



z/OS Performance Spotlight – Some Top Things You May Not Know

aka Peter and Scott's
Tips and Tidbits



Peter Enrico & Scott Chapman

Enterprise Performance Strategies, Inc.

performance.questions@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:

© 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



During this webinar, **Peter Enrico** and **Scott Chapman** will discuss a variety of z/OS performance measurement, analysis, and tuning techniques that may not be commonly known or are not often discussed.

The key objective of this presentation is to provide the attendee with information they can bring back to their shop and conduct some analysis or tuning exercises.

A secondary objective of this session is to help the attendee learn more about the z/OS environment and how things work. This session is sure to be highly educational!

Let's answer some common questions



- We structured this presentation to discuss some common questions we get, as well as some recent recommendations
 - After this presentation, consider these subjects as they relate to your environment
- Topics / questions to be discussed:
 - Why are I/O times in the SMF 30 record so far off from I/O times in other records?
 - How much power is your CEC using, and do you really care?
 - What is the latest info and recommendations regarding the WLM CPU Critical Control?
 - System Recover Boost (SRB) – Are you using it??
 - What are some of your thoughts on I/O avoidance?
 - Is my DDF workload using High Performance DBATs?
 - And if so, how does this affect my WLM goals for DDF transactions?
 - Should I cut SMF 98 and 99 records?
 - Tell me more about these new Java CPU time measurements in SMF 30, and how they relate to the zIIP CPU times in the SMF 30 record.

EPS: We do z/OS performance...



- We are z/OS performance!
- Pivotor
 - Performance reporting and analysis of your z/OS measurements
 - Example: SMF, DCOLLECT, other, etc.
 - Not just reporting, but cost-effective analysis-based reporting based on our expertise
- Performance Educational Workshops (while analyzing your own data)
 - Essential z/OS Performance Tuning
 - Parallel Sysplex and z/OS Performance Tuning
 - WLM Performance and Re-evaluating Goals
- Performance War Rooms
 - Concentrated, highly productive group discussions and analysis
- MSU reductions
 - Application and MSU reduction, as well as contract reviews

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
 - July 27 – 31, 2026 (4 days, excl Wednesday the 29th)
- 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?

Winter / Spring 2026 Webinars



- Free z/OS Performance Educational webinars!
 - The titles for our Winter/Spring 2026 webinars are as follows:
 - *New Year's Resolutions for z/OS Performance and Capacity People*
 - *How WLM Makes Decisions*
 - *What I Learned about VSAM RLS SMF Data*
 - *z/OS Performance Spotlight: Some Top Things You May Not Know*
 - *Building a Strong Foundation When You're New to z/OS Performance*
 - *Wait...Do We Need to Re-evaluate our WLM Goals?*
 - *z15 to z16 to z17 – What has changed?*
 - *Evaluating in the Mainframe Environment*
 - *Managing Workload Manager: Multiple Sysplexes and Asymmetric Sysplexes*
 - *Introduction to z Processor Measurements*
 - *(more to be announced)*
- 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 to show you the results
 - See also: <http://pivotor.com/cursoryReview.html>

Like what you see? Spring / Summer 2025 Webinars



- Free z/OS Performance Educational webinars!
 - 30-60 minutes long
 - 2026 Schedule Being put together now:
 - New Year's Resolutions for z/OS Performance and Capacity People
 - How WLM Makes Decisions
 - What I Learned about VSAM RLS SMF Data
 - Plus more... we're finalizing!
 - 2025 Webinars included:
 - ✓ *Overseeing z/OS Performance Management With Your Outsourcer*
 - ✓ *Back to basics - Processor Consumption Analysis*
 - ✓ *WLM and CPU Critical Control*
 - ✓ *Back to Basics - Evaluating Latent Demand*
 - ✓ *Understanding SMF 98 Locking Measurements*
 - ✓ *Overseeing z/OS Performance Management with Your Outsourcer*
 - ✓ *ETR vs ITR and the Basics of IBM's LSPR*
 - ✓ *Processor Comparison Discussion*
 - ✓ *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>

Like what you see?



- The z/OS Performance Graphs you see here come from Pivotor™
- If you just 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>
- 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

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 c

More Free Things!



● On our web site click on Tools & Resources to access:

- WLM to HTML Tool
 - Get your WLM policy in a useful and usable HTML format
- Our Presentations
 - Lots of great content from the past few years (now even easier/faster to access!)

<https://www.epstrategies.com>

<https://www.pivotor.com/>

(Same site behind both URLs)

The screenshot shows a web browser window with the URL `pivotor.com/content.html`. The website header includes the EPS and PIVOTOR logos, the company name "Enterprise Performance Strategies Inc. Creators of Pivotor®", and a navigation menu with "Tools & Resources" expanded. The expanded menu lists: "WLM to HTML tool", "Our Speaking Schedule", "Our Presentations", "Our YouTube Channel", "Free Cursory Review", and "Free Performance Reporting". The main content area is titled "EPS Papers and Presentation" and contains a list of presentations for the year 2023. Two presentations are visible, each with a "Download" button and a title link.

| 2023 | |
|--------------------------|--|
| Download | Peter Enrico - Key Reports to Evaluate Coupling Facility CPU Utilization |
| Download | Scott Chapman - Understanding How Memory Management Has Evolved in z/OS |

Before we get into the meat of our presentation... some words about Pivotor

Unsurprising recommendation



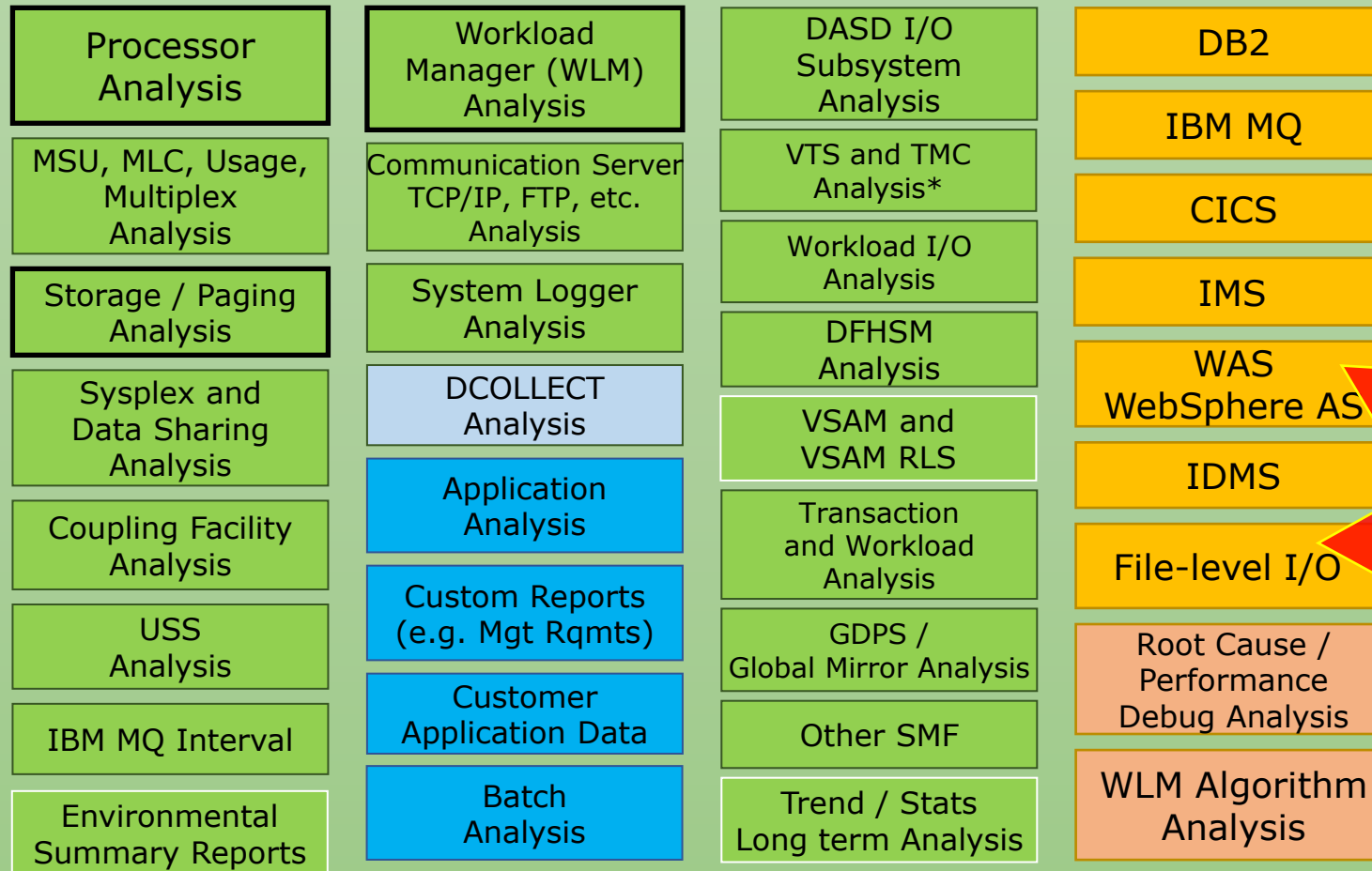
- You need a performance reporting / investigation tool at the ready
 - If all you have is RMF/CMF post processor reports, that's not great
 - And it only helps for a subset of the SMF performance data
 - Whatever you have you should use it on a semi-regular basis to be familiar with it
 - *AND... the costs for such a service and product should not be high!*
 - *Consider Pivotor!*
- We're also seeing more sites outsourcing their z/OS environment
 - There's of course a wide range of arrangements and outcomes, but...
 - Performance management in these environments can suffer
 - *Retaining your own oversight of z/OS performance is a good idea*
 - Have heard from multiple sites where the outsourcer surprisingly didn't provide useful performance reporting
 - *Consider Pivotor!*

Pivotor – Intelligent Reporting



- Pivotor is our data reporting tool & service designed specifically for z/OS performance reporting
 - Designed and used by z/OS performance experts
 - Processes data from SMF, DCOLLECT, and customer sources
 - Contains hundreds of z/OS performance reports “out of the box”
 - Designed to be easy to use and manage
 - Reports are organized into logical and searchable report sets
 - Features include intelligent exceptions, drill down, search, canned analysis, and so much more
 - Built in expanded helps to help foster education

Comprehensive Report Sets for Immediate Performance Analysis



**>3000 reports
"out of the box"**

Across multiple timeframes: daily, weekly, monthly, yearly, rolling n days, etc.

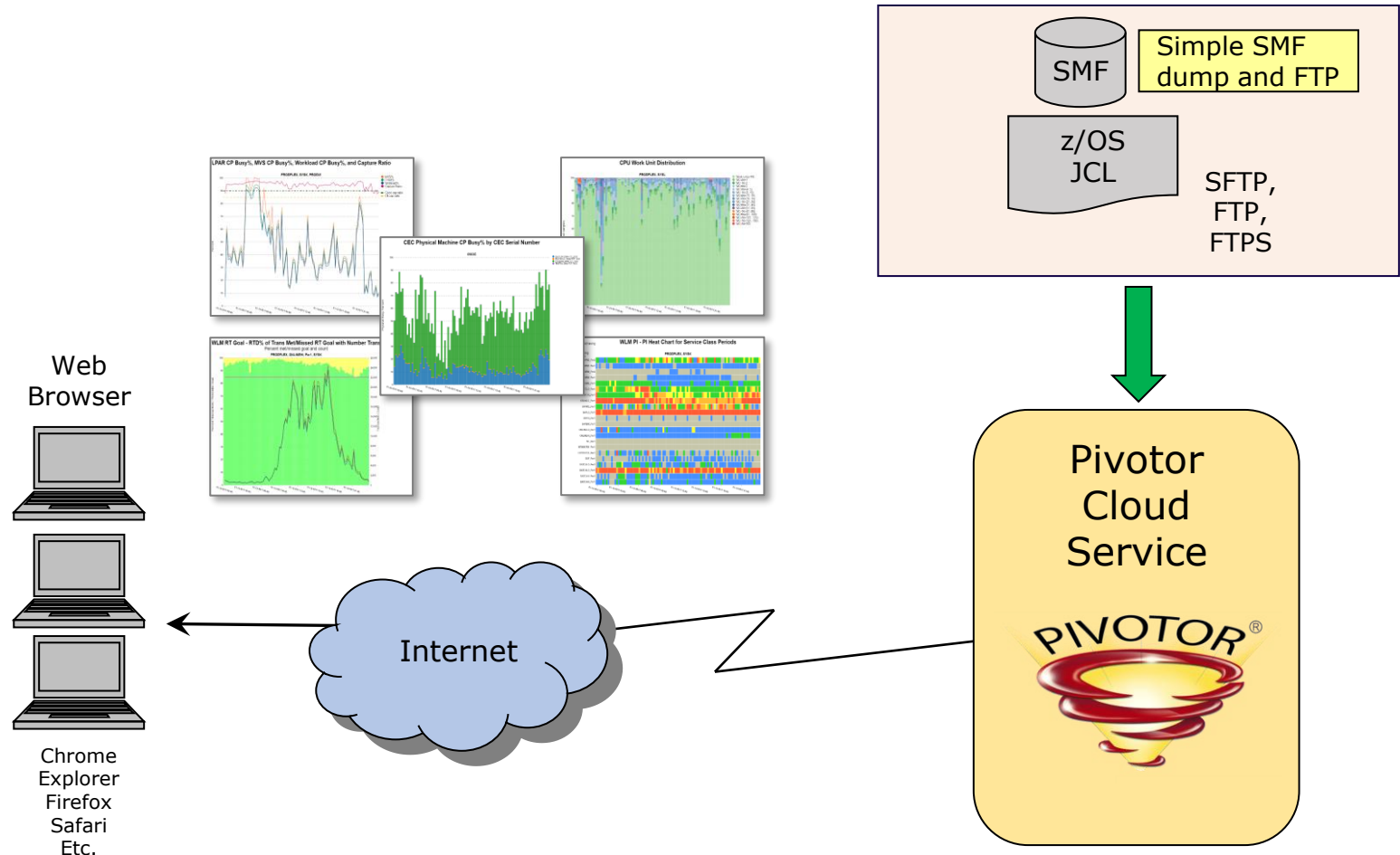
Pivotor Software as a Solution (SaaS)



- Pivotor is offered as both a SaaS or local install
- When SaaS:

SaaS Includes:

- Formal yearly cursory review / discussion
- Ability to ask us performance questions, or for us to look at a particular problem or concern. (support@epstrategies.com)
- We can occasionally look in on your data and performance
- We can participate in performance debug with IBM, or other vendors



Just Some of the Valuable Pivotor Features



- Drill down / Search capability
 - Search by field, value, tag, or keyword
- Automatic aggregation of long-term data, with zoom to detail level
 - Min/max aggregation avoids hiding peaks in long-term averages
- Report overlays
 - When there is a need to look at one report on top of another
- Time filtering
 - When you want to focus on a specific time period within a reportset
- Application Attribution
 - Enables fast and flexible custom reporting based on customer groupings
- Dashboards
 - Multiple customizable dashboards
- Ability to incorporate business metrics into performance reports
 - Capture non-SMF data such as call center calls or bills produced

Just Some of the Valuable Pivotor Features



- Every report rich-view, CSV, table, and image
- Easily share reports
 - Export to PDF or Copy-Paste URL
- Playlist Analysis
 - Playlists ‘walk’ a user through an analysis
- Exception analysis
 - Simple to complex conditions
- Outliers
 - What’s different this week
- Expanded helps
 - Pivotor is a great vehicle for learning about z/OS performance
- Data Explorer
 - For dynamic custom reporting based on logical sets of data
- Batch timeline reports

Why are I/O times in the SMF 30 record so far off from I/O times in other records?

Note: Pivotor has extensive I/O and storage reporting from the DASDPlex level all the down to the dataset / file level.

I/O Time in SMF30 is Wrong



Note: The system adjusts the connect time for FICON DASD to be 1 millisecond per request. This value differs from the channel reported connect time.

- Connect time is (historically) data transfer time
 - FICON's introduction of concurrent active I/Os per channel complicated this measurement
 - Apparently, the complication was too much for the SMF 30 records?
- In the early days of FICON 1ms might have been more reasonable than today
 - Today 0.1-0.2ms might be closer to accurate
- Makes comparing jobs by I/O time is problematic
 - Transfer times can of course vary based on size of the I/Os
 - Which may vary quite a bit for batch work
 - 1ms significantly higher than common FICON DASD total RTs today

Related: I/O is Complicated



- We came across this following up to a question from Scott's "z/OS I/O Performance: Do You Have a Problem?"
 - Thanks for your questions and feedback!
- Summary of that presentation:
 - Some of the measurements are confusing
 - I/O problems can hide in aggregated measurements
 - It's hard (essentially impossible) to know what I/O is really important to you
 - I/O reduction efforts are undervalued today
- See also "I/O, I/O It's Home to Memory We (Should) Go"
 - zHyperLink is fast, but memory is faster!

<https://www.pivotor.com/content.html>

How much power is your CEC using, and do you really care?

How much power is your mainframe using?



- There are now (somewhat surprisingly granular) energy consumption measurements in the SMF 70 records

| SMF 70 Field | Description |
|--------------------------|--|
| SMF70_CPUPower | Accumulated microwatts for all CPU resources allocated to the LPAR during the interval. |
| SMF70_StoragePower | Accumulated microwatts for all storage resources allocated to the LPAR during the interval. |
| SMF70_IOPower | Accumulated microwatts for all I/O resources allocated to the LPAR during the interval. |
| SMF70_CPCTotalPower | Accumulated microwatts for all electrical and mechanical components in the CPC. |
| SMF70_CPCUnassResPower | Accumulated microwatts for all types of resources in the standby or reserved state. |
| SMF70_CPCInfraPower | Accumulated microwatts for all subsystems in the CPC which do not provide CPU, storage, or I/O resources to logical partitions. These include service elements, cooling systems, power distribution, and network switches, among others. |
| SMF70_NumPowerSamples | Number of power samples for the interval. |
| SMF70_PowerPartitionName | The name of the LPAR to which the LPAR-specific power fields apply. |

How much power is your mainframe using?



- Energy consumption measurements in the SMF 70 records
 - Requires z17
 - Recorded at the CEC level
 - IBM has stated that the z17 uses 19% less power than z16
- Audience poll: does anybody care about this data?
 - Other than the abstract/academic point of “that’s interesting trivia”
 - Is somebody pushing you to quantify your datacenter power consumption by device?
 - Do you think this could meaningfully improve the perception of the mainframe?
 - Is the power costs of a z machine a big issue for you?
 - Do other devices, such as storage controllers, record power measurements?
 - That is... should Pivotor bother reporting on this?

System Recover Boost – Are you using it??

A lot of people aren't using it as much as they could...

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
 - Faster subsystem restart
- Requires z15 or later machine
 - + z/OS 2.3 and up on z15, z/OS 2.4 and up on z16
- No additional charge for basic System Recovery Boost
 - Optional 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?



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 | <=5 minutes | No change |

● Recovery Processes Boost

- 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 re-duplexing
- HyperSwap Recovery
 - Recovery from storage controller failover

New with z16:

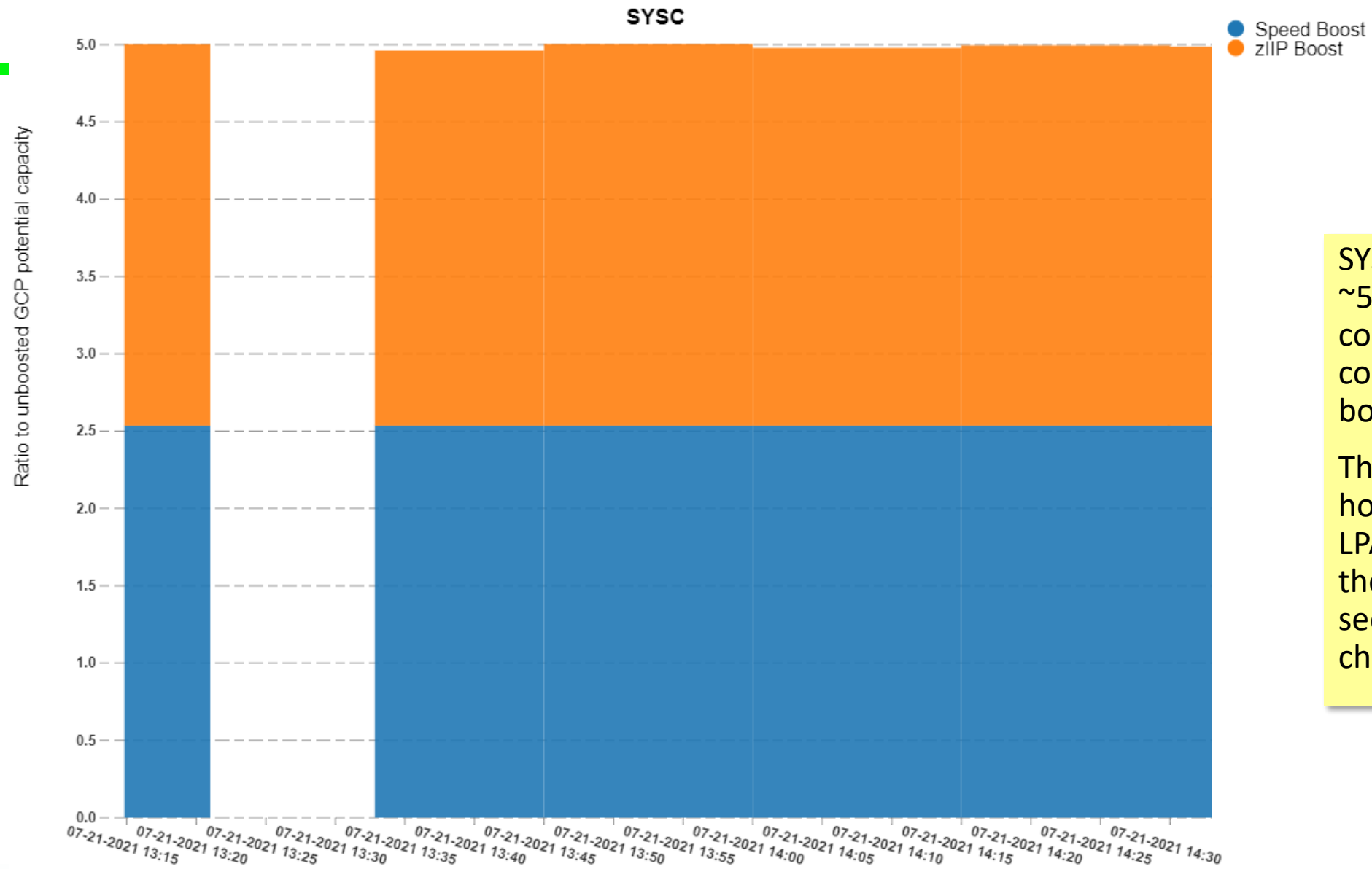
- SVC Dump
- HyperSwap Configuration Load
- **Middleware Region Startup**
- **Dynamic I/O Activation**

More people should be using more SRB



- Boosts you get by default, without really taking action:
 - IPL
 - Recovery process for things like sysplex partitioning
- You must take action to get boosts for:
 - Shutdown
 - Middleware Recovery
 - SVC Dump boosts
- We're not seeing much of the latter—why??
 - If you're on TFP, all the more reason to think about this
- The extra capacity available to a boosted LPAR can be significant!

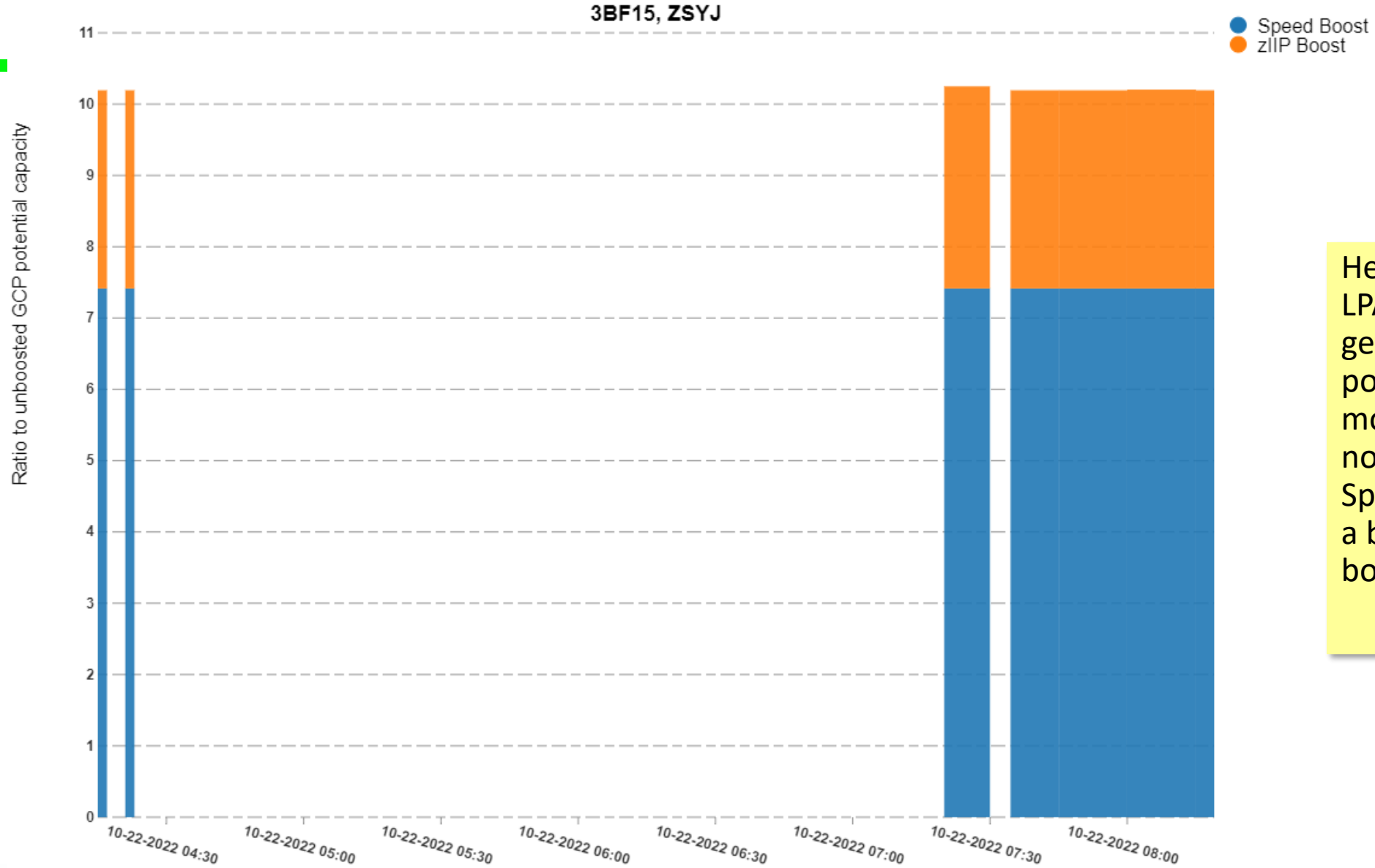
Boost Potential Capacity Ratio



SYSC could have done ~5x the GP work compared to what it could have done unboosted.

This report considers how much the other LPARs are using hence the slight variations you see in the ratios on each chart.

Boost Potential Capacity Ratio



Here's an example of an LPAR on a z15-411 getting access to potentially over 10x more capacity than it normally has access to. Speed boost is obviously a big part of this, but zIIP boost is helping too.

LPARs *not* being boosted



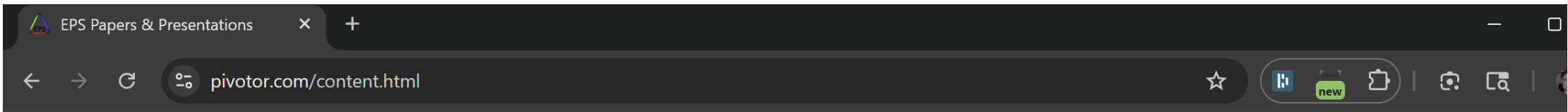
- These LPARs *could* be negatively impacted
- CPU cache effectiveness may be impacted by boosted LPARs (at least theoretically)
 - Speed boosted LPARs do more work per unit of time on the GPs
 - zIIP boosted LPARs may drive more work to zIIPs
 - zIIP caps ignored during boost periods
- Higher physical 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**
 - Or (in theory) if trying to give more capacity to boosted LPAR
- Resource Group caps with sysplex scope don't count work running on boosted LPARs
 - May allow more than expected work to run on un-boosted systems
- Conversely:
 - If boosted systems consume less of GP capacity, might help non-boosted systems

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
- So in some cases, you *may* want to consider setting up automation to make weight changes (particularly for zIIPs) during boost intervals

See also <https://pivotor.com/content.html>



Enterprise Performance Strategies Inc.

Creators of Pivotor®

- Home
- Pivotor
- Workshops
- Consulting
- Webinars
- Tools & Resources ▾
- About ▾

- WLM to HTML tool
- Our Speaking Schedule
- Our Presentations
- Our YouTube Channel
- Free Cursory Review
- Free Performance Reporting

- [Download](#) Scott Chapman - High, Medium, Low: Understanding how HiperDispatch Influences Performance
- [Download](#) Peter Enrico - Intensity! Understanding the Concepts and Usage of Intensity Measurements

2022

- [Download](#) Scott Chapman - WLM Management of DDF Work: What can you do and what has changed?
- [Download](#) Scott Chapman - Understanding System Recovery Boost's Impact on Performance and Performance Reporting
- [Download](#) Peter Enrico - Key Reports to Evaluate z16 Processor Caches



Can you comment with your thoughts on I/O avoidance?

Note: Pivotor has extensive I/O and storage reporting from the DASDPlex level all the down to the dataset / file level.

I/O sometimes worse than we think



- Average I/O RTs have generally looked really good for years
 - Modern I/O subsystems solve many of the problems of the past
- Got more interested in I/O due to the “suffering” reports
 - “High Disk RT” and “Interrupt Delay” showed up more than expected
 - Avoid assuming there are or are not problems in particular areas!
- Averages hide the outliers though, some of which are quite large
 - According to some SMF data, maximums can be 1000x average!
- Interrupt delay a “hidden” increase in I/O response time
- In short: it’s a good idea to do less I/O
 - Data in Memory techniques are still important

Memory and I/O

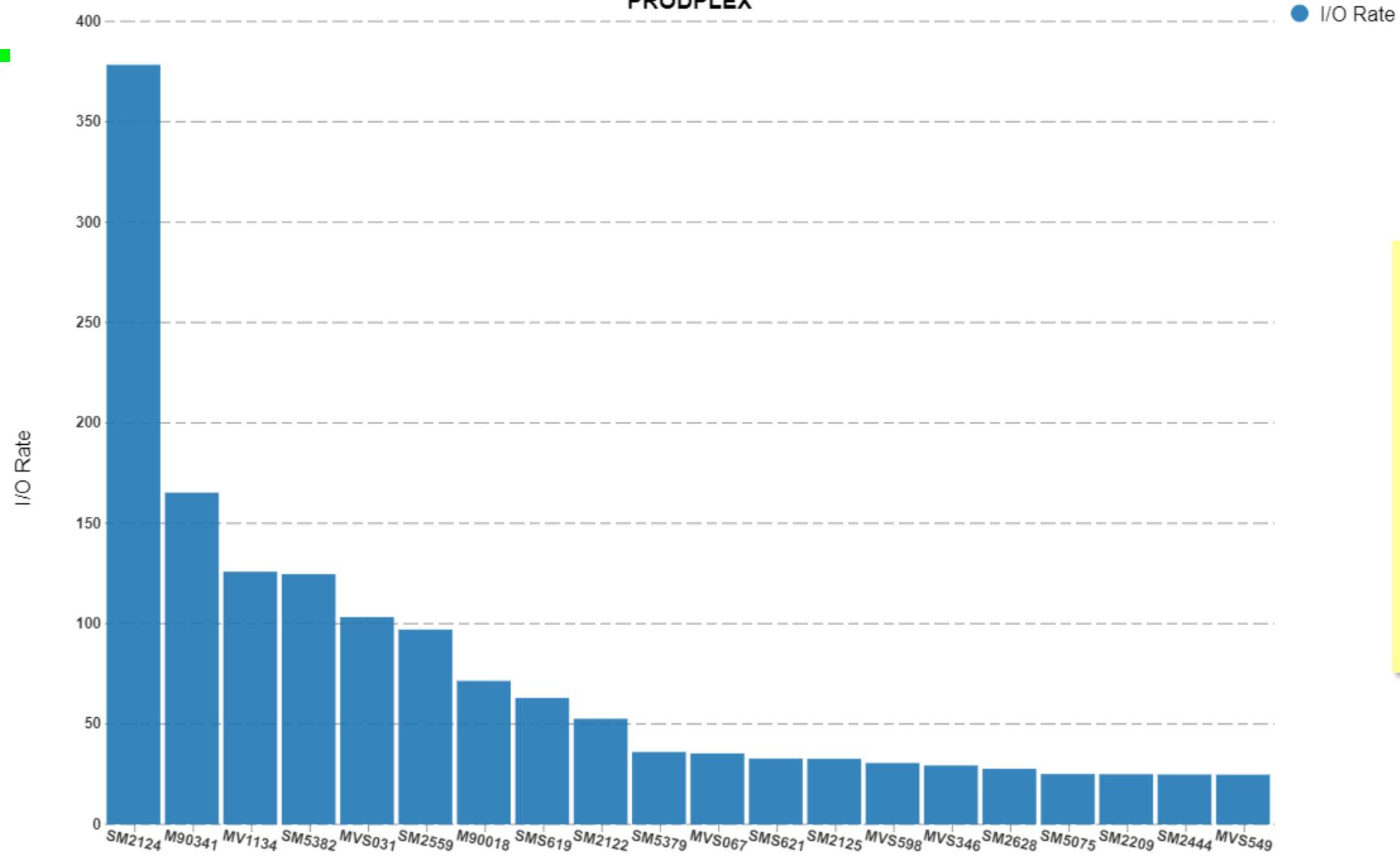


- We see systems with lots of memory free and yet they're doing significant amounts of I/O
- We've been saying for a long while things like "make your BPs bigger"
- But lately we've been trying to look deeper to point out opportunities
 - How much data is really on those busy volumes?
 - Which specific datasets are getting lots of read I/O

LVs with Highest I/O Rates

(Averaged Over Period of Study)

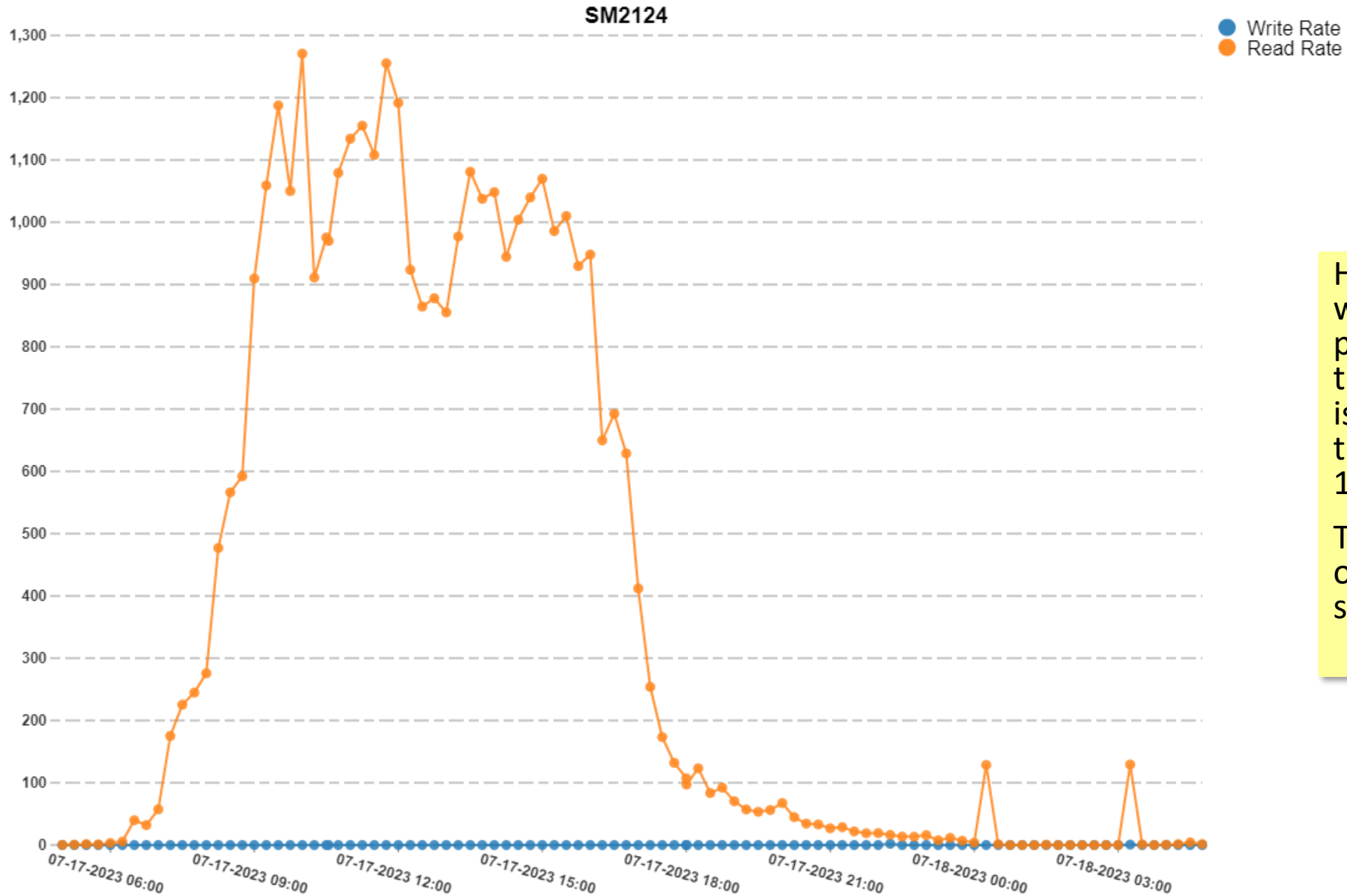
PRODPLEX



This Pivotor report shows the top volumes by I/O rate over the day.

375 IOPS doesn't sound too interesting but note that is an average I/O rate over 24 hours.

Logical DASD Volume Explorer



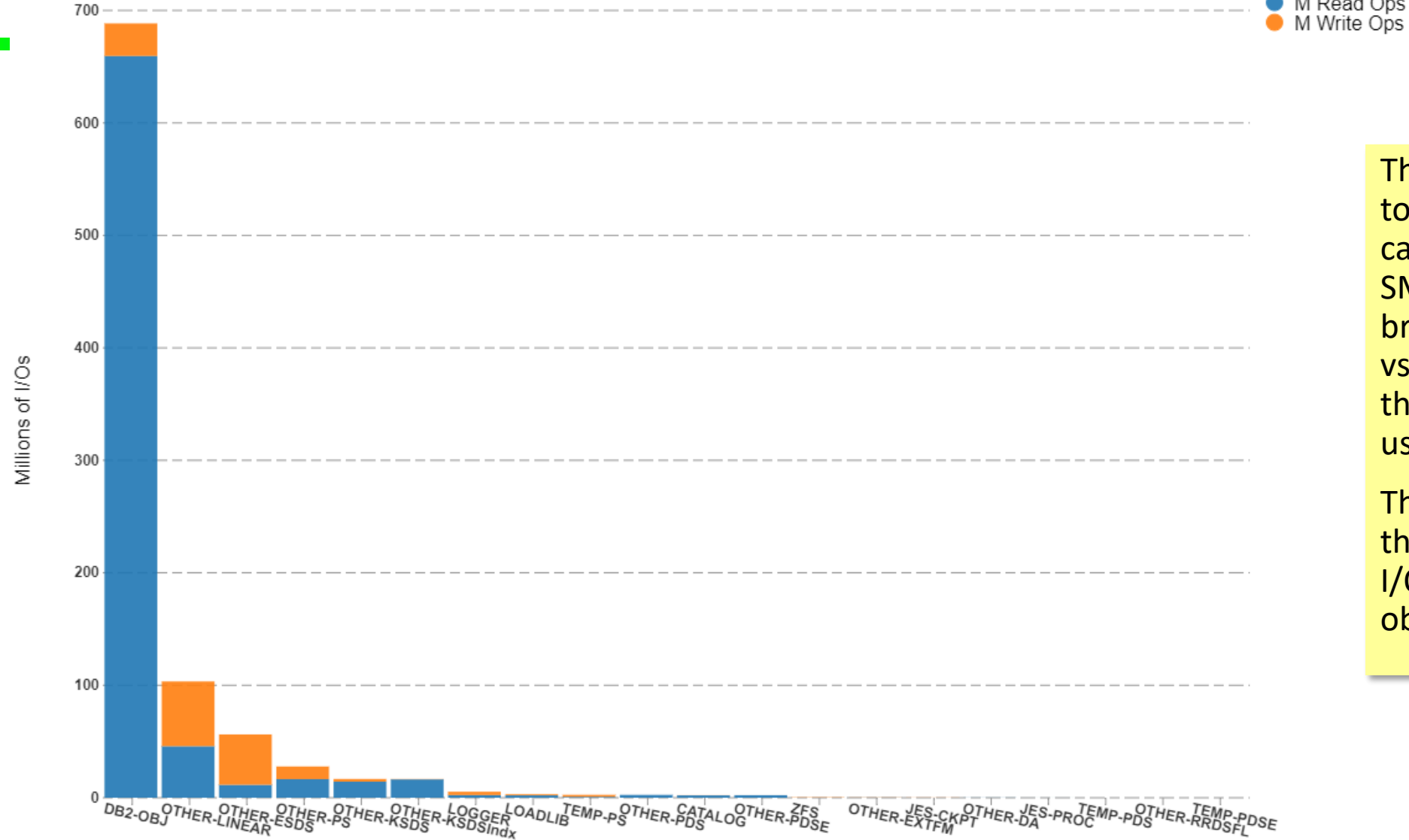
Here's the read and write rate for a particular volume over time. Virtually all the I/O is read I/O, and during the day it is doing over 1000 IOPS.

The kicker: this volume only has 1.5 GB of data stored on it!

Top Dataset I/O Counts by Dataset Usage

Total I/Os for Study Period

SYSA



This reports looks at the total I/O over (in this case) a day from the SMF 42 records and breaks it down by reads vs. writes and by what the dataset is (probably) used for.

This site is not unusual: the vast majority of the I/O is reading from DB2 objects.

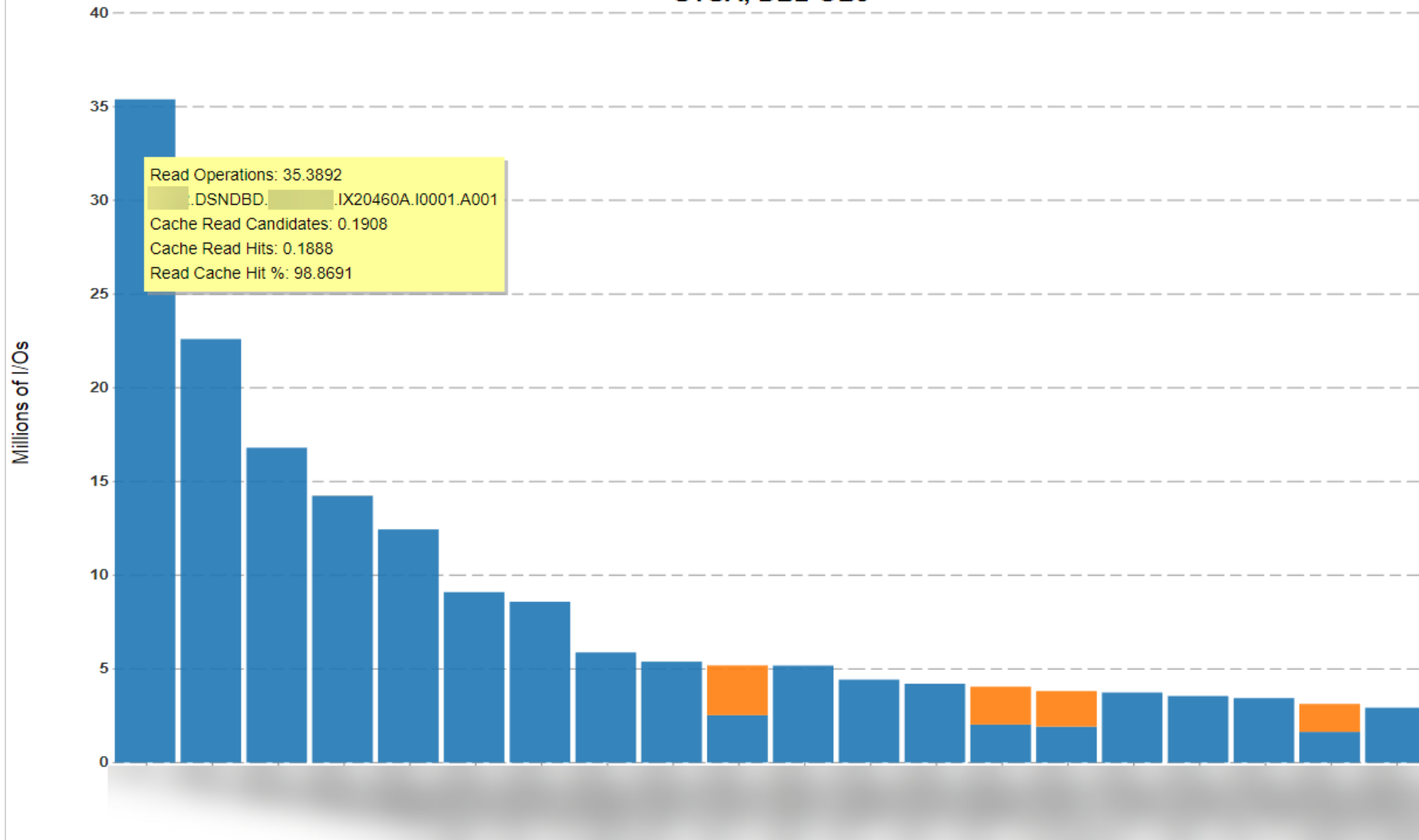
Top Dataset I/O counts by Dataset Usage

Total I/Os for Study Period

SYSA, DB2-OBJ



● Read Operations
● Write Operations



The top dataset appears to be all reads, but oddly, only a tiny fraction of those apparently are flagged as being cache candidates. I'm not sure why that is. But in modern control units all I/O passes through cache.

Top Datasets by Cache Read Hits

2023-07-17



| Date | Usage | DS Name | M Cache Read Hits | Cache Hit Pct | Read MiB | Allocated MiB | Read-Allocated Ratio | Read O... | Write Ops | 42.6 Records | Volume | |
|------------|--------------|---------------|------------------------|---------------|----------|---------------|----------------------|-----------|-----------|--------------|---------|-----|
| Select Fil | Select Filtr | Select Filter | | | | | | | | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | IX20460A.I0001.A001 | 0.189 | 98.869 | 4,201,991.004 | 569.841 | 7,373.973 | 35.389 | 0.000 | 54.000 | 2.0 |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | IX08956B.I0001.A001 | 20.108 | 97.188 | 664,607.695 | 2,361.233 | 281.466 | 22.601 | 0.001 | 251.000 | 8.0 |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | /02.TS06435.J0001.A001 | 1.450 | 98.796 | 1,880,165.453 | 569.841 | 3,299.457 | 16.798 | 0.000 | 245.000 | 5.0 |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | IX00854E.I0001.A001 | 14.170 | 99.960 | 69,569.031 | 1,339.897 | 51.921 | 14.244 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | IX08956B.I0001.A002 | 11.436 | 97.651 | 288,683.320 | 1,159.947 | 248.876 | 12.444 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | .TS01452.J0001.A001 | 8.968 | 99.152 | 71,634.121 | 1,203.719 | 59.511 | 9.097 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS07315.I0001.A001 | 7.206 | 98.403 | 696,717.207 | 4,175.324 | 166.865 | 8.580 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS20310.I0001.A001 | 5.076 | 96.649 | 135,041.418 | 4,720.846 | 28.605 | 5.882 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS08957.J0001.A001 | 4.991 | 99.302 | 228,975.305 | 4,721.657 | 48.495 | 5.390 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS00854.I0001.A014 | 3.724 | 89.910 | 198,572.117 | 1,957.563 | 101.438 | 5.179 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | .TS01451.J0001.A001 | 1.365 | 94.870 | 332,869.598 | 306.401 | 1,086.384 | 4.428 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS07315.I0001.A001 | 3.588 | 99.894 | 499,237.086 | 4,377.969 | 114.034 | 4.210 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS07315.J0001.A001 | 3.121 | 99.978 | 417,840.731 | 4,341.493 | 96.244 | 3.551 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | .TS07292.I0001.A001 | 2.965 | 99.891 | 77,346.273 | 4,722.468 | 16.378 | 3.444 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | .TS17613.I0001.A001 | 0.156 | 94.461 | 345,553.938 | 284.516 | 1,214.534 | 2.930 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | .TS07292.J0001.A002 | 2.255 | 99.791 | 71,282.477 | 4,721.657 | 15.097 | 2.738 | | | |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS02809.J0001.A009 | 2.170 | 89.721 | 77,474.367 | 3,465.251 | 22.358 | 2.688 | 0.003 | 133.000 | 3.0 |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS02813.J0001.A004 | 2.085 | 95.299 | 63,996.520 | 3,465.252 | 18.468 | 2.609 | 0.070 | 338.000 | 7.0 |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS06562.I0001.A001 | 2.316 | 96.014 | 37,477.902 | 1,738.704 | 21.555 | 2.499 | 0.001 | 64.000 | 4.0 |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | TS17820.I0001.A001 | 1.968 | 92.237 | 73,247.258 | 2,691.953 | 27.210 | 2.451 | 0.000 | 51.000 | 4.0 |
| 2023-07-17 | DB2-OBJ | .DSNDBD. | .TS17613.J0001.A001 | 0.062 | 93.959 | 295,867.445 | 284.516 | 1,039.899 | 2.446 | 0.000 | 18.000 | 2.0 |

This table report joins the SMF 42 data with the DCOLLECT data to get the total allocated size (summed across multiple volumes if necessary) of the datasets.

Note there's little write activity and a number of these datasets are only a few GB.

Even if they can't all go into memory, probably some can, saving 10s of millions of I/Os.

I/O – Worse than we think



- While working on my constraint score I saw I/O contributing more than expected to the score
 - Impact to constraint was when I/O RT (excluding connect) > 0.5 ms
 - Seems quite reasonable, but is still quite imperfect
- Also saw a fair bit of impact from Interrupt delay
 - Especially when no high pool processors
- Digging through some details: high response times for particular volumes or datasets seems more common than I would have expected
- Real I/O difficulty: which I/Os *actually* matter???
 - In many situations almost impossible to tell this from the system perspective

Is my DDF workload using High Performance DBATs?

And if so, how does this affect my WLM goals for DDF transactions?

Note: Pivotor can process and report on the DB2 SMF 100 to 102 records to help surface performance and performance problems.

Upending >25 Years of DDF Management



Enclaves allow the *management of individual transactions* flowing through address spaces, something that simply has never been possible before. Since MVS is aware of and has access to each transaction, they can be classified individually and most importantly *each transaction is subject to period switch*. ~~This means that you can separate out the long running CPU killers from the shorter requests in the same manner that most installations already employ to control batch, by period level controls.~~

Each enclave is a single transaction, which starts when the enclave is created and ends when the enclave is deleted. ~~DDF creates an enclave for an incoming request when it detects the first SQL statement and deletes the enclave at SQL COMMIT, thus a DDF enclave transaction consists of a single SQL COMMIT scope.~~

~~In WLM goal mode, all goal types are valid for enclaves.~~

This is now the situation if your DDF work uses High Performance DBATs!

(With DB2 APAR PH34378)

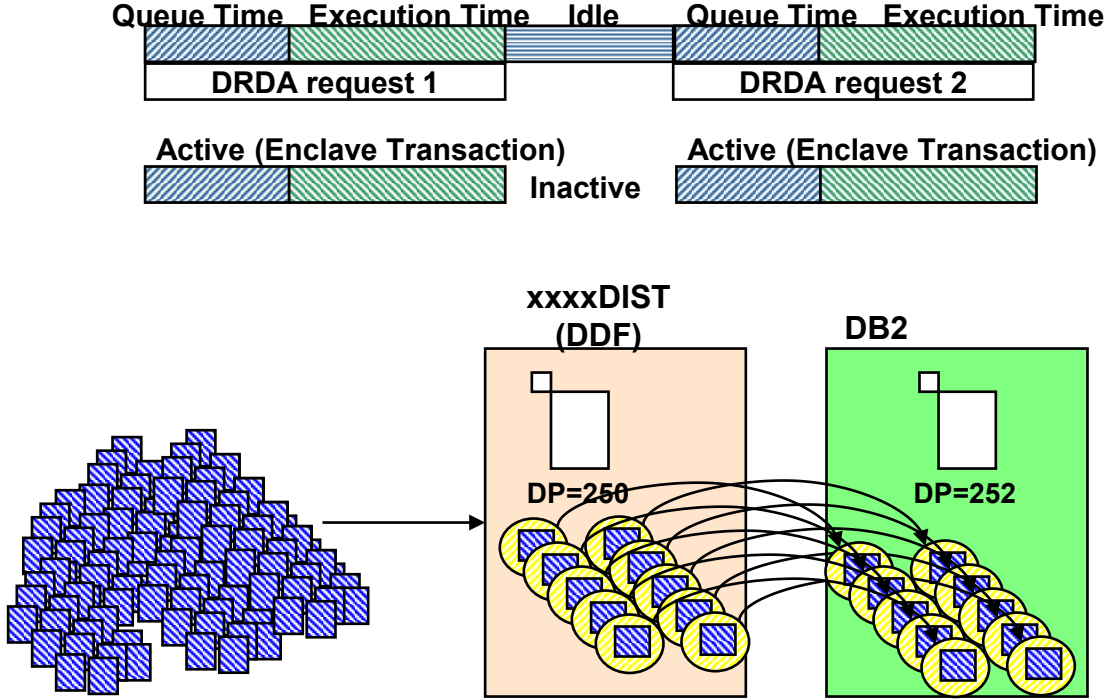
Technically: multi-period service classes & RT goals still are allowed but they will not work as expected!
This could be a surprise when you apply the DB2 maintenance that made this change.

Traditional / Typical DB2 DDF Enclaves



- DDF transactions are managed by WLM as independent enclaves
 - The life of the enclave is the life of the transaction (i.e. 1 tran = 1 ended enclave)

- In this case, 200 transactions come in, and for each transaction an enclave is created
- When each transaction ends, the enclave for that transaction is deleted, and an ended transaction and response time are posted to the WLM Service Class period
- WLM response time goals are best since each transaction is measured and considered by WLM

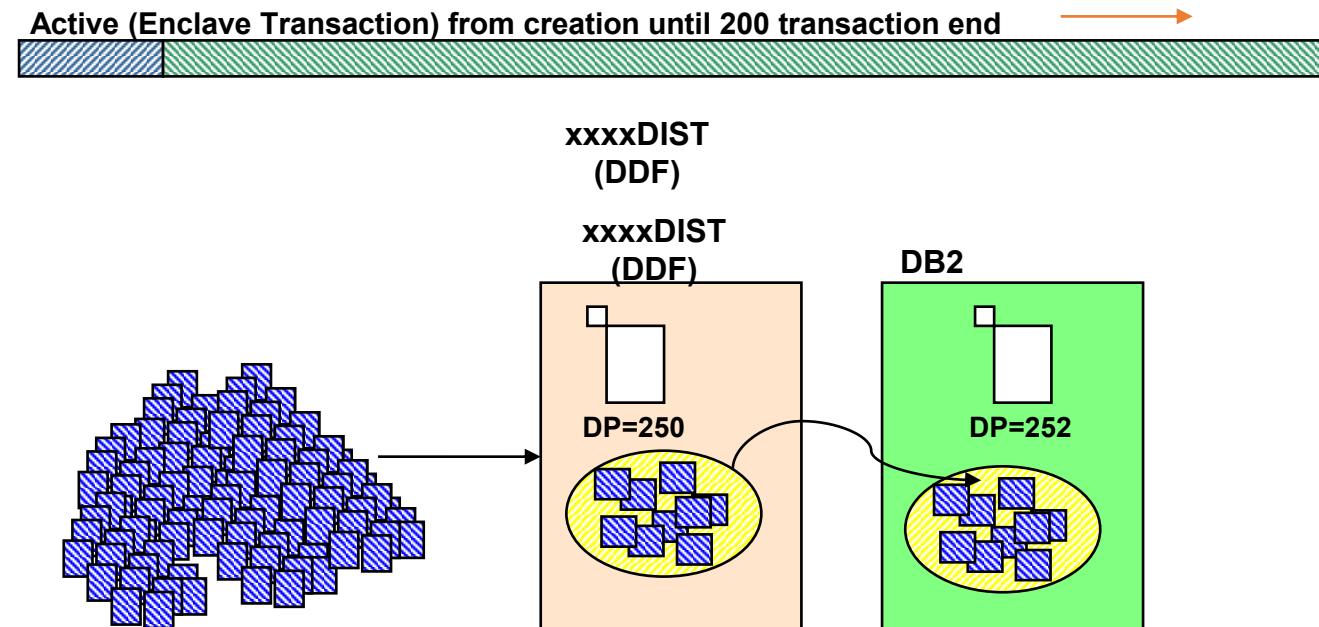


DBAT DB2 DDF Enclaves w/ PH34378



- The DBAT transactions still run in independent enclaves
 - But now the life of the enclave is for as long as up to 200 transactions in the enclave
 - The transactions in the enclave are known as ‘bundled transactions’.
 - Thus, you may have 200 ended bundled transactions, but from a WLM point-of-view, the 200 bundled transactions are part of a single (i.e. 1) longer running enclave.

- In this case, when a new thread is needed, a new thread is created and associated with a new enclave.
- That thread and associated enclave remain active for up to 200 transactions come in, and all are associated with a single enclave.
- After the last of the 200 transactions end, enclave is deleted, and one ended transaction is posted to the WLM service class period. Response time is from start of first transaction to end of last ended transaction
- The many transactions running in the enclave are known as ‘bundled transactions’



New SMF fields in support of Enclaves



- New SMF 72, subtype 3 SMF fields in support of enclaves
 - Fields are useful to help investigate if DB2 High Performance DBATs are being used

| Field Name | Description |
|-----------------|--|
| R723ENCTRXNUM | Number of subsystem transactions processed within enclaves. |
| R723ENCTRXCALLS | Number of times transaction data has been reported by subsystem work managers when deleting an enclave. When zero, no transaction data for enclaves has been provided by the subsystem work manager. |
| R723ENCTRXET | Total execution time, in microseconds, for all subsystem transactions reported in R723ENCTRXNUM |
| R723ENCTRXETS | Sum of squared execution times, in microseconds, for all subsystem transactions reported in R723ENCTRXNUM. |

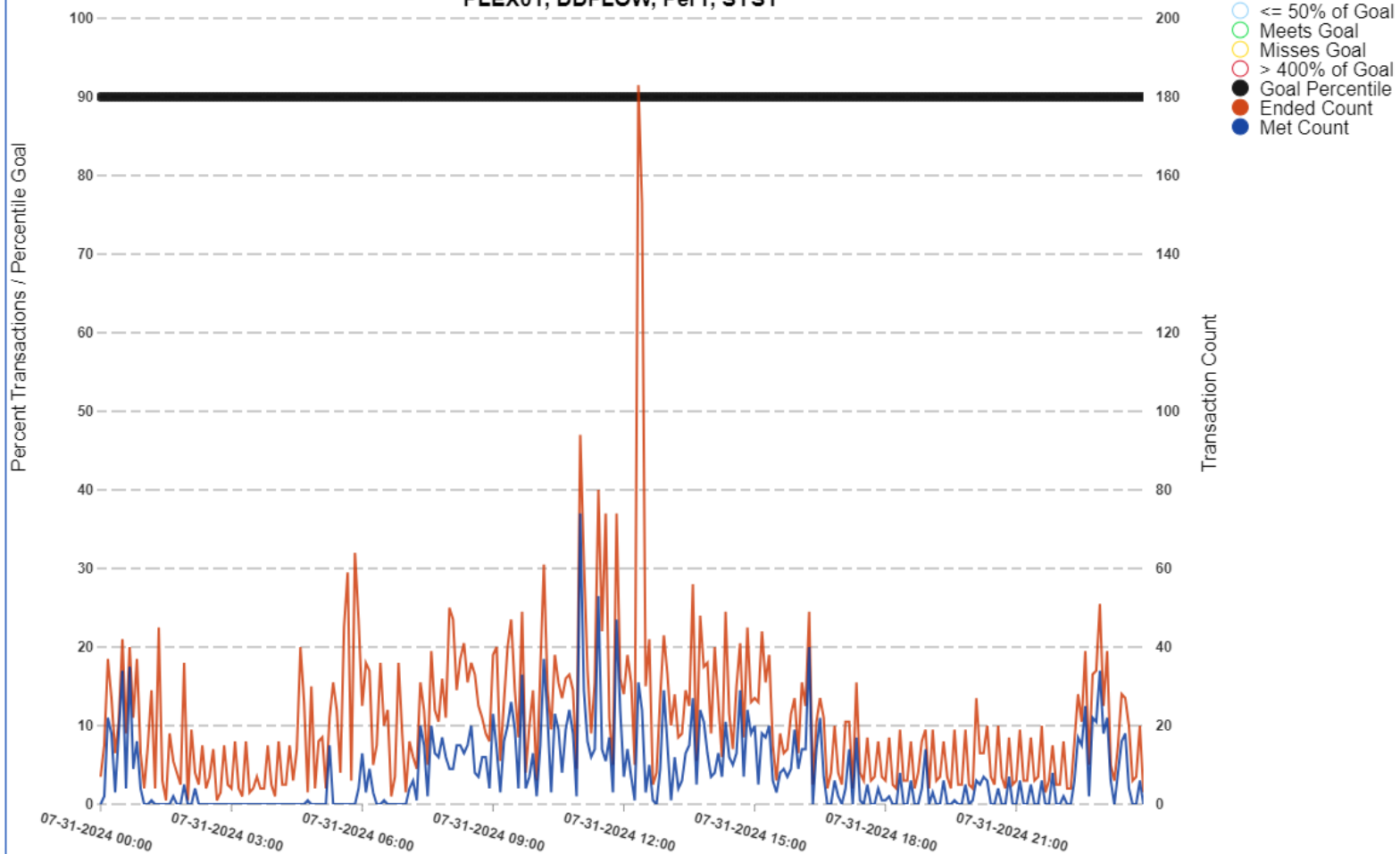
Example: DDFLOW Response Time Distribution



WLM RT Goal - RTD% of Trans Met/Missed RT Goal with Number Trans

Percent met/missed goal and count

PLEX01, DDFLOW, Per1, SYS1



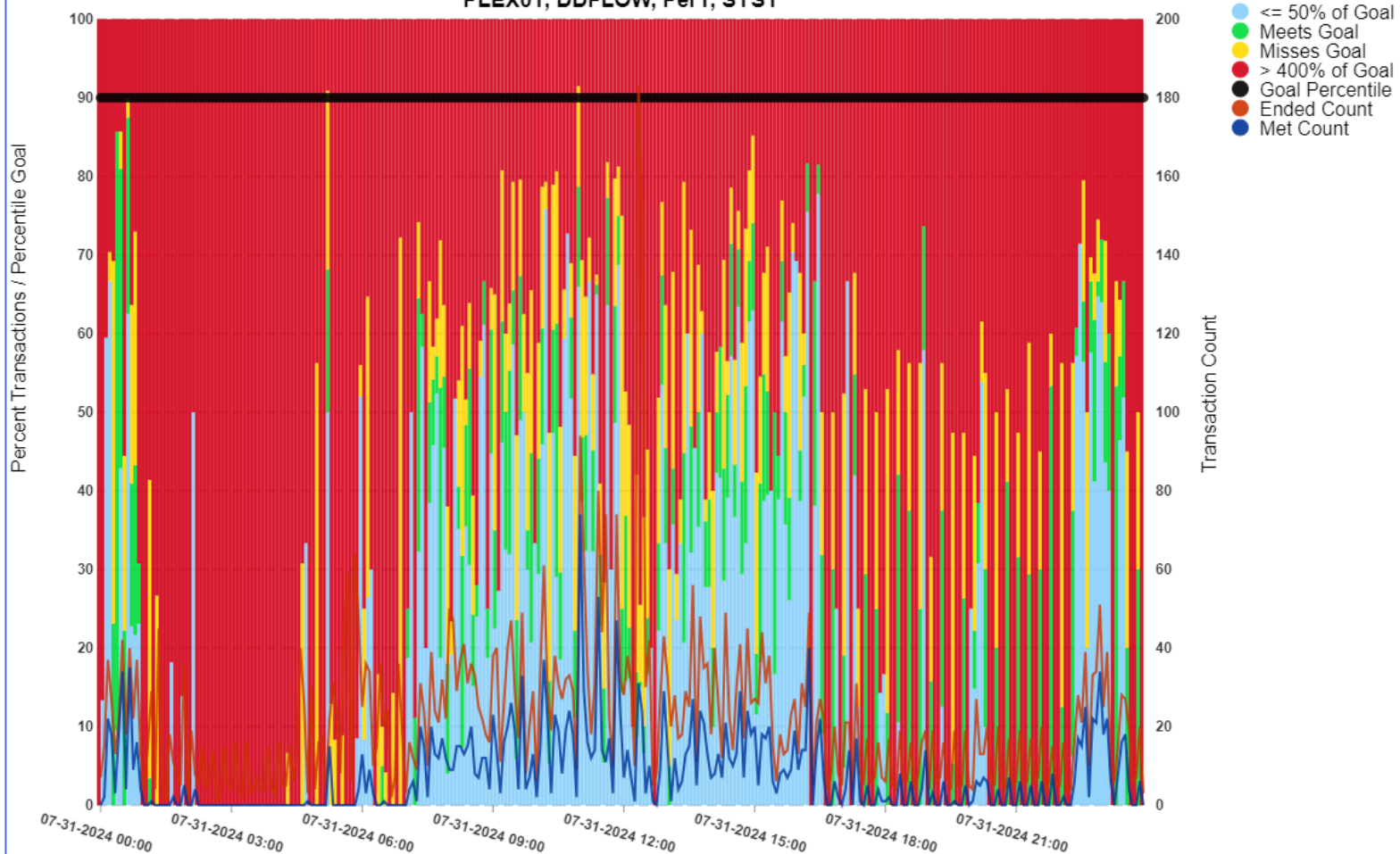
- DDFLOW has a response time goal of 90% complete in 0.3 seconds
- Notice that we have a fair low number of ended transactions
- Notice the large gap between number of ended transactions and the number that met the goal value.

Example: DDFLOW Response Time Distribution



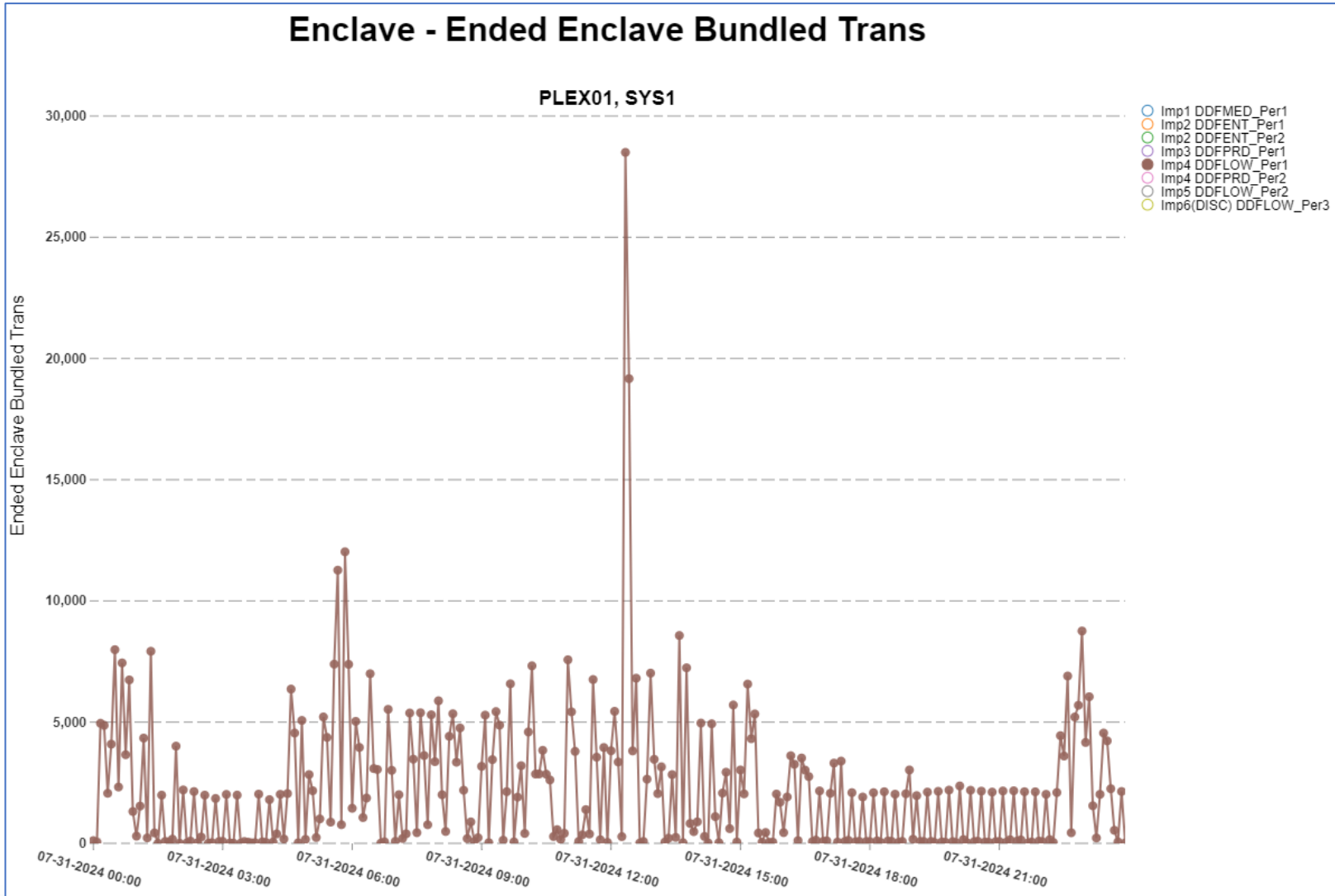
WLM RT Goal - RTD% of Trans Met/Missed RT Goal with Number Trans
Percent met/missed goal and count

PLEX01, DDFLOW, Per1, SYS1



- DDFLOW has a response time goal of 90% complete in 0.3 seconds
- Added to this chart is the percentage of transactions that met and missed the goals
- Not that a large percentage of transactions missed the goal value of 0.3 seconds

Example DDFLOW Ended Bundled Transactions



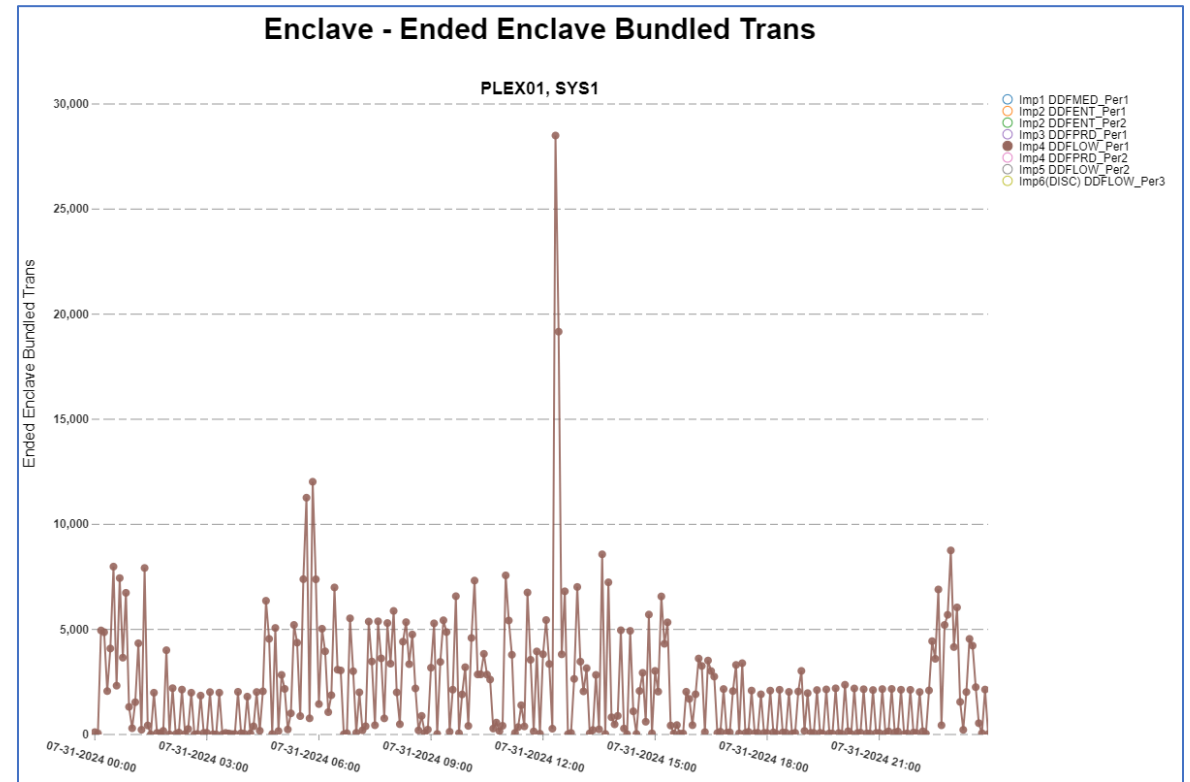
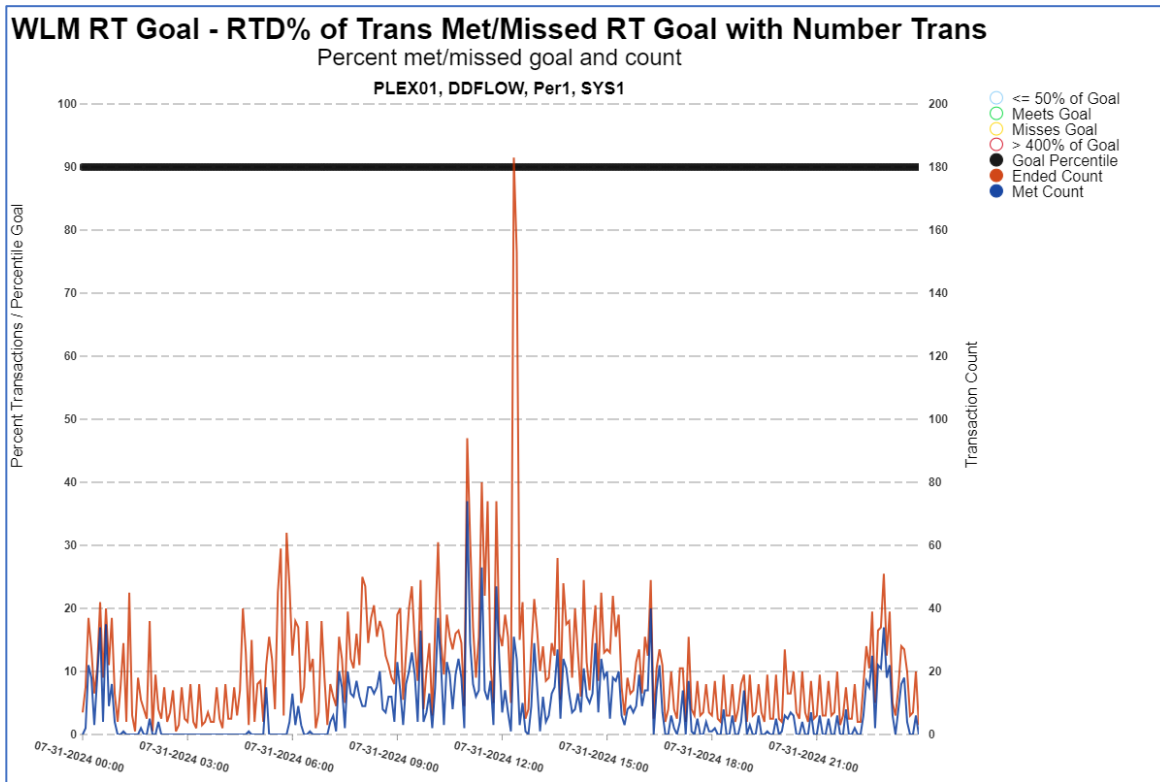
- DDFLOW has a response time goal of 90% complete in 0.3 seconds
- Notice there were thousands of ended bundled transactions

DDFLOW Ended Transactions



Ended WLM Transactions (Enclaves)
(What WLM uses for RT goals)

Ended Bundled Transactions
(Not used by WLM for goal management)

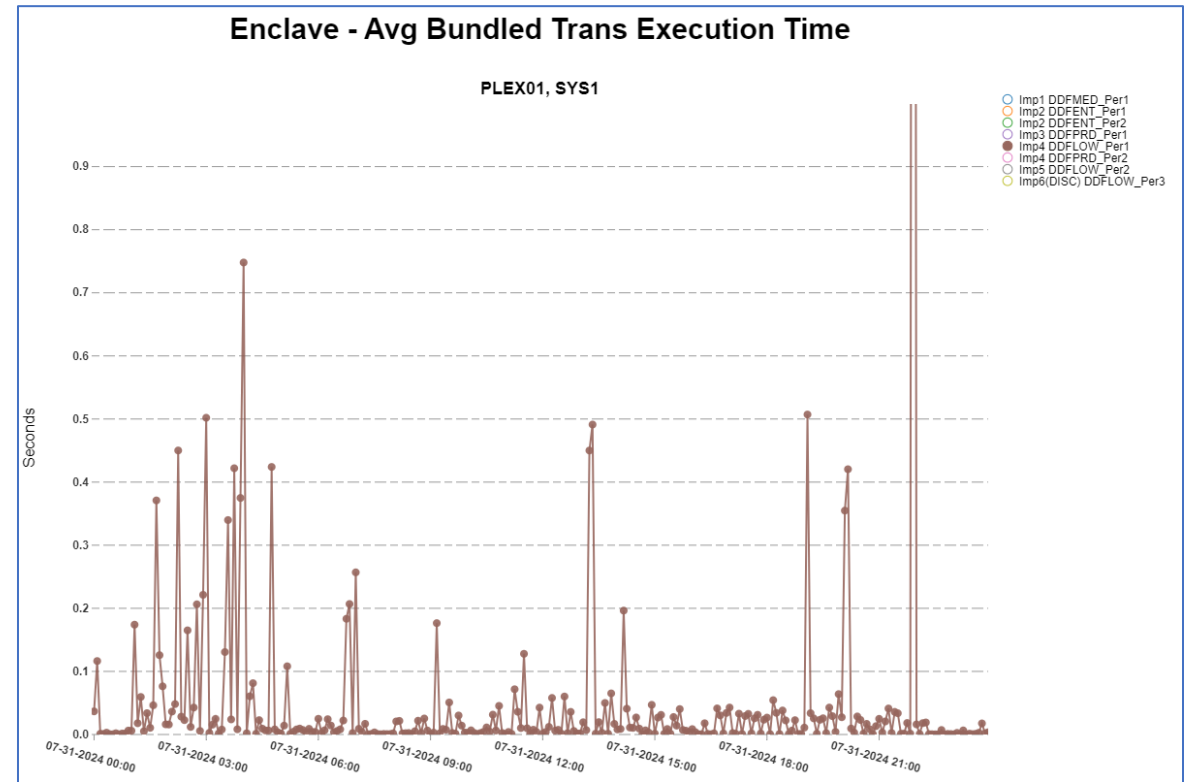
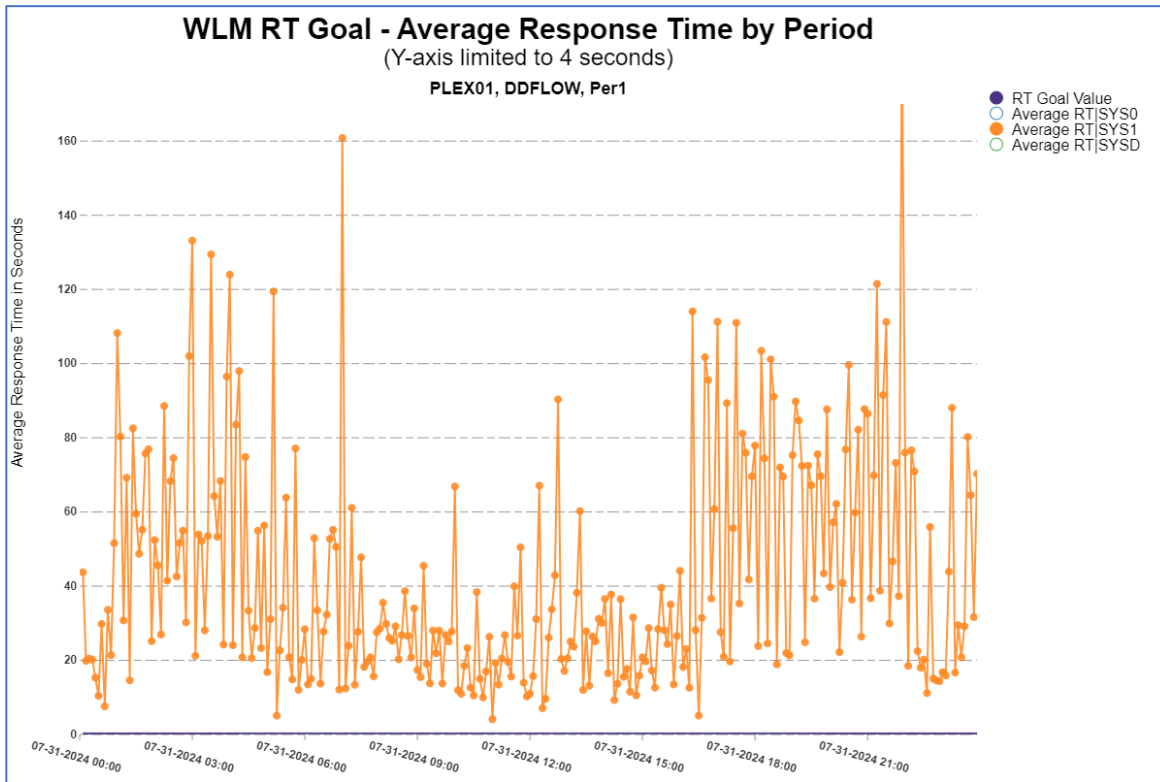


Example: DDFLOW Response Time



Average Response Time of Enclaves
(Used for goal management)

Average Response Time of Bundled Transactions
(Not used for goal management)



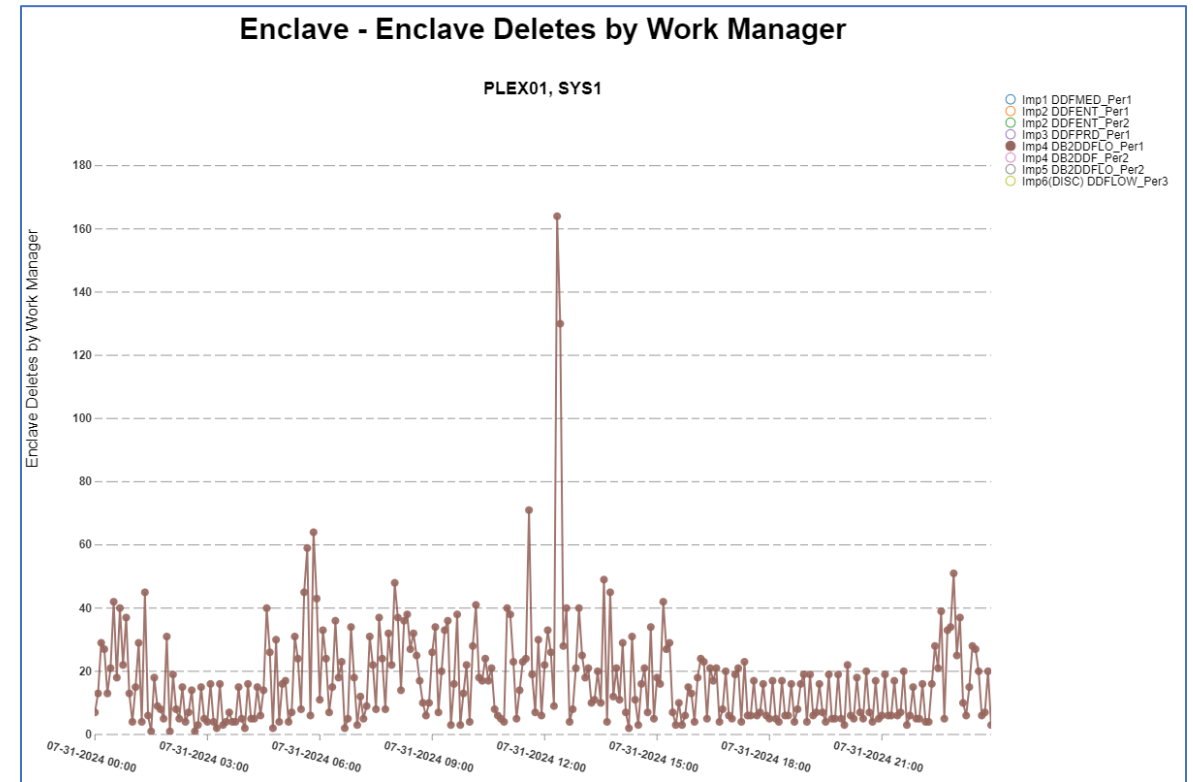
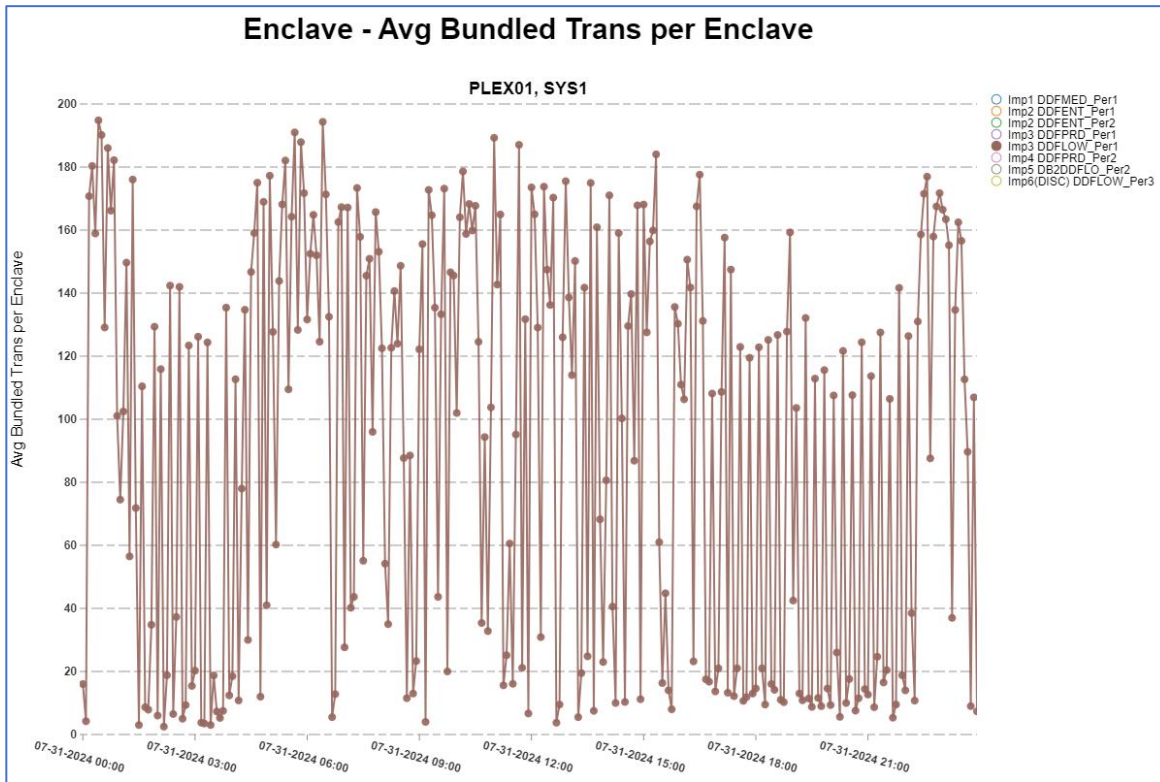
Example: DDFLOW Enclaves for Bundled Transactions



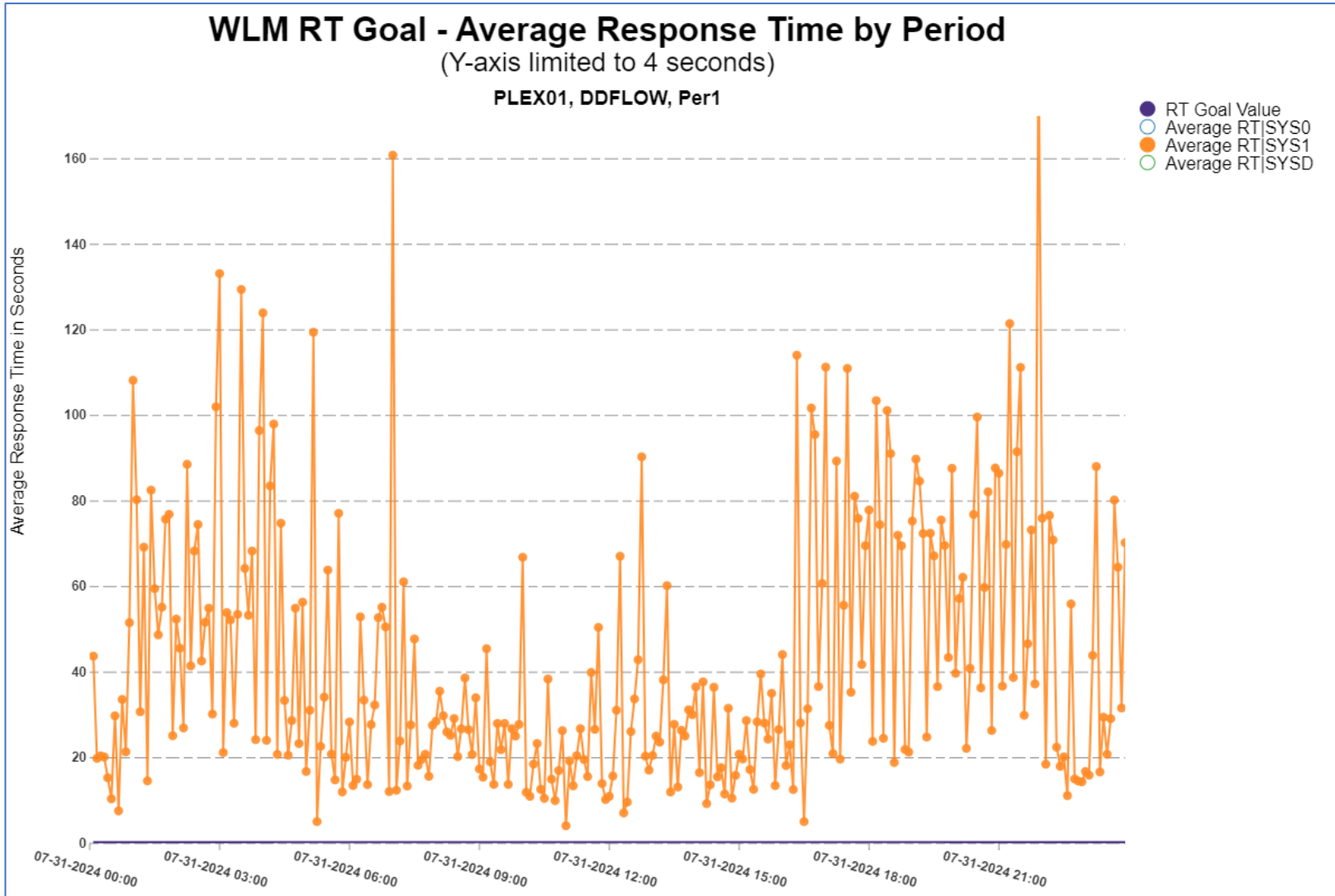
Average Bundled Trans per Enclave
(Note the 200 limit)

Deleted Enclaves for Bundled Trans

(An ended enclave is one WLM ended transaction even though up to 200 DBAT transactions ran in each enclave)

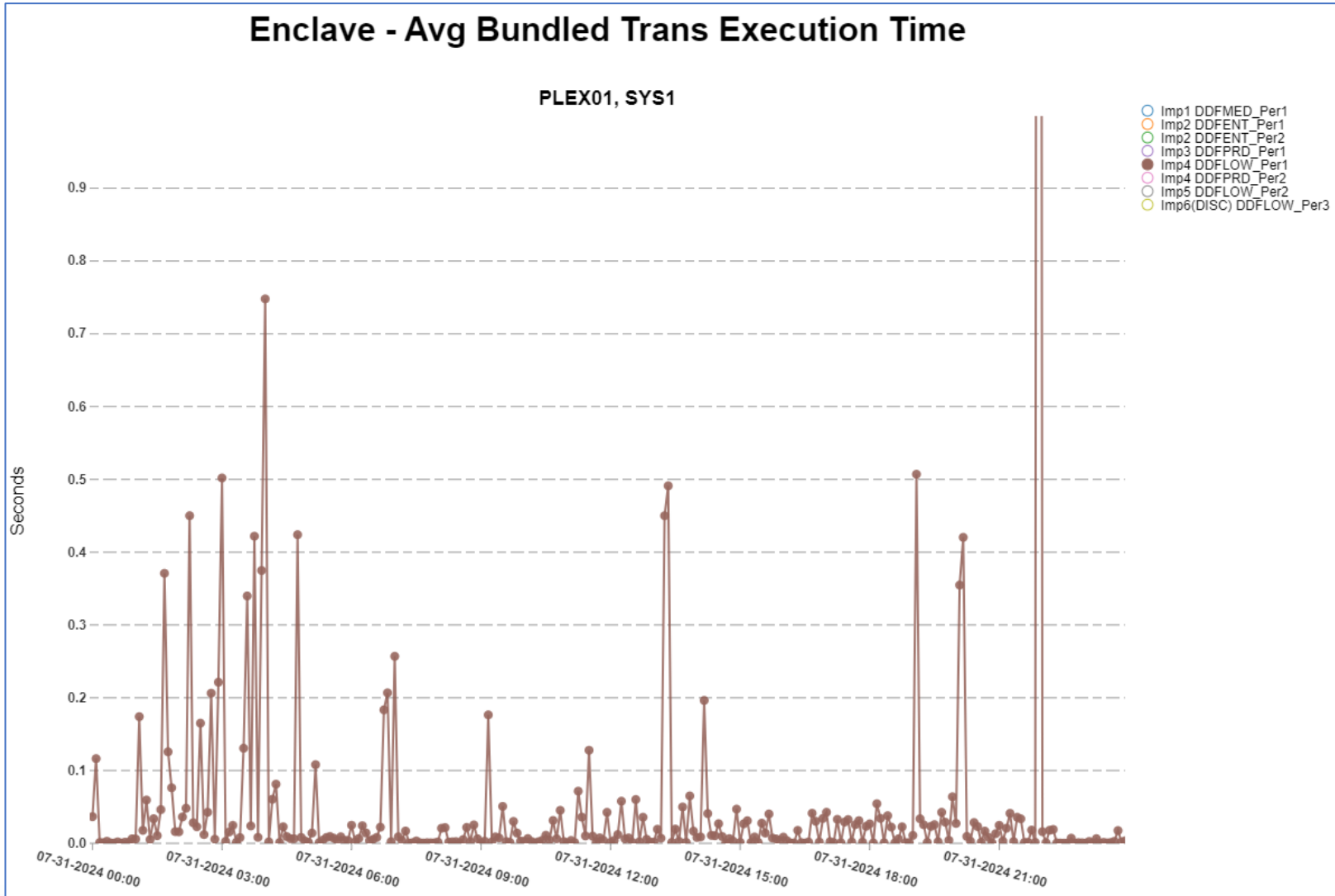


Example: DDFLOW Average Response Time



- DDFLOW has a response time goal of 90% complete in 0.3 seconds
- This chart shows the average response time achieved by the DDFLOW transactions.
- This is the average response time of the enclaves

Example: DDFLOW Bundled Transaction Response Times



- DDFLOW is running a combination of both enclave bundled transactions and non-bundled transactions
- This chart shows the average response time of the bundled transactions
- This response time is not known to WLM, so is not used in goal management

Possible Approaches to DDF Service Classes



- Understand your applications usage of DDF and HiPerf DBATS
 - Are enclaves being used to run bundled transactions?
 - Use the new measurements listed. If non-zero than you are using bundled transactions
- If running HiPerf DBATS, update your WLM setup for DDF
 - Best solution:
 - Work with the DB2 folks to segregate HiPerf DBATs away from other DDF transactions into separate service classes. DBAT transactions should be assigned a velocity goal
 - Mediocre solution:
 - In mixed environments, covert all your DDF service class period goals to velocity
 - Worst solution
 - Do nothing and just use your existing response time goals for the mixed DDF.
 - Adjust goals accordingly
 - Peter's solution (if running regular and HiPerf DBATs in same SC)
 - Convert the service class running both regular DDF trans and HiPerf DBATs transactions into a multiple period service class with both periods at the same importance
 - Period 1: Importance x with short duration with a response time goal
 - Period 2: Importance x, with velocity goal