Key Reports to Evaluate Coupling Facility CPU Utilization

Peter Enrico

Email: Peter.Enrico@EPStrategies.com

Enterprise Performance Strategies, Inc.
3457-53rd Avenue North, #145
Bradenton, FL 34210
http://www.epstrategies.com
http://www.pivotor.com

Voice: 813-435-2297
Mobile: 941-685-6789

z/OS Performance
Education, Software, and
Managed Service Providers

Creators of Pivotor®

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

Key Reports to Evaluate Coupling Facility CPU Utilization

This webinar will walk through and explain several reports that will be useful when evaluating the CPU capacity and usage of coupling facility processors.

Dynamic dispatch, thin interrupts, and the differences of physical CF processor utilization compared to virtual processor utilization will be discussed by Peter Enrico.

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

z/OS Performance workshops available

During these workshops you will be analyzing your own data!

- Essential z/OS Performance Tuning
  - March 20-24, 2023

- Parallel Sysplex and z/OS Performance Tuning
  - May 2-3, 2023

- WLM Performance and Re-evaluating Goals
  - October 2-6, 2023

- Also… please make sure you are signed up for our free monthly z/OS educational webinars! (email contact@epstrategies.com)

Like what you see?

- Free z/OS Performance Educational webinars!
  - The titles for our Fall 2022-2023 webinars are as follows:
    - ✓ *Key Reports to Evaluate z16 Processor Caches*
    - ✓ *Understanding System Recovery Boost's Impact on Performance and Performance Reporting*
    - ✓ *WLM Management of DDF Work: What can you do and what has changed?*
    - ✓ *Intensity! Understanding the Concepts and Usage of Intensity Measurements*
    - ✓ *High, Medium, Low: Understanding how HiperDispatch influences performance in z/OS*
    - ✓ *How and why Pivotor is different than other performance management reporters*
    - ✓ *Putting a lid on XCF*
    - ✓ *Key Reports to Evaluate Usage of Parallel Access Volumes*
    - ✓ *Key Reports to Evaluate Coupling Facility CPU Utilization*
    - *Understanding how memory management has evolved in z/OS*

  - Let me know if you want to be on our mailing list for these webinars

- 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
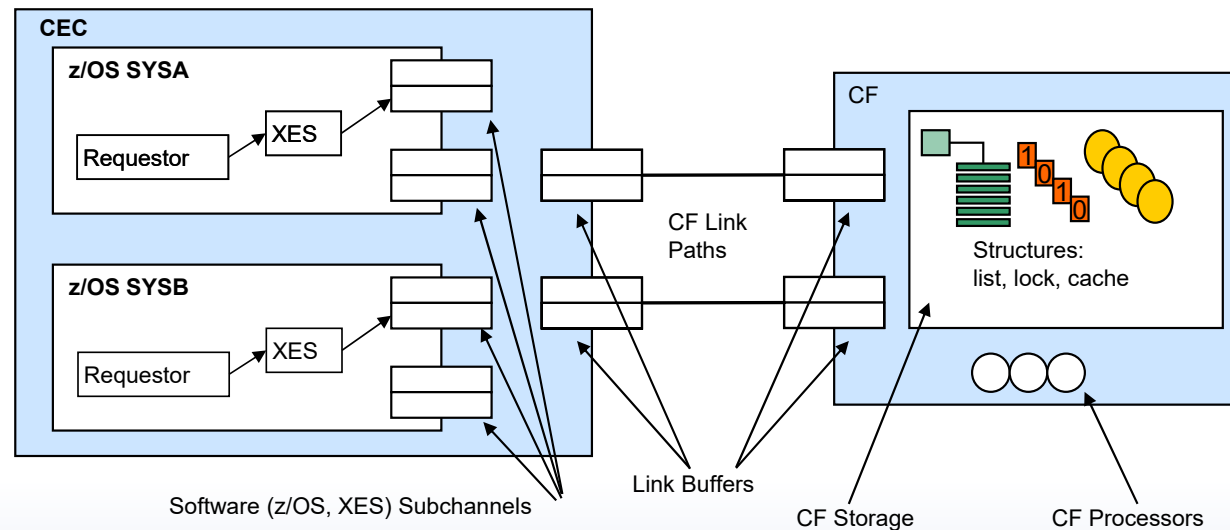
Performance view of CF Requests

☐ **z/OS Processing**

- S/W product making CF request
- XES receives the request and sets up request
- Sub-channel requested
- Path / Link requested Request a path
- Data transfer over link
- On return, S/W processing to handle CF request

☐ **Coupling Facility Processing**

- CF waits for request (various options)
- CF processor to process the request
- List, Lock, Cache structure processing
- Storage for structures
- Duplexing considerations

CEC

z/OS SYSA

Requestor → XES

z/OS SYSB

Requestor → XES

CF Link Paths

CF

Structures: list, lock, cache

Software (z/OS, XES) Subchannels

Link Buffers

CF Storage

CF Processors

Coupling Facility Tuning Questions

- Evaluation of Processor
  - How many CPUs are configured to the coupling facility?

  - Are the coupling facility processors dedicated or shared

  - If coupling facility processors are shared is thin interrupt support turned on?

  - What is the coupling facility processor utilization?

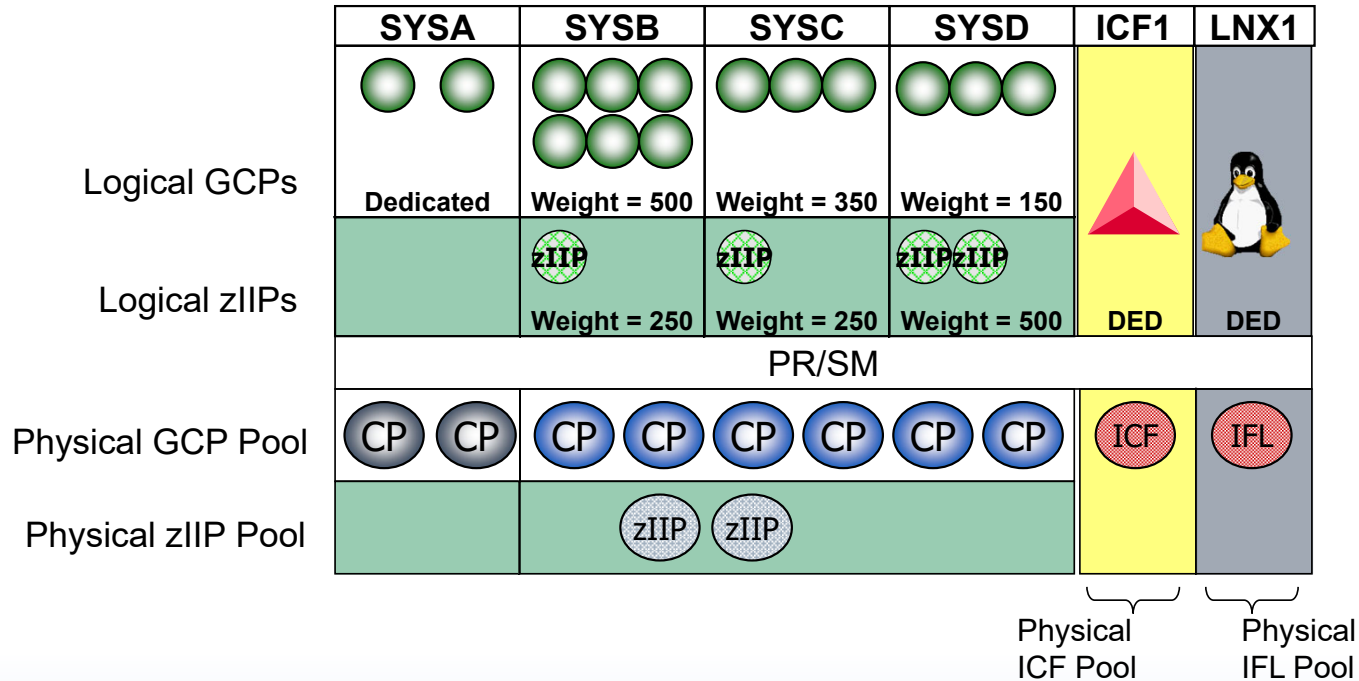 - What is the break down coupling facility utilization by structure?

Request →

Response ←
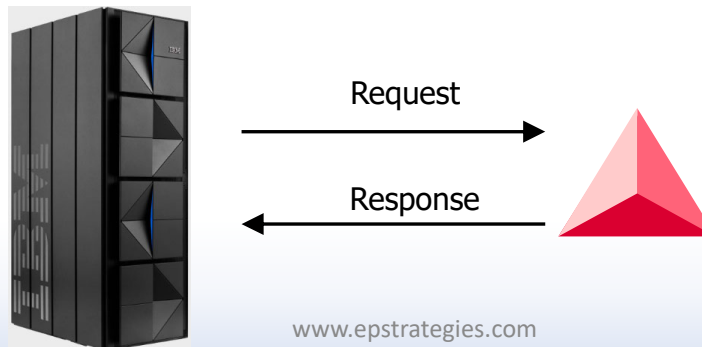
Coupling Facility Processors

- Evaluating CF processor utilization requires an understanding of the CF config

- As a reminder, coupling facilities run

  - In LPARs with ICF processors
    - Shared processors
    - Dedicated processors

  - In LPARs with CP processors
    - Shared processors
    - Dedicated processors (rare to non-existent)

- Coupling facilities can
  - Run on CECs completed separated from the exploiting z/OS images
  - Run on CECs where one of more z/OS exploiting image is also running

zArchitecture Machines Have Pools of Processors

- Each pool of process can be logically configured to a partition with its own weight

|	SYSA	SYSB	SYSC	SYSD	ICF1	LNX1
Logical GCPs	Dedicated	Weight = 500	Weight = 350	Weight = 150		
Logical zIIPs		Weight = 250	Weight = 250	Weight = 500	DED	DED
			PR/SM			
Physical GCP Pool	CP CP	CP CP	CP CP	CP CP	ICF	IFL
Physical zIIP Pool		zIIP zIIP				

Physical ICF Pool   Physical IFL Pool

Parallel Sysplex with External CFs

- "External" or "Standalone" Coupling Facilities dedicated to running CF LPARs

- This configuration was very common in the 1990s when CECs were more capacity-constrained and (slightly) less reliable

- No single point of failure from a processing perspective
  - Planned maintenance can be done non-disruptively as well

- Expense of external CFs typically limits their use to larger environments
  - i.e. likely larger than shown here

- More than 2 CFs can be used in a single Sysplex, but that's rare

CECC — CF 1
CECA — z/OS, z/OS
CECB — z/OS, z/OS
CECD — CF 2
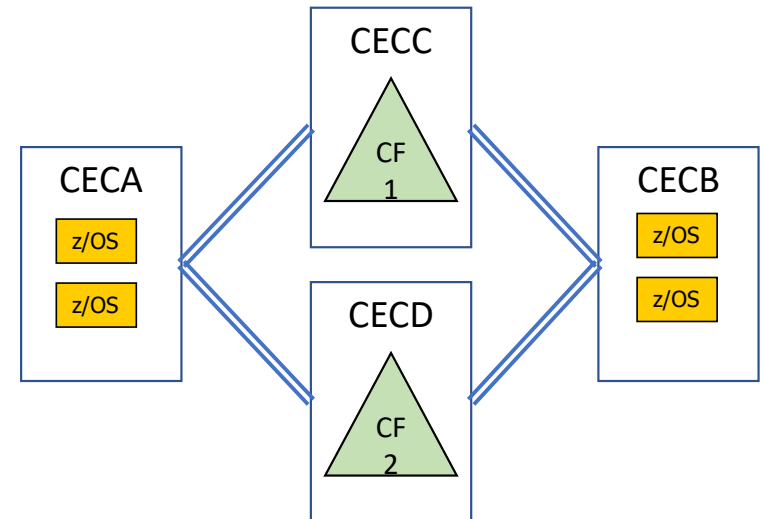
Parallel Sysplex with 1 External CF

- Single external CF + Single internal CF LPAR
  - CF1 normally used, CF2 as backup

- No single point of failure from a processing perspective
  - Planned maintenance can be done non-disruptively as well

- Saves a bit of money compared to having 2 external CFs

- Connections from z/OS to CF on same CEC are internal links
  - Memory to memory transfers facilitated by microcode (no physical connection)
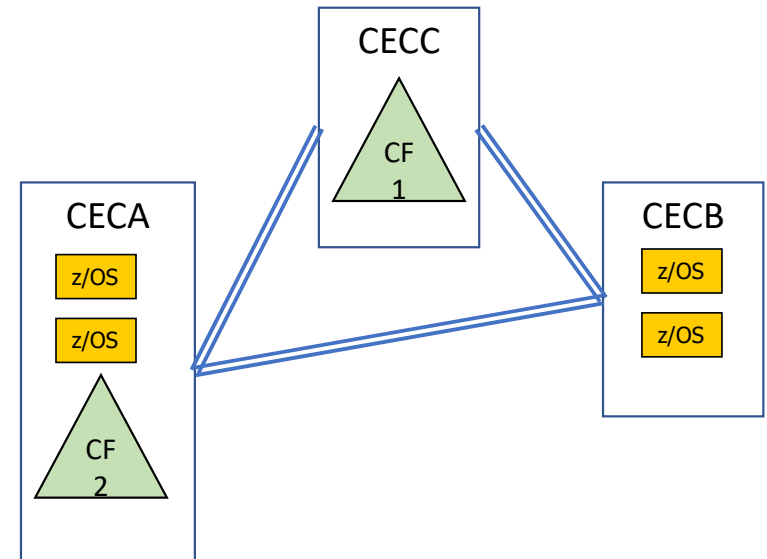
Parallel Sysplex with Internal CFs

- Dual Internal CFs: one on each CEC

- Can have a single point of failure, e.g. a single CEC failure could impact the Sysplex
  - Dual failure of both the CF and the z/OS LPARs that would be needed to rebuild those CF structures
  - CF Structure duplexing is used to address this concern
  - Planned maintenance can be done non-disruptively

- Least expensive way to get to Parallel Sysplex High Availability without a single point of failure
  - Structure Duplexing does add overhead though

- Most common configuration in mid-size environments

CECA
- z/OS
- z/OS
- CF 2
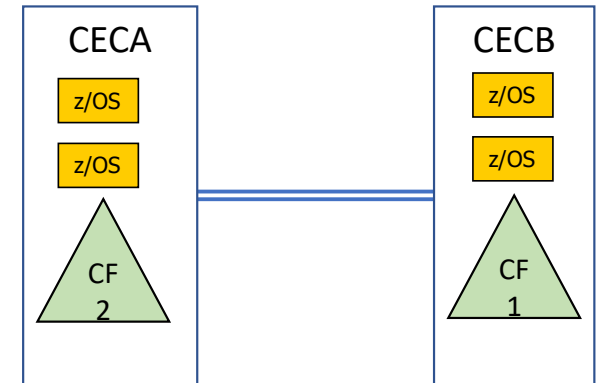
CECB
- z/OS
- z/OS
- CF 1

Parallel Sysplex In A Box

- All LPARs (z/OS and CF) in a single CEC

- The CEC becomes a single point of failure
  - ◦ CF Structure duplexing would only help a CFCC code failure situation, not a CEC-wide failure
  - ◦ Planned maintenance can be done non-disruptively if there are two CF LPARs (highly recommended)

- Least expensive way to get to most of Parallel Sysplex availability benefits for planned maintenance
  - ◦ Although the CEC is a single point of failure, smaller sites that are looking to implement Parallel Sysplex for planned maintenance availability sometimes start here

CECA

z/OS

z/OS

z/OS

z/OS

CF 1

CF 2
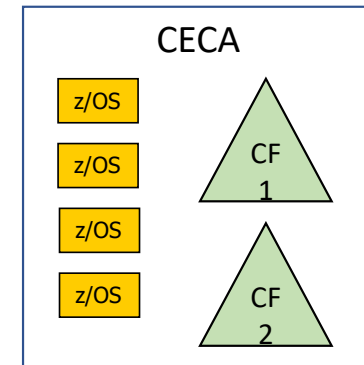
CF with shared CP Engines

- Some installations do use CP engines for coupling links
  - MSUs consumed count towards pricing
  - Variety of reasons why customers do this. One example:

Dedicated Production CF

CF Failure   *Kaboom!*

PR/SM	PR/SM
LPAR z/OS	LPAR z/OS	LPAR z/OS	LPAR z/OS

PR/SM	PR/SM
LPAR z/OS	LPAR z/OS	LPA R	LPA R
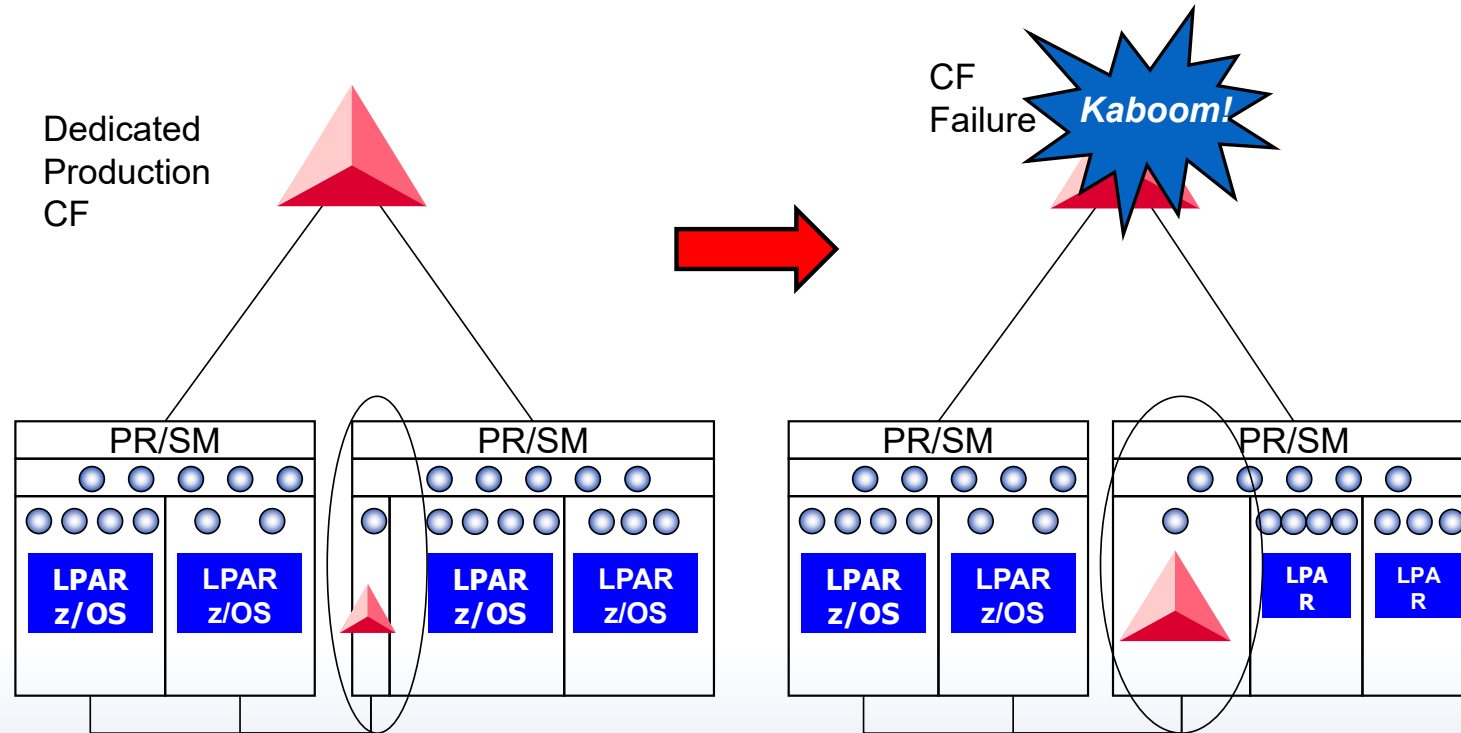
Coupling Facility Processor Measurements

- Used to ensure enough CF processor capacity
  - Major cause of response time objectives being missed
  - Monitor 'Processor Summary' section of 'CF Usage Summary' report

- Information About CF Hardware
  - Model Number, Version, CF Level

- Load on the Coupling Facility processor resource
  - Average CF Utilization (% Busy)
  - Logical Processors: Defined and Effective

```
-----------------------------------------------------------------------------------------------------

   COUPLING FACILITY        3906        MODEL M01       CFLEVEL  23          DYNDISP OFF

   AVERAGE CF UTILIZATION  (% BUSY)        6.2      LOGICAL PROCESSORS:   DEFINED   1    EFFECTIVE  1.000
                                                                         SHARED    0    AVG WEIGHT   0.0
```

Coupling Facility Processor Measurements

- The SMF data does show how much CF CPU is consumed for each structure

- The utilization presented is the percent of busy that each structure used
  - Thus, it is a percentage of a percentage
  - These values have super limited usage

CF processing for this structure is not 56.7% of the CF engine. Instead, it is 56.7% of the time the processor has work to do.

```
--------------------------------------------------------------------------------------------------------
    GENERAL STRUCTURE SUMMARY
    ------------------------------------------------------------------------------------------------
                                            % OF            % OF   % OF   AVG    LST/DIR  DATA     LOCK     DIR REC/
                STRUCTURE               ALLOC CF          #  ALL   CF     REQ/   ENTRIES  ELEMENTS ENTRIES  DIR REC
    TYPE        NAME       STATUS CHG ENC SIZE STOR      REQ  REQ   UTIL   SEC    TOT/CUR  TOT/CUR  TOT/CUR  XI'S

    LIST        DFHCFLS_DAGCPTG  ACTIVE     NO  7M  0.1   7126  9.6   4.4   7.92   1754     1500     N/A      N/A
                                                                                  37       3        N/A      N/A
                IXCSTR1    ACTIVE     NO  13M  0.1  52056 70.3  56.7  57.84  987      950      N/A      N/A
                                                                                  1        18       N/A      N/A
                IXCSTR4    ACTIVE     NO  51M  0.4  11262 15.2  29.0  12.51  10K      10K      N/A      N/A
                                                                                  1        16       N/A      N/A

    LIST        ISTGENERIC14 ACTIVE   NO  12M  0.1   1221  1.6   0.8   1.36   29K      563      4        N/A
                                                                                  3        0        0        N/A
                SYSARC_DC000_RCL ACTIVE NO  7M  0.1   1859  2.5   1.1   2.07   3849     3808     16       N/A
                                                                                  4        0        0        N/A
```

# Coupling Facility Processor Utilization

- Coupling facility processor utilization is reported in two vantage points:

- Viewpoint 1:
  - Examining coupling facility processor utilization of the physical processor
  - Meaning, how busy in the CF image keeping the processor?

- Viewpoint 2:
  - Examining coupling facility processor utilization from a coupling facility point-of-view
  - Meaning, when dispatched, how busy is the CF doing work?

- Key lesson is both of these utilizations need to be examined and understood

(PR/SM) Dispatching

- Dispatch Time
  - Time logical processor is associated with a physical processor

Dispatch Interval

Logical Processor
associated with
a physical processor

Logical Processor
disassociated with
a physical processor
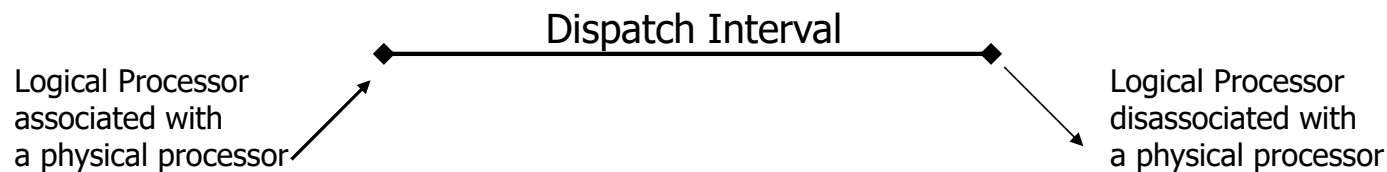
- Wait Completion
  - When z/OS voluntarily gives up the CPU, it is going into a 'wait' state

XXXXXXXXXX
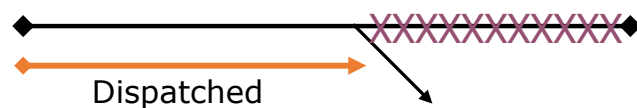
Dispatched

Voluntary Wait and Wait Completion = NO
- z/OS voluntarily gives up the processor
- MVS time equals dispatch time

Wait Wait

Dispatched

Voluntary Wait and Wait Completion = YES
- z/OS voluntarily gives up the processor
- PR/SM keeps logical processor associated with physical

CF LPAR Active Polling with Shared Engines

- DYNDISP=OFF
  - When dispatched, the CF utilizes a polling algorithm to constantly look for work
  - CF LPAR appears to have 100% demand for the processor(s)
  - PR/SM controls CF LPAR time slicing
  - CF always runs until the end of the current PR/SM dispatch interval (time slice)
  - Processor might get taken awhile while there is work out standing
  - CF Response times can be erratic

This CF Time Slice	Other CF Time Slice	This CF Time Slice

| Dispatchable Work | Poll | Dispat ... | | Dispatchable Work | Poll |

Processor Utilization PR/SM Point-of-View

CF CPU - CF Processor (Busy and Wait) Utilization
As a Percentage of Physical Processor
SYSPLX1, CFP01
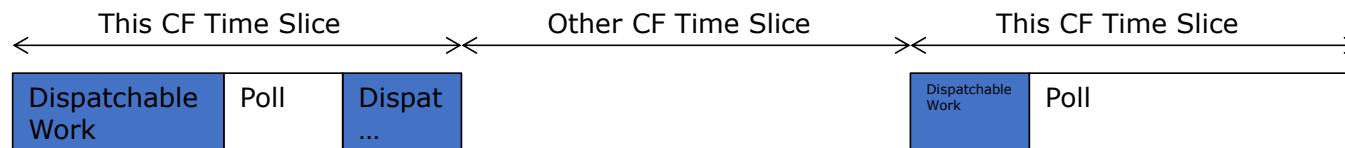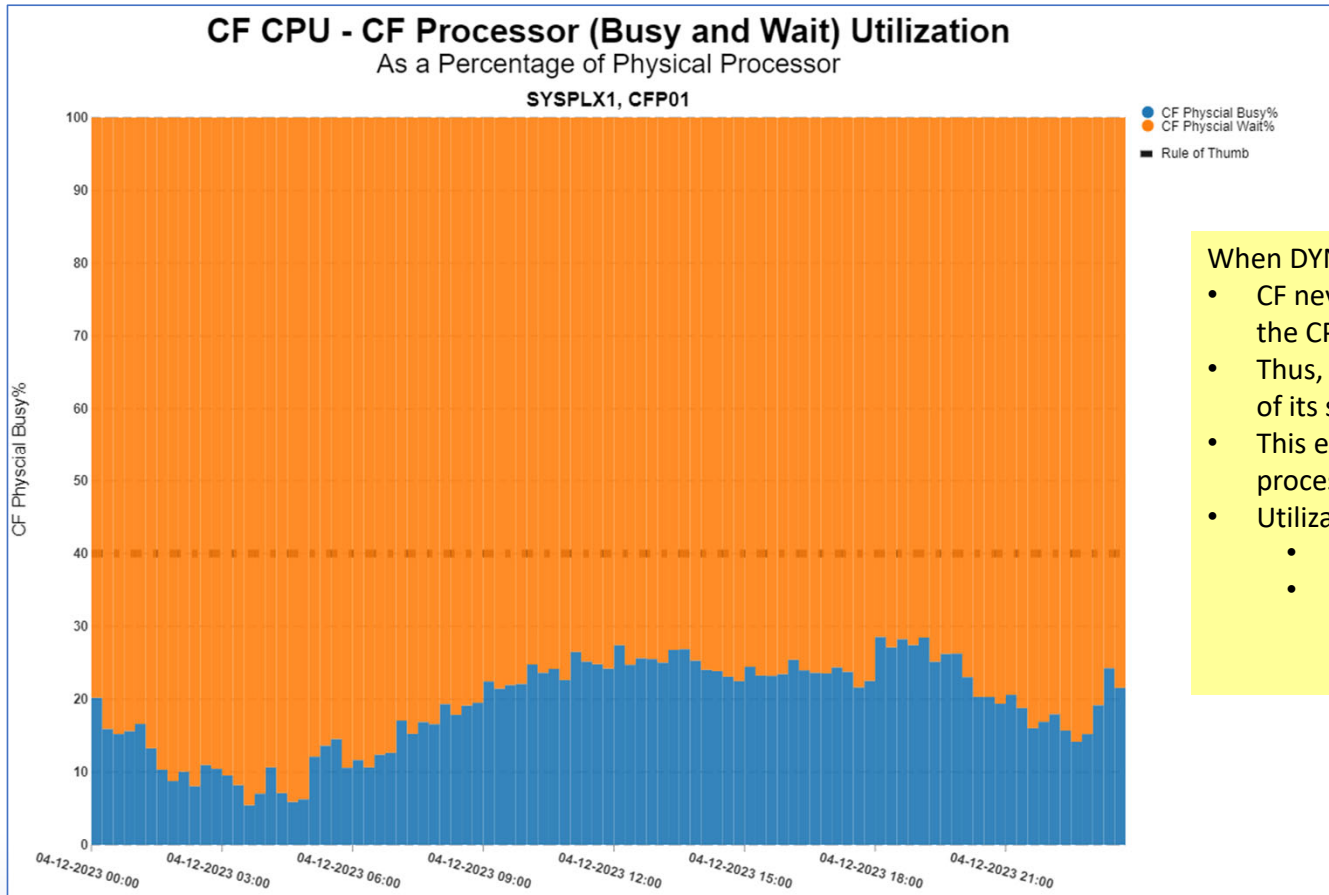
When DYNDISP=OFF
- CF never voluntarily gives up the CPU
- Thus, it is going to want 100% of its share
- This example shows the processor is 100% busy
- Utilization divided up
  - Percent doing work
  - Percent polling looking for work

Processor Utilization PR/SM Point-of-View

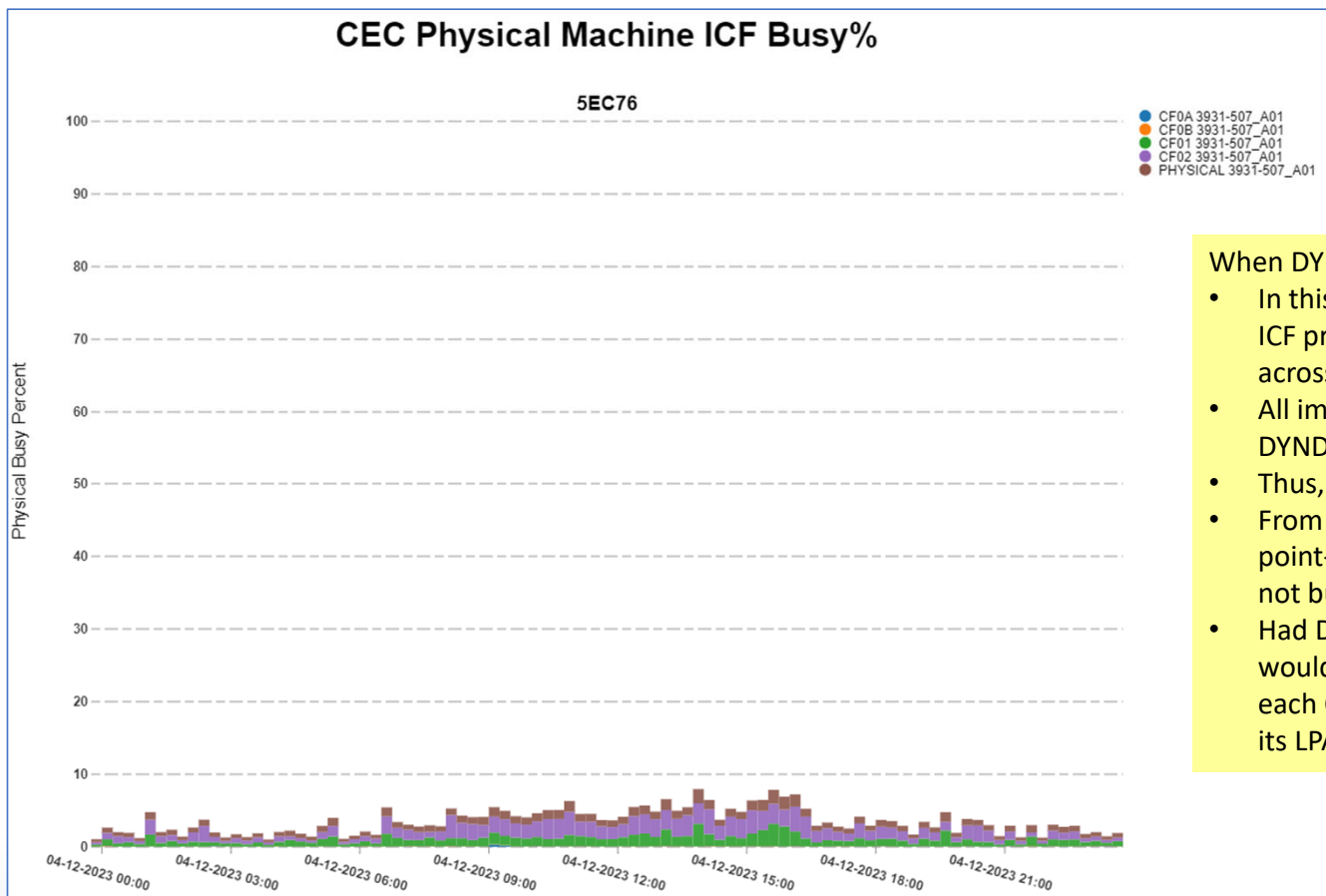**CF CPU - CF Processor Utilization of Dispatch**
(as reported by RMF)

SYSPLX1
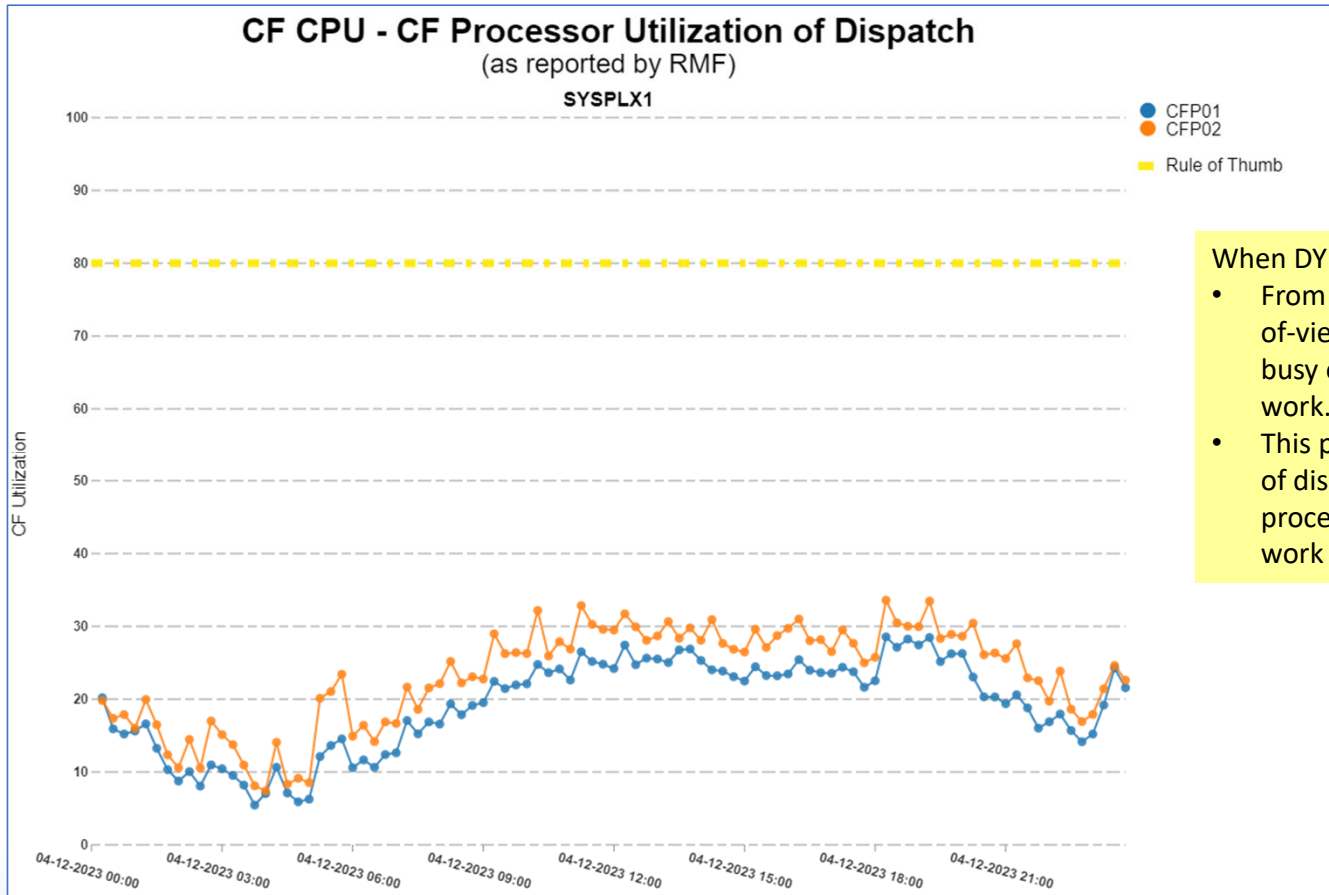
When DYNDISP=OFF
- From a coupling facility point-of-view, the CF is considered busy only when it is doing work.
- This percentage is the percent of dispatch when the processor was busy doing work

CEC Physical Machine ICF Busy%

D54E6

Legend:
- CFP01 8561-602_T01
- CFP02 8561-602_T01
- CFA01 8561-602_T01
- CFA02 8561-602_T01
- CFT01 8561-602_T01
- CFT02 8561-602_T01
- PHYSICAL 8561-602_T01

Y-axis: Physical Busy Percent (0 to 100)

X-axis: 04-11-2023 23:45, 04-12-2023 00:00, 04-12-2023 00:15, 04-12-2023 00:30, 04-12-2023 00:45, 04-12-2023 01:00, 04-12-2023 01:15, 04-12-2023 01:30, 04-12-2023 01:45
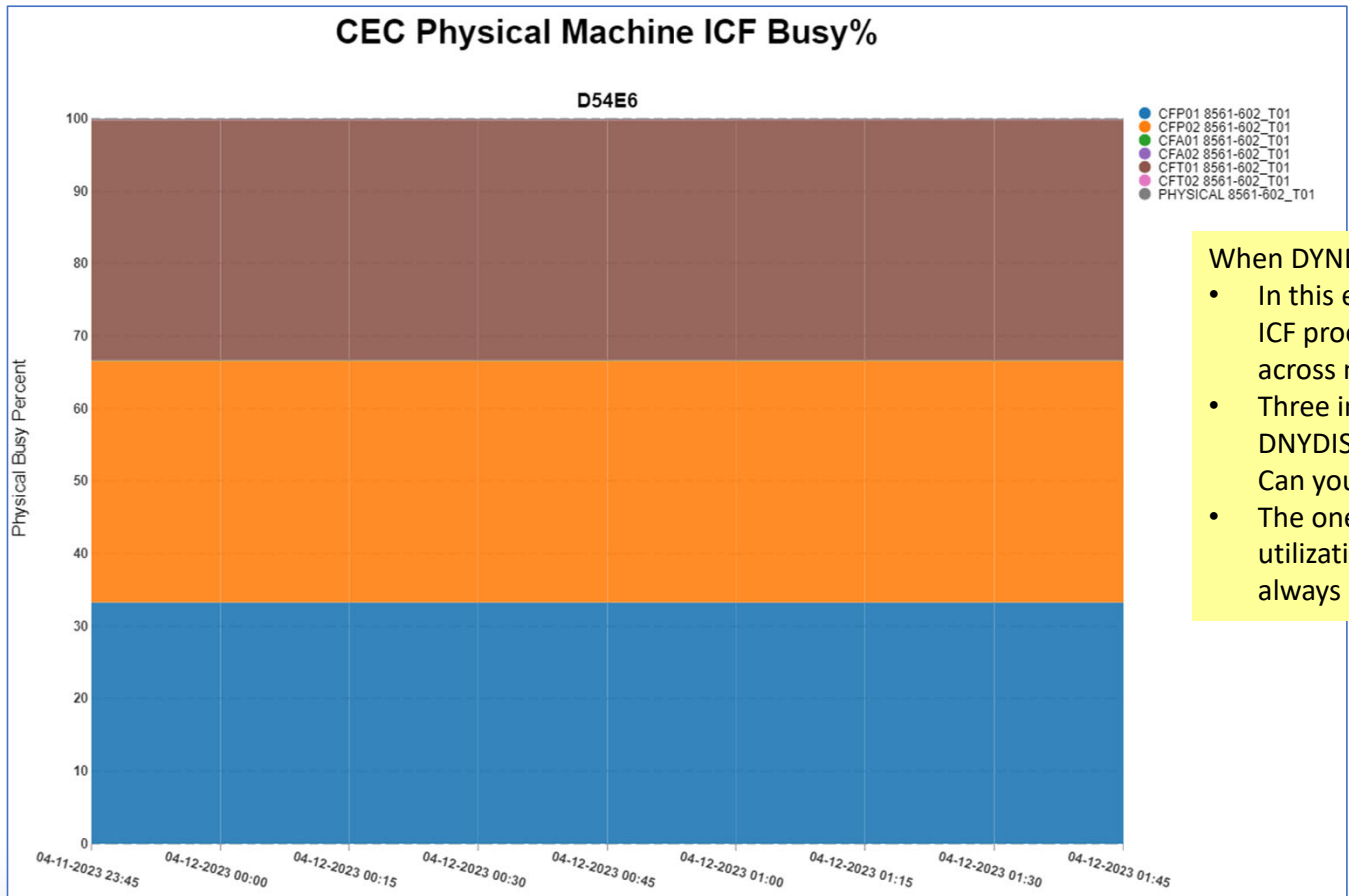
When DYNDISP=OFF
- In this example, there are 2 ICF processors being shared across multiple CF images
- Three images have DNYDISP=OFF
  Can you guess which 3?
- The ones with higher physical utilization because they are always polling

CF LPAR Dynamic Dispatching

- DYNDISP=ON
  - PR/SM and the CF implement a dynamic sleep pattern algorithm to allow processor to be un-dispatched.
  - Polling still happens, but is somewhat limited and duration of polling is dynamic
  - CF can have processor un-dispatched prior to end of current time slice
  - When sleep time is expired, CF re-dispatched
  - Less likely to lose the processor while work is outstanding
  - Results in better responsiveness for shared CF engines
    - But dedicated engines still better

This CF Dispatch Time Slice	Sleep slice	Dispatch Time Slice	Sleep slice

| Dispatchable Work | Poll | Dispatchable Work | Poll | | Wake Wait | Dispatchable Work | Poll |
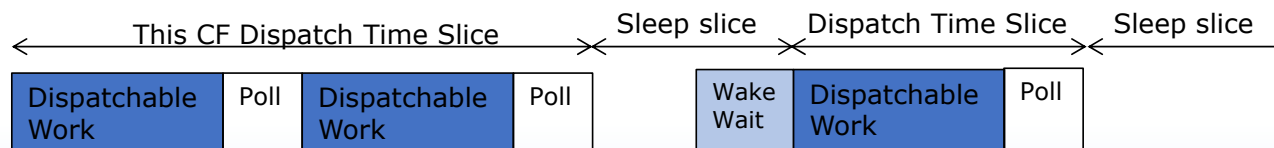
Thin Interrupts

- DYNDISP=THIN
  - Available on CFCC 19+ (z12 GA2)
  - Like DYNDISP=ON but not based on a dynamic sleep pattern algorithm
  - CF will give up processor almost immediately when there is no work to do
  - Gets interrupt to wake up when work arrives
  - Service time close to dedicated engines
    - Dedicated still better, but margin is much thinner
  - Best option when multiple CF LPARs share engines
    - i.e. Use DYNDISP=THIN, not DYNDISP=ON if sharing engines (z12 and later CFs)
    - Default on z15 and above

Processor Utilization PR/SM Point-of-View

CF CPU - CF Processor (Busy and Wait) Utilization
As a Percentage of Physical Processor
PLEX01, CF01
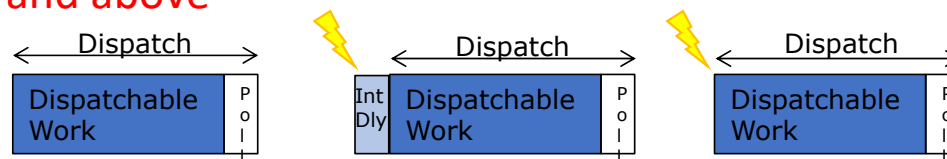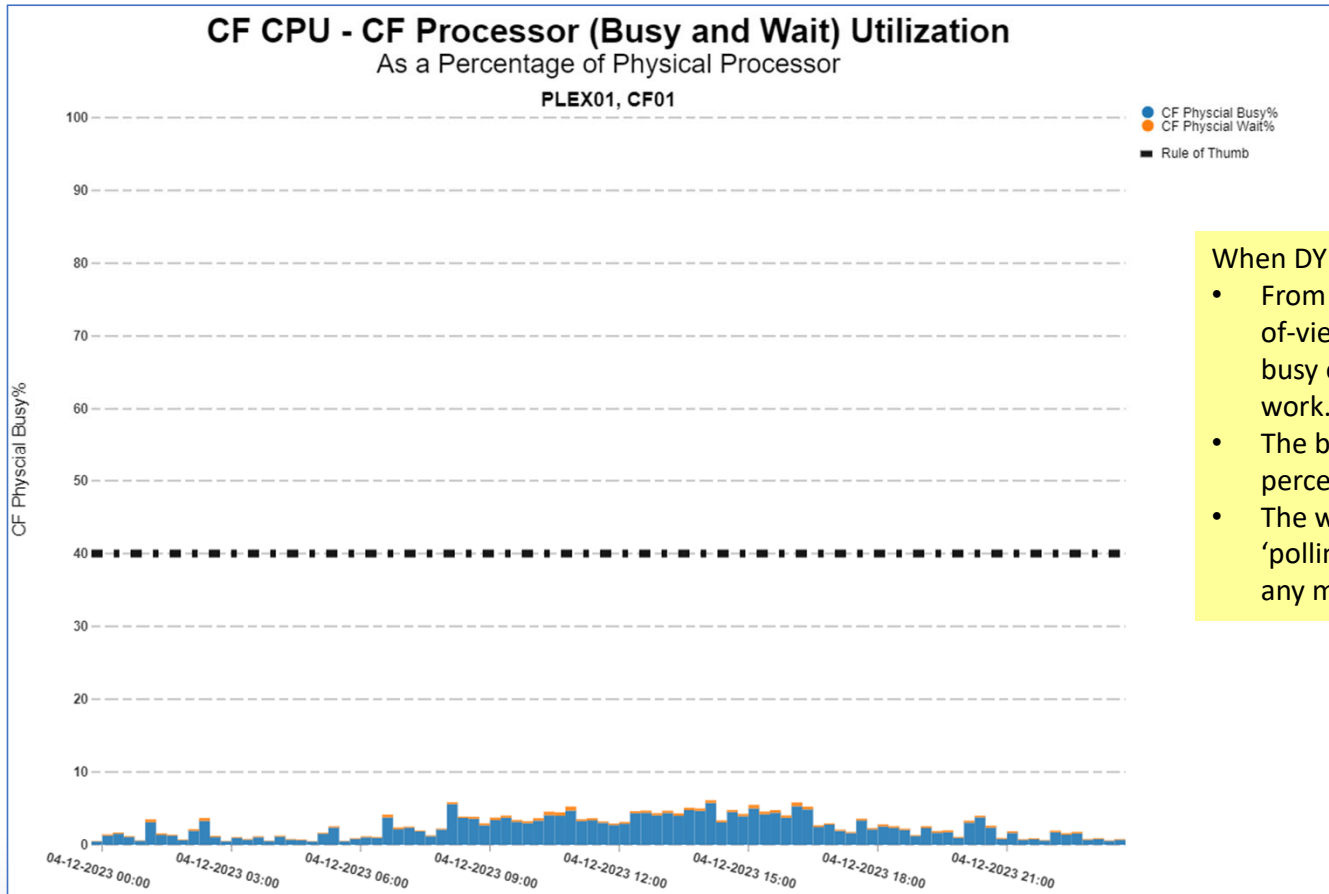
When DYNDISP=THIN
- From a coupling facility point-of-view, the CF is considered busy only when it is doing work.
- The busy percentage is the percentage doing CF work
- The wait percentage is the 'polling'. But not really polling any more

Processor Utilization PR/SM Point-of-View

CF CPU - CF Processor Utilization of Dispatch
(as reported by RMF)
PLEX01
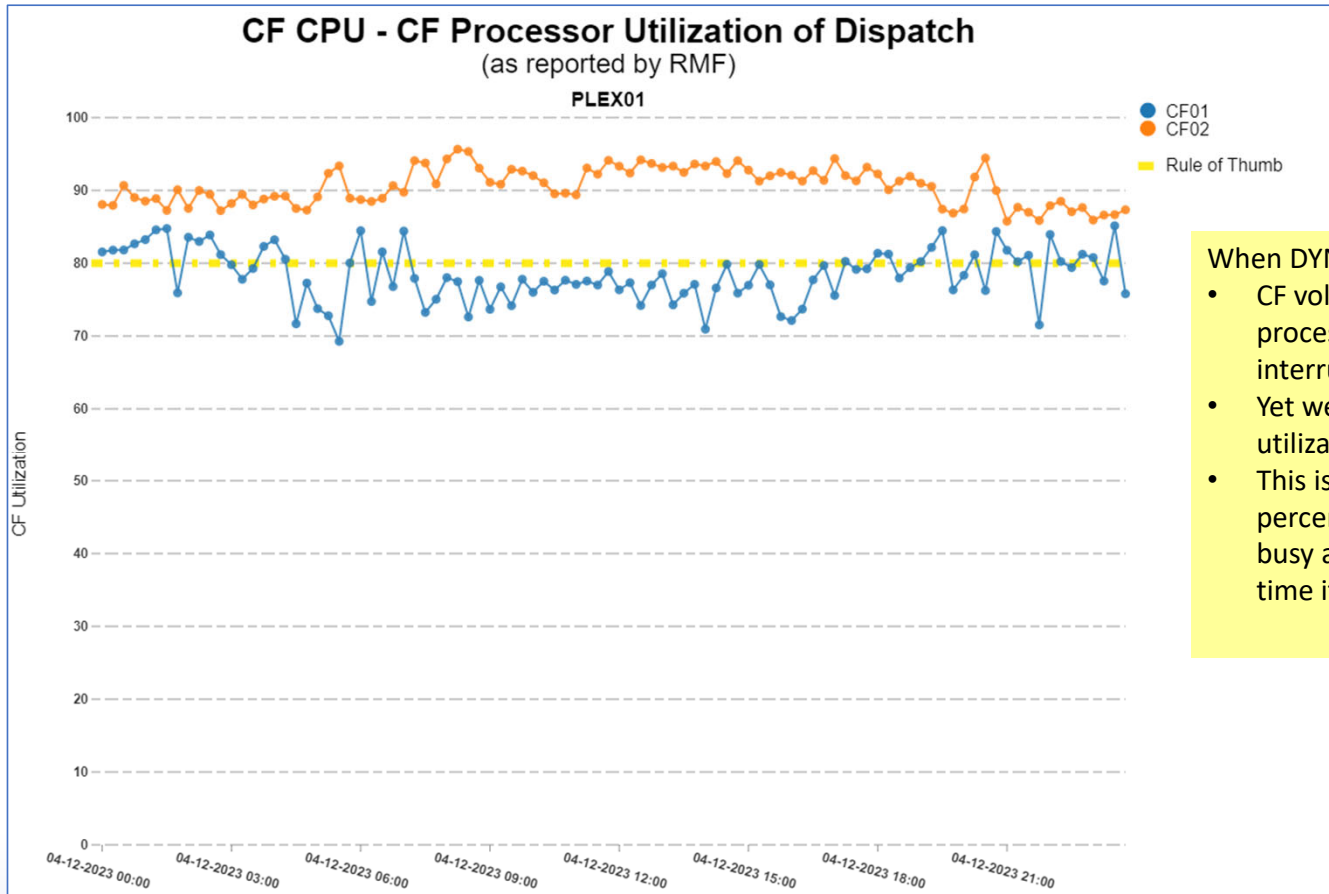
When DYNDISP=THIN
- CF voluntarily gives up the processor and becomes interrupt driven
- Yet we see the CF processor utilization is very high
- This is because this is the percentage the CF was kept busy as a percentage of the time it was dispatched.

CEC Physical Machine ICF Busy%

5EC76

Legend:
- CF0A 3931-507_A01
- CF0B 3931-507_A01
- CF01 3931-507_A01
- CF02 3931-507_A01
- PHYSICAL 3931-507_A01

Y-axis: Physical Busy Percent (0–100)

X-axis timestamps:
04-12-2023 00:00, 04-12-2023 03:00, 04-12-2023 06:00, 04-12-2023 09:00, 04-12-2023 12:00, 04-12-2023 15:00, 04-12-2023 18:00, 04-12-2023 21:00

When DYNDISP=THIN
- In this example, there are 2 ICF processors being shared across multiple CF images
- All images have DYNDISP=THIN
- Thus, no polling at all
- From a PR/SM and machine point-of-view, these ICFs are not busy
- Had DYNDISP=NO, the chart would show 100% busy, and each CF would be as busy as its LPAR weight would allow

Summary: Sharing CF Engines

- Generally, best to dedicate engines to production
- If sharing for engines for dev/test Sysplexes
  - Set weights appropriately
  - Use DYNDISP=THIN
  - If sharing an engine between two CECs for dev/test purposes and you don't need duplexing, consider skewing the weights and putting all activity from each sysplex into one CF:

This only works if you're not doing duplexing for these Sysplexes.

PLEXA's CF activity directed to CFA0, CFA1 used just for backup for planned maintenance

PLEXB's CF activity directed to CFB0, CFB0 used just for backup for planned maintenance

CEC0
- SYA0
- SYB0
- A / B

CEC1
- SYA1
- SYB1
- B / A

Analyzing CF Processor

- Check if CF processor utilization is too high?
    - IBM recommendation is CF should be no more than 40% to 50% busy
        - CPU and Model to understand capacity and expectations
        - CFLEVEL to understand functionality
        - %BUSY to understand load
        - DEFINED to understand what is configured
        - EFFECTIVE to understand what is being used

    - Remember, service times is a function of
        - CF utilization
        - Speed of sending processor verses CF

```
--------------------------------------------------------------------------------------------------------------
PROCESSOR SUMMARY
   --------------------------------------------------------------------------------------------------------------

      COUPLING FACILITY          8561          MODEL T01      CFLEVEL  24          DYNDISP THIN

      AVERAGE CF UTILIZATION  (% BUSY)          69.3      LOGICAL PROCESSORS:   DEFINED   2    EFFECTIVE  0.001
                                                                                SHARED    2    AVG WEIGHT 440.0
```

Analyzing CF on Same CEC as z/OS

- When z/OS and CF are on the same CEC verify if it is DED or shared
  - If shared, CF will try to use all the processor it can get (tight polling)
  - Contention from other partitions will limit what CF can get
  - If shared, make sure the CF is not capped, and give a sufficient weight
  - Otherwise, will elongate CF response times

```
                             P A R T I T I O N   D A T A   R E P O R T
                                                                                    PAGE    2
             z/OS V2R4                SYSTEM ID CPRO            DATE 10/15/2022         INTERVAL 15.00.035
                                      RPT VERSION V2R3 RMF      TIME 09.45.00          CYCLE 1.000 SECONDS
 -
  MVS PARTITION NAME                  ABCPRO        PHYS PROC NUM      9       GROUP NAME    GROUPC        INITIAL CAP   NO
  IMAGE CAPACITY                      189                     CP       3       LIMIT         189           LPAR HW CAP   NO
  NUMBER OF CONFIGURED PARTITIONS      17                     IFL      2       AVAILABLE      20           HW GROUP CAP  NO
  WAIT COMPLETION                     NO                      ICF      2                                   ABS MSU CAP   NO
  DISPATCH INTERVAL                   DYNAMIC                 IIP      2
 -
 ---------- PARTITION DATA ----------------- -- LOGICAL PARTITION PROCESSOR DATA --  -- AVERAGE PROCESSOR UTILIZATION PERCENTAGES --
 0                 ----MSU----   --CAPPING--- --PROCESSOR ----DISPATCH TIME DATA----  LOGICAL PROCESSORS   --- PHYSICAL PROCESSORS ---
   NAME      S  WGT   DEF    ACT  DEF    WLM%  NUM  TYPE  EFFECTIVE       TOTAL        EFFECTIVE    TOTAL   LPAR MGMT  EFFECTIVE  TOTAL
 -
  CFCPROD   A  DED                              1   ICF  00.14.59.999  00.15.00.004    100.0       100.0     0.00       50.00    50.00
  CFCRS     A  DED                              1   ICF  00.14.59.999  00.15.00.004    100.0       100.0     0.00       50.00    50.00
  *PHYSICAL*                                                           00.00.00.025                          0.00                 0.00
            ------                                       ------------  ------------               -----     -----  ----- -----
    TOTAL        0                                       00.29.59.998  00.30.00.034                          0.00       100.0  100.0
```

www.epstrategies.com

Analyzing CF Processor cont…

- If Coupling facility utilization > 50%
  - Monitor spikes in CF utilization
  - Can be done with online CF monitor reports or SMF analysis
  - Usually due to long running commands

  - If spike exists
    - Review Structure Summary
      - Use to narrow down to structure
      - May show for spikes in Sync rates (# REQ TOTAL/SEC)
    - Review Structure Activity Report
      - Use to narrow down to system sending requests
      - Look at Sync rates by system

  - Usually things like DB2 checkpoint processing, DB2 Delete_Name processing, and IMS Detaches

  - May also be due to a batch job
    - OK cause for a spike

Comments from Jamie...
and then
Q & A

Questions about content of webinar?

Of maybe general performance questions?