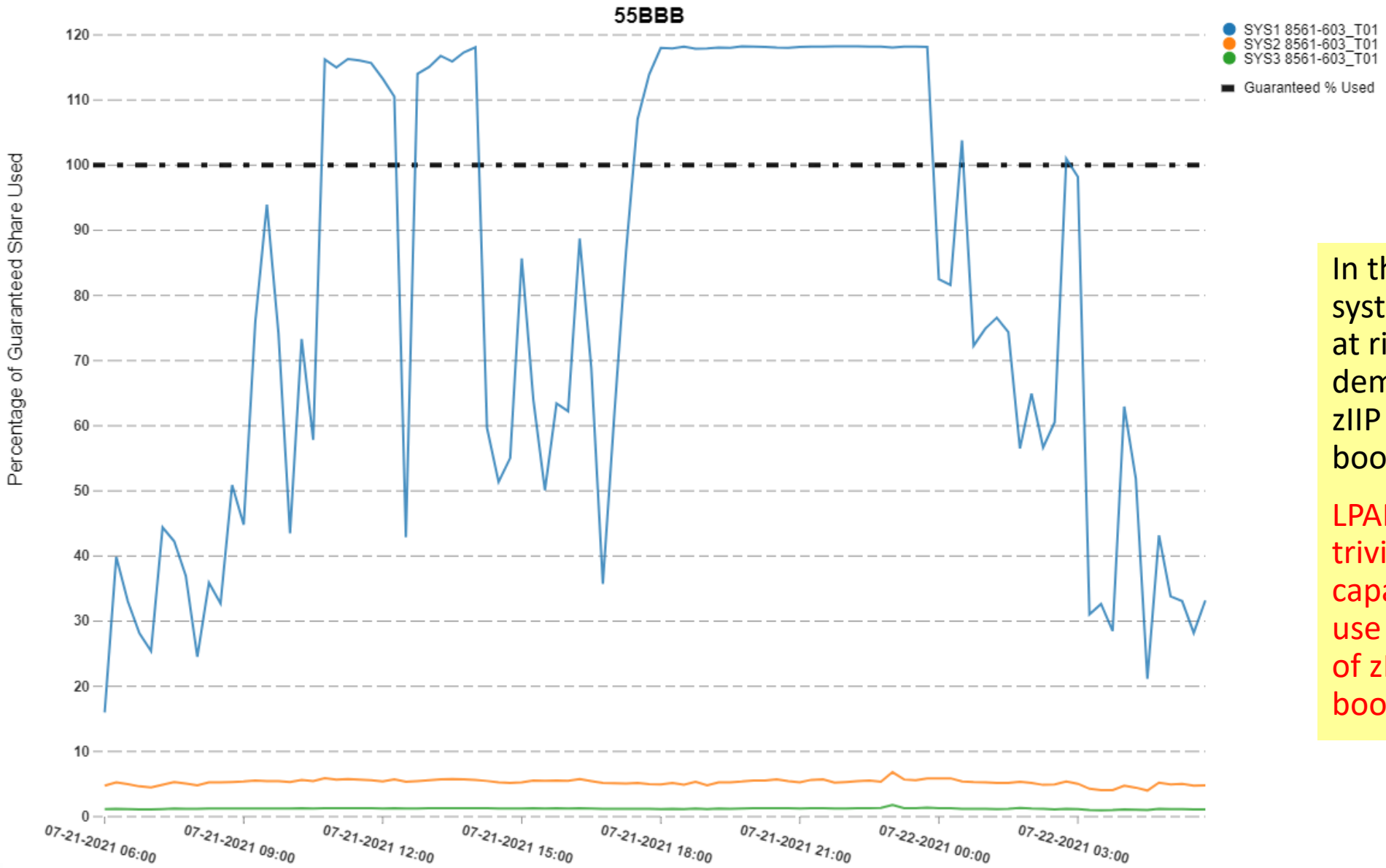
- Boosts happen on a per-LPAR basis
- PR/SM dispatches physical CPs to LPARs as normal
  - Speed boosted LPARs simply get more useful work done in their dispatch interval
- Absent weight changes:
  - zIIP boosted LPARs with low weights may end up using low pool zIIPs
    - May not be able to access that capacity if other LPARs busy
  - Other LPARs using more than their zIIP weight may be limited if they are borrowing weight from the boosted LPAR

CEC Percent zIIP Weight Used



During IPL boost, SYSC did consume more of its zIIP weight but for the peak 15 minute interval didn't quite reach its full weight. But other LPARs were also using less than their full weight.

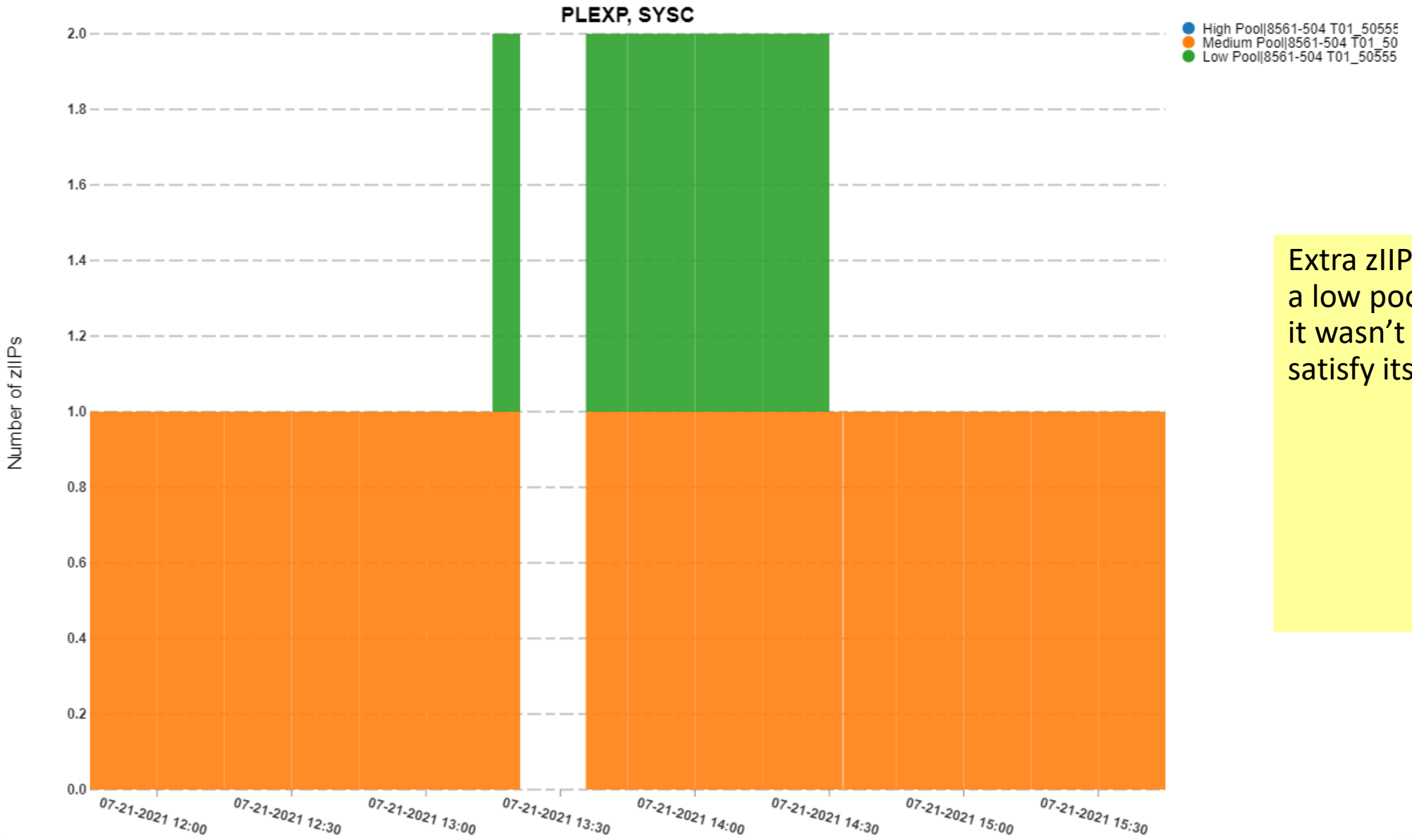
CEC Percent zIIP Weight Used



In this case (different systems), SYS1 might be at risk if SYS2 or SYS3 demanded their entire zIIP weight due to a boost!

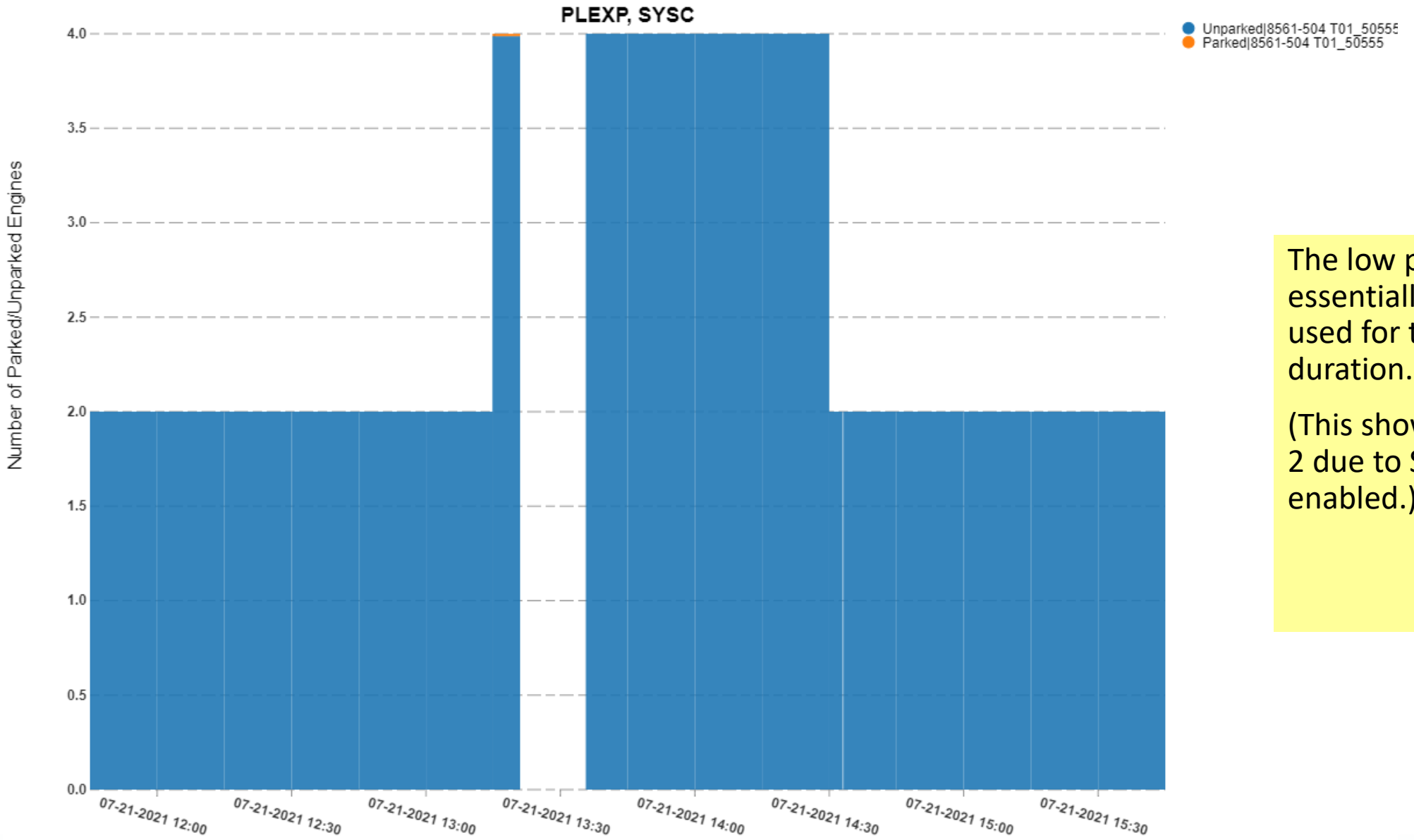
LPARs that normally use trivial amounts of zIIP capacity may suddenly use substantial amounts of zIIP capacity during boost periods!

HiperDispatch zIIP CPU Pooling at End of Interval



Extra zIIP came online as a low pool zIIP because it wasn't needed to satisfy its weight.

HiperDispatch - Parked / Unparked zIIPs



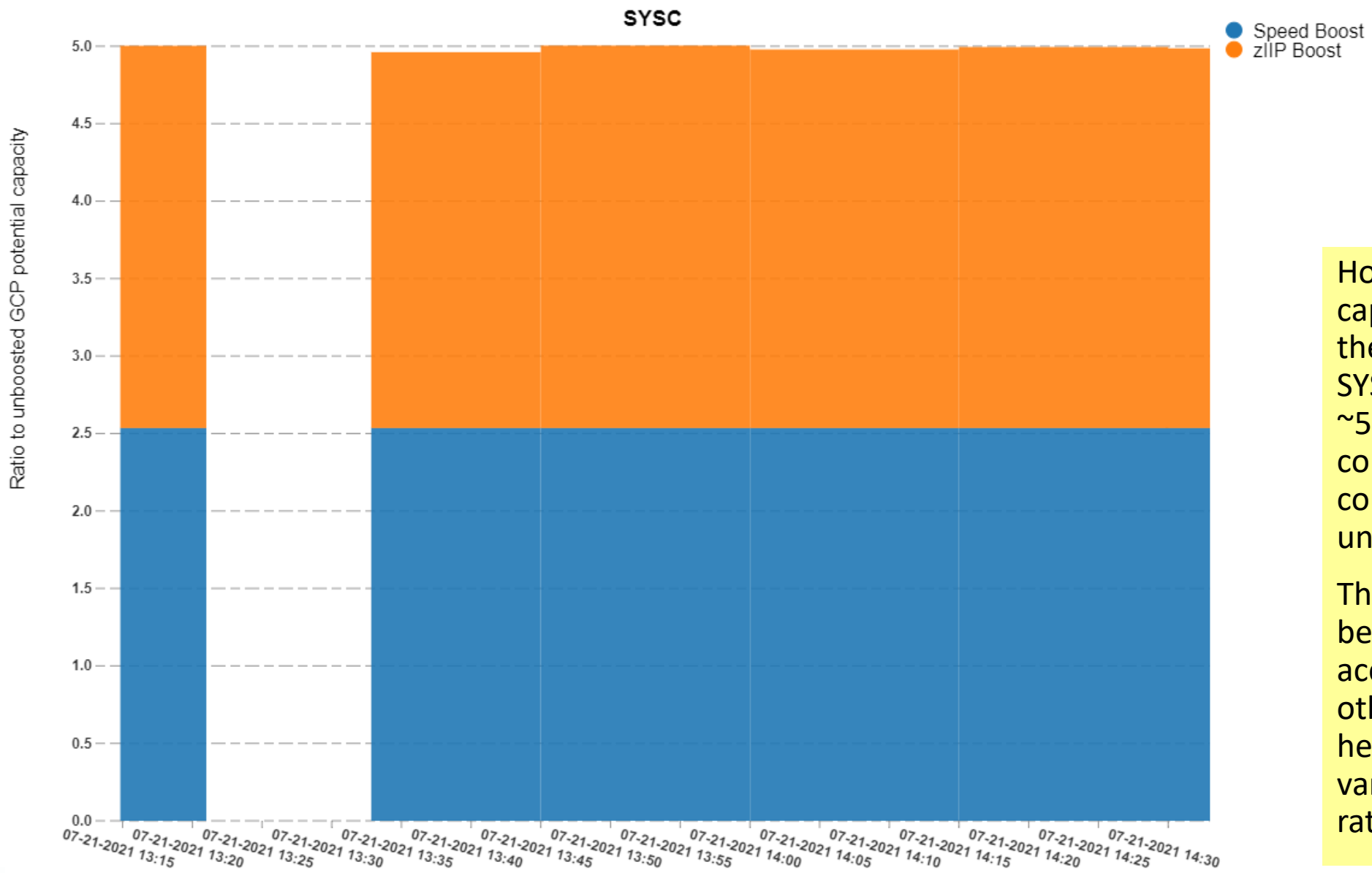
The low pool zIIP was essentially unparked and used for the entire duration.

(This shows 4 instead of 2 due to SMT being enabled.)

SYSC Observations

- In this case SYSC enjoyed the benefits of both Speed and zIIP Boost without impacting the other LPARs
  - Overall zIIP utilization was low, so no significant potential for conflict
  - SYSC also uses relatively little CPU, so it's increased zIIP utilization was pretty minor (just over 3% of the total zIIP capacity at it's peak)
- During boost periods, SYSC:
  - Consumed ***less*** of the machine's GCP capacity
    - (Except for the interval immediately after the IPL)
  - Consumed ***more*** of the machine's zIIP capacity
  - This is partly due to the fact that it's a sysprog test LPAR with little real activity
    - Looking at several examples from customers almost all looked like this: IPLs tend to be scheduled for low utilization times
    - Will be interesting to see an emergency IPL in the middle of the day ☺

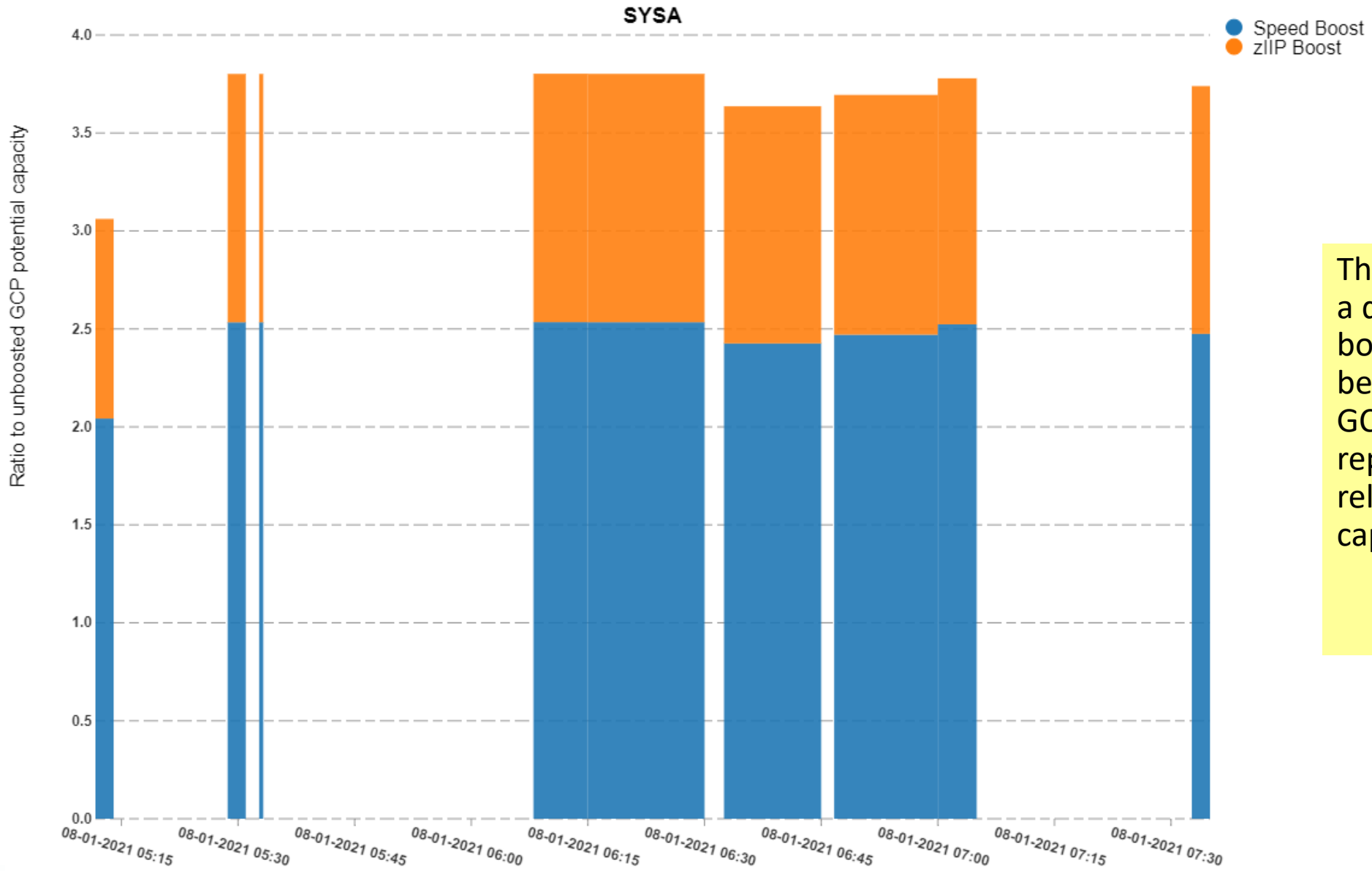
Boost Potential Capacity Ratio



How much extra capacity did we get from the boost? Potentially SYSC could have done ~5x the GP work compared to what it could have done unboosted!

This report was tricky because it takes into account how much the other LPARs are using hence the slight variations you see in the ratios.

Boost Potential Capacity Ratio

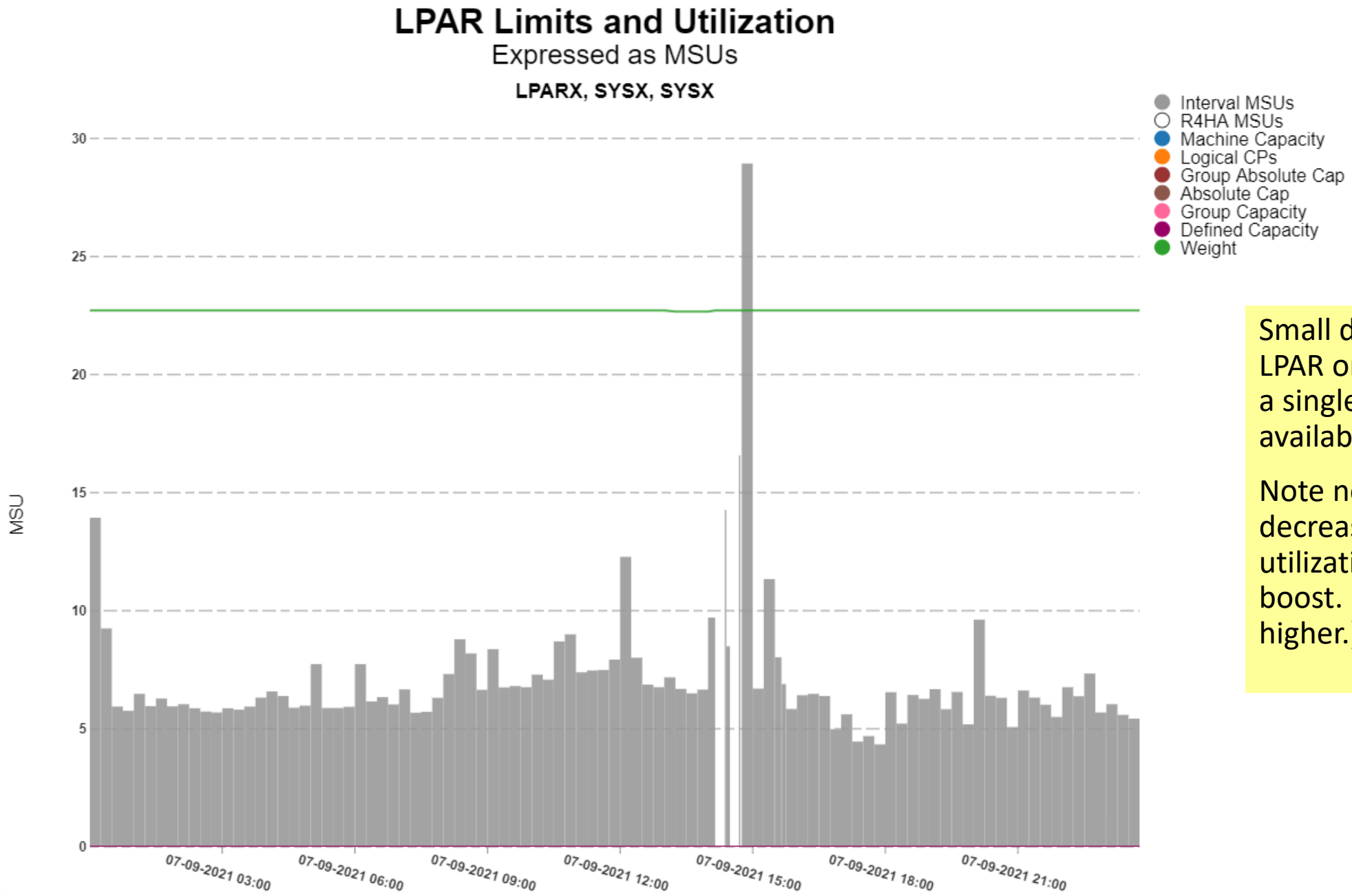


This is the larger LPAR on a different day… its boost ratio is lower because it had more GCPs online, so the zIIPs represented less of a relative increase in capacity.

Things that limit SRB's effectiveness

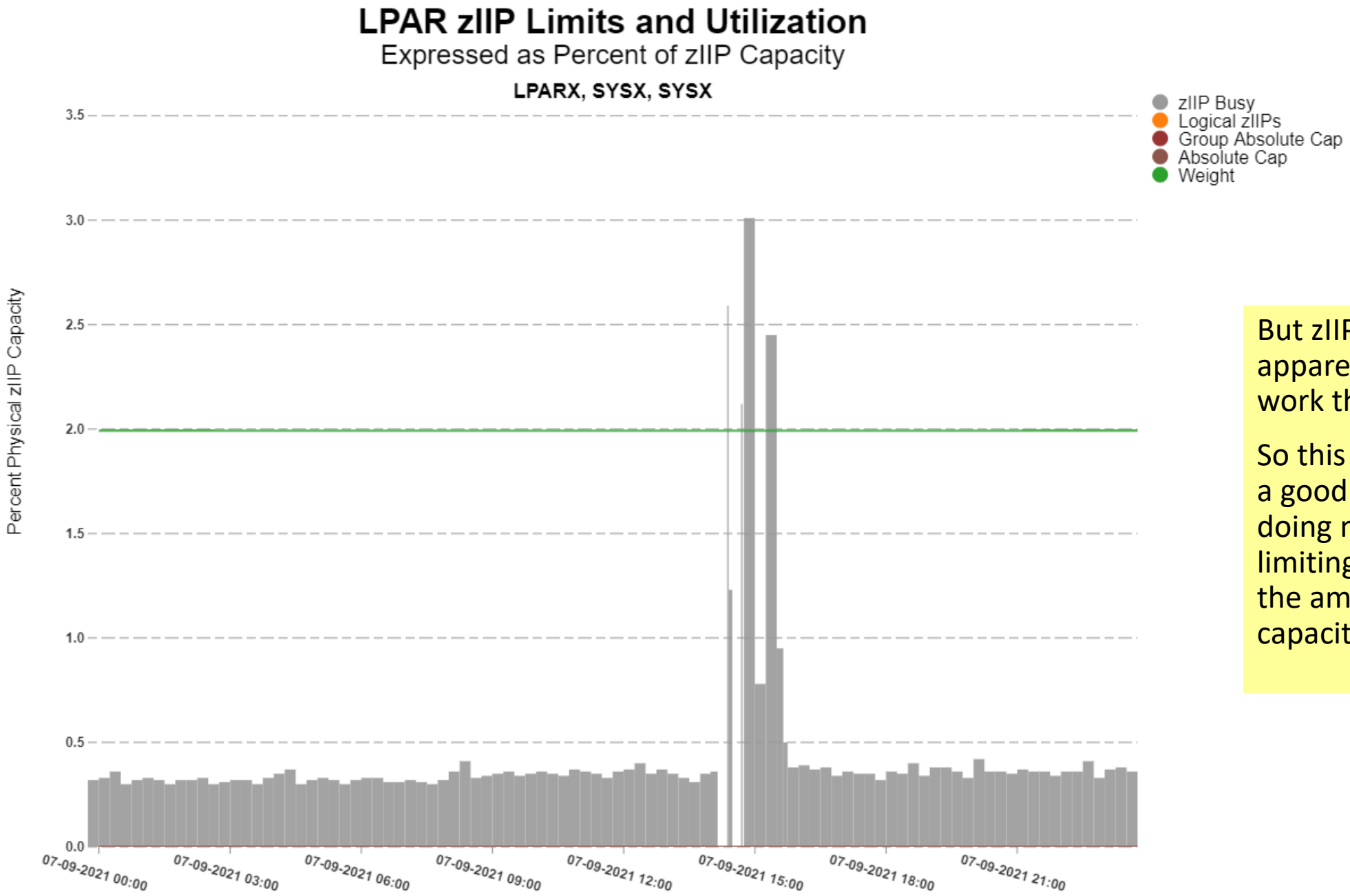
- Full speed GCPs eliminates speed boost
- Busy zIIPs
  - Less available capacity to the boosted LPAR
- Single physical zIIP shared among several LPARs
  - PR/SM will still move the zIIP between LPARs limiting the time the boosted LPAR might be able to use it
- Few or no zIIPs available to the boosted LPAR

Scott's Opinion: In the 2020s, almost every machine should have at least 2 zIIPs!



Small dev/test monoplex LPAR on a 711 with only a single physical zIIP available

Note no significant decrease in GCP utilization during the IPL boost. (Might even be higher.)

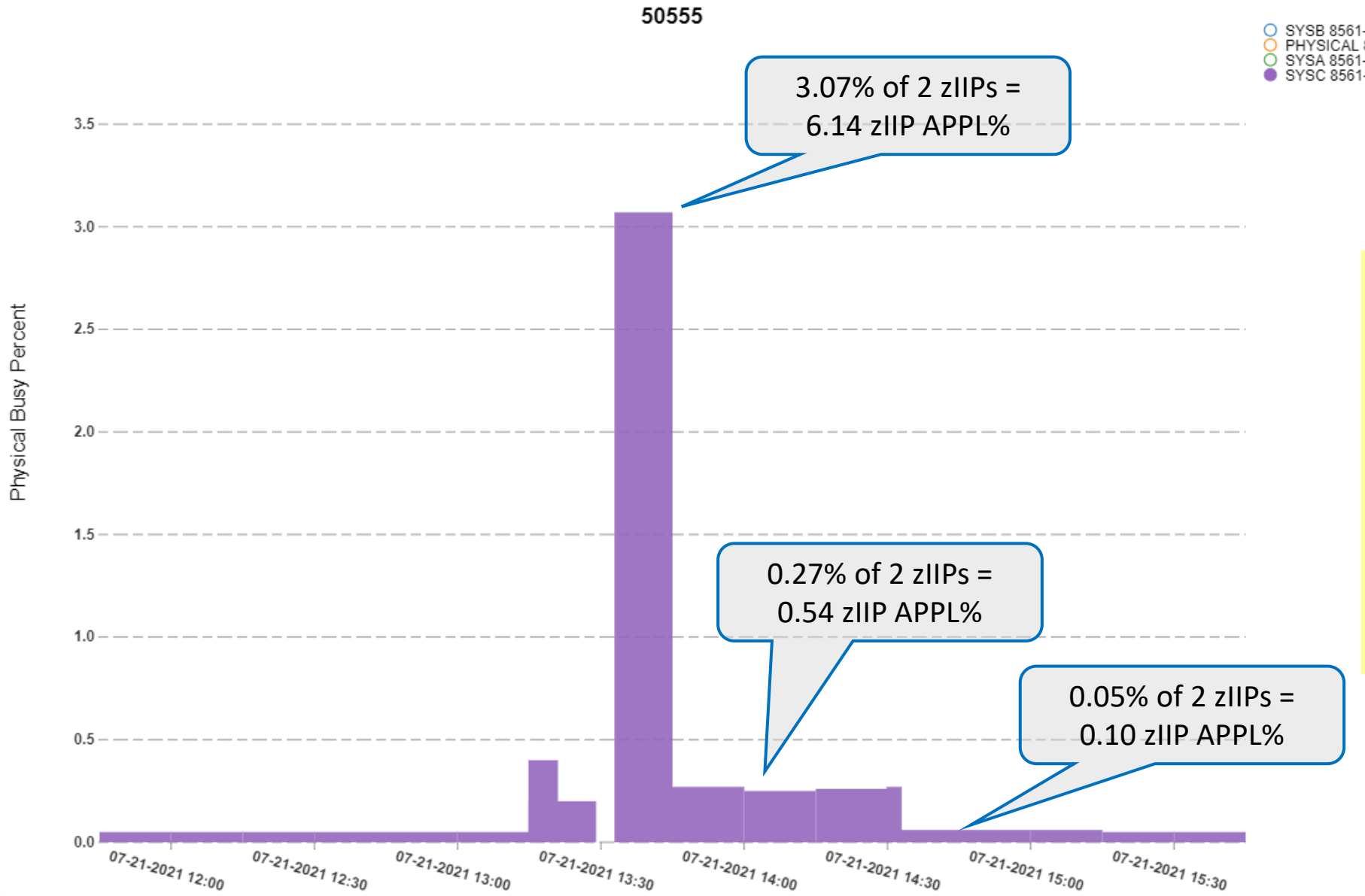


But zIIP Boost did apparently offload some work though.

So this still appears to be a good story: LPAR is doing more work while limiting the increase in the amount of GCP capacity consumed.

How is the SMF data affected?

CEC Physical Machine zIIP Busy%

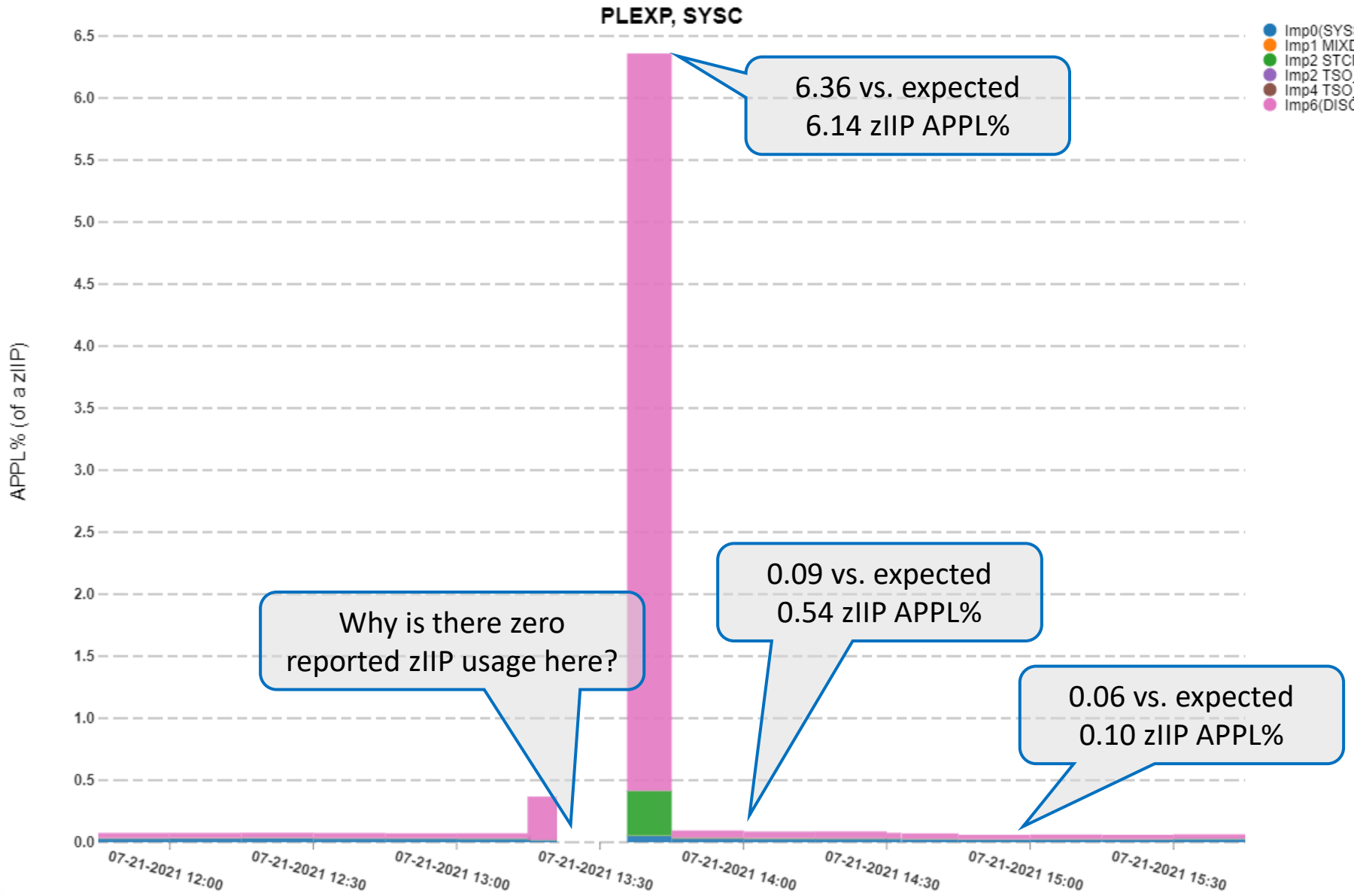


Remember this chart? SYSC using much more zIIP capacity than normal during the boost period. (This from SMF 70)

So what workloads were using the zIIP??

WLM CPU - zIIP APPL% SMT for Service Class

Adjusted for SMT



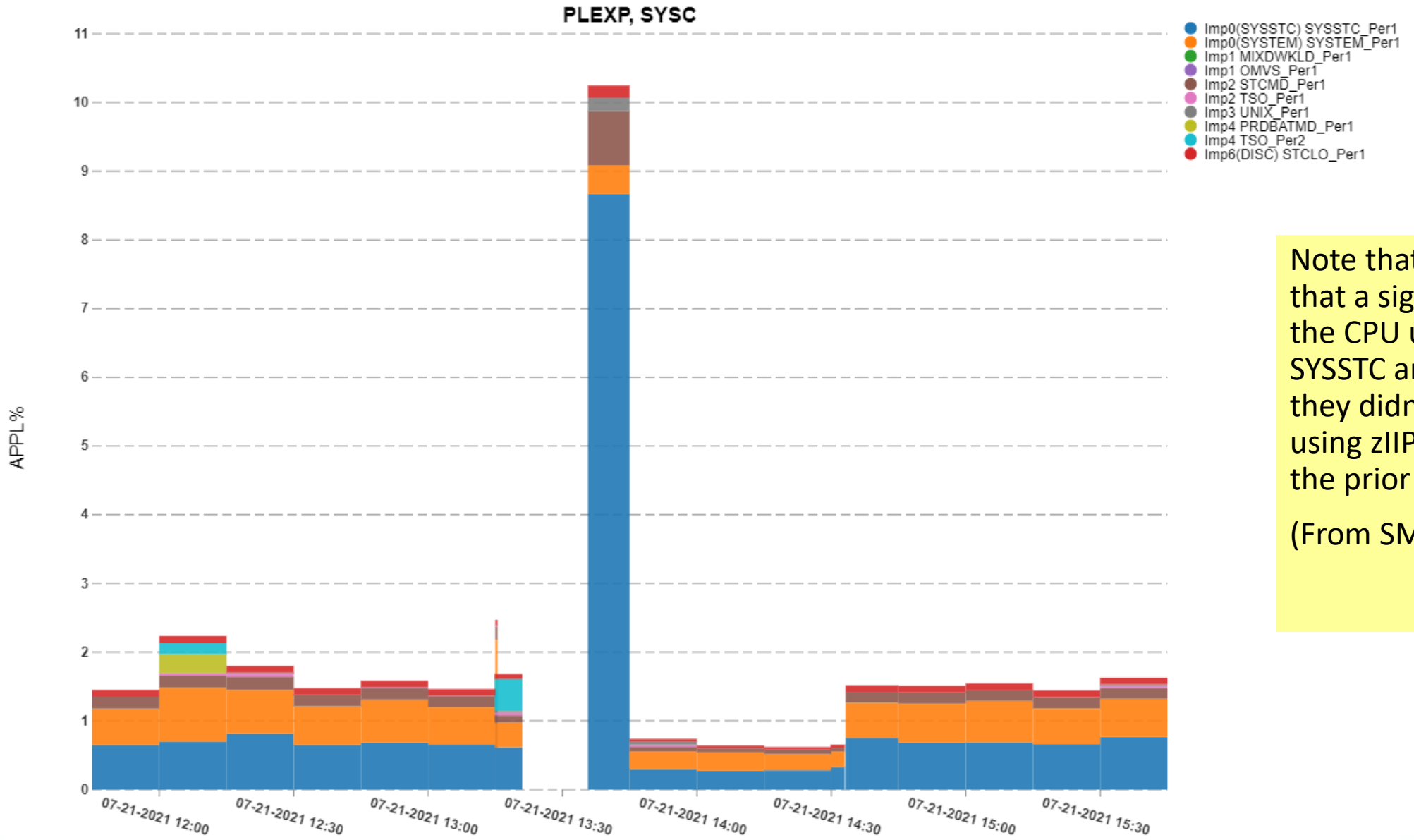
This shows how the SMF 72 doesn't quite match up to the SMF 70.

SMT messes with our zIIP consumption numbers, this particular report tries to compensate for that.

Some of the variance is due to SMT. But not all of it.

WLM CPU - CP APPL% by Service Class

(CP + zAAP on CP + zIIP on CP)

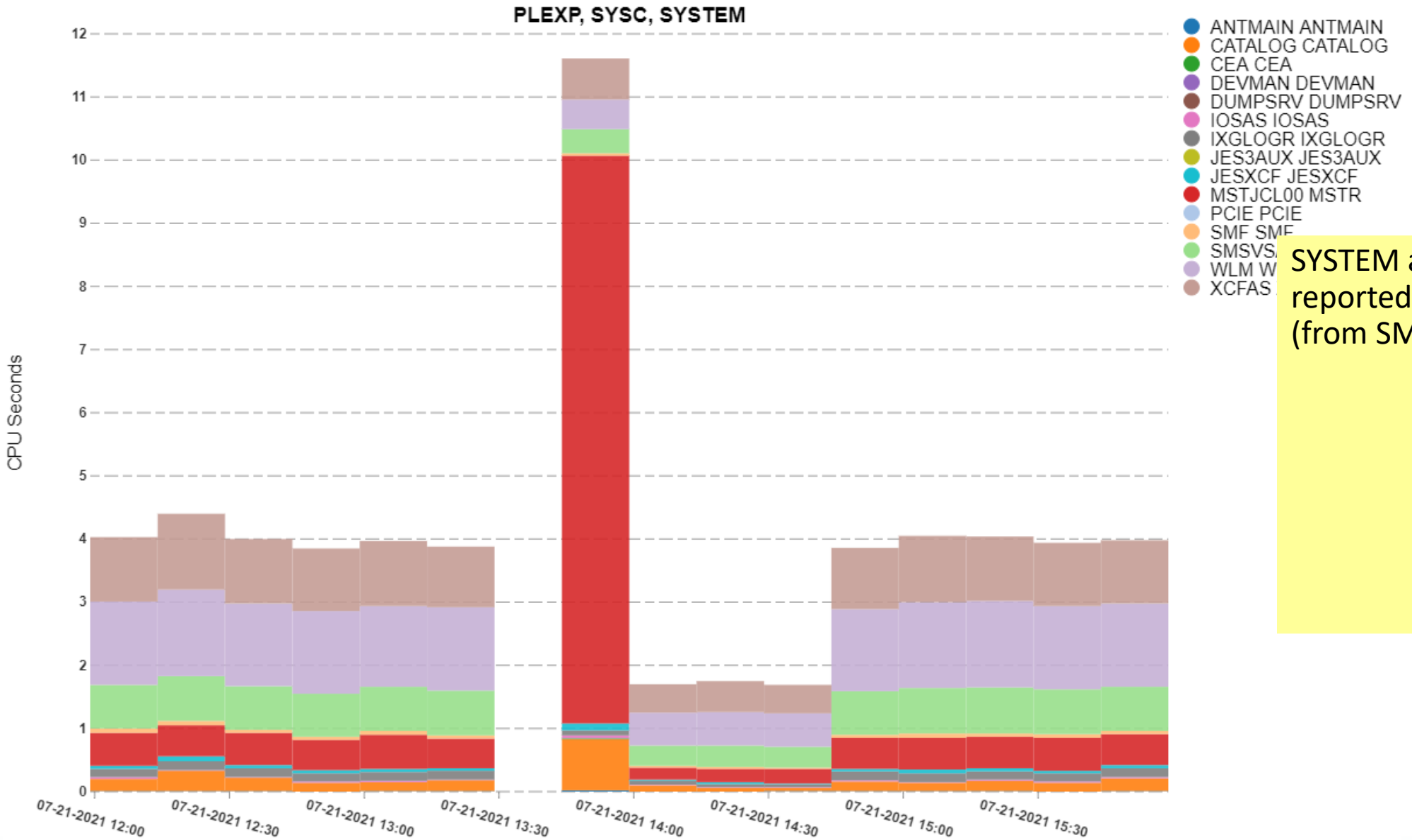


Note that this shows that a significant port of the CPU utilization was SYSSTC and SYSTEM, but they didn't show up using zIIP significantly on the prior chart.

(From SMF 72)

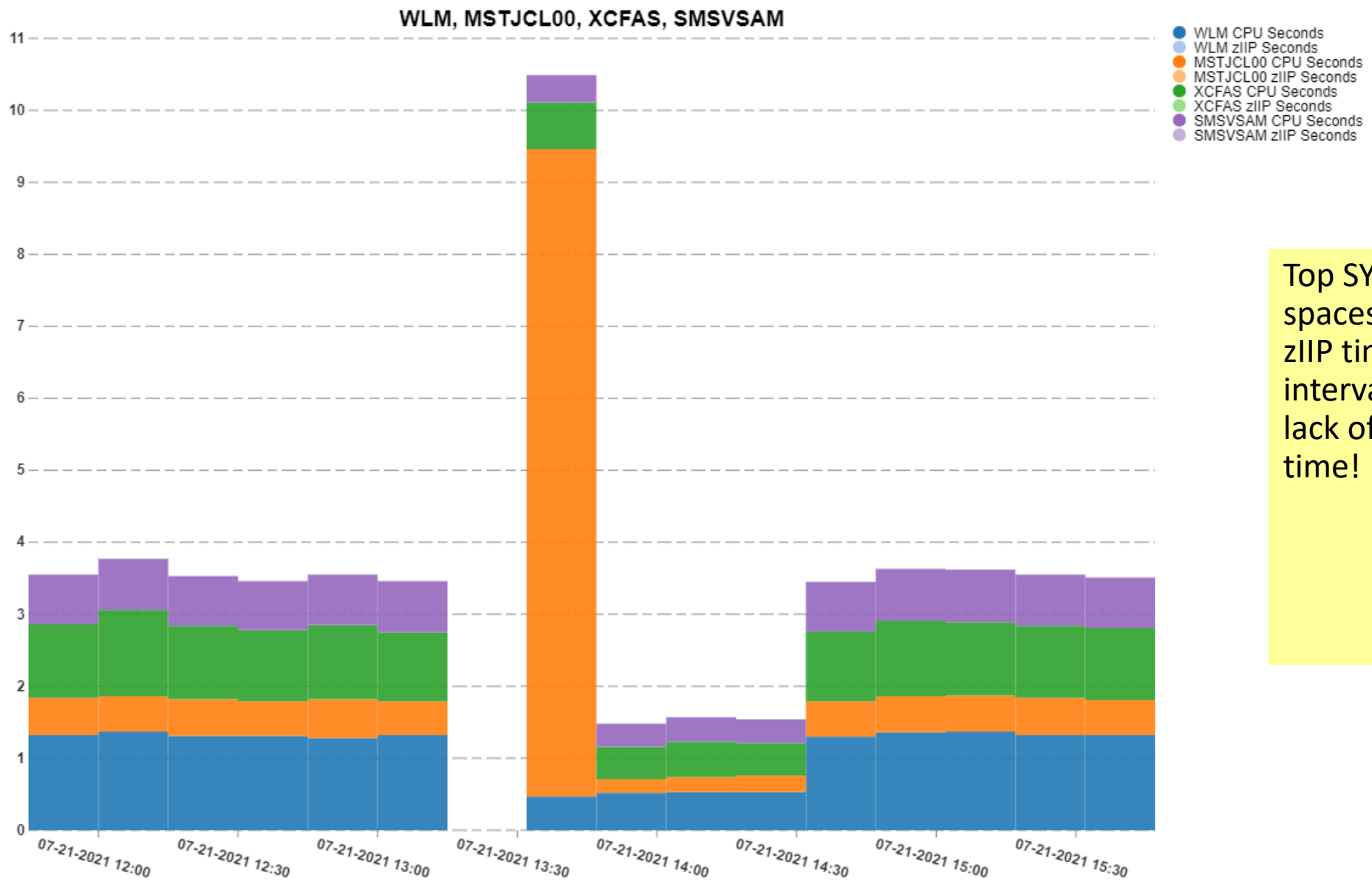
Top Address Space CPU Time for Service Class

Period of Study



SYSTEM address spaces reported GCP utilization (from SMF 30)

Started Tasks Over Time

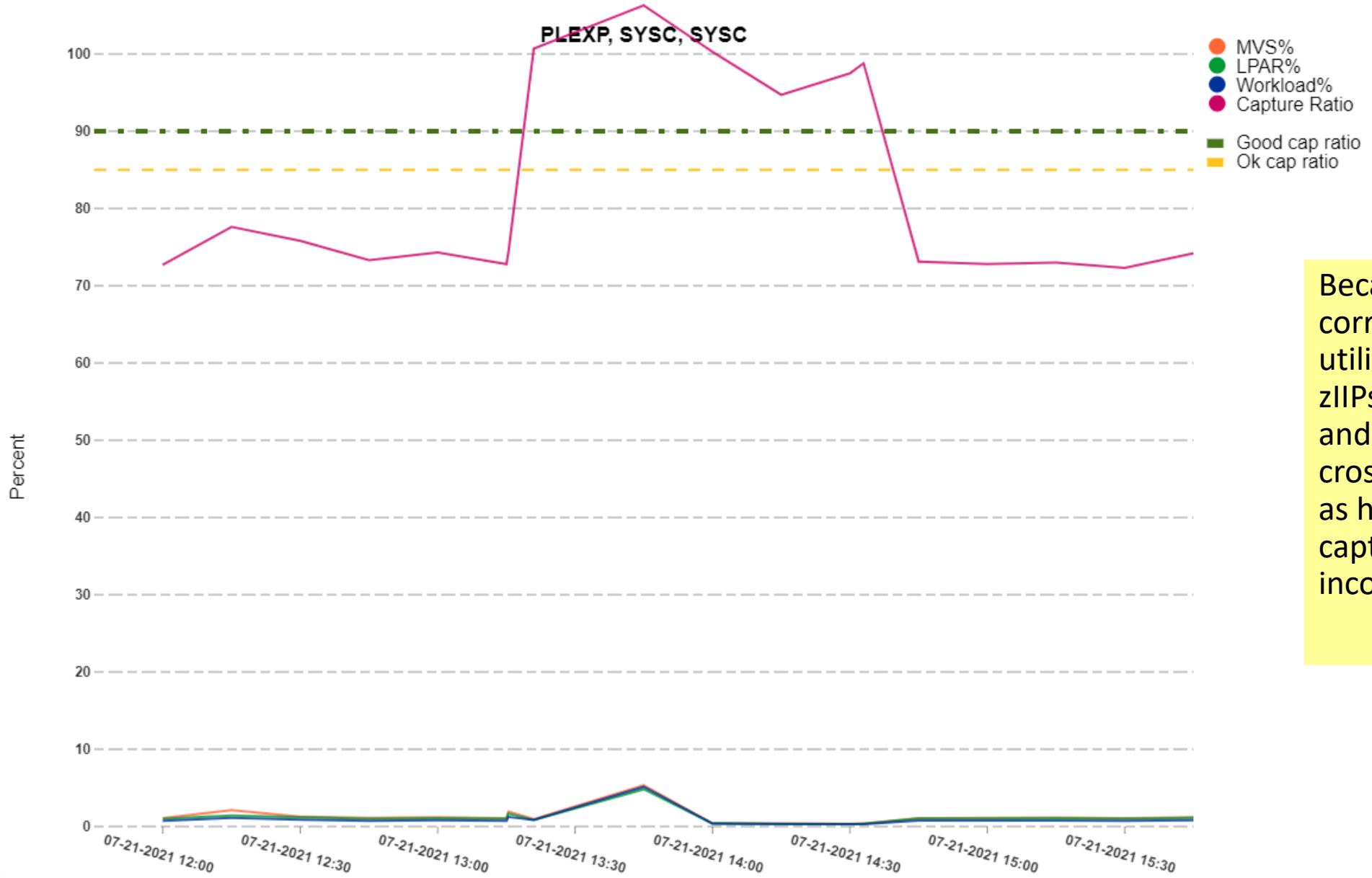


Top SYSTEM address spaces reported GCP and zIIP time from SMF 30 interval records. Note lack of reported zIIP time!

CPU Accounting during boost periods

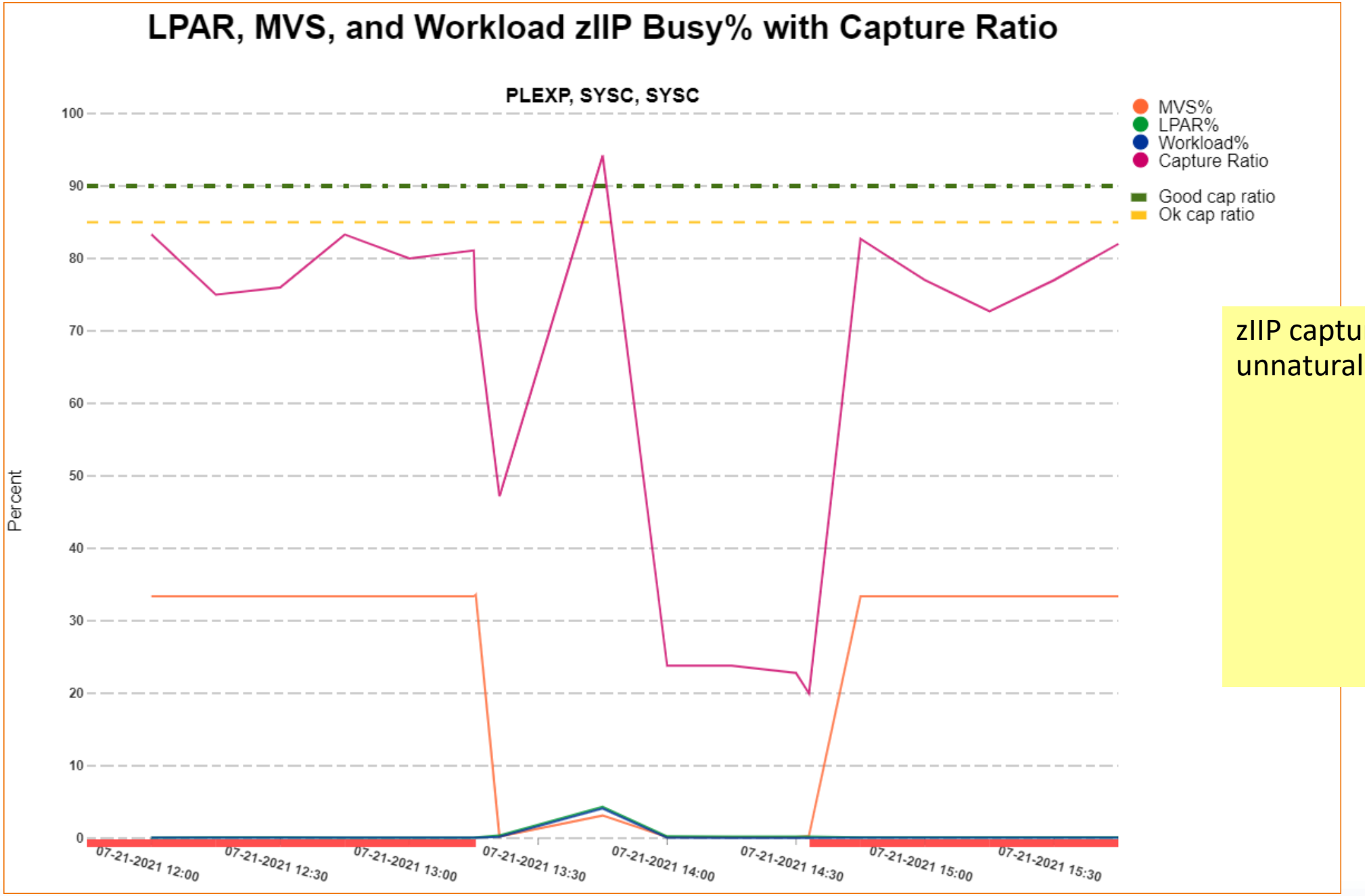
- Generally (across record types), with speed boost:
  - SU/SEC increases
  - Normalization changes to indicate zIIP = CP
  - MSU rating of machine does not change
- SMF 70 – Dispatch times unaffected and report “true” usage
  - I.E. work run on the zIIP increases the zIIP’s effective dispatch time
- SMF 30 & 72 – Records GP-only work on GP, even if it ran on zIIP!
  - GP work running on zIIP in SMT is recorded as MT1ET
  - For sub-capacity machines, CPU times will be lower than non-boosted periods

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



Because SMF 70 correctly records the utilizations of the GPs vs. zIIPs, but the SMF 72 and SMF 70 records crossover from GP to zIIP as having been on GP, capture ratios will be incorrect.

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



zIIP capture ratio goes unnaturally low!

Parting thoughts

SRB: The Good

- Overall, this is goodness: faster shutdown, startup, and recovery
- Sub-capacity systems will see more benefit than full-cap machines
  - It's not hard to come up with scenarios where a boosted LPAR might have access to 5-10x more capacity than normal for GP-only work!
- No additional cost on z15 machines
  - SRB Upgrade Record likely to only be of significant value to large 7xx machines with multiple LPARs to IPL on a single day
- Simple scenarios where IPLs happen at low-utilization times will see benefit while likely not causing any significant concerns
- Despite design goal of no software cost impact, seems like it could potentially reduce costs

SRB: The Bad

- Have to update your shutdown procedures to invoke shutdown boost
- LPAR weights require attention to make sure they're appropriate during boost periods
  - May need to set up automation to change weights when boosted
- Boosted LPARs might impact cache efficiency of non-boosted LPARs
  - In theory; In practice, haven't seen a good example of this
- Recovery Process Boosts result in short intervals where things change
  - Be aware of this if you're analyzing performance!

SRB: The Ugly

- Recorded CPU time during boost periods is… confusing to misleading
  - SMF 30 intervals that cross boost/non-boost intervals especially so
    - Apparently only when not synchronizing your intervals:
      So use SUBSYS(xxx,INTERVAL(SMF,SYNC)) in your SMFPRMxx!
- 2 Minute boosts could be useful in some situations, but may cause confusion when doing performance analysis
  - Performance testing especially impacted!
- Coming off from the boost can be a let down

Recommended To-Do List

- Add call to proc IEASDBS during shutdown to enable boost
  - If using PRESCPU in IEASYSxx, stop boost before final shutdown
- Add reserved zIIPs as appropriate to LPARs to take advantage of zIIP Boost
- Verify that zIIP weights are appropriate during boost periods
- Understand when Recovery Boost periods happen

Not recommended, but…

- What if…
  - You have a small sub-capacity machine that gets a significant CPU boost
  - You have some free zIIP capacity too
  - You have a significant CPU-limited month-end process
- Should you IPL immediately before month-end processing starts?
  - I.E. could boost get you multiple hours of work done in < 1 hour?
- Note: I don’t know if there’s IBM Ts and Cs to prevent this
  - I haven’t heard of any but…
  - It seems likely somebody will start doing proactive IPLs at “interesting” times

Thanks!

Questions/comments:
Scott.Chapman@epstrategies.com