Parallel Sysplex Measurement and Tuning - Hints and Tips

Instructor: 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

Contact, Copyright, and Trademark Notices

Questions?
Send email to Peter at Peter.Enrico@EPStrategies.com, or visit our websites at http://www.epstrategies.com and http://www.pivotor.com.

Reports:
All report examples created by Pivotor® (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®, WebSphere Application Server®, and many others. Other trademarks and registered trademarks may exist in this presentation.
Abstract

Z/OS performance measurement and tuning can sometimes be a big and intimidating area to explore. In the Z/OS environment there is so many measurements available, and there are so many areas to be tuned. Where should one start to become quickly productive.

During this presentation, Peter Enrico will discuss a variety of Coupling Facility and Parallel Sysplex performance measurements and performance tuning recommendations. Attending this session is sure to get any Z/OS performance analyst productive and started on their journey to Z/OS Sysplex optimization.

Performance Workshops Available

During these workshops you will be analyzing your own data!

- **WLM Performance and Re-evaluating of Goals**
  - Instructor: Peter Enrico
  - Scheduled: June 24 - 28, 2013 Raleigh, North Carolina, USA
  - Scheduled: September 09 - 13, 2013 Munich, Germany
  - Scheduled: September 23 - 27, 2013 Indianapolis, Indiana

- **Parallel Sysplex and Z/OS Performance Tuning** (Web / Internet Based!)
  - Instructor: Peter Enrico
  - Scheduled: July 09 - 11, 2013 Web based
  - Scheduled: August 26 - 29, 2013 Web based

- **Essential Z/OS Performance Tuning**
  - Instructor: Peter Enrico and Tom Beretvas

- **Z/OS Capacity Planning and Performance Analysis**
  - Instructor: Ray Wicks
**CF Reports Processing/Discussion Offer !!!**

- **Special Reports Offer!**
  - See your Coupling Facility records in chart and table format
  - Please contact me, Peter Enrico for instructions for sending raw SMF data
    - Send an email to [peter.enrico@epstrategies.com](mailto:peter.enrico@epstrategies.com)
  - Deliverable: Dozens of coupling facility based reports (charts and tables)
    - Configuration / Setup questions
    - Link and general load performance
    - Host effect
    - Processor
    - Storage usage
    - List structure
    - Lock structure
    - Cache structure
    - Duplexing
    - Much more...
    - One-on-one phone call to explain your coupling facility measurements

---

**Performance Analyst View of CF Resource**

- **z/OS Processing**
  - S/W processing to make CF request
  - Request a sub-channel
  - Request a path
  - Data transfer over link
  - On return, S/W processing to handle CF request

- **Coupling Facility Processing**
  - Link time (i.e. time on path)
  - CF busy processing request

---

- Structures:
  - list
  - lock
  - cache

---

http://www.epstrategies.com

© Enterprise Performance Strategies, Inc.  
Some Key CF Measurements - 6
Examples of factors that influence performance

- Performance is heavily dependent on a number of variables:
  - Speed of requesting CPU
    - Faster processor will ‘wait faster’ for a response from a slower processor
  - Type of request – Synchronous versus Asynchronous
  - Busy conditions (Subchannel, path)
  - Time it takes to transmit data to the CF
    - CF link performance
    - Speed of data over link
    - Distances
  - Configuration and speed of CF processor
    - Shared or dedicated CF engines?
  - Structures size deficiencies that may result in additional processing
  - CF structure duplexing

Many Questions Need to be Asked of CF Measurements

- Configuration / Setup questions
- Link and general load performance questions
- Host effect questions
- Processor related questions
- Storage usage related questions
- List structure related questions
- Lock structure related questions
- Cache structure related questions
- Duplexing related questions
Coupling Facility Measurement Questions

Subjects of this Presentation

- What is the structure configuration?
- How many requests to the CF is each structure responsible for, and what sorts of response times are the requests incurring?
- What is the spin component of the host effect?
- Are there cross invalidates due to directory reclaims?
- What percentage of Lock requests are due to false contention?

Sources of CF Measurements

- CMF / RMF Monitor I Coupling Facility Activity report
  - Actually collected by the RMF Monitor III data gatherer
  - Based off of the SMF 74.4 records
  - One record per CF per system in the Sysplex

- CMF / RMF Post Processor Reports divided up into four main sections
  - Coupling Facility Usage Summary
    - Structure Summary
    - Storage Summary
    - Processor Summary
  - Coupling Facility Structure Activity
    - Request Rates
    - Service & Delay Times
    - Other structure activity such as Caching and Locking statistics
  - Coupling Facility Subchannel Activity
    - Request Rates
    - Service & Delay Times
  - Coupling Facility to Coupling Facility communication

Peter Enrico: www.epstrategies.com  © Enterprise Performance Strategies, Inc.
Some Key CF Measurements - 6

---

CF Sysplex Report

- Make sure data from all systems in Sysplex is included

- Coupling Facility measurements are best analyzed when data from all participating systems is analyzed together
  - One SMF record per CF per system in the Sysplex
  - All SMF records for each CF from every system need to be combined
  - If not, some of the measurements are misleading

- Also, make sure if you have multiple CFs that you examine the measurements from each one.

RMF CF Report

SYSA
SMF 74.4
Records

SYSB
SMF 74.4
Records

SYSC
SMF 74.4
Records

---

Some Key CF Measurements - 12

How many requests to the CF is each structure responsible for?
What sorts of response times are the requests incurring?

Synchronous Response Times

Asynchronous Response Times
CF Synchronous Request Processing

- Requesting processor spins waiting for CF request to complete
- Two types of sync requests
  - Those that must continuously run as synchronous
    - Lock requests - XES spins
  - Those that start out as sync
    - But converted to async if doing so helps performance
    - Sync cache/list requests - XES changes to async

Impacts response time of sender, and performance of sending system

CF Asynchronous Request Processing

- Requesting processor can do other processing while the CF request is queued
  - Immediate reply is not required
    - Async cache/list requests - XES queues
    - Serialized access is not required
  - Some async requests start out as sync requests and then converted to async
    - XES changes to async based on busy conditions, and response time conditions

Performance impact much less severe that sync requests
Best Sync and Async Service Times Expected

- Speed of sending processor and CF factor in greatly into service times

- Response time guidelines:
  - **Sync** requests should be in the range of 7 to 20 microseconds (or lower)
  - **Async** requests should be in the range of 80 to 300 microseconds (or lower)

- The ranges are service times take into account the transfer of data
  - Low end of ranges – when there is no transfer of data (such as a GRSLOCK request)
  - High end of ranges – when there is the largest transfer of data allowed
    - Sync request – 4K is largest transfer of data allowed
    - Async request – 64K is largest transfer of data allowed

CF Subchannel Activity Report

- Usage of the Report
  - For measurement and monitoring of subchannel and path utilizations
  - Data reported for all paths for each connected system
  - Insight into synchronous and asynchronous request processing

- Report divided up into the following sections on system basis
  - Subchannel and path configuration and usage
  - Request load and response time information for all paths from each system
  - Delay information

- Measuring and monitoring coupling facility links
  - It is important to ensure that there are enough coupling facility links
    - Helps to decrease service times & queuing times
    - Every CF link is associated with two CF subchannels
    - Data sent to the CF are first loaded into these subchannels

- Possible performance conditions
  - Not enough subchannels
  - Subchannel busy
  - Path busy
Quick Check of Response Times:
Use CF Subchannel Activity Report

```
<table>
<thead>
<tr>
<th>COUPLING FACILITY NAME = P1CF001</th>
</tr>
</thead>
</table>

**SUBCHANNEL ACTIVITY**

<table>
<thead>
<tr>
<th>SYSTEM</th>
<th>TOTAL</th>
<th>-- CF LINKS --</th>
<th>PTH</th>
<th># SERVICE TIME(MIC)</th>
<th>#% OF</th>
<th>DELAYED</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAME</td>
<td>AVG/SEC</td>
<td>TYPE</td>
<td>GEN</td>
<td>USE</td>
<td>BUSY</td>
<td>REQ</td>
</tr>
<tr>
<td>SYS0</td>
<td>49994</td>
<td>CFP</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>31901</td>
</tr>
<tr>
<td></td>
<td>55.5</td>
<td>SUBCH</td>
<td>14</td>
<td>14</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYS1</td>
<td>927793</td>
<td>CFP</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>955481</td>
</tr>
<tr>
<td></td>
<td>1030.9</td>
<td>SUBCH</td>
<td>14</td>
<td>14</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYS2</td>
<td>316464</td>
<td>CFP</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>32352</td>
</tr>
<tr>
<td></td>
<td>351.6</td>
<td>SUBCH</td>
<td>14</td>
<td>14</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYS3</td>
<td>326032</td>
<td>CFP</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>49018</td>
</tr>
<tr>
<td></td>
<td>362.3</td>
<td>SUBCH</td>
<td>14</td>
<td>14</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYS4</td>
<td>907933</td>
<td>CFP</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>391643</td>
</tr>
<tr>
<td></td>
<td>1008.9</td>
<td>SUBCH</td>
<td>14</td>
<td>14</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Quick Check of Response Times:
RMF CF Structure Activity Summary

- **Usage of the Report**
  - For measurement and monitoring of structure activity
  - Data reported for each individual structure
  - Structure activity reported on per system basis and for total CF

- **Report divided up into the following sections on a per structure per system basis**
  - Structure information
  - Request load and response time information
  - Delay information

- **Much of report reads very similar to CF Subchannel Activity**
  - Only data for each individual structure
  - Total # REQ incremented by 1 rather than 2 for
    - Sync non-immediate path busy
    - Async path busy
Quick Check of Response Times:
Use CF Structure Activity Report

COUPLING FACILITY NAME = P1CF001

COUPLING FACILITY STRUCTURE ACTIVITY

STRUCTURE NAME = DSNDB10_LOCK1 TYPE = LOCK  STATUS = ACTIVE

SYSTEM TOTAL # OF -SERV TIME(MIC)- REASON # OF - AVG TIME(MIC)-
NAME AVG/SEC REQ  ALL  AVG  STD_DEV  REQ  DEL  STD_DEV /A

SYS0  8247  SYNC  79K  1.7  20.0  3.9  NO SCH  0  0.0  0.0  0.0  0
  92.05  ASYNC  288  0.1  123.1  136.9  FR NT  0  0.0  0.0  0.0  0
  CHRGD  0  0.0  INCLUDED IN ASYNC  FR CMP  0  0.0  0.0  0.0  0

SYS1  1588K  SYNC  10K  0.1  20.0  1  NO SCH  0  0.0  0.0  0.0  0
  1764  ASYNC  80K  1.7  72.1  16.8  FR NT  0  0.0  0.0  0.0  0
  CHRGD  0  0.0  INCLUDED IN ASYNC  FR CMP  0  0.0  0.0  0.0  0

SYS2  1168K  SYNC  1508K  31.8  20.5  4.2  NO SCH  0  0.0  0.0  0.0  0
  1293  ASYNC  80K  1.7  100.0  16.8  FR NT  0  0.0  0.0  0.0  0
  CHRGD  0  0.0  INCLUDED IN ASYNC  FR CMP  0  0.0  0.0  0.0  0

SYS3  723K  SYNC  706K  14.9  20.3  5.5  NO SCH  0  0.0  0.0  0.0  0
  803.2  ASYNC  17K  0.4  175.1  276.0  FR NT  0  0.0  0.0  0.0  0
  CHRGD  0  0.0  INCLUDED IN ASYNC  FR CMP  0  0.0  0.0  0.0  0

SYS4  1179K  SYNC  1134K  24.0  20.8  4.6  NO SCH  0  0.0  0.0  0.0  0
  1310  ASYNC  45K  0.9  90.3  171.8  FR NT  0  0.0  0.0  0.0  0
  CHRGD  0  0.0  INCLUDED IN ASYNC  FR CMP  0  0.0  0.0  0.0  0

TOTAL  4736K  SYNC  4548K  96.0  20.6  4.5  NO SCH  0  0.0  0.0  0.0  0
  5262  ASYNC  188K  4.0  92.2  132.4  FR NT  0  0.0  0.0  0.0  0

Looking at Volume of Requests SYS1 is Sending to CF

Sync request to CF PINCF002 appears to be largest number of requests.
Number of Sync CF Requests –
Lock Structures requests to SYS1

- DB2 lock structure accounts for most of the sync requests.

Average Sync Request Response Times – Lock Structures

- The average response time of the sync lock requests from SYS1 are a little above our 20 microsecond guideline (but well below our 50 microsecond guideline).
Average Sync Request Response Times – Lock Structures

- Alternative Example showing RTs to Remote CF

In this example, we need many lock structures. Note: two time groupings. Times less than 20ms, and times greater than 20ms. This difference is due to the distance of remote CF. Speed of light still needs to be factored in.

Number of Async CF Requests – Lock Structures

CF Response Time Analysis – Number of Async Requests by Lock Structure
Average Async Request Response Times – Lock Structures

The async response times are within criteria.

What is the spin component of the host effect?

In other words... how many CPU seconds are consumed by the requesting z/OS system while it is ‘spinning’ waiting for the CF to respond to sync requests?
**Performance Cost of Parallel Sysplex**

- **Host CPU effect**
  - Cost to host (z/OS) system and its workloads by having it participate in a parallel Sysplex

- **Host CPU effect varies based on**
  - What portion of the workload is involved in data sharing
  - Access rate to the shared data
  - Hardware configuration for host, CF, and CF links
  - Number of systems in the Sysplex

- **According to IBM, the typical observed performance cost for Parallel Sysplex is:**
  - 3%
    - Cost of multisystem management
    - Cost of resource sharing
  - <10% - Cost of data sharing
  - 0.5% - Incremental cost of adding a new system image to the Sysplex

- **Effects on individual transactions / jobs can vary widely**

---

**Coupling Technology versus Host Processor Speeds**

*Source: The Top 10 Questions, Gary King, IBM*

<table>
<thead>
<tr>
<th>Host</th>
<th>z890 ISC</th>
<th>z890 ICB</th>
<th>z9 BC ISC</th>
<th>z9 BC ICB</th>
<th>z9 EC ISC</th>
<th>z9 EC ICB</th>
<th>z10 ISC</th>
<th>z10 ICB</th>
<th>z10 EC ISC</th>
<th>z10 EC ICB</th>
<th>z196 ISC</th>
<th>z196 PSIFB 12X</th>
<th>z9 BC PSIFB 12X</th>
<th>z9 BC ICB</th>
<th>z9 EC PSIFB 12X</th>
<th>z9 EC ICB</th>
<th>z10 BC ISC</th>
<th>z10 BC ICB</th>
<th>z10 EC ISC</th>
<th>z10 EC ICB</th>
<th>z196 PSIFB 12X</th>
</tr>
</thead>
<tbody>
<tr>
<td>z890 ISC</td>
<td>13%</td>
<td>15%</td>
<td>16%</td>
<td>17%</td>
<td>19%</td>
<td>21%</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z890 ICB</td>
<td>6%</td>
<td>10%</td>
<td>10%</td>
<td>11%</td>
<td>12%</td>
<td>13%</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z990 ISC</td>
<td>13%</td>
<td>14%</td>
<td>14%</td>
<td>15%</td>
<td>16%</td>
<td>17%</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z990 ICB</td>
<td>9%</td>
<td>9%</td>
<td>9%</td>
<td>10%</td>
<td>12%</td>
<td>13%</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z9 BC ISC</td>
<td>12%</td>
<td>13%</td>
<td>14%</td>
<td>15%</td>
<td>16%</td>
<td>17%</td>
<td>19%</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z9 BC ICB</td>
<td>8%</td>
<td>9%</td>
<td>9%</td>
<td>10%</td>
<td>11%</td>
<td>12%</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z9 EC ISC</td>
<td>12%</td>
<td>13%</td>
<td>14%</td>
<td>15%</td>
<td>16%</td>
<td>17%</td>
<td>19%</td>
<td>23%</td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z9 EC ICB</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
<td>9%</td>
<td>10%</td>
<td>11%</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z10 ISC</td>
<td>12%</td>
<td>13%</td>
<td>13%</td>
<td>14%</td>
<td>16%</td>
<td>18%</td>
<td>22%</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z10 ICB</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
<td>9%</td>
<td>10%</td>
<td>11%</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z10 EC ISC</td>
<td>11%</td>
<td>12%</td>
<td>12%</td>
<td>13%</td>
<td>15%</td>
<td>17%</td>
<td>22%</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z10 EC ICB</td>
<td>7%</td>
<td>7%</td>
<td>8%</td>
<td>9%</td>
<td>10%</td>
<td>11%</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z196 ISC</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z196 PSIFB 12X</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z196 PSIFB 12X</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sync Spin Seconds as a component of Host Effect
(One of four primary components)

- One component of the estimated host effect is CPU seconds consumed due to sync immediate Spin
  - Sync Immediate requests cause processor issuing the request to 'spin'
  - How many CPU seconds did the sending z/OS LPAR spend spinning?
    - Logical processor unavailable to other work running in the same LPAR
    - Physical processor that logical processor is dispatched to is unavailable to other LPARs
  - It is helpful to understand capacity consumed to these spinning conditions
    - It part of CPU utilization

Impact response time of sender, and performance of sending system

Using CF Subchannel Activity Report to Calculate CPU Spin Seconds for System

- CPU seconds consumed due to sync immediate Spin can be calculated by the Coupling Facility Subchannel Activity Report
  - Remember to do this exercise for both coupling facilities!

\[
\text{CPU Seconds Spinning} = \frac{\# \text{REQ \ Sync} \times \text{(Sync Service Time)}}{1,000,000}
\]

\[
\text{CPU Seconds Spinning} = \frac{(595,881) \times (22)}{1,000,000} = 13.1 \text{ seconds CPU Spin}
\]

- If you wanted, you can now represent this as a subcomponent of LPAR Busy %

<table>
<thead>
<tr>
<th>SYSTEM</th>
<th>NAME</th>
<th>AVG/SEC TYPE</th>
<th>GEN</th>
<th>USE</th>
<th>BUSY</th>
<th># REQ</th>
<th>AVG</th>
<th>STD_DEV</th>
<th># CF</th>
<th>REQ</th>
<th>Req / DEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYS1</td>
<td>927793 CFP 2 2 0 0</td>
<td>0.0 0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1030.9</td>
<td>SUBCH 14 14 14</td>
<td>115.6 0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Example: Estimated Host Effect Seconds for SYS1

Sync spin seconds for SYS1

Looking at Estimated Host Effect CPU Second by System

Note: I use 900 seconds as a crude guideline since that is the amount of CPU one z/OS processor can consume in 15 minutes. But really should really calculate as a portion of LPAR Busy %.
Example: SYS1 Most Expensive Structures

Shows for periods of study (24 hours) the estimated host CPU seconds for the top most expensive structures for SYS1. Note: These estimated CPU seconds are broken down by reason (of which is sync spin).

Blue show spin seconds portion of host effect.

When Calculated by Structure...

Shows for periods of study (24 hours) the estimated host CPU seconds for the top most expensive structures broken down by system. Note: These values include all components of host effect (of which sync spin is one).
What percentage of Lock requests are due to false contention?

Types of Lock Contention

- **Real Lock Contention**
  - Contention caused by multiple units of work attempting to serialize on the same resource
  - Factors that influence real lock contention:
    - How the locks are being used
    - Amount of time locks are held
    - Degree of data sharing
  - Alleviate real lock contention by tuning the workload (not by tuning the Sysplex or CF structures)
### Types of Lock Contention

- **False Lock Contention**
  - When multiple lock names are hashed to the same lock entry
  - Results in significant excessive processing overhead to resolve
  - Factors that influence false lock contention
    - Size of lock structure
    - Granularity of locking (record, file, block)
    - Concurrent users connected to lock structure
  - Alleviate false lock contention by increasing lock structure size

![Lock Table Diagram]

### Tuning Lock Structures

- **Lock structures can incur two types of contention**
  - True lock contention
    - Multiple units of work attempting to serialize on the same resource
    - Factors that influence real lock contention
      - how the locks are being used and amount of time locks are held
      - degree of data sharing
    - Alleviate real lock contention by tuning the workload
      (not by tuning the sysplex or CF structures)
  - False lock contention
    - When multiple lock names are hashed to the same lock entry
    - Results in significant excessive processing overhead to resolve
    - Factors that influence false lock contention
      - size of lock structure and granularity of locking (record, file, block)
      - concurrent users connected to lock structure
    - Alleviate false lock contention by increasing lock structure size
      - Be careful of very large structures. Impact to rebuild during recovery may outweigh the performance impact of false contention
Tuning Lock Structures

- Use Coupling Facility Structure Summary
  - Look at LOCK ENTRIES (TOT/CUR)
    - Is current approaching total as workload is growing
    - False contention may happen soon in the future as workload continues to grow
    - May need to increase the allocated size
  - All lock requests are synchronous
  - Make sure service time are within guidelines

<table>
<thead>
<tr>
<th>STRUCTURE SUMMARY</th>
</tr>
</thead>
<tbody>
<tr>
<td>STRUCTURE</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>DSNDGP2_LOCK1</td>
</tr>
<tr>
<td>DSNDGP3_LOCK1</td>
</tr>
<tr>
<td>DSNDGP5_LOCK1</td>
</tr>
<tr>
<td>DSNDGT1_LOCK1</td>
</tr>
<tr>
<td>DSNDGT3_LOCK1</td>
</tr>
</tbody>
</table>

- Monitor true lock contention
  - Less than 5% CONT events of REQ TOTAL
  - Lock contention = percent of CONT of Request Total
  - ROT: < .1% for CICS/DBCTL, CICS/VSAM RLS, GRSSTAR
  - ROT: < 1% for IMS-TM/DB2 (less than 5% is probably OK)
  - If missing ROT, look at other areas (like batch) that may be using the locks

<table>
<thead>
<tr>
<th>COUPLING FACILITY STRUCTURE ACTIVITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYSTEM NAME = DSNDGP3_LOCK1</td>
</tr>
<tr>
<td># REQ</td>
</tr>
<tr>
<td>NAME</td>
</tr>
<tr>
<td>SYSA</td>
</tr>
<tr>
<td>1081</td>
</tr>
<tr>
<td>1082</td>
</tr>
<tr>
<td>CHNGD</td>
</tr>
<tr>
<td>SYSB</td>
</tr>
<tr>
<td>1082</td>
</tr>
<tr>
<td>1083</td>
</tr>
<tr>
<td>CHNGD</td>
</tr>
<tr>
<td>CHNGD</td>
</tr>
<tr>
<td>(edited)</td>
</tr>
</tbody>
</table>
Tuning Lock Structures cont...

- Monitor false lock contention
  - Recommend less than 0.5% FALSE CONT of REQ TOTAL
  - False contention = FALSE CONT / REQ TOTAL
  - Make sure it is zero, but take into consideration the increase in rebuild time
  - If greater, then increase structure size (increases number of lock entries)

<table>
<thead>
<tr>
<th>Structure Name</th>
<th>Type</th>
<th>Status</th>
<th># Req</th>
<th>Requests</th>
<th>Delayed Requests</th>
<th>External Request Contention</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYSA 1946K Sync 1946K</td>
<td>Lock</td>
<td>Active</td>
<td>6.7</td>
<td>15.6</td>
<td>8.6</td>
<td>NO SCH</td>
</tr>
<tr>
<td>1082 Async 0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>PR MT</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Chmod 0</td>
<td>0.0</td>
<td>INCLUDE IN ASYNC</td>
<td>PR CHM</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>SYSB 3471K Sync 3471K</td>
<td>Lock</td>
<td>Active</td>
<td>11.9</td>
<td>12.8</td>
<td>7.5</td>
<td>NO SCH</td>
</tr>
<tr>
<td>1082 Async 0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>PR MT</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Chmod 0</td>
<td>0.0</td>
<td>INCLUDE IN ASYNC</td>
<td>PR CHM</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

TOTAL 29048K Sync 29M | 100 | 12.9 | 7.6 | NO SCH | 785 | 0.0 | 9.7 | 14.6 | 0.0 | REQ TOTAL 28M |
| 16138 Async 4497 | 0.0 | 81.5 | 87.0 | PR MT | 0 | 0.0 | 0.0 | 0.0 | 0.0 | REQ DEFERRED 376K |
| Chmod 0 | 0.0 | INCLUDE IN ASYNC | PR CHM | 0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | FALSE CONT 81K |

Lock Structure Events
Lock Structure Events

Are there cross invalidates due to directory reclaims?
General Case of Shared Data Read (Generic example)

- **Initial Read (if the data has never been read before)**
  - Read data and register interest in data into the coupling facility
  - Data may or may not have also been written to the CF upon read
  - Coupling facility always knows which systems have interest in which data
  - Local cache vectors indicate if a system's data is the most recent copy or is invalidated

- **Subsequent read / reference of the data**
  - Local cache vectors used to determine if local buffers contain the most recent copy (and if not the data is re-read).
    - May have been cross invalidated due to an update.
  - Updated page is then gotten from either CF or DASD

General Case of Cross Invalidates Due To Writes (Generic example)

- **To update the data**
  - Updating system must obtain a lock to serialize data update
  - Data is updated in local buffers
  - CF notified of update
  - CF cross invalidates (XI) all the local buffers of the other systems that have registered an interest in this data by marking the data as invalid in the local cache vector tables of each system.
  - Based on installation setup, the updated data may or may not be written immediately to the CF and/or DASD
### General Case Cross Invalidate Due to Directory Reclaim

**Generic example**

- **XI for Directory Reclaim count**
  - The number of times that a directory entry was stolen and XI signals had to be sent to a system because the page for the directory entry was cached in one or more systems.

- **Directory Entry Reclaim count**
  - The number of times that a page name assignment required a coupling facility directory entry to be reclaimed (stolen).

---

### Cross Invalidates Due To Directory Reclaims

- **Cross Invalidates (XIs)**
  - When a data item residing in a local buffer was marked invalid by the coupling facility.
    - To the cache structure user, this means the data item must be re-acquired from DASD or perhaps the coupling facility structure, and interest in the item must be re-registered in the coupling facility structure.

  - There are 5 reasons for Cross Invalids
    - One reason is XI due to Directory Reclaim

- **XI for Directory Reclaim**
  - The number of times that a directory entry was stolen and XI signals had to be sent to a system because the page for the directory entry was cached in one or more systems.

- **Directory Entry Reclaims**
  - The number of times that a page name assignment required a coupling facility directory entry to be reclaimed (stolen).
Coupling Facility Usage Summary for CF

- Provides measurements for total activity to a structure, and usage summary.

### Analyzing Cache Structure Activity Example

#### STRUCTURE NAME = DSNDB10_GBP15     TYPE = CACHE  STATUS = ACTIVE PRIMARY

<table>
<thead>
<tr>
<th>STRUCTURE</th>
<th>STATUS</th>
<th>ALLOC</th>
<th>CF</th>
<th>% OF CHG</th>
<th>SIZE</th>
<th>STOR</th>
<th>REQ</th>
<th>REQ UTIL</th>
<th>SEC</th>
<th>TOT/CUR</th>
<th>TOT/CUR</th>
<th>TOT/CUR XI'S</th>
<th>LOCK</th>
<th>DIR REC</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST LOG_DFHLOG_001</td>
<td>ACTIVE</td>
<td>12M</td>
<td>0.2</td>
<td>18341</td>
<td>0.3</td>
<td>1.2</td>
<td>20.28</td>
<td>2406</td>
<td>9706</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>LOCK DSNDB10_LOCK1</td>
<td>ACTIVE</td>
<td>256M</td>
<td>4.4</td>
<td>4736K</td>
<td>84.7</td>
<td>70.0</td>
<td>5262.0</td>
<td>368K</td>
<td>0</td>
<td>67M</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>LIST LOG_DFHLOG_002</td>
<td>ACTIVE</td>
<td>12M</td>
<td>0.2</td>
<td>10355</td>
<td>0.2</td>
<td>0.6</td>
<td>11.51</td>
<td>2406</td>
<td>9706</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>LOCK DSNDB10_LOCK2</td>
<td>ACTIVE</td>
<td>12M</td>
<td>0.2</td>
<td>38674</td>
<td>0.7</td>
<td>2.0</td>
<td>42.97</td>
<td>2406</td>
<td>9706</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

**COUPLING FACILITY NAME = P1CFALT**

### Coupling Facility Usage Summary for CF

<table>
<thead>
<tr>
<th>SYSTEM NAME</th>
<th>AVG/SEC</th>
<th># OF SERV</th>
<th>TIME(MIC)</th>
<th>REASON</th>
<th># OF SERV</th>
<th>TIME(MIC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYS0</td>
<td>0</td>
<td>55.4</td>
<td>0.0</td>
<td>NO SCH</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>SYS4</td>
<td>475</td>
<td>8.9</td>
<td>0.0</td>
<td>NO SCH</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>SYS3</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
<td>NO SCH</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>SYS2</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
<td>NO SCH</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>SYS1</td>
<td>16323</td>
<td>33.4</td>
<td>0.0</td>
<td>NO SCH</td>
<td>1</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**z/OS V1R11**

© Enterprise Performance Strategies, Inc.

Peter Enrico: www.epstrategies.com

Some Key CF Measurements - 51
Are any Cache Structures doing Directory Reclaims?

For period of study (24 hours) we see that only one structure incurred directory reclaims, and the value of directory reclaims appears high.

The question to now ask is if these reclaims regularly happen during the day?

- If yes, then structure may be too small and these reclaims may be resulting in XIs which would unnecessarily hurt workloads.
- If happen during short period such as batch, then may be due to database maintenance. So much less of a worry.

Example of Directory Reclaims During Batch

In this particular example, the reclaims are happening off shift, during the batch cycle, and all at once. So these reclaims are due to database maintenance, or something similar.
Directory Reclaims in 24 hours
(different Sysplex example)

In this particular example, the reclaims are happening during prime shift.

Directory Reclaims Over Time
(different Sysplex example)
Cross Invalidates Due to Writes (different Sysplex example)

XIs due to writes means one system updating data, so other have to know about that update.

This is what data sharing looks like.

Cross Invalidates Due to Directory Reclaims (different Sysplex example)

The XIs here are XIs that result from directory reclaims.

These could be avoided if directory area of structure is enlarged. Currently unnecessary overhead and invalidation of local buffers.
What is the structure configuration?

Coupling Facility Usage Summary

- Provides an overview of how well the resources in CF are being utilized
  - Report divided up into three primary sections
    - Structure Summary - 1st Section of the report
    - Storage Summary - Middle Section of the report
    - Processor Summary - Final section of the report
Coupling Facility Usage Summary for CF

- Provides measurements for total activity to a structure, and usage summary

---

Example of Summarizing Structure Layout

<table>
<thead>
<tr>
<th>Sysplex</th>
<th>CF Name</th>
<th>Structure Type</th>
<th>Structure Name</th>
<th>Str Duplexed</th>
<th>Str Placement</th>
<th>Structure Alloc MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1PLEX</td>
<td>CF001</td>
<td>Cache</td>
<td>DSNDB10 GBP0</td>
<td>Y Secondary</td>
<td>85</td>
<td>1GB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBP10</td>
<td>Y Secondary</td>
<td>1033</td>
<td>10GB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBP11</td>
<td>Y Secondary</td>
<td>1033</td>
<td>10GB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBP12</td>
<td>Y Secondary</td>
<td>17</td>
<td>17MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBP13</td>
<td>Y Primary</td>
<td>17</td>
<td>17MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBP15</td>
<td>Y Secondary</td>
<td>17</td>
<td>17MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBP16</td>
<td>Y Primary</td>
<td>17</td>
<td>17MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBP19K</td>
<td>Y Secondary</td>
<td>26</td>
<td>26MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBP32K</td>
<td>Y Secondary</td>
<td>26</td>
<td>26MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBP4</td>
<td>Y Secondary</td>
<td>48</td>
<td>48MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 GBRK0</td>
<td>Y Primary</td>
<td>48</td>
<td>48MB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SYSIGGCAS ECS</td>
<td>N</td>
<td>3</td>
<td>N</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sysplex</th>
<th>CF Name</th>
<th>Structure Type</th>
<th>Structure Name</th>
<th>Str Duplexed</th>
<th>Str Placement</th>
<th>Structure Alloc MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>ListU</td>
<td></td>
<td></td>
<td>LOG DHLOG001</td>
<td>N</td>
<td>12</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LOG DHLOG002</td>
<td>N</td>
<td>12</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LOG DHLOG003</td>
<td>N</td>
<td>12</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LOG DHLOG004</td>
<td>N</td>
<td>12</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DSNDB10 LOCK1</td>
<td>N</td>
<td>256</td>
<td>N</td>
</tr>
</tbody>
</table>

---
As a reminder...
Many Questions Need to be Asked of CF Measurements

- Configuration / Setup questions
- Link and general load performance questions
- Host effect questions
- Processor related questions
- Storage usage related questions
- List structure related questions
- Lock structure related questions
- Cache structure related questions
- Duplexing related questions

Thank you!

Please contact me directly if you need more information, or if you want to show me your measurements to get my thoughts.

Email: Peter.Enrico@EPStrategies.com