Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DBE, e-business logo, ESCO, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, VM/ESA, VSE/ESA, Websphere, xSeries, z/OS, zSeries, z/VM

The following are trademarks or registered trademarks of other companies

Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries
LINUX is a registered trademark of Linus Torvalds
UNIX is a registered trademark of The Open Group in the United States and other countries.
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.
Intel is a registered trademark of Intel Corporation
* All other products may be trademarks or registered trademarks of their respective companies.

NOTES:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.
Acknowledgments

- A world of thanks to Brian Wade for pulling together the majority of this information.
1. **Collecting Monwrite Data**: why, and how, to set up and collect CP Monitor records
2. **z/VM is Doing Fine**: This study shows how to recognize a healthy z/VM system.
3. **Emergency Scan**: We get asked about “emergency scan” from time to time. This case study explains it and tells why seeing it is not always an “emergency” in the literal sense.
4. **Why Doesn’t My System Page Faster?**: This system isn’t broken, but the customer didn’t understand its behavior. The case study illustrates why it’s important to know the big picture when trying to discern meaning in measurement data.
5. **Undersized LPAR**: This system is generally short on storage, CPU, and paging. The case study illustrates how to detect it and how to fix it.
6. **PAV and MDC**: This customer called in with a performance PMR and we ended up taking an APAR. See how we put the finger on a CP bug using CP monitordata.
7. **Paging Difficulties**: This system was grossly underconfigured for paging. The case study illustrates what we examined and what changes we recommended. It also illustrates what happens when one relieves a constraint: namely, one usually bumps into another one.
8. **The Grinch that Stole Performance**: This case study illustrates how system performance can change when hardware fails. CP Monitor showed where the problem was and pointed the way for a hardware fix.
9. **HiperSockets Performance**: HiperSockets is thought of as a super high speed connectivity feature. This case study looks at some aspects of that performance and a scenario where it might not be as fast as expected.
Collecting MONWRITE Data

Revision 2008-07-30 BKW

IBM z/VM Performance Evaluation
Brian Wade bkw@us.ibm.com
Bill Bitner bitnerb@us.ibm.com
Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DBE, e-business logo, ESCO, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, VM/ESA, VSE/ESA, Websphere, xSeries, z/OS, zSeries, zVM

The following are trademarks or registered trademarks of other companies

Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries
LINUX is a registered trademark of Linus Torvalds
UNIX is a registered trademark of The Open Group in the United States and other countries
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC
Intel is a registered trademark of Intel Corporation
* All other products may be trademarks or registered trademarks of their respective companies.

NOTES:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.
Agenda

- What is raw monitor data?
- How do I set up to collect it?
- When do I collect it?
- What tools are available to help me collect it?
- How do I package it for transmission?
- How do I study it myself?
- Summary
What is Raw Monitor Data?

- It is unformatted binary data describing system configuration or activity
- Logically, it is a sequence of monitor records
  - Each record comments on some specific aspect of system activity or performance
  - In aggregate they constitute a comprehensive, time-indexed record of system activity
- There are three large classes of monitor records
  - Configuration records: emitted when monitor starts, these describe system configuration
  - Sample records: emitted every so often, these comment on the accumulated activity of an entity (device, user, …)
  - Event records: emitted as needed, these comment on some specific phenomenon that just now occurred
- Some records come from the Control Program and comment on its experience in running the system
- Other records come from guests and comment on their experiences in doing whatever it is they do
- We collect this data using an IBM-supplied utility program called MONWRITE
- During the rest of this presentation, we will call this data MONWRITE data
How Do I Collect MONWRITE Data?

- **By Default the z/VM system is set up with DCSS and user ID named MONWRITE**

- **If somehow skipped, then:**
  - You set up a DCSS where CP will buffer the monitor records it emits
    - CP DEFSEG and SAVESEG commands
  - You tell CP which kinds of records to emit, and how often to emit them, and in fact to begin emitting them
    - CP MONITOR command
  - You set up a guest that drains the DCSS to a disk or a tape via the MONWRITE utility

- **On some occasions, the default DCSS (named MONDCSS) is too small.**

- **You run the guest**

- **You archive the resultant files or tapes, so that you have a long-term historical record of system activity and performance**
When Do I Collect MONWRITE Data?

- Periodically, collect and archive some data during your peak periods, so that you have a historical record
  - Every Tuesday at 10 AM for an hour
  - Month-end processing
  - Whenever you do that really big thing you do

- When directed by IBM
  - PMR, crit sit, ESP, whatever
Tool: Running MONWRITE By Hand

- A great idea, assuming you are not running some other performance product
  - If you know what you are doing, you can do both simultaneously
- Create the DCSS to hold the buffered records
- Set up a guest to run our MONWRITE MODULE (collector)
- Issue some CP MONITOR commands to start CP emitting records
  - Enable all samples
  - Enable all events except seeks and scheduler
  - Use a 1-minute sample interval and a 5-second HFS rate
- In your guest, start MONWRITE to collect the data CP’s emitting
- To stop collecting, type this: MONWSTOP
- You will end up with one MONWRITE file that you can:
  - Archive for the historical record
  - Analyze yourself with z/VM Performance Toolkit
  - Send to IBM so we can look at it
- There is an option for MONWRITE to close the file at regular times of day and a user exit to process the just-closed file.
- Good references:
  - z/VM Performance, chapter 9, “Monitoring Performance Using CP Monitor” – an excellent writeup of every last detail
Tool: Brian’s LINMON Collector

- Based on a modified MONWRITE
- Sets up the DCSS, CP Monitor, etc. on its own, using certain assumptions that are probably safe for many systems
- Can be configured to:
  - Account for presence of another performance product
  - Collect for a while then log off
  - Start a new file every so often
  - Keep only the last N files
Packaging MONWRITE Data For Transmission

- MONWRITE files are binary CMS files, F 4096.
- Just attaching them to an e-mail is NOT recommended.
- We suggest you use the VMARC file archiver that runs on CMS for large files or when several files are being sent.
  - Kind of like “zipping” on a PC (compresses, combines)
    - MONWRITE data is very compressible
    - The data compression and decompression steps assure data integrity
    - Sometimes you also want to send us a console spool, or some QUERY outputs, or whatever
    - You can package everything into one VMARC archive and send us only that
- The standard z/VM Level 2 process for FTPing files calls for COPYFILE (PACK)
  - This is unnecessary for MONWRITE and VMARC files.
- Always, always, always:
  - Move the files in binary (ASCII is a sure showstopper)
  - Do not use FTP’s SITE FIXREC, QUOTE SITE FIXREC, or LOCSITE FIXREC features (error-prone)
- You will FTP your VMARC archive file to IBM’s receiving server in Boulder, CO
  - Testcase.boulder.ibm.com, anonymous, cd /toibm/vm, binary, put, quit
  - Name your file mnemonically and send us a note or update the PMR
  - In PMR and/or note be clear as to what is sent and how packaged
  - See [http://techsupport.services.ibm.com/390/tcprocs.html](http://techsupport.services.ibm.com/390/tcprocs.html) for additional Testcase help
Studying MONWRITE Data

- z/VM Performance Toolkit
- Interactively – possible, but not so useful
- PERFKIT BATCH command – pretty useful
  - Control files tell Perfkit which reports to produce
  - You can then inspect the reports by hand or programmatically
- See z/VM Performance Toolkit Reference for information on how to use PERFKIT BATCH
Other Types of Data Confused with MONWRITE Data

- **Asking for “raw VM monitor” data can be confusing.**
  - Velocity has its own form of raw monitor data and history files, and even a form that attempts to mimic MONWRITE.
  - “VM Monitor” sounds like the “VM:product” often associated with CA products.
  - Performance Toolkit’s history, trend, and summary files do not have the same detail.

- **Be specific when asking for data.**
Summary

- MONWRITE data is the single most comprehensive record of system activity
- It is invaluable in diagnosing performance concerns
- If you ask IBM for performance help, IBM will very likely ask you for MONWRITE data
- Practice collecting and transmitting MONWRITE data when you are not under duress
- Archive your MONWRITE data routinely so that you have a good record of your system’s usual behavior
- Learn to use PERFKIT BATCH to generate reports, and get familiar with a few of the basic reports
Case Study: z/VM Seems To Be Fine

Revision 2008-07-31 BKW

IBM z/VM Performance Evaluation
Brian Wade bkw@us.ibm.com
Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DBE, e-business logo, ESCO, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, VM/ESA, VSE/ESA, Websphere, xSeries, z/OS, zSeries, zVM

The following are trademarks or registered trademarks of other companies

Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries
LINUX is a registered trademark of Linus Torvalds
UNIX is a registered trademark of The Open Group in the United States and other countries. Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.
Intel is a registered trademark of Intel Corporation
* All other products may be trademarks or registered trademarks of their respective companies.

NOTES:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user’s job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.
Question from Customer

- Linux on z/VM on 2086
- Java core and heap dumps
- Linux transaction rollback exceptions
- Linux slow response time
- Can you please take a look
- I have MONWRITE data for you
  - … that’s always tempting, so I looked
Basic Things to Check

- **Do we have enough CPU**
  - FCX225 SYSSUMLG
  - FCX126 LPAR
  - FCX114 USTAT, %CPU
- **Do we have enough storage**
  - FCX114 USTAT, %PGW
  - FCX113 UPAGE, XSTORE and DASD paging
- **Do we have enough SXS storage**
  - FCX264 SXSUTIL
- **Are we spending too much time in the Control Program**
  - FCX225 SYSSUMLG
- **Are we paging OK**
  - FCX109 DEVICE CPOWERD, paging I/O performance
  - FCX103 STORAGE, page blocking factors
  - FCX113 UPAGE, is XSTORE more active than DASD
- **Is I/O performance OK**
  - FCX108 DEVICE
  - FCX111 CACHDBSE
- **Is networking performance OK**
  - Find OSD chpids via FCX161
  - FCX215 FCHANNEL
  - FCX240 VSWITCH
Do We Have Enough CPU: FCX126 and FCX225

FCX126  Run 2008/07/24 12:46:56    LPAR
Logical Partition Activity

Partition Nr.  Upid  #Proc  Weight  Wht-C  Cap  %Load  CPU  %Busy  %Dvh  %Susp  %VMd  %Logd  Type
LMRHA   1    ..    0            NO    0 ...   ...   ...   ...   ...    ... ..
LMRPROD 2    01    2    500     NO  NO ...    0  41.9  41.9 .2  41.7   41.8 ICF
          500     NO    1  52.2  52.2 .2  52.0   52.1 ICF

<-- This is us

Only one LPAR is using these two engines.

("ICF" is misleading - we are investigating this - the PUs are CPs - FCX180 SYSCONF reports correctly.)

We don't see any particularly high percentages.

Also note T/V is nearly perfect.
**CPU and Storage: FCX114 USTAT**

FCX114 Run 2008/07/24 12:46:56

USTAT

Wait State Analysis by User

From 2008/07/24 12:54:38
To 2008/07/24 14:04:38
For 4200 Secs 01:10:00
Result of 12345 Run

```
<table>
<thead>
<tr>
<th>User Id</th>
<th>%ACT</th>
<th>%RUN</th>
<th>%CPU</th>
<th>%LDG</th>
<th>%PGW</th>
<th>%OW</th>
<th>%SM</th>
<th>%FM</th>
<th>%W</th>
<th>%CFW</th>
<th>%FI</th>
<th>%EL</th>
<th>%DM</th>
<th>%QA</th>
<th>%PGA</th>
<th>%LI</th>
<th>%ETH</th>
<th>%Q0</th>
<th>%Q1</th>
<th>%Q2</th>
<th>%Q3</th>
<th>%EO-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;System&lt;</td>
<td>22</td>
<td>21</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>71</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>98</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LMRCH2P</td>
<td>99</td>
<td>29</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>65</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LMRLN2P</td>
<td>99</td>
<td>8</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>86</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LMRWASP</td>
<td>99</td>
<td>64</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DTCVSW2</td>
<td>91</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>91</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LMRLN1P</td>
<td>87</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>92</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PERFSVM</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TCPI P</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>99</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
```
Do We Have Enough Storage: FCX113 UPAGE

FCX113 Run 2008/07/24 12:46:56
UPAGE
User Paging Activity and Storage Utilization

From 2008/07/24 12:54:38
To 2008/07/24 14:04:38
For 4200 Secs 01:10:00
Result of 12345 Run

<table>
<thead>
<tr>
<th>Userid</th>
<th>Owned</th>
<th>Reads</th>
<th>Write</th>
<th>Steals</th>
<th>&gt;2GB</th>
<th>X&gt;MS</th>
<th>MS&gt;X</th>
<th>X&gt;OS</th>
<th>VBS</th>
<th>Resrvd</th>
<th>R&lt;2GB</th>
<th>R&gt;2GB</th>
<th>L&lt;2GB</th>
<th>L&gt;2GB</th>
<th>XSTOR</th>
<th>DASD</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;System&lt;</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.5</td>
<td>.0</td>
<td>.2</td>
<td>.5</td>
<td>.0</td>
<td>117135</td>
<td>0</td>
<td>22167</td>
<td>94880</td>
<td>4</td>
<td>27</td>
<td>9374</td>
<td>39</td>
</tr>
<tr>
<td>BKRBKUP</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>83</td>
<td>0</td>
<td>0</td>
<td>65</td>
<td>0</td>
<td>0</td>
<td>399</td>
<td>0</td>
</tr>
<tr>
<td>BKRCATLG</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>85</td>
<td>0</td>
<td>0</td>
<td>66</td>
<td>0</td>
<td>0</td>
<td>412</td>
<td>0</td>
</tr>
<tr>
<td>DI SKACNT</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>1227</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1227</td>
<td>0</td>
</tr>
<tr>
<td>DTCVSW1</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.1</td>
<td>.1</td>
<td>49</td>
<td>0</td>
<td>1</td>
<td>45</td>
<td>0</td>
<td>1</td>
<td>2614</td>
<td>0</td>
</tr>
<tr>
<td>DTCVSW2</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>273</td>
<td>0</td>
<td>10</td>
<td>297</td>
<td>8</td>
<td>26</td>
<td>2385</td>
<td>0</td>
</tr>
<tr>
<td>EREP</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>78</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1231</td>
<td>0</td>
</tr>
<tr>
<td>LMRDB2P</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.4</td>
<td>4.7</td>
<td>.0</td>
<td>1027k</td>
<td>0</td>
<td>269342</td>
<td>758161</td>
<td>4</td>
<td>68</td>
<td>17870</td>
<td>0</td>
</tr>
<tr>
<td>LMRNL1P</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>3.0</td>
<td>.0</td>
<td>449333</td>
<td>0</td>
<td>116199</td>
<td>333205</td>
<td>0</td>
<td>71</td>
<td>73948</td>
<td>0</td>
</tr>
<tr>
<td>LMRNL2P</td>
<td>.0</td>
<td>.0</td>
<td>.2</td>
<td>.0</td>
<td>.0</td>
<td>.2</td>
<td>.8</td>
<td>.2</td>
<td>501237</td>
<td>0</td>
<td>73352</td>
<td>427956</td>
<td>4</td>
<td>68</td>
<td>20735</td>
<td>847</td>
</tr>
<tr>
<td>LMRWASP</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>2.0</td>
<td>.2</td>
<td>.0</td>
<td>593939</td>
<td>0</td>
<td>28710</td>
<td>565308</td>
<td>48</td>
<td>27</td>
<td>74281</td>
<td>0</td>
</tr>
</tbody>
</table>
**Is SXS OK: FCX264 SXSUTIL**

FCX264  Run 2008/07/24 12:46:56  SXSUTIL

System Execution Space Utilization

From 2008/07/24 12:54:38
To 2008/07/24 14:04:38
For 4200 Secs 01:10:00  Result of 12345 Run

<table>
<thead>
<tr>
<th>Intervals</th>
<th>Total Pages Queues</th>
<th>Total Mapped</th>
<th>Free</th>
<th>Other</th>
<th>Total</th>
<th>Total CPLock</th>
<th>Owned</th>
<th>&lt;2GB</th>
<th>&gt;2GB</th>
<th>Queue</th>
<th>Pages</th>
<th>Mean</th>
<th>524287</th>
<th>513742</th>
<th>10545</th>
<th>3063</th>
<th>812</th>
<th>6485</th>
<th>185</th>
<th>28</th>
<th>0</th>
<th>0</th>
<th>7873</th>
<th>2714</th>
<th>185</th>
<th>1851</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:55:38</td>
<td>524287</td>
<td>513750</td>
<td>10537</td>
<td>3063</td>
<td>812</td>
<td>6485</td>
<td>185</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>7873</td>
<td>2714</td>
<td>185</td>
<td>1851</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:56:38</td>
<td>524287</td>
<td>513756</td>
<td>10531</td>
<td>3063</td>
<td>806</td>
<td>6478</td>
<td>184</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>7860</td>
<td>2719</td>
<td>185</td>
<td>1851</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:57:38</td>
<td>524287</td>
<td>513754</td>
<td>10533</td>
<td>3063</td>
<td>807</td>
<td>6479</td>
<td>184</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>7860</td>
<td>2719</td>
<td>185</td>
<td>1851</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:58:38</td>
<td>524287</td>
<td>513754</td>
<td>10533</td>
<td>3063</td>
<td>806</td>
<td>6480</td>
<td>184</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>7862</td>
<td>2719</td>
<td>185</td>
<td>1851</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Are We Paging OK: FCX109 DEVICE CPOWNED

<table>
<thead>
<tr>
<th>Volume</th>
<th>Area</th>
<th>Used</th>
<th>Rate/s</th>
<th>User</th>
<th>Serv</th>
<th>MLOAD</th>
<th>Block</th>
<th>%Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addr</td>
<td>Devtyp</td>
<td>Serial</td>
<td>Type</td>
<td>Extent</td>
<td>P-Rds</td>
<td>P-Wt</td>
<td>S-Rds</td>
<td>S-Wt</td>
</tr>
<tr>
<td>0D15</td>
<td>3390</td>
<td>LPVPM7 PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0D16</td>
<td>3390</td>
<td>LPVPM8 PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0D17</td>
<td>3390</td>
<td>LPVPM9 PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0D18</td>
<td>3390</td>
<td>LPVPMA PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0D19</td>
<td>3390</td>
<td>LPVPMB PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0E05</td>
<td>3390</td>
<td>LPVPM0 PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0E06</td>
<td>3390</td>
<td>LPVPM2 PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0E07</td>
<td>3390</td>
<td>LPVPMB PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0E08</td>
<td>3390</td>
<td>LPVPML PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0E09</td>
<td>3390</td>
<td>LPVPMS PAGE</td>
<td>1-3338</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Time In The Control Program: FCX225 SYSSUMLG

- We saw this report already
- T/V ~ 1.02
- TV = (CP time + guest time) / guest time
- 1.0 is a perfect T/V
## I/O Performance: FCX108 DEVICE

**FCX108 Run 2008/07/24 12:46:56**

> General I/O Device Load and Performance

<table>
<thead>
<tr>
<th>Addr Type</th>
<th>Label / ID</th>
<th>Links ths</th>
<th>I/O Avoid</th>
<th>Pend</th>
<th>Disc</th>
<th>Conn</th>
<th>Serv</th>
<th>Resp</th>
<th>CLMW</th>
<th>Qued</th>
<th>Busy</th>
<th>READ</th>
<th>Cyls</th>
<th>SSCH</th>
<th>Set/s</th>
<th>Dly/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;&gt; All DASD &lt;&lt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0E21 3390</td>
<td>LPLAML</td>
<td>1 4.6</td>
<td>0</td>
<td>2 2.1</td>
<td>4</td>
<td>3.7</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>737</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0E00 3390</td>
<td>LPVM1 CP</td>
<td>50 2</td>
<td>4</td>
<td>0</td>
<td>2 5.1</td>
<td>1</td>
<td>0</td>
<td>6.3</td>
<td>6.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>0E02 3390</td>
<td>LPVM2 CP</td>
<td>43 2</td>
<td>1</td>
<td>2</td>
<td>3.7</td>
<td>8</td>
<td>4.7</td>
<td>4.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>0E22 3390</td>
<td>LPQM1</td>
<td>1 1.5</td>
<td>0</td>
<td>2</td>
<td>3.2</td>
<td>1</td>
<td>0</td>
<td>4.4</td>
<td>4.4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1910</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0E20 3390</td>
<td>LPQM2</td>
<td>1 2.3</td>
<td>0</td>
<td>2</td>
<td>2.4</td>
<td>9</td>
<td>3.5</td>
<td>3.5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>884</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>0D13 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>2.6</td>
<td>5</td>
<td>3.3</td>
<td>3.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>83</td>
<td>393</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0E23 3390</td>
<td>LPQM2</td>
<td>1 2.7</td>
<td>1</td>
<td>2</td>
<td>2.3</td>
<td>8</td>
<td>3.3</td>
<td>3.3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>870</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>0E19 3390</td>
<td>LPLZM2</td>
<td>1 2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2.5</td>
<td>4</td>
<td>3.1</td>
<td>3.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>52</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0F2C 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>2.0</td>
<td>8</td>
<td>3.0</td>
<td>3.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>803</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0D12 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>2.3</td>
<td>4</td>
<td>2.9</td>
<td>2.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>61</td>
<td>490</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0D0F 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>2.8</td>
<td>0</td>
<td>2</td>
<td>2.0</td>
<td>6</td>
<td>2.8</td>
<td>2.8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>63</td>
<td>108</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0D11 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>1.2</td>
<td>0</td>
<td>2</td>
<td>2.1</td>
<td>4</td>
<td>2.7</td>
<td>2.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>46</td>
<td>399</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0D14 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>2.0</td>
<td>5</td>
<td>2.7</td>
<td>2.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>81</td>
<td>283</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0E26 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>1.6</td>
<td>8</td>
<td>2.6</td>
<td>2.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>415</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0D2D 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>1.3</td>
<td>0</td>
<td>2</td>
<td>2.2</td>
<td>2</td>
<td>4.2</td>
<td>2.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>17</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0E28 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1.4</td>
<td>8</td>
<td>2.4</td>
<td>2.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>71</td>
<td>697</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>0D00 3390</td>
<td>RM1LPV</td>
<td>0 0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1.6</td>
<td>2</td>
<td>2.3</td>
<td>2.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>0D10 3390</td>
<td>LMRDB2P</td>
<td>0 2</td>
<td>2.2</td>
<td>0</td>
<td>2</td>
<td>1.6</td>
<td>5</td>
<td>2.3</td>
<td>2.3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>17</td>
<td>420</td>
<td>0</td>
<td>...</td>
</tr>
</tbody>
</table>

OE21 LPLAML and 0E00 LPVM1 a little slow, but I/O rates are so low...worth studying the workload.
FCX111 CACHDBSE for Interesting Volumes

```
FCX111 Run 2008/07/24 12:46:56

CACHDBSE
Load and Performance for Cached Disks

From 2008/07/24 12:54:38                                        To   2008/07/24 14:04:38
For  4200 Secs 01:10:00                Result of 12345 Run

Device            Overall CU-Cache Performance           Split                 DIR ADDR VOLSER   IO/S %READ %RDHIT %WRHIT ICL/S BYP/S   IO/S %READ %RDHIT
------------------- 108 line(s) not displayed --------------------
38  0E00 LPVRM1     .5     4    100     83    .0    .0 .3     4    100 (N)
     .2     0      0 (F)

------------------- 42  line(s) not displayed --------------------
38  0E21 LPLAM1    4.6     0     23     81    .0    .0 3.2     0     7 (N)
     1.4     0     85 (S)
     .0     0      0 (F)

0E21 LPLAM1 is almost all writes (see FCX108), so this data isn't surprising.
```
Networking Performance: FCX161 and FCX215

**FCX161 Run 2008/07/24 12:46:56**

LCHANNEL

Channel Load and Channel Busy Distribution

From 2008/07/24 12:54:38
To 2008/07/24 14:04:38
For 4200 Secs 01:10:00
Result of 12345 Run

<table>
<thead>
<tr>
<th>CHPID</th>
<th>Chan-Group</th>
<th>&lt;%-Busy&gt;</th>
<th>Channel %Busy Distribution 12:54:38-14:04:38 ---&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Hex)</td>
<td>Descr Qual Shrd Cur Ave 0-10 11-20 21-30 31-40 41-50 51-60 61-70 71-80 81-90 91-100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>OSD 00 Yes 15 10 77 23 0 0 0 0 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00</td>
<td>OSD 00 Yes 0 0 100 0 0 0 0 0 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>OSD 00 Yes 0 0 100 0 0 0 0 0 0 0 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>OSD 00 Yes 0 0 100 0 0 0 0 0 0 0 0 0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FCX215 Run 2008/07/24 12:46:56**

FCHANNEL

FICON Channel Load

<table>
<thead>
<tr>
<th>Channel</th>
<th>Bus Work &lt;Data Units&gt; Work &lt;Data Units&gt; &lt;Transfer Rate-&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path</td>
<td>Cycle Units Write Read Units Write Read &lt; (Bytes/s) -&gt;</td>
</tr>
<tr>
<td>ID</td>
<td>Shrd T_BCy T_WUn T_DUW T_DUR L_WUn L_DUW L_DUR Write/s Read/s</td>
</tr>
<tr>
<td>11</td>
<td>Yes 0 10 0 0 0 0 0 0 0 513697 127155</td>
</tr>
<tr>
<td>50</td>
<td>Yes 0 0 0 0 0 0 0 0 0 259435 25383</td>
</tr>
<tr>
<td>40</td>
<td>Yes 0 0 0 0 0 0 0 0 0 250804 25338</td>
</tr>
</tbody>
</table>

< 1 MB/sec altogether
Networking Performance: FCX240 VSWITCH

<table>
<thead>
<tr>
<th>VSWITCH Activity</th>
<th>From 2008/07/24 12:54:38</th>
<th>To 2008/07/24 14:04:38</th>
<th>For 4200 Secs 01:10:00</th>
<th>Result of 12345 Run</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Q Time</th>
<th>Outbound/s</th>
<th>Inbound/s</th>
<th>Signal/s</th>
<th>Intrpts</th>
</tr>
</thead>
<tbody>
<tr>
<td>S Out</td>
<td>Bytes &lt;- Packets- &gt;</td>
<td>Bytes &lt;- Packets-&gt;</td>
<td>&lt;- issued/s -&gt;</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Addr Name</th>
<th>Controlr V</th>
<th>Sec</th>
<th>T_Bit</th>
<th>T_Pack</th>
<th>T_Disc</th>
<th>R_Bit</th>
<th>R_Pack</th>
<th>R_Disc</th>
<th>Write</th>
<th>Read</th>
<th>Sync</th>
<th>Rcv/s</th>
<th>Pro/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; System</td>
<td>8 300</td>
<td>502784</td>
<td>463.6</td>
<td>0</td>
<td>116135</td>
<td>351.0</td>
<td>0</td>
<td>165.0</td>
<td>0</td>
<td>0</td>
<td>244.4</td>
<td>238.8</td>
<td></td>
</tr>
<tr>
<td>02F2</td>
<td>DTCVSW 8</td>
<td>300</td>
<td>502784</td>
<td>463.6</td>
<td>0</td>
<td>116135</td>
<td>351.0</td>
<td>0</td>
<td>165.0</td>
<td>0</td>
<td>0</td>
<td>244.4</td>
<td>238.8</td>
</tr>
</tbody>
</table>
Summary

- There doesn’t seem to be anything wrong with this z/VM
- It’s worth looking inside the Linux guests
- I referred the questioner to Boeblingen
Case Study: Emergency Scan

Revision 2008-07-18 BKW

IBM z/VM Performance Evaluation
Brian Wade bkw@us.ibm.com
Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DBE, e-business logo, ESCO, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, VM/ESA, VSE/ESA, Websphere, xSeries, z/OS, zSeries, zVM

The following are trademarks or registered trademarks of other companies

Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries
LINUX is a registered trademark of Linus Torvalds
UNIX is a registered trademark of The Open Group in the United States and other countries
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.
Intel is a registered trademark of Intel Corporation
* All other products may be trademarks or registered trademarks of their respective companies.

NOTES:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.
Question from Customer

- My system seems to have a high percentage of emergency scan
- Application performance doesn’t seem bothered
- Should I be worried?
Graph from Customer
Finding a Memory Frame

- **Pass 1**: tries to be friendly to dispatched users
  - Unreferenced shared address space pages
  - Long-term-dormant users
  - Eligible-list users
  - Dispatch-list users’ unreferenced pages down to WSS

- **Pass 2**: a little more aggressive... like pass 1 except:
  - Avoids shared address spaces
  - Will take from dispatch-list users down to their SET RESERVE

- **Emergency scan**: anything we can find

- Bit of a misnomer

- Want to know more? Read the prologue of HCPALD
Is Emergency Scan A Sign of Duress?

- Not alone, no.
- Evaluate some other things too.
  - Are free frame lists routinely zero? (FCX254 AVAILLOG)
  - Is system T/V high? (FCX225 SYSSUMLG)
  - Are we spinning significantly on any locks? (FCX265 LOCKLOG)
  - Does USTAT show users in page wait? (FCX114 USTAT)
  - Is an eligible list forming? (FCX100 CPU)
  - Are MDC hits satisfactory? (FCX103 STORAGE, FCX108 DEVICE)
  - Do you have plenty of SXS space? (FCX264 SXSUTIL)
  - Is DASD page rate > XSTORE page rate? (FCX143 PAGELOG)
  - Are there queues at paging DASD? (FCX109 DEVICE CPOWNED)
  - Is paging MLOAD OK? (FCX109 DEVICE CPOWNED)
  - Is paging blocking factor OK? (FCX103 STORAGE)
  - Is paging space too full? (FCX109 DEVICE CPOWNED)
  - Does application performance seem OK? (you tell me)
Storage Management and VDISKs

- Referenced VDISK pages are avoided in Pass 1
- This customer realized he had a lot of VDISK for Linux swap space
- If those VDISK pages are used often, they will tend to stick and be ejectable by only emergency scan
Customer Experiment

- Second system identical to first, except Linux swaps to minidisk instead of VDISK
- Defined two 2 GB VDISKs (half of his central) and began writing to them
- Watched what happened to emergency scan while the VDISKs were around
- Results graphed on next page
Second System, Effect of Experimental VDISKs

Source data: Storage
Summary

- Try to look at system as a whole
- Whether applications seem debilitated is probably the most meaningful indicator
Case Study: Why Doesn’t My System Page Faster?
Revision 2008-07-30 BKW

IBM z/VM Performance Evaluation
Brian Wade bkw@us.ibm.com
Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DBE, e-business logo, ESCO, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, VM/ESA, VSE/ESA, Websphere, xSeries, z/OS, zSeries, zVM

The following are trademarks or registered trademarks of other companies

Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries
LINUX is a registered trademark of Linus Torvalds
UNIX is a registered trademark of The Open Group in the United States and other countries
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.
Intel is a registered trademark of Intel Corporation

* All other products may be trademarks or registered trademarks of their respective companies.

NOTES:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change or withdrawal without notice.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.
Question from Customer

- “z/VM pages extremely slowly”
- Inactive Linux guest is paged in at only about 1000 pages per second
- 12 3390-9 paging packs, 2 LCUs, with 6 FICON chpids
- During busy periods of running 30 guests, he sees 6000 pages per second
- Customer thinks this single guest should page in much faster
- He devised a 300 MB thrasher that reproduced the behavior
- He sent us lots of charts and graphs
- I asked for MONWRITE data
Customer Sent MONWRITE Data

- User LIN102 is running the 300 MB thrasher
- It touched 64,000 pages in 61 seconds (1049/sec)
- The interesting time period is 15:12:30 to 15:13:20
- He used MONITOR SAMPLE 10 SEC (brilliant!)
- I ran his data through PERFKIT BATCH
- I looked at some interesting reports for that period
## User Configuration

**FCX226 UCONF - user configuration report**

<table>
<thead>
<tr>
<th>Userid</th>
<th>SVM</th>
<th>CPUs</th>
<th>Mode</th>
<th>Mode</th>
<th>Relative Absolute Value/ % Share Limit</th>
<th>DSP</th>
<th>Fair</th>
<th>Size Reserved</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIN102</td>
<td>No</td>
<td>1</td>
<td>EME</td>
<td>V=V</td>
<td>100</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

A virtual uniprocessor with one process (thread) running the memory initializer.

**Implications:**

1. Memory initializer will touch pages serially.
2. Page faults will happen serially.
### Activity on Paging DASD

<table>
<thead>
<tr>
<th>Device Descr.</th>
<th>Mdisk Pa-</th>
<th>Rate/s</th>
<th>Time (msec)</th>
<th>Req.</th>
<th>Percent</th>
<th>SEEK</th>
<th>Addr</th>
<th>Type</th>
<th>Label/ID</th>
<th>Links</th>
<th>ths</th>
<th>I/O Avoi d</th>
<th>Pend</th>
<th>Disc</th>
<th>Conn</th>
<th>Serv</th>
<th>Resp</th>
<th>CUM</th>
<th>Qued</th>
<th>Busy</th>
<th>READ</th>
<th>Cy</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>9F11 3390</td>
<td>VSPPG8 CP</td>
<td>0 6  25.5</td>
<td>0 2  3.9  4.1 4.1</td>
<td>0 0</td>
<td>10</td>
<td>0</td>
<td>131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A062 3390</td>
<td>VSPPG5 CP</td>
<td>0 6  25.0</td>
<td>0 2  3.3  3.5 3.5</td>
<td>0 0</td>
<td>9</td>
<td>100</td>
<td>2580</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A02D 3390</td>
<td>VSPPG3 CP</td>
<td>0 6  27.4</td>
<td>0 2  3.1  3.4 3.4</td>
<td>0 0</td>
<td>9</td>
<td>100</td>
<td>505</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9F41 3390</td>
<td>VSPPG8 CP</td>
<td>0 6  29.8</td>
<td>0 2  3.0  3.2 3.2</td>
<td>0 0</td>
<td>10</td>
<td>100</td>
<td>753</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A03D 3390</td>
<td>VSPPG2 CP</td>
<td>0 6  35.4</td>
<td>0 2  2.9  3.1 3.1</td>
<td>0 0</td>
<td>11</td>
<td>100</td>
<td>832</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9F01 3390</td>
<td>VSPPG7 CP</td>
<td>0 6  38.0</td>
<td>0 2  2.8  3.0 3.0</td>
<td>0 0</td>
<td>11</td>
<td>0</td>
<td>1174</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9F5A 3390</td>
<td>VSAPAG CP</td>
<td>0 6  40.9</td>
<td>0 2  2.7  2.9 2.9</td>
<td>0 0</td>
<td>12</td>
<td>100</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A05D 3390</td>
<td>VSPPG6 CP</td>
<td>0 6  38.9</td>
<td>0 2  2.7  2.9 2.9</td>
<td>0 0</td>
<td>11</td>
<td>100</td>
<td>1446</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A01B 3390</td>
<td>VSPPG4 CP</td>
<td>0 6  32.3</td>
<td>0 2  2.5  2.7 2.7</td>
<td>0 0</td>
<td>9</td>
<td>100</td>
<td>2670</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9F21 3390</td>
<td>VSPPG9 CP</td>
<td>0 6  45.6</td>
<td>0 2  2.2  2.4 2.4</td>
<td>0 0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9F51 3390</td>
<td>VSPPGC CP</td>
<td>0 6  48.5</td>
<td>0 2  2.2  2.4 2.4</td>
<td>0 0</td>
<td>12</td>
<td>100</td>
<td>2971</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
<td><strong>387.3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Eleven paging devices:**

1. Each in the neighborhood of 10% busy, all reads
2. Each showing response time of about 3.1 msec
### Who Else is Doing Paging Activity?

<table>
<thead>
<tr>
<th>Spaces</th>
<th>Page Rate</th>
<th>Page</th>
<th>Page Migration</th>
<th>Nr of Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Userid</td>
<td>Owned</td>
<td>Reads</td>
<td>Write</td>
<td>Steals</td>
</tr>
<tr>
<td>&gt;System&lt;</td>
<td>.0</td>
<td>2.3</td>
<td>1.6</td>
<td>7.2</td>
</tr>
</tbody>
</table>

**User Data:**

<table>
<thead>
<tr>
<th>Userid</th>
<th>Owned</th>
<th>Reads</th>
<th>Write</th>
<th>Steals</th>
<th>&gt;2GB</th>
<th>X&gt;M</th>
<th>M6&gt;X</th>
<th>X&gt;DS</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIN102</td>
<td>.0</td>
<td>75.8</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>35.2</td>
<td>4.5</td>
<td>.0</td>
<td></td>
</tr>
</tbody>
</table>

\[44 \times 2.3 = 101 \text{ pages read/sec altogether.}\]

\[\text{LIN102 accounts for 76\% of this, 76 pages read/sec.}\]
What We Know So Far

- Each paging I/O takes about 3.1 msec
- One single-threaded application in one guest is responsible for most of the paging I/Os
- This means we should see about \( \frac{1000}{3.1} = 323 \) SSCH ops for paging per second
- We actually saw 387/sec, but remember other guests are paging slightly
- Because one single-threaded guest is responsible for most of the paging I/O, the paging device utilizations should add to about 100%
- They actually add to 115%, but remember other guests are paging slightly
What Did I Tell The Customer?

- LIN102’s page reading speed is limited by its single-threaded nature and the speed of the paging DASD.

- Your system pages at higher rates when 30 guests are running because with multiple guests you can generate concurrent page reads. You have multiple paging exposures too and so you can parallelize paging I/O.

- Your 11 paging exposures look like they could support \((1100\%/115\%) = 9.5\) such thrashers concurrently.

- But from FCX109 DEVICE CPOWNED, I see your page space is about 15% full so I wouldn’t try more than four of them at once.
## Something Interesting About LIN102

FCX163  Run 2008/05/19 12:18:57  UPAGELOG LI N102

User Paging Activity

---

From 2008/05/15 15:10:10
To 2008/05/15 15:15:50
For 340 Secs 00:05:40

---

Page Data Log for User LIN102

<table>
<thead>
<tr>
<th>Interval</th>
<th>Spaces</th>
<th>&lt;Page Rate&gt;</th>
<th>Page</th>
<th>&lt;Page Migration&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>End Time</td>
<td>Owned</td>
<td>Reads</td>
<td>Write</td>
<td>Steals</td>
</tr>
<tr>
<td>15:12:40</td>
<td>0</td>
<td>437</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15:12:50</td>
<td>0</td>
<td>534</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15:13:00</td>
<td>0</td>
<td>440</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15:13:10</td>
<td>0</td>
<td>313</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15:13:20</td>
<td>0</td>
<td>473</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Avg** 439 232

Thrasher touched 1049/sec altogether.

1. 439/sec read from disk
2. 232/sec read from XSTORE
3. 378/sec resident
A Note on User States

FCX164  Run  2008/05/19 12:18:57         USTATLOG LI N102

From 2008/05/15 15:10:10                                 To   2008/05/15 15:15:50
For    340 Secs 00:05:40

________________________________________________________

Wait State Data Log for User LI N102

<table>
<thead>
<tr>
<th>Interval</th>
<th>%ACT</th>
<th>%RUN</th>
<th>%CPU</th>
<th>%LDG</th>
<th>%PGW</th>
<th>%OW</th>
<th>%M</th>
<th>%TI</th>
<th>%CF</th>
<th>15:12:30</th>
<th>15:12:40</th>
<th>15:12:50</th>
<th>15:13:00</th>
<th>15:13:10</th>
<th>15:13:20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15:12:30</td>
<td>15:12:40</td>
<td>15:12:50</td>
<td>15:13:00</td>
<td>15:13:10</td>
<td>15:13:20</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15:12:40</td>
<td>15:12:50</td>
<td>15:12:50</td>
<td>15:13:00</td>
<td>15:13:10</td>
<td>15:13:20</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15:12:50</td>
<td>15:12:50</td>
<td>15:12:50</td>
<td>15:13:00</td>
<td>15:13:10</td>
<td>15:13:20</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15:13:00</td>
<td>15:13:00</td>
<td>15:13:00</td>
<td>15:13:10</td>
<td>15:13:10</td>
<td>15:13:20</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15:13:10</td>
<td>15:13:10</td>
<td>15:13:10</td>
<td>15:13:10</td>
<td>15:13:10</td>
<td>15:13:20</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15:13:20</td>
<td>15:13:20</td>
<td>15:13:20</td>
<td>15:13:20</td>
<td>15:13:20</td>
<td>15:13:20</td>
</tr>
</tbody>
</table>

Customer said this means LI N102 “is in page wait 100% of the time”. This is not correct.

It means 100% of the times we looked, LI N102 was in a page wait. We looked only once every two seconds (FCX149 MONSET).

After all, LI N102 was also touching pages.
Summary

- Customer became absorbed with z/VM measurements but forgot what his workload does.
- Knowledge of the workload’s behavior is crucial in understanding why the system performs the way it does.
- Customer was very good at collecting raw monitor data appropriate for the diagnosis task at hand.
- Fun question that was not too difficult to answer.
Case Study: Undersized LPAR

Revision 2008-07-28 BKW

IBM z/VM Performance Evaluation
Virg Meredith  vmeredit@us.ibm.com
Question from Customer

- Why do my workloads run so slowly?
### Customer’s Configuration

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>System model</strong></td>
<td>2094-606</td>
</tr>
<tr>
<td><strong>Processors</strong></td>
<td>2 IFL</td>
</tr>
<tr>
<td><strong>SYSGEN storage</strong></td>
<td>19968 (19.5 GB)</td>
</tr>
<tr>
<td><strong>XSTORE</strong></td>
<td>4096 (4 GB)</td>
</tr>
<tr>
<td><strong>Page slots</strong></td>
<td>24641k (94 GB)</td>
</tr>
<tr>
<td><strong>Paging devices</strong></td>
<td>25</td>
</tr>
<tr>
<td><strong>Logged-on virtual</strong></td>
<td>68 GB</td>
</tr>
</tbody>
</table>
What We Saw in Customer’s Data

- Long queues and long response times for paging devices
- Possibility for processor contention during peak hours
BEFORE: Customer’s FCX109 DEVICE CPOWNED

<table>
<thead>
<tr>
<th>Device Name</th>
<th>Type</th>
<th>Extent</th>
<th>% P-Rds</th>
<th>P-Wrt</th>
<th>S-Rds</th>
<th>S-Wrt</th>
<th>Total</th>
<th>%RSCH feres</th>
<th>Lngth / Page</th>
<th>Time</th>
<th>Resp</th>
<th>Page</th>
<th>Size</th>
<th>Alloc</th>
</tr>
</thead>
<tbody>
<tr>
<td>7904 3390</td>
<td>520PG7</td>
<td>PAGE</td>
<td>0-3338</td>
<td>57</td>
<td>21.7</td>
<td>16.9</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>791B 3390</td>
<td>520PG4</td>
<td>PAGE</td>
<td>0-3338</td>
<td>59</td>
<td>23.0</td>
<td>18.0</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>7921 3390</td>
<td>520PGL4</td>
<td>PAGE</td>
<td>0-3338</td>
<td>59</td>
<td>22.7</td>
<td>17.7</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>7922 3390</td>
<td>520PGL5</td>
<td>PAGE</td>
<td>0-3338</td>
<td>60</td>
<td>23.1</td>
<td>18.0</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>792C 3390</td>
<td>520PGL4</td>
<td>PAGE</td>
<td>803060</td>
<td>22</td>
<td>25.7</td>
<td>19.6</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>792D 3390</td>
<td>520PGLC</td>
<td>PAGE</td>
<td>803060</td>
<td>22</td>
<td>25.5</td>
<td>19.7</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>7934 3390</td>
<td>520PGB</td>
<td>PAGE</td>
<td>803060</td>
<td>22</td>
<td>25.5</td>
<td>19.6</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>79FC 3390</td>
<td>520PGL0</td>
<td>PAGE</td>
<td>0-3338</td>
<td>56</td>
<td>21.5</td>
<td>16.8</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
### BEFORE: Customer’s FCX225 SYSSUMLG

**System Performance Summary by Time**

From 2008/04/30 09:50:08 to 2008/04/30 23:53:33

For 50604 Secs 14:03:24

<table>
<thead>
<tr>
<th>Interval</th>
<th>Pct</th>
<th>Cap-</th>
<th>On-</th>
<th>Pct Log-</th>
<th>+RSCH Resp in PGI N+ Read+ Pages</th>
<th>Non- Quick</th>
<th>Non- Quick</th>
</tr>
</thead>
<tbody>
<tr>
<td>End Time</td>
<td>Busy</td>
<td>T/V</td>
<td>ture line</td>
<td>Busy</td>
<td>ged Activ</td>
<td>/s</td>
<td>msec</td>
</tr>
<tr>
<td>&gt;&gt;Mean&lt;&lt;</td>
<td>46.7</td>
<td>1.25</td>
<td>.8177</td>
<td>2.0</td>
<td>....</td>
<td>31</td>
<td>26</td>
</tr>
<tr>
<td>10:00:08</td>
<td>71.3</td>
<td>1.45</td>
<td>.7064</td>
<td>2.0</td>
<td>....</td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td>10:10:08</td>
<td>68.0</td>
<td>1.38</td>
<td>.7363</td>
<td>2.0</td>
<td>....</td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td>10:20:08</td>
<td>71.6</td>
<td>1.25</td>
<td>.8138</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>26</td>
</tr>
<tr>
<td>10:30:08</td>
<td>37.4</td>
<td>1.17</td>
<td>.8732</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>10:40:08</td>
<td>28.6</td>
<td>1.28</td>
<td>.7995</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>10:50:08</td>
<td>28.0</td>
<td>1.25</td>
<td>.8183</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>11:00:08</td>
<td>32.6</td>
<td>1.21</td>
<td>.8433</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>27</td>
</tr>
<tr>
<td>11:10:08</td>
<td>31.6</td>
<td>1.22</td>
<td>.8385</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>11:20:08</td>
<td>39.4</td>
<td>1.26</td>
<td>.8079</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>11:30:08</td>
<td>37.0</td>
<td>1.18</td>
<td>.8603</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>11:40:08</td>
<td>54.7</td>
<td>1.16</td>
<td>.8757</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>11:50:08</td>
<td>57.1</td>
<td>1.22</td>
<td>.8318</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>12:00:08</td>
<td>47.7</td>
<td>1.37</td>
<td>.7433</td>
<td>2.0</td>
<td>....</td>
<td>30</td>
<td>26</td>
</tr>
</tbody>
</table>
BEFORE: Customer’s FCX126 LPAR

<table>
<thead>
<tr>
<th>LPAR Data, Collected in Partition VMLNX1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor type and model                : 2094-606</td>
</tr>
<tr>
<td>Nr. of configured partitions:           : 7</td>
</tr>
<tr>
<td>Nr. of physical processors:             : 9</td>
</tr>
<tr>
<td>Dispatch interval (msec)                : dynamic</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Partition Nr.</th>
<th>Upld</th>
<th>#Proc</th>
<th>Weight</th>
<th>Vlt-C</th>
<th>Cap</th>
<th>%Load</th>
<th>CPU</th>
<th>%Busy</th>
<th>%Hd</th>
<th>%Susp</th>
<th>%Vmd</th>
<th>%Logd</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF01A</td>
<td>1</td>
<td>01</td>
<td>1</td>
<td>900</td>
<td>NO</td>
<td>NO</td>
<td>0</td>
<td>98.7</td>
<td>98.7</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>ICF</td>
</tr>
<tr>
<td>CF01B</td>
<td>2</td>
<td>02</td>
<td>1</td>
<td>100</td>
<td>NO</td>
<td>NO</td>
<td>0</td>
<td>.7</td>
<td>.7</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>ICF</td>
</tr>
<tr>
<td>... CPs removed for readability ...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VMLNX1</td>
<td>6</td>
<td>13</td>
<td>2</td>
<td>500</td>
<td>NO</td>
<td>NO</td>
<td>0</td>
<td>47.2</td>
<td>47.2</td>
<td>1.9</td>
<td>46.7</td>
<td>47.6</td>
<td>IFL</td>
</tr>
<tr>
<td>VMLNX3</td>
<td>7</td>
<td>14</td>
<td>2</td>
<td>500</td>
<td>NO</td>
<td>NO</td>
<td>0</td>
<td>1.4</td>
<td>1.4</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>IFL</td>
</tr>
</tbody>
</table>

General LPAR mgmt overhead 0.3
Overall physical load 51.5

If VMLNX3 ever gets hungry, there is going to be a CPU constraint.
What We Recommended The Customer Change

- **More central storage**
  - Calculated new amount based on:
    - User pages resident on DASD (“before” FCX113 UPAGE report)
    - Understanding that memory comes in 32 GB increments

- **More and faster paging devices**
  - One would think if we were adding storage to reduce paging, we wouldn’t have to tinker with the paging configuration too
  - However, we knew the customer wanted to grow his logged-on virtual
  - Also, customer reconfigured his [non-IBM] DASD to improve I/O response time

- **1 additional processor**
  - Seems sufficient based on old FCX225 SYSSUMLG report
## Customer’s New Configuration

<table>
<thead>
<tr>
<th></th>
<th>Old</th>
<th>New</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>System model</td>
<td>2094-606</td>
<td>2094-705</td>
<td>-</td>
</tr>
<tr>
<td>Processors</td>
<td>2</td>
<td>3</td>
<td>+50%</td>
</tr>
<tr>
<td>SYSGEN storage</td>
<td>19968 (19.5 GB)</td>
<td>52736 (52 GB)</td>
<td>+267%</td>
</tr>
<tr>
<td>XSTORE</td>
<td>4096 (4 GB)</td>
<td>4096 (4 GB)</td>
<td>0</td>
</tr>
<tr>
<td>Page slots</td>
<td>24641k (94 GB)</td>
<td>72121k (275 GB)</td>
<td>+293%</td>
</tr>
<tr>
<td>Paging devices</td>
<td>25</td>
<td>40</td>
<td>+160%</td>
</tr>
<tr>
<td>Logged-on virtual</td>
<td>68 GB</td>
<td>87 GB</td>
<td>+28%</td>
</tr>
</tbody>
</table>
Measurement After The Change

- 60% reduction in user pages on DASD (FCX113 UPAGE)
- No queuing for paging devices (FCX109 DEVICE COWNED)
- No more user page waits (FCX114 USTAT)
- No processor constraint during the peak hour (FCX225 SYSSUMLG)
### Page / SPOOL Allocation Summary

<table>
<thead>
<tr>
<th>Page slot available</th>
<th>72121k</th>
<th>SPOOL slots available</th>
<th>600840</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAGE slot utilization</td>
<td>4%</td>
<td>SPOOL slot utilization</td>
<td>54%</td>
</tr>
<tr>
<td>T-Disk cylinders avail.</td>
<td>......</td>
<td>DUMP slots available</td>
<td>0</td>
</tr>
<tr>
<td>T-Disk space utilization</td>
<td>...%</td>
<td>DUMP slot utilization</td>
<td>..%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>&lt; Device Descr. -&gt;</th>
<th>&lt;------------- Rate/s -------------&gt;</th>
<th>User</th>
<th>Serv</th>
<th>MLOAD</th>
<th>Block %Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume Area</td>
<td>Area Used</td>
<td>SSCH</td>
<td>Inter</td>
<td>Queue</td>
<td>Time</td>
</tr>
<tr>
<td>Addr</td>
<td>Devtyp</td>
<td>Serial</td>
<td>Type</td>
<td>Extent</td>
<td>Used</td>
</tr>
<tr>
<td>790D 3390</td>
<td>520SPL</td>
<td>SPOOL</td>
<td>1-3338</td>
<td>54</td>
<td>.0</td>
</tr>
<tr>
<td>7957 3390</td>
<td>520PAG</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
<tr>
<td>7958 3390</td>
<td>520PG3</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
<tr>
<td>7959 3390</td>
<td>520PG5</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
<tr>
<td>795A 3390</td>
<td>520PG7</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
<tr>
<td>795B 3390</td>
<td>520PG9</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
<tr>
<td>795C 3390</td>
<td>520PGB</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
<tr>
<td>795D 3390</td>
<td>520PGD</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
<tr>
<td>795E 3390</td>
<td>520PGF</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
<tr>
<td>795F 3390</td>
<td>52PG1</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
<tr>
<td>7960 3390</td>
<td>52PG13</td>
<td>PAGE</td>
<td>803060</td>
<td>4</td>
<td>.1</td>
</tr>
</tbody>
</table>
**AFTER: Customer’s FCX225 SYSSUMSLG**

```
<table>
<thead>
<tr>
<th>Interval</th>
<th>Pct</th>
<th>Cap- On-</th>
<th>Pct Log-</th>
<th>+RSCH Resp</th>
<th>i n PQI N+ Read+ Pages</th>
<th>Non- Quick</th>
<th>Non- Quick End Time</th>
<th>Busy</th>
<th>T/V</th>
<th>Triv</th>
<th>Triv</th>
<th>Triv</th>
<th>Triv</th>
<th>Triv</th>
<th>Triv</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:45:25</td>
<td>14.0</td>
<td>1.05</td>
<td>.9836</td>
<td>3.0</td>
<td>.0</td>
<td>1.8</td>
<td>2.3</td>
<td>.0</td>
<td>837</td>
<td>2.980</td>
<td>.963</td>
<td>3.78</td>
<td>1.71</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td>11:55:25</td>
<td>14.5</td>
<td>1.05</td>
<td>.9838</td>
<td>3.0</td>
<td>.0</td>
<td>1.2</td>
<td>3.7</td>
<td>.0</td>
<td>821</td>
<td>2.878</td>
<td>1.018</td>
<td>3.88</td>
<td>1.71</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td>12:05:25</td>
<td>13.5</td>
<td>1.05</td>
<td>.9853</td>
<td>3.0</td>
<td>.0</td>
<td>6.3</td>
<td>7.1</td>
<td>.0</td>
<td>928</td>
<td>2.867</td>
<td>.975</td>
<td>3.71</td>
<td>1.80</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>12:25:25</td>
<td>14.8</td>
<td>1.06</td>
<td>.9822</td>
<td>3.0</td>
<td>.0</td>
<td>3.0</td>
<td>4.0</td>
<td>.0</td>
<td>730</td>
<td>2.779</td>
<td>.943</td>
<td>3.83</td>
<td>1.79</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>12:35:25</td>
<td>13.6</td>
<td>1.05</td>
<td>.9830</td>
<td>3.0</td>
<td>.0</td>
<td>4.3</td>
<td>4.9</td>
<td>.0</td>
<td>852</td>
<td>3.054</td>
<td>.953</td>
<td>3.71</td>
<td>1.64</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>12:45:25</td>
<td>14.1</td>
<td>1.05</td>
<td>.9823</td>
<td>3.0</td>
<td>.0</td>
<td>2.7</td>
<td>3.8</td>
<td>.0</td>
<td>817</td>
<td>2.687</td>
<td>.887</td>
<td>3.99</td>
<td>1.77</td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td>12:55:25</td>
<td>13.0</td>
<td>1.05</td>
<td>.9825</td>
<td>3.0</td>
<td>.0</td>
<td>4.5</td>
<td>5.5</td>
<td>.0</td>
<td>783</td>
<td>2.895</td>
<td>.900</td>
<td>3.76</td>
<td>1.76</td>
<td>.74</td>
<td></td>
</tr>
<tr>
<td>13:05:25</td>
<td>13.5</td>
<td>1.05</td>
<td>.9820</td>
<td>3.0</td>
<td>.0</td>
<td>4.5</td>
<td>5.5</td>
<td>.0</td>
<td>783</td>
<td>2.536</td>
<td>.923</td>
<td>4.10</td>
<td>1.81</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>13:15:25</td>
<td>13.5</td>
<td>1.05</td>
<td>.9832</td>
<td>3.0</td>
<td>.0</td>
<td>4.5</td>
<td>5.5</td>
<td>.0</td>
<td>783</td>
<td>2.779</td>
<td>.943</td>
<td>3.83</td>
<td>1.79</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>13:25:25</td>
<td>13.6</td>
<td>1.06</td>
<td>.9826</td>
<td>3.0</td>
<td>.0</td>
<td>4.5</td>
<td>5.5</td>
<td>.0</td>
<td>756</td>
<td>2.832</td>
<td>.997</td>
<td>3.88</td>
<td>1.78</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td>13:35:25</td>
<td>13.6</td>
<td>1.06</td>
<td>.9826</td>
<td>3.0</td>
<td>.0</td>
<td>4.5</td>
<td>5.5</td>
<td>.0</td>
<td>771</td>
<td>2.676</td>
<td>.990</td>
<td>3.77</td>
<td>1.88</td>
<td>.74</td>
<td></td>
</tr>
<tr>
<td>13:45:25</td>
<td>13.7</td>
<td>1.05</td>
<td>.9829</td>
<td>3.0</td>
<td>.0</td>
<td>4.5</td>
<td>5.5</td>
<td>.0</td>
<td>793</td>
<td>2.975</td>
<td>.929</td>
<td>3.79</td>
<td>1.71</td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td>13:55:25</td>
<td>14.6</td>
<td>1.05</td>
<td>.9838</td>
<td>3.0</td>
<td>.0</td>
<td>4.5</td>
<td>5.5</td>
<td>.0</td>
<td>749</td>
<td>2.733</td>
<td>.964</td>
<td>3.63</td>
<td>1.85</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td>14:05:25</td>
<td>13.8</td>
<td>1.06</td>
<td>.9830</td>
<td>3.0</td>
<td>.0</td>
<td>4.5</td>
<td>5.5</td>
<td>.0</td>
<td>766</td>
<td>2.921</td>
<td>.887</td>
<td>3.87</td>
<td>1.78</td>
<td>.80</td>
<td></td>
</tr>
</tbody>
</table>
```
AFTER: Customer's FCX126 LPAR

FCX126  Run 2008/06/16 13:51:28    LPAR
Logical Partition Activity

From 2008/06/13 11:35:25
To   2008/06/13 15:35:25
For  14400 Secs 04:00:00                  "This is a performance report for SYSTEM XYZ"

______________________________________________________________

LPAR Data, Collected in Partition VMLNX1

Processor type and model    : 2094-705
Nr. of configured partitions:     7
Nr. of physical processors  :       9
Dispatch interval (msec)    : dynamic

Partition Nr. Upid #Proc Weight Wait-C Cap %Load CPU %Busy %Ovhd %Susp %VMld %Logld Type
CF01A      1    01     1    900     NO  NO ...   0  98.2  98.2 ...   ...    ... ICF
CF01B      2    02     1    100     NO  NO ...   0   1.1   1.1 ...   ...    ... ICF
...CPs deleted for readability ...
VMLNX1     6    13     3    500     NO  NO ...   0  14.3  14.3 .4  14.0   14.1 IFL
VMLNX3     7    14     3    500     NO  NO ...   0  2.9   2.9 ...   ...    ... IFL

General LPAR mgmt overhead                     .4
Overall physical load                        41.0
Summary

- System was running “as it was designed to run” before the changes
- Workloads ran as expected after the changes
PAV and MDC
Case Study

Brian Wade/Endicott/IBM
May 23, 2007
Customer Called IBM

- I have disk I/O problems on z/VM 5.2
- When I turn on MDC, my system slows down
- OK, nobody panic or speculate
- Send us some raw monitor data...
  - For MDC off, when things are good
  - For MDC on, when things are not so good
- Customer sent two very descriptive sets of data
- So good, in fact, that we easily replicated the customer's problem on GDLSPRF3
All five minidisks are on the same RDEV.
Without MDC

z/VM 5.2, MDC OFF, without the fix, excerpt from typical FCX108 (DEVICE) report

<table>
<thead>
<tr>
<th>Addr</th>
<th>Type</th>
<th>Label/ID</th>
<th>Links</th>
<th>ths</th>
<th>I/O Avoid</th>
<th>Pend</th>
<th>Disc</th>
<th>Conn</th>
<th>Serv</th>
<th>Resp</th>
<th>CUWt</th>
<th>Qued</th>
</tr>
</thead>
<tbody>
<tr>
<td>E700</td>
<td>3390</td>
<td>LDB307</td>
<td>15</td>
<td>4</td>
<td>564</td>
<td>.0</td>
<td>.3</td>
<td>.2</td>
<td>1.2</td>
<td>1.7</td>
<td>.0</td>
<td>.0</td>
</tr>
<tr>
<td>E7FC -&gt;E700</td>
<td>LDB307</td>
<td>15</td>
<td>4</td>
<td>543</td>
<td>.0</td>
<td>.3</td>
<td>.2</td>
<td>1.2</td>
<td>1.7</td>
<td>1.7</td>
<td>.0</td>
<td>.0</td>
</tr>
<tr>
<td>E7FD -&gt;E700</td>
<td>LDB307</td>
<td>15</td>
<td>4</td>
<td>541</td>
<td>.0</td>
<td>.4</td>
<td>.2</td>
<td>1.2</td>
<td>1.8</td>
<td>1.8</td>
<td>.0</td>
<td>.0</td>
</tr>
<tr>
<td>E7FE -&gt;E700</td>
<td>LDB307</td>
<td>15</td>
<td>4</td>
<td>539</td>
<td>.0</td>
<td>.4</td>
<td>.2</td>
<td>1.2</td>
<td>1.8</td>
<td>1.8</td>
<td>.0</td>
<td>.0</td>
</tr>
</tbody>
</table>

Things to notice:
- E700 with three PAV aliases
- Four users doing I/O to the volume
- Aggregate volume I/O rate is (564+543+541+539) = 2187 IOs/sec

From this report, how do we know...
- PAV is correctly configured for the volume?
- PAV is functioning correctly?
- MDC is turned off for this volume?
But With MDC...

z/VM 5.2, MDC ON, without the fix

<-- Device Descr. -->  Mdisk Pa- <-Rate/s-> <-------- Time (msec) --------> Req.
Addr  Type  Label/ID  Links ths  I/O  Avoid  Pend  Disc  Conn  Serv  Resp  CUWt  Qued
E700  3390  LDB307    15   4   564  90.6   .3   1.3  1.7  1.9   .0  2.9
E7FC ->E700 LDB307   15   4   .5   .0   .3   1.5  2.1  2.3   .0   .0
E7FD ->E700 LDB307   15   4   .3   .0   .3   1.4  2.0  2.2   .0   .0
E7FE ->E700 LDB307   15   4   .3   .0   .3   1.3  2.2  2.4   .0   .0

Things to ask ourselves:
- Is MDC really on for this volume?
- What is the MDC-on volume I/O rate? \((564+90.6+.5+.3+.3) = 655.7\) IOs/sec
- Is it correct that the volume I/O rate should go down with MDC on?
- Is the CP I/O subsystem functioning correctly? Why or why not?
Things That Look Suspicious

- The volume I/O rate should not go down substantially when MDC comes online.
  - If some other limit is holding the applications back, the volume I/O rate should stay about the same.
  - If nothing else is holding the applications back, the volume I/O rate should increase.

- The CP I/O subsystem appears not to be functioning correctly.
  - I/Os are happening almost exclusively on the base RDEV.
  - The aliases appear to be doing almost no work.
  - There is queueing at the base RDEV.
A Visit to Development

- I went to see Bill Stephens (virtual I/O and MDC expert)

- He felt MDC's I/Os should be PAV-able

- But investigation revealed...
  - There are bits CP sets for its own I/Os to tell the real I/O layer whether to try to PAV the I/O...
  - but MDC was forgetting to set these bits...
  - (in fact, nowhere did CP ever set those bits!)
  - thus all I/Os originating in MDC were being forced to the base...
  - thus MDC was failing to exploit the volume's PAV capability

- VM64199 repaired CP MDC so that its I/Os are PAV-able
MDC ON, With The Fix

z/VM 5.2, MDC ON, with the fix

Things to notice:
- MDC is functioning (there are avoided I/Os)
- Aggregate I/O rate is \( (442+402+421+415+410) = 2090 \) I/Os/sec
- About one-fifth of the I/Os are being avoided... makes sense
- Connect time is up compared to MDC off (1.2 to 1.5) -- IOs are bigger
- I/Os are spreading across base and aliases
- No queueing at the base device
It Turns Out...

- Our System Test group saw this behavior too, during z/VM 5.2 test
- They thought it was expected that the I/O rate would go down when MDC was ON, and we can't entirely blame them
- The only clue anything is wrong is that there is a queue at the base RDEV -- I doubt System Test would see that
- All of our PAV measurements were done with MDC OFF, of course
Status

- APAR VM64199, UM32047 (z/VM 5.2), UM32048 (z/VM 5.3)
- No word yet on whether it will be HIPER (thus RSU) for z/VM 5.2
- Will be on the GA RSU for z/VM 5.3
- Our golden driver U0PK2518 **does** have this fix
A Case Study
Customer Calls In

- My system isn't running fast, but it isn't paging either
- My application formats lots of VDISKs... aren't they in memory? Shouldn't this be fast?
- I have raw monitor data... will you take a look?
- Customer sent raw monitor file 20070501 MD111606
- He says his workload uses disk volumes 1240-59 and 16C0-E3
- I took a look-see
Basic System Summary

Look at those T/V ratios! What is CP doing?
Think About the Application

- Customer says he is formatting VDISKs
- VDISKs are address spaces
- We page them when storage gets tight
- We do seem to be spending a lot of time in CP
- Let's see if DEVICE CPOWNED shows us anything
From 11:16 to 12:37 the paging devices have queues on average? Let's look at some INTERIM reports and see what we see...
Look at that pending time on the paging volumes!
High pending time usually means channel contention...
Configuration

From FCX131 DEVCONF:

1240-1259  0008-0021  3390-3 (E)  67 69 . . . . . .  2105-E8  Online
16C0-16E3  0050-0073  3390-3 (E)  67 69 . . . . . .  2105-E8  Online

From FCX161 LCHANNEL:

67  ESCON  00 Yes  15 6 93 7
69  ESCON  00 Yes  16 10 73 27

Two ESCON chpids for all this paging DASD? I don't think so...
Recommendation

- Customer added four ESCON chpids
- (Why didn't he add FICON? Who knows...)
- He was quiet for a while, and then...
He's Baa-aaack

Load and Performance of CP Owned Disks

Page / SPOOL Allocation Summary

<table>
<thead>
<tr>
<th>PAGE slots available</th>
<th>SPOOL slots available</th>
</tr>
</thead>
<tbody>
<tr>
<td>51540k</td>
<td>4257606</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>T-Disk cylinders avail.</th>
<th>DUMP slots available</th>
<th>DUMP slot utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>0</td>
<td>..%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Volume Area</th>
<th>Area Used</th>
<th>Rate/s</th>
<th>User</th>
<th>Serv</th>
<th>MLOAD</th>
<th>Block % Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addr</td>
<td>Devtp</td>
<td>Serial</td>
<td>Type</td>
<td>Extent</td>
<td>% P-Rds</td>
<td>P-Wt</td>
</tr>
<tr>
<td>16D5 3390</td>
<td>XXPG0A</td>
<td>PAGE</td>
<td>0-3338</td>
<td>88</td>
<td>21.7</td>
<td>19.1</td>
</tr>
<tr>
<td>16D6 3390</td>
<td>XXPG0B</td>
<td>PAGE</td>
<td>0-3338</td>
<td>88</td>
<td>20.5</td>
<td>17.2</td>
</tr>
<tr>
<td>16D8 3390</td>
<td>XXPG0C</td>
<td>PAGE</td>
<td>0-3338</td>
<td>88</td>
<td>22.7</td>
<td>18.1</td>
</tr>
<tr>
<td>16D9 3390</td>
<td>XXPG0D</td>
<td>PAGE</td>
<td>0-3338</td>
<td>87</td>
<td>21.1</td>
<td>18.5</td>
</tr>
<tr>
<td>16DB 3390</td>
<td>XXPG0E</td>
<td>PAGE</td>
<td>0-3338</td>
<td>87</td>
<td>22.3</td>
<td>20.0</td>
</tr>
<tr>
<td>16DC 3390</td>
<td>XXPG0F</td>
<td>PAGE</td>
<td>0-3338</td>
<td>86</td>
<td>21.9</td>
<td>17.7</td>
</tr>
<tr>
<td>16DD 3390</td>
<td>XXPG0G</td>
<td>PAGE</td>
<td>0-3338</td>
<td>86</td>
<td>22.0</td>
<td>18.2</td>
</tr>
<tr>
<td>16DE 3390</td>
<td>XXPG0H</td>
<td>PAGE</td>
<td>0-3338</td>
<td>86</td>
<td>21.4</td>
<td>19.6</td>
</tr>
<tr>
<td>16DF 3390</td>
<td>XXPG0I</td>
<td>PAGE</td>
<td>0-3338</td>
<td>84</td>
<td>22.1</td>
<td>19.6</td>
</tr>
<tr>
<td>16EO 3390</td>
<td>XXPG0J</td>
<td>PAGE</td>
<td>0-3338</td>
<td>83</td>
<td>20.4</td>
<td>17.6</td>
</tr>
<tr>
<td>5805 3390</td>
<td>CF5805</td>
<td>PAGE</td>
<td>810000</td>
<td>12</td>
<td>46.5</td>
<td>41.9</td>
</tr>
<tr>
<td>9F23 3390</td>
<td>XPC2</td>
<td>PAGE</td>
<td>0-3338</td>
<td>99</td>
<td>18.6</td>
<td>18.1</td>
</tr>
<tr>
<td>9F24 3390</td>
<td>XPG3</td>
<td>PAGE</td>
<td>0-3338</td>
<td>99</td>
<td>19.2</td>
<td>17.5</td>
</tr>
<tr>
<td>9F25 3390</td>
<td>XPG4</td>
<td>PAGE</td>
<td>0-3338</td>
<td>99</td>
<td>18.6</td>
<td>17.4</td>
</tr>
<tr>
<td>9F26 3390</td>
<td>XPG5</td>
<td>PAGE</td>
<td>0-3338</td>
<td>99</td>
<td>20.9</td>
<td>17.9</td>
</tr>
<tr>
<td>C09E 3390</td>
<td>PC09B</td>
<td>PAGE</td>
<td>0-3338</td>
<td>100</td>
<td>22.4</td>
<td>19.2</td>
</tr>
<tr>
<td>D007 3390</td>
<td>CF007</td>
<td>PAGE</td>
<td>896800</td>
<td>17</td>
<td>46.1</td>
<td>40.7</td>
</tr>
<tr>
<td>D008 3390</td>
<td>CF008</td>
<td>PAGE</td>
<td>896800</td>
<td>17</td>
<td>42.2</td>
<td>39.7</td>
</tr>
<tr>
<td>D00D 3390</td>
<td>JSPG04</td>
<td>PAGE</td>
<td>896800</td>
<td>20</td>
<td>42.9</td>
<td>39.0</td>
</tr>
</tbody>
</table>

I removed 25 100% full 3990-3's from this excerpt!
So What's His Problem

- 40 3390-3 paging volumes nearly full
- 4 3390-9 paging volumes have the free space
- We can do only one I/O at a time to those gigantic model 9's
- Get rid of those mod 9's and add a lot of mod 3's
- He's working on it
A Case Study
The Grinch That Stole Performance

From Performance Toolkit USTAT FCX114 Report January 5:

<table>
<thead>
<tr>
<th>%CPU</th>
<th>%LDG</th>
<th>%PGW</th>
<th>%IOW</th>
<th>%SIM</th>
<th>%TIW</th>
<th>%CFW</th>
<th>%TI</th>
<th>%EL</th>
<th>%DM</th>
<th>%IOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>2</td>
<td>10</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>51</td>
<td>8</td>
</tr>
</tbody>
</table>

From Performance Toolkit DEVICE FCX108 Report January 5:

<table>
<thead>
<tr>
<th>Addr</th>
<th>I/O Avoid</th>
<th>Pend</th>
<th>Disc</th>
<th>Conn</th>
<th>Serv</th>
<th>Resp</th>
<th>Resp</th>
<th>CUWt</th>
<th>Qued</th>
<th>Busy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1742</td>
<td>26.7</td>
<td>0</td>
<td>1.3</td>
<td>18.4</td>
<td>4.7</td>
<td>24.5</td>
<td>69.0</td>
<td>.0</td>
<td>1.2</td>
<td>65.4</td>
</tr>
</tbody>
</table>

Went to check Toolkit CACHEXT FCX177 Report for control unit cache stats, but it didn’t exist!
It is a good thing I keep historical data -- let's go back and see what's going on...
When Did We Last See Cache?

From Performance Toolkit DEVICE FCX108 Report:

<table>
<thead>
<tr>
<th>Addr</th>
<th>I/O Avoid</th>
<th>Pend</th>
<th>Disc</th>
<th>Conn</th>
<th>Serv</th>
<th>Resp</th>
<th>CUWt</th>
<th>Qued</th>
<th>Busy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec8</td>
<td>41.0</td>
<td>0.0</td>
<td>0.3</td>
<td>0.2</td>
<td>2.0</td>
<td>2.6</td>
<td>2.9</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Jan5</td>
<td>26.7</td>
<td>0.0</td>
<td>1.3</td>
<td>18.4</td>
<td>4.7</td>
<td>24.5</td>
<td>69.0</td>
<td>0.0</td>
<td>1.2</td>
</tr>
</tbody>
</table>

From Performance Toolkit CACHEXT FCX177 Dec. 8th Report:

<table>
<thead>
<tr>
<th>Cache SCMBK N-Seq</th>
<th>Seq</th>
<th>FW</th>
<th>Read</th>
<th>Tot</th>
<th>RdHt</th>
<th>Wrt</th>
<th>DFW</th>
<th>CFW</th>
</tr>
</thead>
<tbody>
<tr>
<td>53.0</td>
<td>41.0</td>
<td>52.3</td>
<td>0</td>
<td>0.6</td>
<td>99</td>
<td>99</td>
<td>96</td>
<td>96</td>
</tr>
</tbody>
</table>
q dasd details 1742

1742 CUTYPE = 3990-EC, DEVTYPE = 3390-06, VOLSER=USE001

CACHE DETAILS: CACHE NVS CFW DFW PINNED CONCOPY
-Subsystem F Y Y - Y N
-Device Y - - Y N N

DEVICE DETAILS: CCA = 02, DDC = 02

DUPLEX DETAILS: SIMPLEX

Pinned data! Yikes! I had never seen that before!
## Performance Toolkit Device Details

**FCX110**  **CPU 2003**  **GDLVM7**  **Interval INITIAL. - 13:08:47**  **Remote Data**

**Detailed Analysis for Device 1742 ( SYSTEM )**

- **Device type**: 3390-2  
  - Function pend.: 0.8ms  
  - Device busy: 27%
- **VOLSER**: USE001  
  - Disconnected: 20.3ms  
  - I/O contention: 0%
- **Nr. of LINKs**: 404  
  - Connected: 5.4ms  
  - Reserved: 0%
- **Last SEEK**: 1726  
  - Service time: 26.5ms  
  - SENSE SSCH: ...
- **SSCH rate/s**: 10.5  
  - Response time: 26.5ms  
  - Recovery SSCH: ...
- **Avoided/s**:  
  - CU queue time: 0.0ms  
  - Throttle del/s: ...

**Status**: SHARABLE

- **Path(s) to device 1742**: 0A 2A 4A
- **Channel path status**: ON ON ON ON

<table>
<thead>
<tr>
<th>Device</th>
<th>Overall CU-Cache Performance</th>
<th>Split</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIR ADDR VOLSER</td>
<td>IO/S %READ %RDHIT %WRHIT ICL/S BYP/S</td>
<td>IO/S %READ %RDHIT 'NORMAL' I/O only</td>
</tr>
<tr>
<td>08 1742 USE001</td>
<td>0 0 0 0 0 0 0</td>
<td></td>
</tr>
</tbody>
</table>
### Performance Toolkit Device Details

<table>
<thead>
<tr>
<th>MDISK Extent</th>
<th>Userid</th>
<th>Addr</th>
<th>Status</th>
<th>LINK</th>
<th>MDIO/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>101 - 200</td>
<td>EDLSFS</td>
<td>0310</td>
<td>WR</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>201 - 500</td>
<td>EDLSFS</td>
<td>0300</td>
<td>WR</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>501 - 600</td>
<td>EDLSFS</td>
<td>0420</td>
<td>WR</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>601 - 1200</td>
<td>EDLSFS</td>
<td>0486</td>
<td>WR</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>1206 - 1210</td>
<td>RAID</td>
<td>0199</td>
<td>owner</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BRIANKT</td>
<td>0199</td>
<td>RR</td>
<td>5</td>
<td>0.0</td>
</tr>
<tr>
<td>1226 - 1525</td>
<td>DATABASE</td>
<td>0465</td>
<td>owner</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>K007641</td>
<td>03A0</td>
<td>RR</td>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td>1526 - 1625</td>
<td>DATABASE</td>
<td>0269</td>
<td>owner</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BASILEMM</td>
<td>0124</td>
<td>RR</td>
<td>25</td>
<td>0.0</td>
</tr>
<tr>
<td>1626 - 1725</td>
<td>DATABASE</td>
<td>0475</td>
<td>owner</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SUSANF7</td>
<td>0475</td>
<td>RR</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>1726 - 2225</td>
<td>DATABASE</td>
<td>0233</td>
<td>owner</td>
<td>366</td>
<td>10.5</td>
</tr>
</tbody>
</table>
Solution

- Use **Q PINNED** CP command to check for what data is pinned.
- Discussion with Storage Management team.
- Moved data off string until corrected.

Pinned data is **very** rare, but when it happens it is serious.
Case Study: HiperSockets Seem Slow

Revision 2008-09-10 BKW

IBM z/VM Performance Evaluation
Brian Wade bkw@us.ibm.com
Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DBE, e-business logo, ESCO, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, VM/ESA, VSE/ESA, Websphere, xSeries, z/OS, zSeries, zVM

The following are trademarks or registered trademarks of other companies

Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries
LINUX is a registered trademark of Linus Torvalds
UNIX is a registered trademark of The Open Group in the United States and other countries.
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.
Intel is a registered trademark of Intel Corporation
* All other products may be trademarks or registered trademarks of their respective companies.

NOTES:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.
Question from Customer

- **My system seems to have:**
  - Long ping times from my z/VM partition to my z/OS partition
  - Long transaction times from my z/VM partition to my z/OS partition
- **Seems related to my use of HiperSockets to connect z/VM to z/OS**
  - When I use a real OSA to connect the partitions, I don’t have these problems
  - When I drive the z/OS server from external AIX boxes, I don’t have the long transaction times
- **Customer sent MONWRITE data**
- **I took a look**
CEC and LPAR Configuration, from MONWRITE Data

- 2094 with:
  - 6 CPs
  - 4 ICFs
  - 12 IFLs
  - 4 zIIPs
- Several z/VM partitions, all shared IFL 12-ways, but only one of these partitions is active
- Several z/OS partitions, all shared, with varying logical PU configurations, that use the CPs and zIIPs
- Two coupling partitions that use the ICFs, dedicated
Workload Configuration, from Customer

- Linux on z/VM is the origin
- Over to z/OS database server via real HiperSocket
- Back to Linux on z/VM via real HiperSocket
- Transaction ends
Long Response Time? Let’s Hunt z/VM Constraints

- FCX126 LPAR and FCX225 SYSSUMLG – no IFL constraints found
- FCX225 SYSSUMLG and FCX109 INTERIM DEVICE CPOWNED – no paging found – everything fits in central
- FCX108 DEVICE DASD revealed the active user volumes, and FCX168 DEVLOG showed good service time and small to no queues
- FCX215 INTERIM FCHANNEL – no FICON adapter CPU problems; FCX108 DEVICE DASD – no pending time concerns
- FCX112 USER revealed the big CPU users, and FCX162 USERLOG showed very low T/V and no CPU peaks
- FCX231 INTERIM HIPSOCK showed <10 msgs/sec and 600 data units/message – seems small
- I couldn’t find a z/VM constraint
### FCX126 LPAR: General View of CPU Busy (z/VM)

<table>
<thead>
<tr>
<th>Partition Nr.</th>
<th>Upid</th>
<th>#Proc</th>
<th>Weight</th>
<th>Wait-C</th>
<th>Cap</th>
<th>%Load</th>
<th>CPU %Busy</th>
<th>%Ovhd</th>
<th>%Susp</th>
<th>%VMd</th>
<th>%Logld</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>XXX1</td>
<td>11</td>
<td>14</td>
<td>12</td>
<td>200</td>
<td>NO</td>
<td>NO</td>
<td>...</td>
<td>0</td>
<td>20.5</td>
<td>1</td>
<td>3</td>
<td>20.3</td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>1</td>
<td>21.9</td>
<td>.2</td>
<td>3</td>
<td>21.7</td>
<td></td>
<td>21.7</td>
<td>21.7</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>2</td>
<td>21.5</td>
<td>.2</td>
<td>3</td>
<td>21.3</td>
<td></td>
<td>21.3</td>
<td>21.3</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>3</td>
<td>22.0</td>
<td>.2</td>
<td>3</td>
<td>21.7</td>
<td></td>
<td>21.8</td>
<td>21.8</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>4</td>
<td>22.1</td>
<td>.2</td>
<td>3</td>
<td>21.9</td>
<td></td>
<td>21.9</td>
<td>21.9</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>5</td>
<td>22.2</td>
<td>.2</td>
<td>3</td>
<td>22.0</td>
<td></td>
<td>22.0</td>
<td>22.0</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>6</td>
<td>22.0</td>
<td>.2</td>
<td>3</td>
<td>21.8</td>
<td></td>
<td>21.8</td>
<td>21.8</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>7</td>
<td>22.1</td>
<td>.2</td>
<td>3</td>
<td>21.8</td>
<td></td>
<td>21.9</td>
<td>21.9</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>8</td>
<td>21.9</td>
<td>.1</td>
<td>3</td>
<td>21.7</td>
<td></td>
<td>21.8</td>
<td>21.8</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>9</td>
<td>22.2</td>
<td>.2</td>
<td>3</td>
<td>22.0</td>
<td></td>
<td>22.0</td>
<td>22.0</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>10</td>
<td>22.1</td>
<td>.2</td>
<td>3</td>
<td>21.8</td>
<td></td>
<td>21.9</td>
<td>21.9</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>NO</td>
<td>11</td>
<td>21.9</td>
<td>.2</td>
<td>3</td>
<td>21.6</td>
<td></td>
<td>21.7</td>
<td>21.7</td>
<td>IFL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The other partitions using IFLs are not running.

I don't see a problem here.
FCX126 LPAR: CPU Busy, z/OS Partitions

<table>
<thead>
<tr>
<th>Partition Nr.</th>
<th>Upid</th>
<th>#Proc</th>
<th>Weight</th>
<th>Cap</th>
<th>%Load</th>
<th>CPU %Busy</th>
<th>%Ovhd</th>
<th>%Susp</th>
<th>%VMld</th>
<th>%Logld</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>XXXA</td>
<td>8</td>
<td>03</td>
<td>8</td>
<td>190</td>
<td>NO</td>
<td>NO</td>
<td>5.9</td>
<td>NO</td>
<td>NO</td>
<td>28.2</td>
<td>.2</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>1</td>
<td>28.1</td>
<td>NO</td>
<td>NO</td>
<td>28.1</td>
<td>NO</td>
<td>NO</td>
<td>28.0</td>
<td>.2</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>2</td>
<td>28.1</td>
<td>NO</td>
<td>NO</td>
<td>28.1</td>
<td>NO</td>
<td>NO</td>
<td>28.1</td>
<td>.2</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>3</td>
<td>28.0</td>
<td>NO</td>
<td>NO</td>
<td>28.0</td>
<td>NO</td>
<td>NO</td>
<td>28.0</td>
<td>.2</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>6</td>
<td>10.6</td>
<td>NO</td>
<td>NO</td>
<td>10.6</td>
<td>NO</td>
<td>NO</td>
<td>10.6</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>7</td>
<td>10.5</td>
<td>NO</td>
<td>NO</td>
<td>10.5</td>
<td>NO</td>
<td>NO</td>
<td>10.5</td>
<td>.0</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>22</td>
<td>10.5</td>
<td>NO</td>
<td>NO</td>
<td>10.5</td>
<td>NO</td>
<td>NO</td>
<td>10.5</td>
<td>.0</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>23</td>
<td>10.6</td>
<td>NO</td>
<td>NO</td>
<td>10.6</td>
<td>NO</td>
<td>NO</td>
<td>10.6</td>
<td>.1</td>
</tr>
<tr>
<td>XXXB</td>
<td>9</td>
<td>04</td>
<td>8</td>
<td>190</td>
<td>NO</td>
<td>NO</td>
<td>5.8</td>
<td>NO</td>
<td>NO</td>
<td>26.5</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>1</td>
<td>26.5</td>
<td>NO</td>
<td>NO</td>
<td>26.5</td>
<td>NO</td>
<td>NO</td>
<td>26.5</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>2</td>
<td>26.4</td>
<td>NO</td>
<td>NO</td>
<td>26.4</td>
<td>NO</td>
<td>NO</td>
<td>26.4</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>3</td>
<td>26.4</td>
<td>NO</td>
<td>NO</td>
<td>26.4</td>
<td>NO</td>
<td>NO</td>
<td>26.4</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>6</td>
<td>11.5</td>
<td>NO</td>
<td>NO</td>
<td>11.5</td>
<td>NO</td>
<td>NO</td>
<td>11.5</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>7</td>
<td>11.5</td>
<td>NO</td>
<td>NO</td>
<td>11.5</td>
<td>NO</td>
<td>NO</td>
<td>11.5</td>
<td>.0</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>22</td>
<td>11.5</td>
<td>NO</td>
<td>NO</td>
<td>11.5</td>
<td>NO</td>
<td>NO</td>
<td>11.5</td>
<td>.0</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>NO</td>
<td>23</td>
<td>11.5</td>
<td>NO</td>
<td>NO</td>
<td>11.5</td>
<td>NO</td>
<td>NO</td>
<td>11.5</td>
<td>.0</td>
</tr>
<tr>
<td>XXXC</td>
<td>10</td>
<td>05</td>
<td>2</td>
<td>20</td>
<td>NO</td>
<td>NO</td>
<td>0</td>
<td>NO</td>
<td>NO</td>
<td>.0</td>
<td>.0</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>NO</td>
<td>1</td>
<td>.0</td>
<td>NO</td>
<td>NO</td>
<td>.0</td>
<td>NO</td>
<td>NO</td>
<td>.0</td>
<td>.0</td>
</tr>
<tr>
<td>YYYA</td>
<td>13</td>
<td>11</td>
<td>9</td>
<td>600</td>
<td>NO</td>
<td>NO</td>
<td>23.5</td>
<td>NO</td>
<td>NO</td>
<td>70.5</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>NO</td>
<td>1</td>
<td>70.5</td>
<td>NO</td>
<td>NO</td>
<td>70.5</td>
<td>NO</td>
<td>NO</td>
<td>70.5</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>NO</td>
<td>2</td>
<td>70.4</td>
<td>NO</td>
<td>NO</td>
<td>70.4</td>
<td>NO</td>
<td>NO</td>
<td>70.4</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>NO</td>
<td>3</td>
<td>70.2</td>
<td>NO</td>
<td>NO</td>
<td>70.2</td>
<td>NO</td>
<td>NO</td>
<td>70.2</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>NO</td>
<td>4</td>
<td>64.9</td>
<td>NO</td>
<td>NO</td>
<td>64.9</td>
<td>NO</td>
<td>NO</td>
<td>64.9</td>
<td>.0</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>NO</td>
<td>5</td>
<td>64.9</td>
<td>NO</td>
<td>NO</td>
<td>64.9</td>
<td>NO</td>
<td>NO</td>
<td>64.9</td>
<td>.0</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>NO</td>
<td>6</td>
<td>69.3</td>
<td>NO</td>
<td>NO</td>
<td>69.3</td>
<td>NO</td>
<td>NO</td>
<td>69.3</td>
<td>.1</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>NO</td>
<td>22</td>
<td>64.9</td>
<td>NO</td>
<td>NO</td>
<td>64.9</td>
<td>NO</td>
<td>NO</td>
<td>64.9</td>
<td>.0</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>NO</td>
<td>23</td>
<td>65.0</td>
<td>NO</td>
<td>NO</td>
<td>65.0</td>
<td>NO</td>
<td>NO</td>
<td>65.0</td>
<td>.1</td>
</tr>
</tbody>
</table>

None of these partitions look really busy.

But we will see shortly that there is in fact a problem.

Perfkit doesn’t reveal it very well.

But it’s there.
How Do HiperSockets work?

- Synchronous data transfer between partitions
- Two PUs meet (rendezvous) in firmware
- Firmware copies data from one PU’s buffers to the other PU’s buffers
- Firmware runs on the PUs that did the calls
- Works fine in CPU-rich environments
- Not so well in CPU-constrained environments
How Are The z/OS CPUs Doing?

- 6 real standard CPs
- 4 z/OS partitions have 14 logical standard CPs altogether
- Good thing MONWRITE data captures LPAR’s view of the partitions’ consumptions
- For each sample interval, add up those 14 logical standard CPs’ utilizations to see how much of the 6 real CPs they’re using altogether
  - Post-process the FCX126 INTERIM LPAR reports
  - Requires a little Rexx, but I could use the exercise
- While we’re at it, do this for every engine type
CEC View, PU Utilization by Type

D090408, PU Busy by PU Type

Percent PU Pool Busy

Time Stamps
Findings

- Real standard CPs are saturated
- Real HiperSockets require readily accessible cycles to work well
- This is a CPU-constrained environment
- Explains why real OSA was better
- Recommended either:
  - Adding more standard CPs, or
  - Tuning z/OS partitions to reduce CP resource they need
- Also recommended a z/OS expert look at RMF reports to see how else the CP constraint was impacting z/OS function